NOTAS PARA EL CURSO DE INTRODUCCION A LOS PROCESOS ESTOCÁSTICOS. 2o. SEMESTRE DE 2003.

Mario Wschebor
E mail: wschebor@cmat.edu.uy

February 27, 2004

1 MARTINGALAS CON TIEMPO DISCRETO.

1.1 Esperanzas condicionales

Definición.

$(\Omega, \mathfrak{A}, P)$ espacio de probabilidad
$X : \Omega \to \overline{\mathcal{R}}$ variable aleatoria, $X \in L^1$.
$\mathfrak{F}$ sub-σ-algebra de $\mathfrak{A}$.

Se define, para cada $A \in \mathfrak{F}$,

$$\mu(A) = E\left(X.1_A\right)$$

μ es una medida finita signada sobre $\mathfrak{F}$, absolutamente continua con respecto a $P/\mathfrak{F}$. Por la tanto, existe la derivada (de Radon-Nykodim) de μ con respecto a $P/\mathfrak{F}$, que denotamos $\frac{d\mu}{dP}$.

Se define la "esperanza condicional de la variable aleatoria X dada la σ-álgebra $\mathfrak{F}$":

$$E\left(X/\mathfrak{F}\right) = \frac{d\mu}{dP}.$$

es decir que $E\left(X/\mathfrak{F}\right)$ es la única función (módulo la probabilidad P) de $\Omega \to \overline{\mathcal{R}}$ tal que

1. $E\left(X/\mathfrak{F}\right)$ es $\mathfrak{F}$-medible,
2. $$\int_A X \; dP = \int_A E\left(X/\mathfrak{F}\right) \; dP \quad \forall \; A \in \mathfrak{F}$$

Propiedades de las esperanzas condicionales.

1. $X \rightsquigarrow E(X/\mathfrak{F})$ es lineal y monótona.

2. Si $\mathfrak{F} = \{\phi, \Omega\}$, entonces $E(X/\mathfrak{F}) = E(X)$ c.s. [notar que $E[E(X/\mathfrak{F})] = E(X)$].

3. Más en general, si $\mathfrak{F} \subset \mathfrak{G}$ son sub σ-álgebras de $\mathfrak{A}$, se verifica que c.s.

$$E(X/\mathfrak{F}) = E[E(X/\mathfrak{G})/\mathfrak{F}]$$

4. $X_n \geq 0$, $X_n \uparrow X$ $c.s. \Longrightarrow E(X_n/\mathfrak{F}) \uparrow E(X/\mathfrak{F})$ $c.s.$

5. Si X es $\mathfrak{F}$-medible y $XY \in L^1$, entonces

$$E(XY/\mathfrak{F}) = XE(Y/\mathfrak{F}) \quad c.s.$$

   Para probar esto, hacerlo primero cuando $Y \geq 0$, $X = 1_B$, $b \in \mathfrak{F}$. Luego extender.

6. X independiente de $\mathfrak{F}$ (i.e. las variables aleatorias X y 1_A son independientes $\forall\ A \in \mathfrak{F}$) $\Longrightarrow$

$$E(X/\mathfrak{F}) = E(X).$$

7. **(caso finito)**

   Sea $\{A_1, ..., A_m\}$ una partición medible de Ω (es decir que $A_1, ..., A_m \in \mathfrak{A}$, son 2 a 2 disjuntos y $\bigcup_{j=1}^{j=m} A_j = \Omega$). Suponemos que $P(A_j) > 0$, $j = 1, ..., m$ y denotamos $\mathfrak{F} = \sigma(A_1, ..., A_m) =$ mínima σ-álgebra de partes de Ω que contiene a los conjuntos $A_1, ..., A_m =$ familia de todas las posibles uniones de los conjuntos dados $A_1, ..., A_m$.

   Sea $X : \Omega \to \mathcal{R}$ una variable aleatoria que toma un número finito de valores, $\{x_1, ..., x_n\}$.

   Entonces,

$$E(X/\mathfrak{F})(\omega) = \sum_{k=1}^{n} x_k.P(X = x_k/A_j) \quad \forall \omega \in A_j \tag{1}$$

   donde si $P(F) > 0$, $P(E/F) = \frac{P(E \cap F)}{P(F)}$.

Es decir que sobre cada conjunto A_j, $E\left(X/\mathfrak{F}\right)$ es constante y asume el valor indicado; es claro que esa función es $.\mathfrak{F}$-medible. De modo que para probar (1) hay que probar que

$$\int_A X\ dP = \int_A \sum_{j=1}^{m} 1_{A_j} \sum_{k=1}^{n} x_k.P\left(X = x_k/A_j\right)\ dP \quad \forall\ A \in \mathfrak{F}.$$

Basta hacerlo para $A = A_j$. Es inmediato.

En particular, si Y es una nueva variable aleatoria que toma un número finito de valores (2 a 2 $\neq$) $y_1, ..., y_m$ y $A_j = \{Y = y_j\}$, usaremos la notación

$$E(X/Y = y_j) = \sum_{k=1}^{n} x_k.P\left(X = x_k/Y = y_j\right)$$

donde el primer miembro se debe interpretar como $E\left(X/\mathfrak{F}\right)(\omega)$ con $Y(\omega) = y_j$.

Observaciones.

- Si en lugar de una partición *finita* consideramos una partición *numerable* de Ω, todo lo anterior vale sin cambios significativos.
- Heurísticamente, el significado de $E(X/Y = y_j)$ es el del valor esperado de la variable aleatoria X, "cuando sabemos ocurrió el evento de que la variable aleatoria Y toma el valor y_j", es decir, el valor esperado de una variable aleatoria que toma los mismos valores posibles que X, o sea $x_1, ..., x_n$, pero no con las probabilidades originales (que son $P\left(X = x_k\right), k = 1, ..., n$) sino con las probabilidades condicionales $P\left(X = x_k/Y = y_j\right)$. Claro que esto es posible si $P(Y = y_j) > 0$, aplicando la definición elemental de probabilidad condicional.

  En cambio, si la variable aleatoria Y no es discreta (es decir, si la distribución de Y no es puramente atómica) tenemos dificultades, porque el evento $\{Y = y_j\}$ que figura como condición, puede tener probabilidad nula. Podríamos tratar de aproximar ese evento por eventos "próximos" con probabilidad pequeña, escribir la probabilidad condicional como cociente y luego pasar al límite; ésta es la idea de derivar la probabilidad, aunque la forma rigurosa de hacerlo en el caso general es mediante el teorema de Radon-Nykodim. Éste es el motivo de la definición formal de esperanza condicional que dimos al principio.

8. **(pareja de variables aleatorias con densidad conjunta)**

Consideremos como espacion de probabilidad el espacio producto $E_1 \times E_2$, donde $E_1 \simeq \mathcal{R}$, $E_2 \simeq \mathcal{R}$, y una medida P sobre la σ-álgebra de Borel, absolutamente continua con respecto a la medida de Borel, con densidad $f(x, y)$.

Denotamos mediante X, Y las variables aleatorias coordenadas (es decir, $X(x, y) = x, Y(x, y) = y$), que son variables aleatorias a valores reales, cuyas distribuciones de probabilidad tienen densidades respectivas

$$p_X(x) = \int_{-\infty}^{+\infty} f(x, y)dy, \quad p_Y(y) = \int_{-\infty}^{+\infty} f(x, y)dx$$

(las "densidades marginales").

Sea $\mathfrak{F}$ la sub σ-álgebra de la σ-álgebra de Borel de $E_1 \times E_2$ engendrada por todos los conjuntos de la forma $E_1 \times B$, donde B es un Boreliano en E_2. Observar que una función definida en $E_1 \times E_2$, medible con respecto a $\mathfrak{F}$, es sólo función de la segunda coordenada y. Eso ocurre, por lo tanto con $E\left(X/\mathfrak{F}\right)$ y tenemos derecho a escribir

$$\varphi(y) = E\left(X/\mathfrak{F}\right)(x, y).$$

Entonces, para cada pareja de reales y_1, y_2, $y_1 < y_2$:

$$E\left(X 1_{[y_1,y_2]}(Y)\right) = E\left[E\left(X/\mathfrak{F}\right) 1_{[y_1,y_2]}(Y)\right] \tag{2}$$

usando 2. y 4. ut supra (observar que la variable aleatoria $1_{[y_1,y_2]}(Y)$ es $\mathfrak{F}$-medible).

Reemplazando en ambos miembros de (2), se tiene la igualdad:

$$\int_{-\infty}^{+\infty} dx \int_{y_1}^{y_2} x.f(x, y)dy = \int_{-\infty}^{+\infty} dx \int_{y_1}^{y_2} \varphi(y).f(x, y)dy$$

que implica (Teorema de Fubini mediante):

$$\varphi(y) = \frac{1}{p_Y(y)} \int_{-\infty}^{+\infty} x.f(x, y)dx = \int_{-\infty}^{+\infty} x.f(x/y)dx \quad \text{con} \quad f(x/y) = \frac{f(x, y)}{p_Y(y)}$$

donde la igualdad vale módulo P_Y, la distribución de probabilidad de Y.

$f(x/y)$ es la "densidad condicional de X dado que $Y = y$".

9. **(desigualdad de Jensen para esperanzas condicionales)**

Sean $\Phi : \mathcal{R}^1 \to \mathcal{R}^1$ convexa, X una variable aleatoria en L^1 y tal que $Y = \Phi \circ X \in L^1$.

Entonces

$$\Phi\left[E\left(X/\mathfrak{F}\right)\right] \le E\left(\Phi\left(X\right)/\mathfrak{F}\right) \quad c.s. \tag{3}$$

Para probar (3), como Φ es convexa, para cada $x \in \mathcal{R}$ podemos encontrar $\lambda(x)$ (coeficiente angular de una recta de apoyo al gráfico de Φ en x), de modo que

$$\Phi(y) \ge \Phi(x) + \lambda(x)(y - x) \quad \forall\, y \in \mathcal{R} \tag{4}$$

Más aún, podemos elegir la función λ de modo que sea Borel-medible, ya que una elección posible es

$$\lambda(x) = \lim_{n\to+\infty} \frac{\Phi(x + \frac{1}{n}) - \Phi(x)}{\frac{1}{n}}$$

y para cada n, la función

$$x \rightsquigarrow \frac{\Phi(x + \frac{1}{n}) - \Phi(x)}{\frac{1}{n}}$$

es continua y, por lo tanto, Borel medible. (Recordar que el límite puntual de funciones medibles es medible).

Reemplazando en (4) y por $X(\omega)$ y x por $E\left(X/\mathfrak{F}\right)(\omega)$ se obtiene:

$$\Phi(X) \ge \Phi\left(E\left(X/\mathfrak{F}\right)\right) + \lambda\left(E\left(X/\mathfrak{F}\right)\right)\left[X - E\left(X/\mathfrak{F}\right)\right]$$

(donde, como es habitual, hemos suprimido "ω" en ambos miembros). Ahora tomamos esperanza condicional dada $\mathfrak{F}$ en ambos miembros y usamos que la variable aleatoria

$$\omega \rightsquigarrow \lambda\left(E\left(X/\mathfrak{F}\right)(\omega)\right)$$

es $\mathfrak{F}$-medible (composición de una función $\mathfrak{F}$-medible con una función Borel-medible), conjuntamente con la propiedad 4. ut-supra. Esto prueba (3).$\square$

10. $X_n \to X$ en $L^1 \implies E\left(X_n/\mathfrak{F}\right) \to E\left(X/\mathfrak{F}\right)$ en L^1.

Aplicando la desigualdad de Jensen:

$$E\left[\left|E\left(X_n/\mathfrak{F}\right) - E\left(X/\mathfrak{F}\right)\right|\right] \le E\left[E\left(\left|X_n - X\right|/\mathfrak{F}\right)\right] = E\left(\left|X_n - X\right|\right) \to 0.$$

11. **(La esperanza condicional como proyección ortogonal en L^2)**

Sea S el subespacio de $L^2 = L^2(\Omega, \mathfrak{A}, P)$ de las variables aleatorias a valores reales que son $\mathfrak{F}$-medibles. Es obvio que S es cerrado.

Sea π_S la proyección ortogonal de L^2 sobre S.

Entonces, si $X \in L^2$ se tiene

$$\pi_S(X) = E\left(X/\mathfrak{F}\right) \tag{5}$$

Observar que si $X \in L^2$, entonces también $X \in L^1$, de modo que la esperanza condicional está bien definida. Notar además que $E\left(X/\mathfrak{F}\right) \in L^2$ ya que:

$$E\left(\left[E\left(X/\mathfrak{F}\right)\right]^2\right) \leq E\left(E\left(X^2/\mathfrak{F}\right)\right) = E\left(X^2\right) < \infty$$

(5) se sigue de que

$$E\left[\left(X - E\left(X/\mathfrak{F}\right)\right) Y\right] = 0$$

para toda variable aleatoria $Y \in S$.

1.2 Filtraciones. Definición de martingala discreta.

Definición. $(\Omega, \mathfrak{A}, P)$ espacio de probabilidad.

En lo que sigue, una "filtración" $\{\mathfrak{F}_n\}_{n \in T}$ es una sucesión de sub σ-álgebras de $\mathfrak{A}$, creciente, i.e. $\mathfrak{F}_n \subset \mathfrak{F}_{n+1} \forall n$. Por un largo trecho, supondremos que el conjunto T es el de los naturales o el de los enteros; más adelante, tendremos que modificar una serie de puntos para encarar los problemas cuando el parámetro varía en los reales, o en algún otro conjunto de índices.

Usamos las notaciones

$$\mathfrak{F}_\infty = \bigvee_n \mathfrak{F}_n$$

y

$$\mathfrak{F}_{-\infty} = \bigcap_n \mathfrak{F}_n$$

para representar, respectivamente, la σ-álgebra engendrada por la unión de las $\mathfrak{F}_n$'s y la σ-álgebra intersección de las $\mathfrak{F}_n$'s.

Definiciónes.

Un proceso estocástico definido en el mismo espacio de probabilidad (o lo que es lo mismo en nuestro caso, una sucesión de variables aleatorias)

$\{X_n\}_{n\in T}$ es **adaptado** a la filtración $\{\mathfrak{F}_n\}_{n\in T}$ si X_n es $\mathfrak{F}_n$-medible para todo $n \in T$.

El proceso estocástico $\{X_n\}_{n\in T}$ es **previsible** con respecto a la filtración $\{\mathfrak{F}_n\}_{n\in T}$ si X_n es $\mathfrak{F}_{n-1}$-medible para todo $n \in T$.

Un proceso estocástico $\{X_n\}_{n\in T}$ es una **submartingala** (respectivamente **supermartingala**, **martingala**) con respecto a la filtración $\{\mathfrak{F}_n\}_{n\in T}$ si cumple las siguientes condiciones:

1. $\{X_n\}_{n\in T}$ es adaptado a $\{\mathfrak{F}_n\}_{n\in T}$.
2. $X_n \in L^1 \quad \forall n \in T$.
3. casi seguramente $E(X_{n+1}/\mathfrak{F}_n) \geq X_n$ (respectivamente $\leq$, $=$)

Observaciones.

- En la definición la condición 2. puede reemplazarse por $X_n^+ \in L^1$ (respectivamente $X_n^- \in L^1$, $X_n \in L^1$) $\forall n \in T$.

- Un caso especial importante, es aquél en que el proceso estocástico $\{X_n\}_{n\in T}$ está dado y se define la filtración mediante $\mathfrak{G}_n = \sigma(X_m : m \leq n)$, la mínima σ-álgebra con respecto a la cual las variables aleatorias $\{X_m : m \leq n\}$ son medibles. Es decir que $\mathfrak{G}_n$ contiene la información generada por el proceso hasta el instante n. La filtración $\{\mathfrak{G}_n\}_{n\in T}$ así definida, se denomina la "filtración engendrada por el proceso dado".

1.2.1 Propiedades iniciales y ejemplos.

1. **(Paseo al azar)**

   Sea $\xi_1, \xi_2, ...$ una sucesión de variables aleatorias independientes, $\xi_n \in L^1, E(\xi_n) = 0 \ \forall n$.

   Definimos
   $$X_0 = 0, \ X_n = \xi_1 + ... + \xi_n \quad \text{para} \quad n \geq 1$$
   $$\mathfrak{F}_0 = \{\phi, \Omega\}, \quad \mathfrak{F}_n = \sigma(\xi_m : m \leq n) \quad \text{para} \quad n \geq 1$$

   Entonces, $\{X_n\}$ es una $\mathfrak{F}_n$-martingala.

2. Si en el ejemplo anterior se reemplaza $E(\xi_n) = 0$ por $E(\xi_n) \geq 0$, se obtiene una submartingala.

3. Si $\{X_n\}, \{Y_n\}$ son $\mathfrak{F}_n$-martingalas y a,b números reales, entonces $\{aX_n + bY_n\}$ es también $\mathfrak{F}_n$-martingala. Del mismo modo, si $\{X_n\}, \{Y_n\}$ son $\mathfrak{F}_n$-submartingalas (respectivamente supermartingalas) y a,b números reales no negativos, entonces $\{aX_n + bY_n\}$ es también $\mathfrak{F}_n$-submartingala (resp. supermartingala).

4. Sea $\{X_n\}$ una $\mathfrak{F}_n$-martingala y $\Phi : \mathcal{R} \to \mathcal{R}$ una función convexa. Se define $Y_n = \Phi(X_n)$.

   Entonces, si $Y_n \in L^1 \ \forall n$, resulta que $\{Y_n\}$ es una $\mathfrak{F}_n$-submartingala.

   En particular, si $X_n \in L^p \ \forall n$, $p \geq 1$, entonces $\{|X_n|^p\}$ es una $\mathfrak{F}_n$-submartingala.

5. Sea $\{X_n\}$ una $\mathfrak{F}_n$-*sub*martingala y $\Phi : \mathcal{R} \to \mathcal{R}$ una función convexa *y creciente*. Se define $Y_n = \Phi(X_n)$. Entonces, si $Y_n \in L^1 \ \forall n$, resulta que $\{Y_n\}$ es una $\mathfrak{F}_n$-submartingala.

   En particular, $\{X_n^+\}$ es una $\mathfrak{F}_n$-submartingala.

6. Sea $Y \in L^1$ y $\{\mathfrak{F}_n\}_{n\in T}$ una filtración.

   Definimos
$$X_n = E(Y/\mathfrak{F}_n) \quad (n \in T) \tag{6}$$

   Se prueba fácilmente que $\{X_n\}_{n\in T}$ es una $\{\mathfrak{F}_n\}_{n\in T}$-martingala.

   Más adelante consideraremos la pregunta inversa, a saber ¿en qué condiciones, dada una $\{\mathfrak{F}_n\}_{n\in T}$-martingala $\{X_n\}_{n\in T}$, se puede encontrar una variable aleatoria Y tal que (6) se cumpla para todo n? Y más aún, ¿cómo se puede describir Y?

7. **(Derivación en un espacio abstracto)**

   - $(X, \mathfrak{F}, \mu)$ espacio de probabilidad.
   - $\pi_n = \{A_1, ..., A_{m_n}\}$ partición medible finita de X, creciente con n en el sentido de que π_{n+1} es un refinamiento de π_n.
   - $\mathfrak{F}_n = \sigma(\pi_n) \subset \mathfrak{F}_{n+1}$
   - Sea ν otra medida signada finita sobre $(X, \mathfrak{F}, \mu)$, $\nu \ll \mu$ y μ_n, ν_n las restricciones de μ, ν respectivamente a $\mathfrak{F}_n$.

     Es claro que
$$L_n = \frac{d\nu_n}{d\mu_n} = \sum_{j=1}^{m_n} \frac{\nu(A_j)}{\mu(A_j)} 1_{A_j}$$

con la convención de que $\frac{\nu(A_j)}{\mu(A_j)} = 0$ is $\mu(A_j) = 0$.

Veamos que $\{L_n\}$ es una $\{\mathfrak{F}_n\}$-martingala.

Es claro que L_n es $\mathfrak{F}_n$-medible, ya que las funciones $\mathfrak{F}_n$-medibles son justamente las que son constantes sobre los conjuntos de la partición que define a $\mathfrak{F}_n$.

Hay que ver que $E(L_{n+1}/\mathfrak{F}_n) = L_n$ c.s., es decir, que

$$E(L_{n+1}1_A) = E(L_n 1_A) \quad \forall\ A \in \mathfrak{F}_n \tag{7}$$

Basta verificar (7) para $A = A_j$ $(j = 1, ..., m_n)$, ya que cada $A \in \mathfrak{F}_n$ es una unión finita disjunta de esos A_j.

Pero si $A_j \in \mathfrak{F}_n$, entonces, dado que $\in \mathfrak{F}_{n+1}$ es un refinamiento de $\in \mathfrak{F}_n$, se puede escribir

$$A_j = \bigcup_{k=1}^{k=\nu_j} A'_k \quad \text{con} \quad A'_k \in \mathfrak{F}_{n+1},\ \ 2 \text{ a } 2 \text{ disjuntos}$$

con lo cual

$$E\left(L_{n+1}1_{A_j}\right) = \sum_{k=1}^{k=\nu_j} E\left(\frac{\nu(A'_k)}{\mu(A'_k)}1_{A'_k}\right) = \sum_{k=1}^{k=\nu_j} \nu(A'_k) = \nu(A_j) = E\left(L_n 1_{A_j}\right).$$

Una pregunta natural es bajo qué condiciones L_n converge cuando $n \to \infty$ y en ese caso, si el límite es la derivada de Radon-Nydodim $\frac{d\nu}{d\mu}$, dando una interpretación de ésta similar a la derivación ordinaria en un espacio euclidiano.

1.3 Compensador de una sucesión adaptada de variables aleatorias. Descomposición de Doob elemental (caso discreto).

Sea $(\Omega, \mathfrak{A}, P)$ un espacio de probabilidad, $\{\mathfrak{F}_n\}_{n=0,1,2,...}$ una filtración definida en él y $\{X_n\}_{n=0,1,2,...}$ una sucesión de variables aleatorias adaptada a $\{\mathfrak{F}_n\}_{n=0,1,2,...}$, $X_n \in L^1\ \forall n$.

Proposición. $\exists$ una única sucesión de variables aleatorias $\left\{\widetilde{X}_n\right\}_{n=0,1,2,...}$ tal que:

1. $\widetilde{X}_0 = 0$.
2. $\left\{\widetilde{X}_n\right\}_{n=0,1,2,...}$ es previsible

3. $\left\{X_n - \widetilde{X}_n\right\}_{n=0,1,2,\ldots}$ es una $\{\mathfrak{F}_n\}$-martingala.

Demostración. Definimos:

$$\begin{aligned}\widetilde{X}_0 &= 0\\ \widetilde{X}_n &= \sum_{j=1}^{n} E\left(X_j - X_{j-1}/\mathfrak{F}_{j-1}\right) \quad n \geq 1\end{aligned}$$

1. y 2. son obvias. En cuanto a 3., si $n \geq 1$:

$$\begin{aligned}& E\left[\left(X_{n+1} - \widetilde{X}_{n+1}\right) - \left(X_n - \widetilde{X}_n\right)/\mathfrak{F}_n\right]\\ = \;& E\left[X_{n+1} - X_n/\mathfrak{F}_n\right] - \sum_{j=1}^{n+1} E\left(X_j - X_{j-1}/\mathfrak{F}_{j-1}\right) + \sum_{j=1}^{n} E\left(X_j - X_{j-1}/\mathfrak{F}_{j-1}\right) = 0\end{aligned}$$

Si $n = 0$, se verifica de manera igualmente trivial.

En cuanto a la unicidad, supongamos que $\left\{\widetilde{X}_n\right\}$ y también $\left\{\widehat{X}_n\right\}$ verifican 1.,2.,3.

Entonces,

$$\begin{aligned}E\left(X_1 - \widetilde{X}_1/\mathfrak{F}_0\right) &= X_0 - \widetilde{X}_0 = X_0 \quad c.s.\\ E\left(X_1 - \widehat{X}_1/\mathfrak{F}_0\right) &= X_0 - \widehat{X}_0 = X_0 \quad c.s.\end{aligned}$$

por lo que

$$\widehat{X}_1 - \widetilde{X}_1 = E\left(\widehat{X}_1 - \widetilde{X}_1/\mathfrak{F}_0\right) = 0 \quad c.s.$$

ya que $\widehat{X}_1 - \widetilde{X}_1$ es $\mathfrak{F}_0$-medible. El mismo tipo de argumento permite probar por inducción completa que $\widehat{X}_n - \widetilde{X}_n = 0$ c.s. $\forall n.\square$

Definición. $\left\{\widetilde{X}_n\right\}_{n=0,1,2,\ldots}$ se denomina el *compensador (previsible)* de la sucesión $\{X_n\}_{n=0,1,2,\ldots}$.

Notas y ejemplo básico.

- $\widetilde{X}_n = 0 \; \forall n$ si y sólo si el proceso dado $\{X_n\}$ es una $\{\mathfrak{F}_n\}$-martingala.
- $\left\{\widetilde{X}_n\right\}$ es creciente (en sentido amplio) si y sólo si el proceso dado $\{X_n\}$ es una $\{\mathfrak{F}_n\}$-submartingala.

- Sea $\{X_n\}_{n=0,1,2,...}$ una $\{\mathfrak{F}_n\}_{n=0,1,2,...}$-martingala, y supongamos además que $X_n \in L^2\ \forall n$. Entonces $Y_n = X_n^2$ $(n = 0, 1, 2, ...)$ es una submartingala y el compensador es, para $n \geq 1$:

$$\begin{aligned}
\widetilde{Y}_n &= \sum_{j=1}^{n} E\left(Y_j - Y_{j-1}/\mathfrak{F}_{j-1}\right) = \sum_{j=1}^{n} E\left(X_j^2 - X_{j-1}^2/\mathfrak{F}_{j-1}\right) \\
&= \sum_{j=1}^{n} E\left(\left(X_j - X_{j-1}\right)^2/\mathfrak{F}_{j-1}\right).
\end{aligned}$$

  De modo que $\left\{X_n^2 - \widetilde{Y}_n\right\}$ es una martingala.

- Un caso particular de lo anterior es el siguiente.

  Sea $X_0 = 0$, $X_n = \xi_1 + ... + \xi_n$ $(n \geq 1)$, donde $\xi_1, \xi_2, ...$ es una sucesión de variables aleatorias independientes, $E(\xi_n) = 0$, $Var(\xi_n) = \sigma_n^2$.

  Entonces, el compensandor de X_n^2 es $\sum_{j=1}^{n} \sigma_j^2$, es decir que $\left\{X_n^2 - \sum_{j=1}^{n} \sigma_j^2\right\}$ es una martingala con respecto a la filtración engendrada por $\{\xi_n\}$. Obsérvese que en este caso el compensador es determinístico.

La situación es ligeramente distinta si consideramos un proceso indizado en $\mathcal{Z}$ en lugar de $\mathcal{N}$. Vemos la modificación con respecto a la proposición anterior en la proposición siguiente:

Proposición.

Sea $(\Omega, \mathfrak{A}, P)$ un espacio de probabilidad, $\{\mathfrak{F}_n\}_{n\in\mathcal{Z}}$ una filtración definida en él y $\mathcal{X} = \{X_n\}_{n\in\mathcal{Z}}$ una submartingala adaptada a $\{\mathfrak{F}_n\}_{n\in\mathcal{Z}}$.

Suponemos además que

$$\sup_{n<0} E\left(X_n^-\right) < \infty.$$

Entonces, $\exists$ una martingala $\mathcal{M} = \{M_n\}_{n\in\mathcal{Z}}$ y un proceso $\mathcal{A} = \{A_n\}_{n\in\mathcal{Z}}$ previsible, integrable y creciente, tales que

$$\mathcal{X} = \mathcal{M} + \mathcal{A} \tag{8}$$

Si además $\lim\limits_{n\to-\infty} A_n = 0$, entonces la descomposición (8) es única.

Si además es $X_n \geq 0\ \forall n \implies E(A_n) \leq E(X_n)\ \forall n$.

Demostración.

Ponemos

$$Y_j = E\left(X_j - X_{j-1}/\mathfrak{F}_{j-1}\right) \geq 0$$

Para $n < m$:

$$E\left(\sum_{n<j\leq m} Y_j\right) = E\left(X_m - X_n\right) \leq E\left(X_m\right) + \sup_{n\leq m} E\left(X_n^-\right)$$

Por lo tanto (Beppo-Levi),

$$\exists \lim_{n\downarrow -\infty} \sum_{n<j\leq m} Y_j = A_m \in L^1$$

y $\mathcal{A} = \{A_n\}_{n\in\mathcal{Z}}$ es previsible, integrable y creciente. Poniendo $M_n = X_n - A_n$, se verifica que $\mathcal{M}$ es una martingala.

Si hubiera dos descomposiciones, $\mathcal{X} = \mathcal{M}^{(i)} + \mathcal{A}^{(i)}$ $(i = 1, 2) \Longrightarrow \mathcal{A}^{(2)} - \mathcal{A}^{(1)} = \mathcal{M}^{(1)} - \mathcal{M}^{(2)}$ es a la vez una martingala y es previsible $\Longrightarrow$

$$A_{n+1}^{(2)} - A_{n+1}^{(1)} = E\left(A_{n+1}^{(2)} - A_{n+1}^{(1)}/\mathfrak{F}_n\right) = A_n^{(2)} - A_n^{(1)}$$

donde la primera igualdad es por ser previsible y la segunda por ser martingala. Es decir que $A_n^{(2)} - A_n^{(1)}$ no depende de n y como tiende a cero en $-\infty$, es idénticamente nula.

Finalmente, si $X_n \geq 0 \implies X_n^- = 0$ y lo indicado resulta de la acotación de $E\left(A_m\right)$.$\square$

1.4 Tiempos de parada.

Definiciones.

Como antes, sea $\{\mathfrak{F}_n\}_{n\in T}$ una filtración ($T = \mathcal{N}$ ó $T = \mathcal{Z}$).

Una variable aleatoria $\tau : \Omega \to T \cup \{\infty\}$ es un *tiempo de parada* - o un $\{\mathfrak{F}_n\}$-tiempo de parada) si

$$\{\tau = n\} \in \mathfrak{F}_n \quad \forall n$$

(esto incluye el caso $n = \infty$; hay una cierta imprecisión aquí en el caso $T = \mathcal{Z}$, en el que se debe agregar $+\infty$ y también -∞, pero esto no será una fuente de problemas).

También se define la "σ-álgebra de los sucesos anteriores a τ":

$$\mathfrak{F}_\tau = \{A \in \mathfrak{F}_\infty : \forall n < \infty \text{ se cumple que } A \cap \{\tau = n\}\} \in \mathfrak{F}_n$$

Ejemplo básico (tiempo que demora un proceso estocástico en llegar a un conjunto).

- $\{X_n\}_{n=0,1,2,...}$ sucesión de variables aleatorias a valores reales.
- $\mathfrak{F}_n = \sigma\{X_m : m \leq n\}$ filtración engendrada por la sucesión.
- $a \in \mathcal{R}$

Se define el "tiempo que demora el proceso en llegar al conjunto $[a, +\infty)$" mediante

$$\begin{aligned} \tau_a &= \inf\{n : X_n \geq a\} \quad si\ \{\} \neq \phi \\ \tau_a &= +\infty \quad si\ \{\} = \phi \end{aligned}$$

1.4.1 Propiedades de los tiempos de parada.

(a) Una variable aleatoria τ a valores en $T \cup \{\infty\}$ es un tiempo de parada si y sólo si $\{\tau \leq n\} \in \mathfrak{F}_n\ \forall n$.

(b) $\mathfrak{F}_\tau$ es una σ-álgebra.

(c) Generalización del ejemplo anterior.

- $\{X_n\}_{n=0,1,2,...}$ sucesión de variables aleatorias a valores en un espacio medible $(E, \mathfrak{E})$.
- $A \in \mathfrak{E}$
-

$$\begin{aligned} \tau_A &= \inf\{n : X_n \in A\} \quad si\ \{\} \neq \phi \\ \tau_A &= +\infty \quad si\ \{\} = \phi \end{aligned}$$

(d) $\tau \equiv n$ es tiempo de parada

(e) Si τ_1, τ_2 son $\{\mathfrak{F}_n\}$-tiempos de parada, entonces $\tau_1 \vee \tau_2$ y $\tau_1 \wedge \tau_2$ también son $\{\mathfrak{F}_n\}$-tiempos de parada.

(f) Si τ_1, τ_2 son $\{\mathfrak{F}_n\}$-tiempos de parada, $\tau_1 \leq \tau_2$, entonces $\mathfrak{F}_{\tau_1} \subset \mathfrak{F}_{\tau_2}$.

En efecto, si $A \in \mathfrak{F}_{\tau_1} \Longrightarrow A \in \mathfrak{F}_\infty$, $A \cap \{\tau_1 = n\}$. Por lo tanto:

$$A \cap \{\tau_2 = n\} = \bigcup_{m \leq n} [A \cap \{\tau_2 = n\} \cap \{\tau_1 = m\}] \in \mathfrak{F}_n$$

ya que en la unión, $A \cap \{\tau_1 = m\} \in \mathfrak{F}_m \subset \mathfrak{F}_n$.

(g) **(Proceso estocástico parado)**

Sean $\mathcal{X} = \{X_n\}_{n\in T}$ un proceso $\{\mathfrak{F}_n\}_{n\in T}$-adaptado y τ un tiempo de parada con respecto a esta filtración. Se define el "proceso estocástico parado en τ":$\mathcal{X}^\tau = \{X_n^\tau\}_{n\in T}$ mediante:

$$X_n^\tau(\omega) = X_{n\wedge\tau(\omega)}(\omega)$$

Proposición.

Si $\{X_n\}_{n=0,1,2,..}$ es $\{\mathfrak{F}_n\}_{n=0,1,2,...}$-martingala (resp. supermartingala, submartingala) entonces $\mathcal{X}^\tau$ es $\{\mathfrak{F}_n\}_{n=0,1,2,...}$ martingala (resp. supermartingala, submartingala).

Demostración.

Probamos, lo que es más general, que si $\{\xi_n\}_{n=1,2,...}$ es un proceso estocástico predecible, acotado y ≥ 0, la misma conclusión del enunciado vale para el proceso $\{Y_n\}_{n=0,1,2,..}$ definido mediante:

$$\begin{aligned} Y_0 &= X_0 \\ Y_n &= Y_{n-1} + \xi_n (X_n - X_{n-1}) \quad \text{para } \ n \geq 1 \end{aligned}$$

La proposición resultará de aplicar esto a $\xi_n = 1_{\{\tau\geq n\}}$, porque en ese caso se verifica fácilmente que $Y_n = X_n^\tau$.

Para el caso martingala, tenemos que ver que para $n \geq 1$:

$$E(Y_n/\mathfrak{F}_{n-1}) = Y_{n-1} \quad c.s. \tag{9}$$

Para los otros casos, es similar. Pero:

$$E[\xi_n (X_n - X_{n-1})/\mathfrak{F}_{n-1}] = \xi_n E[(X_n - X_{n-1})/\mathfrak{F}_{n-1}] = 0 \quad c.s.$$

lo que prueba (9).□

1.5 Desigualdades para martingalas, convergencia, leyes de grandes números.

Teorema (del muestreo opcional de Doob).

Sea $\mathcal{X} = \{X_n\}_{n\in 0,1,2,...}$ una $\{\mathfrak{F}_n\}_{n\in 0,1,2,..}$-martingala (resp. supermartingala, submartingala).

Consideramos una sucesión (finita) creciente de tiempos de parada $\tau_1 \leq \tau_2 \leq ... \leq \tau_n$, que suponemos uniformemente acotados superiormente, en el sentido de que existe una constante C tal que $\tau_k \leq C \ \ \forall k = 1, ..., n$.

Entonces, $\{X_{\tau_k}\}_{k=1,...,n}$ es una $\{\mathfrak{F}_{\tau_k}\}_{k=1,...,n}$-martingala (resp. supermartingala, submartingala).

Demostración.

Basta probar el enunciado para $n = 2$.

Empecemos por probar que

$$E\left(X_{\tau_2}\right) = E\left(X_{\tau_1}\right) \quad (\text{resp.} \ \leq, \geq) \tag{10}$$

Aplicamos lo visto en la demostración de la proposición anterior, con la misma notación, poniendo $\xi_n = 1_{\{\tau_1 < n \leq \tau_2\}} = 1_{\{n \leq \tau_2\}} - 1_{\{n \leq \tau_1\}}$ que es predecible.

Entonces, $\{Y_n\}_{n=0,1,2,\dots}$ es una $\{\mathfrak{F}_n\}_{n=0,1,2,\dots}$-martingala (resp. supermartingala, submartingala), con (¡verificar!):

$$Y_n = Y_0 + \left(X_{n\wedge\tau_2} - X_{n\wedge\tau_1}\right)$$

Por lo tanto,

$$E\left(X_{n\wedge\tau_2} - X_{n\wedge\tau_1}\right) = 0 \quad \forall n$$

Poniendo $n \geq C$, resulta (10).

Probemos ahora que

$$E\left(X_{\tau_2}/\mathfrak{F}_{\tau_1}\right) = E\left(X_{\tau_1}\right) \quad (\textit{resp.} \leq, \geq) \quad \textit{c.s.}$$

Es decir, se trata de probar que $\forall\ B \in \mathfrak{F}_{\tau_1}$ se verifica

$$E\left(X_{\tau_2} 1_B\right) = E\left(X_{\tau_1} 1_B\right) \quad (\textit{resp.} \leq, \geq) \tag{11}$$

Veamos el caso de la igualdad; los otros son similares.

Definimos, para $i = 1, 2$:

$$\tau_{i,B} = \tau_i 1_B + C 1_{B^C}$$

$\tau_{1,B}$ es un tiempo de parada, ya que si $n \geq C$, $\{\tau_{1,B} \leq n\} = \Omega$, y si $n < C$, entonces $\{\tau_{1,B} \leq n\} = \{\tau_1 \leq n\} \cap B \in \mathfrak{F}_n$ por ser $B \in \mathfrak{F}_{\tau_1}$.

Análogamente, como $B \in \mathfrak{F}_{\tau_1} \subset \mathfrak{F}_{\tau_2}$, si $n < C$, entonces $\{\tau_{2,B} \leq n\} = \{\tau_2 \leq n\} \cap B \in \mathfrak{F}_n$.

Además es obvio que $\tau_{1,B} \leq \tau_{2,B}$, de modo que se puede aplicar (10) a la pareja de tiempos de parada $\tau_{1,B}, \tau_{2,B}$ y resulta $E(X_{\tau_{1,B}}) = E(X_{\tau_{2,B}})$, es decir que

$$E\left(X_{\tau_2} 1_B\right) + E\left(X_C 1_{B^C}\right) = E\left(X_{\tau_1} 1_B\right) + E\left(X_C 1_{B^C}\right)$$

lo que prueba (11).□

Teorema (desigualdades de Doob).

Sea $\{X_n\}_{0\le n\le N}$ una submartingala con respecto a la filtración $\{\mathfrak{F}_n\}_{0\le n\le N}$ y $\lambda > 0$.

Entonces:

(i)

$$\lambda P\left(\max_{0\le n\le N} X_n \ge \lambda\right) \le E\left(X_N 1_{\left\{\max_{0\le n\le N} X_n \ge \lambda\right\}}\right) \le E\left(X_N^+\right) \le E\left(|X_N|\right) \tag{12}$$

$$\lambda P\left(\min_{0\le n\le N} X_n \le -\lambda\right) \le -E(X_0)+E\left(X_N 1_{\left\{\min_{0\le n\le N} X_n \le -\lambda\right\}}\right) \le E\left(X_0^-\right)+E\left(X_N^+\right) \tag{13}$$

(ii) Si $\{X_n\}_{0\le n\le N}$ es martingala y $E\left(|X_n|^p\right) < \infty \quad \forall n = 0, ..., N$ y para algún $p \ge 1$, entonces

$$P\left(\max_{0\le n\le N} |X_n| \ge \lambda\right) \le \frac{1}{\lambda^p} E\left(|X_N|^p\right) \tag{14}$$

(iii) Si además de las hipótesis de (ii) es $p > 1$, entonces:

$$E\left(\max_{0\le n\le N} |X_n|^p\right) \le \left(\frac{p}{p-1}\right)^p E\left(|X_N|^p\right) \tag{15}$$

Demostración.

(i) Probamos (12). (13) es similar.

Sea

$$\begin{aligned} \tau &= \min\{n : 0 \le n \le N, X_n \ge \lambda\} \qquad \text{si } \{\} \ne \phi \\ &= N \qquad \text{si } \{\} = \phi \end{aligned}$$

En virtud del teorema del muestreo opcional de Doob, $\{X_\tau, X_N\}$ es una submartingala con respecto a la filtración $\{\mathfrak{F}_\tau, \mathfrak{F}_N\}$. Por lo tanto,

$$\begin{aligned} E\left(X_N\right) &\ge E\left(X_\tau\right) = E\left(X_\tau 1_{\left\{\max_{0\le n\le N} X_n \ge \lambda\right\}}\right) + E\left(X_N 1_{\left\{\max_{0\le n\le N} X_n < \lambda\right\}}\right) \\ &\ge \lambda P\left(\max_{0\le n\le N} X_n \ge \lambda\right) + E\left(X_N 1_{\left\{\max_{0\le n\le N} X_n < \lambda\right\}}\right) \end{aligned}$$

Pasando el último término al primer miembro resulta (12).

(ii) Aplicar (i) a la submartingala $\{|X_n|^p\}$.

(iii) Sea $M = \max_{0\le n\le N} |X_n|$. Entonces:

$$\begin{aligned} E\left(M^p\right) &= p\int_0^{+\infty} \lambda^{p-1} P(M \ge \lambda)\, d\lambda \le p\int_0^{+\infty} \lambda^{p-2} E(|X_N|\, 1_{\{M\ge\lambda\}})\, d\lambda \\ &= \frac{p}{p-1} E\left(|X_N|\, M^{p-1}\right) \end{aligned}$$

ya que $(p-1)\int_0^{+\infty}\lambda^{p-2}1_{\{M\geq\lambda\}}d\lambda = M^{p-1}$.

Aplicando la desigualdad de Hölder:

$$E\left(M^p\right) \leq \frac{p}{p-1}\left[E\left(|X_N|^p\right)\right]^{\frac{1}{p}}\left[E\left(M^p\right)\right]^{\frac{p-1}{p}}$$

Pasando el último factor al primer miembro, resulta (iii).□

1.5.1 Teorema de convergencia de martingalas.

1. $\{X_n\}_{n=0,1,2,..}$ submartingala, $sup_n E\left(X_n^+\right) < \infty \Longrightarrow c.s.\quad \exists \lim_{n\to+\infty} X_n = X_\infty,\ X_\infty \in L^1$.

2. Bajo las condiciones de 1., se puede extender $\{X_n\}_{n=0,1,2,..}$ a una submartingala definida en $\{0,1,2...\}\cup\{+\infty\}$ agregando $+\infty \rightsquigarrow X_\infty$, si y sólo si, $\{X_n^+\}_{n=0,1,2,..}$ es uniformemente integrable.

3. Sea $Y \in L^1$ y $\{\mathfrak{F}_n\}_{n=0,1,2,...}$ una filtración.

   Entonces, $X_n = E\left(Y/\mathfrak{F}_n\right)$ es una martingala adaptada a $\{\mathfrak{F}_n\}_{n=0,1,2,...}$, uniformemente integrable y existe $\lim_{n\to+\infty} X_n = X_\infty$, casi seguramente y en L^1. Además, $X_\infty = E\left(Y/\mathfrak{F}_\infty\right)$.

4. $\mathcal{X} = \{X_n\}_{n=0,-1,-2,..}$ submartingala, $\inf_n E\left(X_n\right) > -\infty \Longrightarrow \mathcal{X}$ es uniformemente integrable y $\exists\ \lim_{n\to-\infty} X_n = X_{-\infty},\ c.s.$ y en L^1.

La demostración se basa en el lema siguiente, debido a Doob.

Lema (sobre cruces).

Sea $\{X_n\}_{n=0,1,2,..}$una submartingala con respecto a la filtración $\{\mathfrak{F}_n\}_{n=0,1,2,..}$ y a,b dos números reales, $a<b$.

Entonces:

$$E\left(U_N^X(a,b)\right) \leq \frac{1}{b-a}E\left[\left(X_N-a\right)^+ - \left(X_0-a\right)^+\right] \tag{16}$$

Aquí, $U_N^X(a,b) =$ número de cruces hacia arriba de la banda $[a,b]$ entre 0 y N, que se puede definir formalmente de la manera siguiente:

Para cada trayectoria $\{X_n\}_{n=0,1,2,..}$, definimos por recurrencia la sucesión

$$\begin{aligned}
\tau_0 &= 0 \\
\tau_1 &= \min\{n : X_n \le a\} \\
\tau_2 &= \min\{n : n \ge \tau_1,\ X_n \ge b\} \\
&\quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\tau_{2k+1} &= \min\{n : n \ge \tau_{2k},\ X_n \le a\} \\
\tau_{2k+2} &= \min\{n : n \ge \tau_{2k+1},\ X_n \ge b\}
\end{aligned}$$

donde se usa la convención de que $min(\phi) = +\infty$, y si un cierto τ es igual a $+\infty$, desde allí en adelante los τ sucesivos también valen $+\infty$.

Entonces,

$$U_N^X(a,b) = \max\{k : \tau_{2k} \le N\}$$

Demostración del lema.

Consideramos el nuevo proceso $\mathcal{Y} = \{Y_n\}$ definido por $Y_n = (X_n - a)^+$, que también es una submartingala, ya que la función $x \rightsquigarrow x^+$ es convexa y creciente (en sentido amplio).

Es claro que

$$U_N^X(a,b) = U_N^Y(0, b-a)$$

Definimos la sucesión $\tau_0, \tau_1, \tau_2, ...$ como antes, pero para el proceso $\mathcal{Y}$ en lugar de $\mathcal{X}$. Es inmediato que es una sucesión creciente de tiempos de parada.

Sea $\tau'_n = \tau_n \wedge N$. $\{\tau'_n\}_{n=0,1,2,...}$ es también una sucesión creciente de tiempos de parada, acotada por N. Elegimos además el natural K de tal modo que $2K > N$. Dado que mientras que los τ_n son finitos, forman una sucesión estrictamente creciente, tiene que ser $\tau_{2K} > N$ y, por lo tanto, $\tau'_{2K} = N$.

Se tiene:

$$Y_N - Y_0 = \sum_{n=1}^{2K} \left(Y_{\tau'_n} - Y_{\tau'_{n-1}}\right) = \sum_{n=1}^{K} \left(Y_{\tau'_{2n}} - Y_{\tau'_{2n-1}}\right) + \sum_{n=0}^{K-1} \left(Y_{\tau'_{2n+1}} - Y_{\tau'_{2n}}\right)$$

El primer término está minorado por $(b-a)U_N^Y(0, b-a)$. Tomando esperanzas:

$$E(Y_N) - E(Y_0) \ge (b-a)E\left(U_N^Y(0, b-a)\right) + \sum_{n=0}^{K-1} \left[E(Y_{\tau'_{2n+1}}) - E(Y_{\tau'_{2n}})\right]$$

Por el teorema del muestreo opcional de Doob, $\left\{Y_{\tau'_n}\right\}$ es una submartingala y cada término en la última suma es ≥ 0. Se deduce que

$$E\left[(X_N - a)^+ - (X_0 - a)^+\right] = E(Y_N) - E(Y_0) \geq (b-a)E\left(U_N^Y(0, b-a)\right)$$

Esto prueba (16).□

Disgresión. (Sobre integrabilidad uniforme).

A título de repaso, recordemos que la sucesión de variables aleatorias $\{Z_n\}_{n=0,1,2,..}$, definidas en el espacio de probabilidad $(\Omega, \mathfrak{A}, P)$, $Z_n \in L^1$, es *uniformemente integrable* si

$$\sup_n E\left(|Z_n| \, 1_{\{|Z_n| \geq a\}}\right) \to 0$$

cuando $a \to +\infty$.

El lector verificará que si la sucesión $\{Z_n\}$ es acotada en L^p para algún $p > 1$, entonces es uniformemente integrable. También que si, casi seguramente, $Z_n \to Z_\infty$ cuando $n \to +\infty$, entonces $Z_n \to Z_\infty$ en L^1 si y sólo si $\{Z_n\}$ es uniformemente integrable. Mostrar un ejemplo en el cual esto último se verifica, pero no se cumplen las condiciones del teorema de convergencia dominada de Lebesgue.

Demostración del teorema de convergencia.

Para la parte 1. se trata de probar que

$$P(\underline{\lim} X_n < \overline{\lim} X_n) = 0.$$

Sea $U_\infty^Y(a,b) = \lim_{N \to +\infty} U_N^Y(a,b)$. Entonces:

$$\left\{\underline{\lim} X_n < \overline{\lim} X_n\right\} = \bigcup_{r,s \in Q, \ r<s} \left\{U_\infty^Y(r,s) = +\infty\right\}$$

Usando Beppo Levi y el lema sobre cruces (para r, s fijos, $r < s$:

$$E\left(U_\infty^Y(r,s)\right) = \lim_{N \to +\infty} E\left(U_N^Y(r,s)\right) \leq \underline{\lim}_{N \to +\infty} \frac{1}{s-r} E\left[(X_N - r)^+ - (X_0 - r)^+\right] < \infty$$

La finitud se deduce de la hipótesis. Por lo tanto, $U_\infty^Y(r,s) < \infty$ con probabilidad 1. Esto prueba que $c.s. \exists \lim_{n \to +\infty} X_n = X_\infty$. Para ver que $X_\infty \in L^1$, i.e. $E(|X_\infty|) < \infty$, notar que

$$E(|X_n|) = 2E(X_n^+) - E(X_n) \leq 2E(X_n^+) - E(X_0)$$

donde la desigualdad se sigue de que $\{X_n\}$ es submartingala. Usar ahora la hipótesis y el lema de Fatou.

Para probar 2. supongamos que $\{X_n^+\}$ es uniformemente integrable. Entonces, $\forall a > 0$, *c.s.* $X_m \vee (-a) \to X_\infty \vee (-a)$ no sólo casi seguramente sino también en L. Por lo tanto, para cada n fijo, $E\left(X_m \vee (-a)/\mathfrak{F}_n\right) \to E\left(X_\infty \vee (-a)/\mathfrak{F}_n\right)$ en L^1 cuando $m \to +\infty$ y casi seguramente sobre una subsucesión $\{m_k\}$. Como $\{X_n \vee (-a)\}$ es submartingala,

$$E\left(X_\infty \vee (-a)/\mathfrak{F}_n\right) = \lim_{k\to+\infty} E\left(X_{m_k} \vee (-a)/\mathfrak{F}_n\right) \geq X_n \vee (-a) \quad c.s.$$

Haciendo $a \uparrow +\infty$, resulta

$$E\left(X_\infty/\mathfrak{F}_n\right) \geq X_n \quad c.s.$$

Recíprocamente, si $E\left(X_\infty/\mathfrak{F}_n\right) \geq X_n \quad c.s.\ \forall n$, entonces (desigualdad de Jensen)

$$E\left(X_\infty^+/\mathfrak{F}_n\right) \geq \left[E\left(X_\infty/\mathfrak{F}_n\right)\right]^+ \geq X_n^+ \quad c.s.$$

y como $\{X_n^+ \geq x\}$ es $\mathfrak{F}_n$-medible,

$$E\left(X_n^+ 1_{\{X_n^+\geq x\}}\right) \leq E\left(X_\infty^+ 1_{\{X_n^+\geq x\}}\right) \tag{17}$$

Se trata de probar que $\{X_n^+\}$ es uniformemente integrable, es decir que $E\left(X_n^+ 1_{\{X_n^+\geq x\}}\right) \to 0$ cuando $x \to +\infty$, uniformemente en n. En virtud de (17) y de que $X_\infty^+ \in L^1$, basta probar que $P(X_n^+ \geq x) \to 0$ cuando $x \to +\infty$, uniformemente en n. Esto se deduce de ($x > 0$):

$$P(X_n^+ \geq x) \leq \frac{1}{x}E(X_n^+) \leq \frac{1}{x}E(X_\infty^+).$$

La demostración de 3. es análoga. Para probar que $X_\infty = E(Y/\mathfrak{F}_\infty)$ basta ver que $E(X_\infty 1_B) = E(Y 1_B)\ \forall B \in \mathfrak{F}_\infty$. Para esto, observar que la familia $\{B : E(X_\infty 1_B) = E(Y 1_B)\}$ es un álgebra y también una clase monótona que contiene a $\mathfrak{F}_n\ \forall n$, por lo que contiene a $\mathfrak{F}_\infty$.

La parte de convergencia casi segura en la demostración de 4. es igual a la análoga de 1. La parte de la integrabilidad uniforme se ve como sigue.

$E(X_n)$ decrece cuando n decrece. La hipótesis dice que $lim_{n\to-\infty}E(X_n)$ es finito. Sea n_0 tal que $E(X_n) > E(X_{n_0}) - \varepsilon$ si $n \leq n_0$. Entonces, si $n \leq n_0$ y $x > 0$:

$$\begin{aligned}
E\left(|X_n|\, 1_{\{|X_n|\geq x\}}\right) &= E\left(X_n 1_{\{X_n\geq x\}}\right) + E\left(X_n 1_{\{X_n>-x\}}\right) - E\left(X_n\right) \\
&\leq E\left(X_{n_0} 1_{\{X_n\geq x\}}\right) + E\left(X_{n_0} 1_{\{X_n>-x\}}\right) - E\left(X_{n_0}\right) + \varepsilon \\
&= E\left(X_{n_0} 1_{\{X_n\geq x\}}\right) - E\left(X_{n_0} 1_{\{X_n\leq -x\}}\right) + \varepsilon \\
&\leq E\left(|X_{n_0}|\, 1_{\{|X_n|\geq x\}}\right) + \varepsilon
\end{aligned}$$

Análogamente a lo anterior, falta ver que $P(|X_n| \geq x) \to 0$ cuando $x \to +\infty$, uniformemente en n. Esto se deduce de:

$$\begin{aligned} P(|X_n| &\geq x) \leq \frac{1}{x}E(|X_n|) = \frac{1}{x}\left[2E(X_n^+) - E(X_n)\right] \\ &\leq \frac{1}{x}\left[2E(X_0^+) - E(X_{n_0}) + \varepsilon\right]. \end{aligned}$$

□

1.5.2 Aplicación. Teorema de Wald.

Teorema (Wald).

Sea $S_0 = 0$, $S_n = \xi_1 + ... + \xi_n$ $(n \geq 1)$, donde $\xi_1, \xi_2, ...$ es una sucesión de variables aleatorias independientes, $E(\xi_n) = \mu, E(|\xi_n|) = \widetilde{\mu}$ $\forall n$.

Denotamos por $\{\mathfrak{F}_n\}_{n=0,1,2,..}$ la filtración engendrada por $\{\xi_n\}_{n=^o1,2,...}$ $\mathfrak{F}_0 = \{\phi, \Omega\}$.

Sea τ un tiempo de parada con respecto a $\{\mathfrak{F}_n\}_{n=0,1,2,..}$ y supongamos que $E(\tau) < \infty$.

Entonces,

(a)

$$E(S_\tau) = \mu E(\tau)$$

(b) Si además $Var(\xi_n) = \sigma^2 < \infty$, entonces

$$E\left[(S_\tau - \tau\mu)^2\right] = \sigma^2 E(\tau)$$

Demostración. $\{S_n - n\mu\}_{n=0,1,...}$ es una $\{\mathfrak{F}_n\}_{n=0,1,2,..}$-martingala $\Longrightarrow$ (usando el teorema del muestreo opcional) que también $\{S_{n\wedge\tau} - (n \wedge \tau)\mu\}_{n=0,1,...}$ es una $\{\mathfrak{F}_n\}_{n=0,1,2,..}$-martingala $\Longrightarrow$

$$E\left[S_{n\wedge\tau} - (n \wedge \tau)\mu\right] = 0$$

o sea que

$$E(S_{n\wedge\tau}) = \mu\ E(n \wedge \tau) \tag{18}$$

Queremos hacer $n \uparrow +\infty$. El segundo miembro tiende a $E(\tau)$. En cuando al primero, si las variables aleatorias ξ_n son todas ≥ 0, se pasa al límite en el primero aplicando el teorema de Beppo Levi y eso prueba (a).

En caso contrario, se aplica el teorema a la sucesión $\{|\xi_n|\}_{n=1,2,...}$ en lugar de $\{\xi_n\}_{n=1,2,...}$ Si ponemos $\widetilde{S}_0 = 0$, $\widetilde{S}_n = |\xi_1| + ... + |\xi_n|$ $(n \geq 1)$, resulta entonces

$$E\left(\widetilde{S}_\tau\right) = \widetilde{\mu}E(\tau) < \infty$$

y como $|S_{n\wedge\tau}| \leq \widetilde{S}_\tau \ \forall n$, podemos aplicar el teorema de convergencia dominada para pasar al límite en el primer miembro de (18).

La demostración de (b) es similar, teniendo en cuenta que $\left\{(S_n - n\mu)^2 - n\sigma^2\right\}$ es martingala.□

Ejemplo.

Se considera un paseo al azar (p, q). Esto significa que en el ejemplo anterior, las variables aleatorias independientes de la sucesión $\{\xi_n\}$ tienen todas la misma distribución y es la siguiente: $P(\xi_n = 1) = p, P(\xi_n = -1) = q, p + q = 1$.

Sean además a, b dos enteros positivos y $\tau_{a,b} = \min\{n : S_n = a \text{ ó } S_n = b\}$.

$\tau_{a,b}$ es un tiempo de parada con respecto a la filtración engendrada por el proceso.

Entonces:

1. $P(\tau_{a,b} < \infty) = 1$, lo que significa que la probabilidad de que el proceso quede eternamente comprendido estrictamente entre las barreras $-a$ y b es igual a cero.

2. Denotamos

   $$\pi_a = P(S_{\tau_{a,b}} = -a), \quad \pi_b = P(S_{\tau_{a,b}} = b) \quad \text{donde } \pi_a + \pi_b = 1$$

   π_a (resp. π_b) es la probabilidad de que el proceso llegue antes a la barrera $-a$ (resp. b) que a la barrera b)(resp. $-a$).

   Entonces:

   $$E(S_{\tau_{a,b}}) = -a\pi_a + b(1 - \pi_a) = (p - q)E(\tau_{a,b})$$

   Esta igualdad resultará del teorema de Wald, una vez que probemos que $\tau_{a,b} \in L^1$, lo cual implicará a su vez la validez de la afirmación hecha en 1.

3. Para el caso en que $p = q = 1/2$ se deduce que $\pi_a = \frac{b}{a+b}$.

En el ejemplo anterior, podemos modificar las cosas de modo que los cálculos combinatorios que hacen posible la obtención de las fórmulas sin recurrir al teorema de Wald, resulten muy complicados o imposibles. Por ejemplo, supongamos que en cada paso, el salto del paseo al azar pueda tomar 3 valores en lugar de 2, a saber $-1, 0, 1$ y que

$$P(\xi_n = 1) = p_n, \quad P(\xi_n = -1) = q_n, \quad P(\xi_n = 0) = r_n$$

$p_n + q_n + r_n = 1$. Además, suponemos que $\inf(p_n) > 0, \inf(q_n) > 0$.

Las fórmulas anteriores permanecen válidas, incluyendo $\pi_a = \frac{b}{a+b}$ cuando $\frac{1}{n}\sum(p_n - q_n) \to 0$.

También podemos modificar el teorema de Wald, permitiendo que las medias de los saltos no sean todas iguales, i.e. $E(\xi_n) = \mu_n$, con las hipótesis

$$\frac{\mu_1 + ... + \mu_n}{n} \to \mu, \quad \left\{ \frac{|\mu_1 + ... + \mu_n|}{n} \right\} \textit{ acotada}$$

La fórmula (a) del teorema de Wald permanece válida.

El único punto que queda pendiente es que $\tau_{a,b} \in L^1$. Esto es un caso particular de la proposición siguiente, que tiene interés en si misma, por sus numerosas aplicaciones.

Proposición.

Sea $\{\xi_n\}_{n=1,2,..}$ una sucesión de variables aleatorias reales, independientes con igual distribución F. Suponemos que F no se reduce a un átomo en el origen. Entonces,

$$P(\overline{\lim}_{n\to+\infty} |S_n| = +\infty) = 1$$

Más aún, si $P(\xi_1 > 0) > 0$ y $P(\xi_1 < 0) > 0$, entonces $P(\overline{\lim}_{n\to+\infty} S_n = +\infty, \underline{\lim}_{n\to-\infty} S_n = -\infty) = 1$.

Si $P(\xi_1 > 0) > 0$ y $P(\xi_1 < 0) = 0$, es inmediato que casi seguramente $\{S_n\}$ es monótona creciente y tiende a $+\infty$.

Demostración.

Basta demostrar que $\forall a, b$ reales, $a < b$, se tiene $P(\tau_{a,b} < \infty) = 1$. Probamos un resultado más fuerte, que implica que

$$E\left[(\tau_{a,b})^k\right] < \infty \ \forall k = 1, 2, .. \tag{19}$$

Sean $\delta, \varepsilon > 0$ tales que $P(\xi_1 \geq \delta) \geq \varepsilon$ y $P(\xi_1 \leq -\delta) \geq \varepsilon$ y ν tal que $\delta.\nu \geq a + b$.

Se tiene:

$$\begin{aligned} P(\tau_{a,b} &> n.\nu) \\ &\leq P\left(|S_\nu| < a + b, |S_{2\nu} - S_\nu| < a + b, ..., \left|S_{n\nu} - S_{(n-1)\nu}\right| < a + b\right) \\ &= \left[P\left(|S_\nu| < a + b\right)\right]^n \\ &= \left[1 - P\left(|S_\nu| \geq a + b\right)\right]^n \end{aligned}$$

Por otra parte,

$$P(S_\nu \geq a+b) \geq P(\xi_1 \geq \delta, ..., \xi_\nu \geq \delta) = [P(\xi_1 \geq \delta)]^\nu \geq \varepsilon^\nu$$

$$P(S_\nu \leq -(a+b)) \geq P(\xi_1 \leq -\delta, ..., \xi_\nu \leq -\delta) = [P(\xi_1 \leq -\delta)]^\nu \geq \varepsilon^\nu$$

Por lo tanto,

$$P(\tau_{a,b} > n.\nu) \leq \eta^n$$

donde $0 < \eta < 1$ y η depende de la pareja a, b de partida. En conclusión,

$$P(\tau_{a,b} > m) \leq \eta_1^m \tag{20}$$

para $m \geq m_0$, donde η_1 es una nueva constante, $0 < \eta_1 < 1$.

Es fácil ver que (20) implica (19). En efecto, para cada entero positivo k,

$$E\left[(\tau_{a,b})^k\right] = \sum_{j=1}^{\infty} P(\tau_{a,b} \geq j^{\frac{1}{k}}) \leq \sum_{j=1}^{\infty} (\eta_1)^{\left[j^{\frac{1}{k}}\right]}$$

donde [] denota la parte entera. La última serie es convergente.□

1.5.3 Leyes fuertes de los grandes números para sumas de variables aleatorias independientes.

Teorema (Kolmogorov). Sea $\{Y_n\}_{n=0,1,...}$ una sucesión de variables aleatorias reales, independientes, con los dos primeros momentos finitos, $E(Y_n) = 0$, $Var(Y_n) = \sigma_n^2$.

Si se cumple que

$$\sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} < \infty \tag{21}$$

entonces se tiene

$$P\left(\lim_{n\to+\infty} \frac{Y_1 + ... + Y_n}{n} = 0\right) = 1 \tag{22}$$

(en otros términos, casi seguramente, los promedios $\frac{Y_1+....+Y_n}{n} \to 0$ cuando $n \to +\infty$).

Demostración.

Se define la sucesión de variables aletorias

$$X_0 = 0, \quad X_n = \sum_{j=1}^{n} \frac{Y_j}{j} \text{ si } n \geq 1$$

y se considera la filtración $\{\mathfrak{F}_n\}$ engendrada por el proceso $\{Y_n\}$.

Es inmediato que $\{X_n\}_{n=0,1,...}$ es una $\{\mathfrak{F}_n\}_{n=0,1,...}$-martingala y que verifica la parte 1. del teorema de convergencia de martingalas, ya que

$$E\left(X_n^+\right) \leq E\left(|X_n|\right) \leq \left[E\left(X_n^2\right)\right]^{\frac{1}{2}}$$

y

$$E\left(X_n^2\right) = \sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} < \infty$$

por hipótesis. Por lo tanto, casi seguramente la sucesión de variables aleatorias $\{X_n\}_{n=0,1,...}$ converge, i.e. para casi todo ω, existe el $\lim_{n\to+\infty} X_n$ y es finito. La conclusión del teorema se sigue del siguiente lema (de Kronecker).□

Lema. Sea $\{y_n\}_{n=1,2,...}$ una sucesión de números reales y $\{x_n\}_{n=0,1,...}$ definida mediante

$$x_0 = 0, \quad x_n = \sum_{j=1}^{n} \frac{y_j}{j} \text{ si } n \geq 1$$

Entonces, $\{x_n\}$ convergente implica que $\frac{y_1+....+y_n}{n} \to 0$ cuando $n \to +\infty$.

Demostración. Se tiene que

$$\begin{aligned} & \frac{y_1 + ... + y_n}{n} \\ = & \frac{(x_1 - x_0) + 2(x_2 - x_1) + ... + (x_n - x_{n-1})}{n} \\ = & x_n - \frac{x_1 + ... + x_{n-1}}{n} \to 0 \end{aligned}$$

ya que la convergencia de la sucesión implica la de los promedios, al mismo límite.□

Observación. En el enunciado del teorema anterior, si la esperanza de las variables aleatorias Y_n en lugar de ser nula es igual a μ, entonces la misma conclusión vale, sólo que casi seguramente, es $\frac{Y_1+....+Y_n}{n} = \mu$. En efecto, en la demostración, basta cambiar Y_n por $Y_n - \mu$.

A continuación, probamos una ley fuerte de grandes números para sumas de variables independientes, en la que no se requiere - como en el teorema anterior - que los sumandos sean de cuadrado integrable. En cambio, les pediremos que tengan todos la misma distribución.

Teorema.

Sea $\{Y_n\}_{n=0,1,...}$ una sucesión de variables aleatorias reales, independientes, con igual distribución y esperanza finita μ.

Entonces,

$$P\left(\lim_{n\to+\infty}\frac{Y_1+....+Y_n}{n}=\mu\right)=1 \tag{23}$$

Demostración. Basta considerar el caso $\mu = 0$. Denotamos por F la función de distribución (común) de las Y_n's.

Introducimos la nueva sucesión de variables aleatorias independientes

$$\widetilde{Y}_n = Y_n 1_{\{|Y_n|<n\}}.$$

Probamos:

a) para casi todo ω, $\exists\, n_0(\omega)$ tal que si $n \geq n_0(\omega)$ entonces $Y_n(\omega) = \widetilde{Y}_n(\omega)$.

b) definimos

$$Z_n = \widetilde{Y}_n - v_n \quad \text{con} \quad v_n = E(\widetilde{Y}_n)$$

Entonces la sucesión $\{Z_n\}$ cumple las hipótesis del teorema anterior, y por lo tanto, casi seguramente $\frac{Z_1+....+Z_n}{n} \to 0$ cuando $n \to +\infty$.

c) $\frac{v_1+....+v_n}{n} \to 0$ cuando $n \to +\infty$ y por lo tanto, también casi seguramente $\frac{\widetilde{Y}_1+....+\widetilde{Y}_n}{n} \to 0$ cuando $n \to +\infty$.

d) a) y c) implica la tesis.

a) es consecuencia del lema de Borel-Cantelli. Se trata de probar que la serie $\sum_{n=1}^{\infty} P(\widetilde{Y}_n \neq Y_n) < \infty$. Tenemos

$$\sum_{n=1}^{\infty} P(\widetilde{Y}_n \neq Y_n) = \sum_{n=1}^{\infty} P(|Y_n| \geq n) = \sum_{n=1}^{\infty} P(|Y_1| \geq n)$$

Además:

$$\sum_{n=1}^{\infty} P(Y_1 \geq n) \leq \int_0^{+\infty} P(Y_1 \geq x)\, dx = E(Y_1^+) < \infty$$

Del mismo modo resulta que $\sum_{n=1}^{\infty} P(Y_1 \leq -n) < \infty$. Esto prueba a).

Para b) tenemos que verificar que

$$\sum_{n=1}^{\infty} \frac{Var(Z_n)}{n^2} < \infty \tag{24}$$

Es claro que $Var(Z_n) = E(\widetilde{Y}_n^2) - \upsilon_n^2 \leq E(\widetilde{Y}_n^2)$. Por lo tanto:

$$\sum_{n=1}^{\infty} \frac{Var(Z_n)}{n^2} \leq \sum_{n=1}^{\infty} \frac{E\left(Y_n^2 1_{\{|Y_n|<n\}}\right)}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n^2} \int_{(-n,n)} x^2 F(dx)$$

Integrando por partes:

$$\begin{aligned}
& \sum_{n=1}^{\infty} \frac{1}{n^2} \int_0^n x^2 F(dx) \\
= & \sum_{n=1}^{\infty} \frac{1}{n^2} \left[-x^2 \left[1 - F(x)\right] |_0^n + 2 \int_0^n x \left[1 - F(x)\right] \, dx \right] \\
\leq & \; 2 \sum_{n=1}^{\infty} \frac{1}{n^2} \int_0^n x(1 - F(x) \, dx \leq 2 \sum_{n=1}^{\infty} \frac{1}{n^2} \sum_{j=1}^{n} j \left[1 - F(j-1)\right] \\
= & \; 2 \sum_{j=1}^{\infty} j \left[1 - F(j-1)\right] \sum_{n=j}^{\infty} \frac{1}{n^2} \leq (const) \sum_{j=1}^{\infty} \left[1 - F(j-1)\right] \\
\leq & \; (const) \left[1 - F(0) + \int_0^{+\infty} \left[1 - F(x)\right] \, dx \right] \\
= & \; (const) \left[1 - F(0) + E(Y_1^+)\right] < \infty
\end{aligned}$$

Análogamente resulta que $\sum_{n=1}^{\infty} \frac{1}{n^2} \int_{-n}^{0} x^2 F(dx) < \infty$. Esto prueba (24).

Falta probar c), que se deduce de $\upsilon_n \to 0$ cuando $n \to +\infty$. En efecto

$$\upsilon_n = E(\widetilde{Y}_n) = E(\widetilde{Y}_n - Y_n) = -E\left(Y_n.1_{\{|Y_n| \geq n\}}\right) = -E\left(Y_1.1_{\{|Y_1| \geq n\}}\right)$$

que implica

$$|\upsilon_n| \leq E\left(|Y_1|.1_{\{|Y_1| \geq n\}}\right) \downarrow 0$$

en virtud del teorema de Lebesgue de convergencia dominada, ya que $Y_1 \in L^1$ y casi seguramente, $|Y_1|.1_{\{|Y_1| \geq n\}} \downarrow 0$ cuando $n \to +\infty$, ya que $|Y_1|$ es finita casi seguramente.$\square$

Observación general. Recordemos que la convergencia fuerte implica la convergencia en probabilidad, es decir, que si $\{X_n\}_{n=0,1,2,..}$ es una sucesión de variables aleatorias tal que casi seguramente $\exists \lim_{n \to +\infty} X_n = X_\infty$, entonces $\forall \ \varepsilon > 0$ se cumple que $P(|X_n - X_\infty| \geq \varepsilon) \to 0$ cuando $n \to +\infty$. Naturalmente, la convergencia en probabilidad es cierta bajo condiciones más débiles que aquéllas en las que podemos asegurar la convergencia casi segura.

No habremos de extendernos sobre este punto, pero habrá de ocurrir que probemos convergencia en probabilidad directamente. Por ejemplo, es inmediato que si $X_n \to X_\infty$ en L^p para algún $p > 0$, entonces $X_n \to X_\infty$ en probabilidad, ya que

$$P(|X_n - X_\infty| \geq \varepsilon) \leq \frac{1}{\varepsilon^p} E\left(|X_n - X_\infty|^p\right).$$

2 Convergencia débil de medidas de probabilidad en $\mathcal{R}$.

2.1 Definición y primeras propiedades.

En lo que sigue, todas las medidas de probabilidad son medidas de Borel en $\mathcal{R}^1$. Con frecuencia haremos el abuso de notación de llamar a las medidas y a sus funciones de distribución, con la misma letra.

En lo que sigue, para simplificar la exposición, denominaremos "medidas de probabilidad" a las medidas de Borel F tales que la medida de todo el espacio $F(\mathcal{R})$ es *menor o igual* que 1. Si $F(\mathcal{R}) = 1$, diremos que F es una medida de probabilidad "propia" y si $F(\mathcal{R}) < 1$, diremos que es "defectiva". Asimismo, si I es el intervalo $(a, b]$ en la recta, la notación $F(I)$ es, como es habitual, la medida del intervalo semiabierto $(a, b]$ o también $F(b) - F(a)$, donde ahora llamamos F a la función de distribución.

Definición 1. Sea $\{F_n\}_{n=01,2,...}$ y F respectivamente, una sucesión de medidas y una medida en la recta. Diremos que $\{F_n\}$ "converge débilmente" a F (notación: $F_n \Longrightarrow F$) si $F_n(I) \to F(I)$ para todo intervalo I en cuyos extremos la función F es continua.

Definición 2. Sea $\{F_n\}_{n=01,2,...}$una sucesión de medidas de probabilidad propias. Diremos que la convergencia débil de F_n a F es propia, si F es propia.

Ejemplos.

1. Sea $F_n = \delta_{x_n}$ $(n = 1, 2, ...)$, donde δ_a denota la medida de Dirac en el punto a de la recta real, es decir que $\delta_a(\{a\}) = 1$ y $\delta_a(\mathcal{R} \setminus \{a\}) = 0$.

   Entonces, $x_n \to x$ implica que $\delta_{x_n} \to \delta_x$.

2. Sea $\psi : \mathcal{R} \to \mathcal{R}^+$ una función Borel-medible, $\int_{-\infty}^{+\infty} \psi(x)dx = 1$.

   Se define la sucesión de medidas de probabilidad $F_n(dx) = \frac{1}{a_n}\psi\left(\frac{x}{a_n}\right)dx$, donde $a_n \downarrow 0$.

   Entonces, $F_n \Longrightarrow \delta_0$.

3. Sea $F_n(dx) = f_n(x)dx$, donde la sucesión $\{f_n\}$ es de funciones no-negativas, $\int_{-\infty}^{+\infty} f_n(x)dx = 1 \ \forall n$.

   Si $f_n \to f$, uniformemente en cada compacto de $\mathcal{R}$, entonces $F_n \Longrightarrow F$, donde $F(dx) = f(x)dx$.

4. Sea $F_n(dx) = 1_{\{n,n+1\}}(x)\ dx$. Entonces $F_n \Longrightarrow 0$, donde 0 indica la medida idénticamente nula.

Propiedades básicas.

1. $F_n \Longrightarrow F$ implica que $F(\mathcal{R}) \leq 1$.

2. $F_n \Longrightarrow F$ si y sólo si $\forall\ f \in C_0(\mathcal{R})$ se cumple que $\int_{\mathcal{R}} f\ dF_n \to \int_{\mathcal{R}} f\ dF$.

   ($C_0(\mathcal{R})$ denota el espacion de las funciones de $\mathcal{R}$ en $\mathcal{R}$, continuas y que tienen límite cero en $+\infty$ y en $-\infty$).

3. $F_n \Longrightarrow F$ propia, si y sólo si $\forall\ f \in C_b(\mathcal{R})$ se cumple que $\int_{\mathcal{R}} f\ dF_n \to \int_{\mathcal{R}} f\ dF$.

   ($C_b(\mathcal{R})$ denota el espacion de las funciones de $\mathcal{R}$ en $\mathcal{R}$, continuas y acotadas).

4. $F_n \Longrightarrow F$ propia, si y sólo si, $F_n(x) \to F(x)$ para todo x de continuidad de F y $F(\mathcal{R}) = 1$.

5. Sea $\{F_n\}_{n=01,2,...}$una sucesión de medidas de probabilidad. Entonces, siempre existe una subsucesión que converge débilmente.

Demostraciones.

1. inmediato

2. Supongamos que $F_n \Longrightarrow F$.

   La función de distribución F es monótona creciente (en sentido amplio); por lo tanto el conjunto $D = \{x : F \text{ no es continua en } x\}$ es numerable.

   Sea $f \in C_0(\mathcal{R})$. Dado $\varepsilon > 0$ se puede encontrar una función en escalera:

$$g_\varepsilon(x) = \sum_{i=1}^{m} f(a_i) 1_{(a_i,b_i]}(x)$$

que verifica que los puntos a_i, b_i $(i = 1, ..., m)$ son puntos de continuidad de F, $a_1 < b_1 = a_2 < b_2 = ... = a_m < b_m$ y que además

$$\|f - g_\varepsilon\|_\infty < \varepsilon$$

donde $\|h\|_\infty = \sup_{x \in \mathcal{R}} |h(x)|$.

Por lo tanto:

$$\begin{aligned}
& \left| \int_{\mathcal{R}} f\ dF_n - \int_{\mathcal{R}} f\ dF \right| \\
\leq & \int_{\mathcal{R}} |f - g_\varepsilon|\ dF_n + \int_{\mathcal{R}} |f - g_\varepsilon|\ dF + \left| \int_{\mathcal{R}} g_\varepsilon\ dF_n - . \int_{\mathcal{R}} g_\varepsilon\ dF \right| \\
\leq & \ 2.\varepsilon + \left| \int_{\mathcal{R}} g_\varepsilon\ dF_n - . \int_{\mathcal{R}} g_\varepsilon\ dF \right| \\
= & \ 2.\varepsilon + \sum_{i=1}^{m} f(a_i) \left[F_n\left((a_i, b_i]\right) - F\left((a_i, b_i]\right) \right]
\end{aligned}$$

Haciendo $n \to +\infty$, cada término en la suma tiende a cero, de modo que $\overline{lim}_{n \to +\infty} \left| \int_{\mathcal{R}} f\ dF_n - \int_{\mathcal{R}} f\ dF \right| \leq 2.\varepsilon$. Como $\varepsilon > 0$ es arbitrario, eso prueba que $\lim_{n \to +\infty} \left[\int_{\mathcal{R}} f\ dF_n - \int_{\mathcal{R}} f\ dF \right] = 0$.

Recíprocamente, supongamos que $\forall\ f \in C_0(\mathcal{R})$ se cumple que $\int_{\mathcal{R}} f\ dF_n \to \int_{\mathcal{R}} f\ dF$ y sea $I = (a, b]$ con a, b puntos de continuidad de F.

Para $\varepsilon > 0$ dado, elegimos $\delta > 0$ de tal modo que

$$F(a + \delta) - F(a), F(a) - F(a - \delta), F(b + \delta) - F(b), F(b) - F(b - \delta)$$

sean todos $<\varepsilon$.

Consideremos las dos funciones continuas f y g (con gráfico trapezoidal) definidas de la manera siguiente:

- f vale 1 en $[a, b]$, vale 0 en $[a - \delta, b + \delta]^C$ y su gráfico es un segmento de recta en los intervalos $[a - \delta, a]$ y $[b, b + \delta]$.
- g vale 1 en $[a + \delta, b - \delta]$, vale 0 en $[a, b]^C$ y su gráfico es un segmento de recta en los intervalos $[a, a + \delta]$ y $[b - \delta, b]$. Entonces, dado que

  $$g(x) \leq 1_I(x) \leq f(x) \quad \forall x \in \mathcal{R}$$

  se sigue que

  $$\int_{\mathcal{R}} g\ dF_n \leq F_n(I) \leq \int_{\mathcal{R}} f\ dF_n \tag{25}$$

Es obvio que $f, g \in C_0(\mathcal{R})$ y que, por lo tanto, el primer miembro y el tercero de (25) tienden respectivamente a

$$\int_{\mathcal{R}} g \; dF \geq F(I) - 2\varepsilon$$

e

$$\int_{\mathcal{R}} f \; dF \leq F(I) + 2\varepsilon$$

Esto prueba que $\overline{\lim}_{n\to+\infty} |F_n(I) - F(I)| \leq 2\varepsilon$ y termina la prueba de la propiedad 2.

3. Probemos previamente la propiedad siguiente, que tiene interés en si misma:

   $F_n \Longrightarrow F$ propia, si y sólo si (a) $F_n \Longrightarrow F$ y (b) $\forall\varepsilon > 0 \; \exists \; A$ tal que $F_n (\{x : |x| > A\}) < \varepsilon \; \forall n$.

   En efecto, supongamos que $F_n \Longrightarrow F$ propia. Se trata de probar (b). Dado $\varepsilon > 0$, elegimos $A_1 > 0$, de modo que $-A_1$ y A_1 sean ambos puntos de continuidad de F y suficientemente grande como para que

$$F(A_1) - F(-A_1) > 1 - \frac{\varepsilon}{2}$$

   lo cual es posible porque F es propia. Entonces, como hay convergencia débil, $F_n(A_1) - F_n(-A_1) \to F(A_1) - F(-A_1)$ y $\exists \; n_0$ tal que $F_n(A_1) - F_n(-A_1) > 1 - \varepsilon \; \forall \; n \geq n_0$. Para que $F_n(A) - F_n(-A) > 1 - \varepsilon$ se cumpla para todo n, basta con elegir $A > A_1$ y además suficientemente grande para verificarla para $n = 1, ..., n_0 - 1$. Finalmente, notar que $F_n (\{x : |x| > A\}) \leq 1 - [F_n(A) - F_n(-A)]$.

   Recíprocamente, supongamos que se cumplen (a) y (b). Si elegimos en (b) $A > 0$ de modo que $-A$ y A sean puntos de continuidad de F, podemos pasar al límite en $F_n (\{x : |x| \leq A\}) > 1 - \varepsilon$, obteniendo $F(\mathcal{R}) \geq F (\{x : |x| \leq A\}) > 1 - \varepsilon$. Esto prueba que $F(\mathcal{R}) = 1$.

   Probemos ahora la propiedad 3. Supongamos que se cumple que $F_n \Longrightarrow F$ propia y sea $\varepsilon > 0$. Sea $A > 0$ punto de continuidad de F tal que $F_n(A) - F_n(-A) > 1 - \varepsilon \; \forall \; n$ y construimos una función continua (cuyo gráfico es trapezoidal) $g_A : \mathcal{R} \to [0, 1]$, que vale 1 en el intervalo $[-A, A]$, 0 fuera de algún intervalo de la forma $[-B, B]$, con $B > A$, y tiene por gráfico un segmento de recta en los intervalos $[-B, -A]$ y $[A, B]$.

Sea $f \in C_b(\mathcal{R})$. Es claro que $f.g_A \in C_0(\mathcal{R})$ y se tiene:

$$\begin{aligned}
&\left|\int_{\mathcal{R}} f\ dF_n - \int_{\mathcal{R}} f\ dF\right| \\
\leq\ &\left|\int_{\mathcal{R}} f\ g_A dF_n - \int_{\mathcal{R}} f g_A\ dF\right| + \left|\int_{\mathcal{R}} f\ (1-g_A) dF_n\right| + \left|\int_{\mathcal{R}} f\ (1-g_A) dF\right| \\
\leq\ &\left|\int_{\mathcal{R}} f\ g_A dF_n - \int_{\mathcal{R}} f g_A\ dF\right| + 2.\varepsilon. \|f\|_\infty
\end{aligned}$$

Haciendo $n \to +\infty$, el primer término del último miembro tiende a cero, y resulta $\overline{\lim}_{n\to+\infty} \left|\int_{\mathcal{R}} f\ dF_n - \int_{\mathcal{R}} f\ dF\right| \leq 2.\varepsilon. \|f\|_\infty$.

Recíprocamente, si $\forall\ f \in C_b(\mathcal{R})$ se cumple que $\int_{\mathcal{R}} f\ dF_n \to \int_{\mathcal{R}} f\ dF$, es claro que se cumple la definición de convergencia débil y tomando $f = 1$, resulta que F es propia.

4. A cargo del lector.

5. El lector que conozca la dualidad en los espacios de Banach, reconocerá un caso particular del teorema de Banach-Alaoglou. Damos una demostración directa más elemental, basada en el proceso diagonal de Cantor.

   Como para cada x, $\{F_n(x)\}$ es una sucesión acotada de números reales, podemos construir una subsucesión $\{F_{n_k}\}$ con la propiedad de que $F_{n_k}(x)$ converge cuando $k \to +\infty$ para todo racional x. Sea $\widetilde{F}(x) = \lim_{k\to+\infty} F_{n_k}(x)$ $(x \in \mathcal{Q})$ y definamos $F(x) = \inf\left\{\widetilde{F}(y) : y > x, y \in \mathcal{Q}\right\}$. Se verifica que F es creciente en sentido amplio y continua por la derecha y que $F_{n_k} \Longrightarrow F$.

2.2 Transformada de Fourier (función característica) de medidas de probabilidad en $\mathbf{R}^1$.

Sea μ una medida de probabilidad boreliana en $\mathcal{R}$, posiblemente defectiva. Se define su transformada de Fourier (o función característica), $\varphi : \mathcal{R} \to \mathcal{R}$ mediante

$$\varphi(t) = \int_{\mathcal{R}} \exp(itx)\ \mu(dx).$$

Si μ es la distribución de probabilidad de una variable aleatoria real ξ diremos - por extensión - también que φ es la transformada de Fourier (o la función característica) de ξ y denotaremos $\varphi = \varphi_\xi$.

2.2.1 Propiedades básicas y ejemplos.

1. $\varphi(0) = \mu(\mathcal{R})$, $|\varphi(t)| \leq \mu(\mathcal{R}) \leq 1 \ \forall t \in \mathcal{R}$.
2. φ es uniformemente continua.
3. Si a, b son constantes reales y ξ es una variable aleatoria real, entonces $\varphi_{a\xi+b} = \exp(itb)\ \varphi_\xi(at)$
4. Si ξ, η son variables aleatorias independientes, entonces $\varphi_{\xi+\eta}(t) = \varphi_\xi(t).\varphi_\eta(t)$
5. **Ejemplos**.
   - ξ con distribución normal $(0,1) \implies \varphi_\xi(t) = \exp(-\frac{t^2}{2})$.
   - ξ con distribución uniforme en el intervalo $(-1,1) \implies \varphi_\xi(t) = \frac{sen\ t}{t}$.
   - ξ con distribución de Cauchy (i.e., con densidad $\frac{1}{\pi(1+x^2)}$) $\implies \varphi_\xi(t) = \exp(-|t|)$.
   - ξ con distribución de Poisson con parámetro $\lambda > 0 \implies \varphi_\xi(t) = \exp\left[-\lambda(1-e^{it})\right]$.
   - ξ con densidad $\frac{1-\cos x}{\pi x^2} \implies \varphi_\xi(t) = (1-|t|) \vee 0$.
6. **Unicidad, continuidad, inversión**.

   La identidad fundamental es la fórmula de Parseval (pars).

   Sean F, G distribuciones de probabilidad en la recta y φ, ψ sus respectivas transformadas de Fourier.Entonces $\forall t \in \mathcal{R}$:

   $$\int_{-\infty}^{+\infty} e^{-i\zeta t}\varphi(\zeta)\ G(d\zeta) = \int_{-\infty}^{+\infty} \psi(x-t)\ F(dx) \tag{26}$$

   La demostración es inmediata, mediante aplicación del teorema de Fubini.

   Teorema (unicidad).$F = G$ si y sólo si $\varphi(t) = \psi(t)\ \forall t \in \mathcal{R}$.

   Demostración. Es fácil ver que basta considerar el caso en que ambas distribuciones son propias.

   Aplicamos (26) tomando para G la distribución normal $(0, \sigma^2)$, con lo cual $\psi(t) = \exp(-\frac{\sigma^2 t^2}{2})$.Entonces

   $$\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-i\zeta t}\varphi(\zeta)\exp\left(-\frac{\zeta^2}{2\sigma^2}\right)\ d\zeta = \int_{-\infty}^{+\infty} \frac{\sigma}{\sqrt{2\pi}}\exp\left(-\frac{\sigma^2(x-t)^2}{2}\right)\ F(dx) \tag{27}$$

Sean ξ, η_σ dos variables aleatorias independientes, ξ con distribución F y η_σ con distribución normal $(0,\sigma^2)$. El segundo miembro de (27) es la densidad de la variable aleatoria $\xi + \eta_{1/\sigma}$.

Por lo tanto, si a, b son puntos de continuidad de F, se verifica que

$$\begin{aligned}
& \frac{1}{2\pi}\int_a^b dt \int_{-\infty}^{+\infty} e^{-i\zeta t}\varphi(\zeta)\exp\left(-\frac{\zeta^2}{2\sigma^2}\right)\, d\zeta \\
= & \int_a^b dt \int_{-\infty}^{+\infty} \frac{\sigma}{\sqrt{2\pi}}\exp\left(-\frac{\sigma^2(x-t)^2}{2}\right)\, F(dx) = P(\xi + \eta_{1/\sigma} \in [a,b]) \\
= & \ P(\xi \in [a,b]).
\end{aligned}$$

cuando $\sigma \to +\infty$, porque $\eta_{1/\sigma} \to 0$ en L^2.

Es decir que φ determina los incrementos de F en todos los intervalos cuyos extremos son puntos de continuidad de F, y por lo tanto, en todos los intervalos. Esto prueba que φ determina a F completamente y, al mismo tiempo, da una fórmula de inversión (cálculo de F a partir de φ), que por el momento no es muy cómoda y es la siguiente.

$$F(b) - F(a) = \lim_{\sigma\to+\infty} \frac{1}{2\pi}\int_a^b dt \int_{-\infty}^{+\infty} e^{-i\zeta t}\varphi(\zeta)\exp\left(-\frac{\zeta^2}{2\sigma^2}\right)\, d\zeta \quad (28)$$

para a, b, $a < b$, puntos de continuidad de F. *En particular*, si $\varphi \in L^1$, se deduce que F tiene densidad continua f y que ésta se calcula mediante la fórmula (más agradable):

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-i\zeta x}\varphi(\zeta)\, d\zeta \quad (29)$$

En efecto, verificar que si $\varphi \in L^1$, entonces se puede pasar al límite bajo el signo integral en (28) aplicando el teorema de convergencia dominada y resulta

$$F(b) - F(a) = \frac{1}{2\pi}\int_a^b dt \int_{-\infty}^{+\infty} e^{-i\zeta t}\varphi(\zeta)\, d\zeta$$

para a, b, $a < b$, puntos de continuidad de F y, por lo tanto, para toda pareja a, b. Esto prueba que F es absolutamente continua con respecto a la medida de Lebesgue y que la densidad está dada por (29).$\square$

Teorema (Lévy-Cramér). Sea $\{F_n\}_{n=1,2,\ldots}$ una sucesión de distribuciones de probabilidad propias en $\mathcal{R}$ $\{\varphi_n\}_{n=1,2,\ldots}$ la sucesión de las transformadas de Fourier respectivas.

Entonces, $F_n \Longrightarrow F$ (propia) si y sólo si $\{\varphi_n\}_{n=1,2,\ldots}$ converge *puntualmente* y la función límite φ es continua en $t = 0$. En ese caso, φ es la transformada de Fourier de F y $\{\varphi_n\}$ converge uniformemente en cada compacto.

Demostración.

Supongamos que $F_n \Longrightarrow F$ (propia). Como la función $x \rightsquigarrow \exp(itx) = e_t(x)$ es continua y acotada, se deduce que para cada $t \in \mathcal{R}$:

$$\varphi_n(t) = \int_{-\infty}^{+\infty} \exp(itx)\ F_n(dx) \to \int_{-\infty}^{+\infty} \exp(itx)\ F(dx) = \varphi(t) \quad (30)$$

y también φ es la transformada de Fourier de F.

Como además, si I es un intervalo compacto de la recta, la familia de funciones $\{e_t(.)\}_{t\in I}$ es equicontinua, se deduce que la convergencia en (30) es uniforme en cada compacto y φ es continua (¡probar cada una de estas afirmaciones!).

Probemos el recíproco, es decir, ahora supondremos que $\{\varphi_n\}_{n=1,2,\ldots}$ converge *puntualmente* y la función límite φ es continua en $t = 0$. Según hemos visto, toda sucesión de distribuciones contiene una subsucesión que converge débilmente. Por lo tanto, para probar que $\{F_n\}$ converge débilmente, alcanzará con probar que todas las subsucesiones de ella que convergen débilmente, tienen el mismo límite, en cuyo caso, éste será el límite débil de la sucesión entera.

Sea entonces $\{F_{n_k}\}$ una subsucesión que converge débilmente y $F_{n_k} \Longrightarrow F^*$ (obsérvese que, a priori, F^* podría ser defectiva). Veremos que F^* es independiente de cuál es la subsucesión considerada y, más aún, que F^* es propia y que tiene a la función límite φ por transformada de Fourier.

Aplicamos la igualdad (27), reemplazando F por F_{n_k} y φ por φ_{n_k}. Poniendo $t = 0$, se obtiene:

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} \varphi_{n_k}(\zeta)\exp\left(-\frac{\zeta^2}{2\sigma^2}\right)\, d\zeta = \int_{-\infty}^{+\infty} \frac{\sigma}{\sqrt{2\pi}}\exp\left(-\frac{\sigma^2 x^2}{2}\right)\, F_{n_k}(dx)$$

En el segundo miembro, el integrando pertenece a $C_0(\mathcal{R})$ por lo tanto, podemos pasar al límite cuando $k \to +\infty$ reemplazando F_{n_k} por F^*.

En el primer miembro, podemos aplicar el teorema de convergencia dominada, ya que $\varphi_{n_k}(\zeta) \to \varphi(\zeta)$ para cada ζ y el valor absoluto del integrando está acotado por $\exp\left(-\frac{\zeta^2}{2\sigma^2}\right)$, que esta en $L^1(\mathcal{R})$. Por lo tanto,

$$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}\varphi(\zeta)\exp\left(-\frac{\zeta^2}{2\sigma^2}\right)\, d\zeta = \int_{-\infty}^{+\infty}\exp\left(-\frac{\sigma^2 x^2}{2}\right)\, F^*(dx) \tag{31}$$

Hacemos ahora $\sigma \downarrow 0$. El segundo miembro tiende a $F^*(\mathcal{R})$, usando el teorema de Beppo Levi.

Como φ es continua en $\zeta = 0$, dado $\varepsilon > 0$ podemos elegir $\delta > 0$, de modo que $|\zeta| < \delta \implies |\varphi(\zeta) - 1| < \varepsilon$ (notar que $\varphi(0) = 1$).

Entonces:

$$\begin{aligned}
& \left|\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}\varphi(\zeta)\exp\left(-\frac{\zeta^2}{2\sigma^2}\right)\, d\zeta - 1\right| \\
= \quad & \left|\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}[\varphi(\zeta) - 1]\exp\left(-\frac{\zeta^2}{2\sigma^2}\right)\, d\zeta\right| \\
\leq \quad & \varepsilon + \frac{2}{\sigma\sqrt{2\pi}}\int_{|\zeta|\geq\delta}\exp\left(-\frac{\zeta^2}{2\sigma^2}\right)\, d\zeta \\
= \quad & \varepsilon + 2.P(|\sigma\xi| \geq \delta)
\end{aligned}$$

donde ξ es una variable aleatoria normal $(0,1)$. Se deduce que el primer miembro de (31) tiende a 1 cuando $\sigma \to 0$. Por lo tanto $F^*(\mathcal{R}) = 1$.

Pero entonces, la convergencia $F_{n_k} \Longrightarrow F^*$ es propia y podemos aplicarle el directo de este mismo teorema, es decir que $\left\{\varphi_{n_k}\right\}$ converge puntualmente a la transformada de Fourier de F^*. Pero, por hipótesis, $\left\{\varphi_{n_k}\right\}$ converge puntualmente a la función φ, de modo que φ no es otra que la transformada de Fourier de F^* y por la vista unicidad, F^* no depende de la subsucesión. Esto termina la demostración.$\square$

7. **Propiedades complementarias de las transformadas de Fourier de medidas de probabilidad**.

- $\int_{-\infty}^{+\infty} |x|^m\, F(dx) < \infty$ para un cierto natural $m \implies \varphi$ es de clase C^m y

$$\varphi^{(m)}(t) = \int_{-\infty}^{+\infty}(ix)^m \exp(itx)F(dx).$$

Se deduce que si F es la distribución de probabilidad de una variable aleatoria ξ, entonces

$$\varphi^{(m)}(0) = i^m E\left(\xi^m\right)$$

En particular, si $E\left(\xi^2\right) < \infty$, entonces $\varphi'(0) = i.E(\xi)$ y $\varphi"(0) = -E(\xi^2)$.

Para probar esto, proceder progresivamente para $m = 1, 2, ...$, escribiendo el cociente incremental y pasando al límite mediante aplicación del teorema de convergencia dominada.

Utilizar la desigualdad (cuya demostración también queda a cargo del lector), válida para $m = 0, 1, 2, ...$ y t real:

$$\left| e^{it} - \left(1 + \frac{it}{1!} + \frac{(it)^2}{2!} + ... + \frac{(it)^m}{m!}\right) \right| \leq \frac{|t|^{m+1}}{(m+1)!} \tag{32}$$

Una manera de probar esta desigualdad es observar que si $g_m(t)$ es la expresión cuyo módulo figura en el primer miembro, entonces:

$$g_m(t) = i \int_0^t g_{m-1}(s)\ ds \quad si\ m \geq 1, \quad g_0(t) = i \int_0^t e^{is} ds$$

- Un recíproco de la propiedad anterior para $m = 2$.

  Si $\exists\ \varphi''(0)$, entonces $\int_{-\infty}^{+\infty} x^2 F(dx) < \infty$.

  Para probar esto, escribir $\varphi = u + iv$, u, v reales. Se tiene:

$$\frac{1 - u(h)}{h^2} = \int_{-\infty}^{+\infty} \frac{1 - \cos(hx)}{h^2 x^2} x^2 F(dx)$$

  Usando que $u'(0) = 0$ y el teorema del valor medio $(0 < \theta < 1)$:

$$\left| \frac{1 - u(h)}{h^2} \right| = \left| \frac{u'(\theta h)}{h} \right| \leq \left| \frac{u'(\theta h)}{\theta h} \right| \to \left| u''(0) \right| \quad \text{cuando } h \to 0.$$

  Aplicando el lema de Fatou:

$$\frac{1}{2} \int_{-\infty}^{+\infty} x^2 F(dx) \leq \underline{\lim}_{h \to 0} \int_{-\infty}^{+\infty} \frac{1 - \cos(hx)}{h^2 x^2} x^2 F(dx) \leq \left| u''(0) \right|$$

  Esto prueba la afirmación.$\square$

- (Lema de Riemann-Lebesgue).

  $f \in L^1, \varphi(t) = \int_{\mathcal{R}} \exp(itx)\ f(x)\ dx \Longrightarrow\ \varphi(t) \to 0$ cuando $|t| \to +\infty$.

- (Fórmula de Plancherel).

  Sea $\varphi \in L^2$ la transformada de Fourier de la densidad f. Entonces $\int_{-\infty}^{+\infty} f^2(x)dx = \frac{1}{2\pi}\int_{-\infty}^{+\infty} |\varphi(t)|^2\, dt$.

  Para probar esto, observar que $|\varphi(t)|^2$ es la transformada de Fourier de la densidad $g(x) = \int_{-\infty}^{+\infty} f(x-y)f(-y)dy$ y aplicar la fórmula de inversión (en $x=0$): $g(0) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} |\varphi(t)|^2\, dt$.

  Más aún, el lector verificará que si $\varphi \in L^2$ es la transformada de Fourier de una medida de probabilidad μ, entonces μ es absolutamente continua con respecto a la medida de Lebesgue y su densidad f verifica lo anterior. Para esto, sea Y una variable aleatoria real, con distribución de probabilidad μ y ξ una variable aleatoria con distribución normal $(0,1)$, independiente de Y. Escribir la densidad f_σ de la variable aleatoria $Y + \sigma\xi$, donde $\sigma > 0$ y luego pasar al límite cuando $\sigma \to 0$.

- (Propiedades aritméticas).

  Sea φ la transformada de Fourier de una distribución de probabilidad propia F.

  Entonces hay tres posibilidades mutuamente excluyentes:

  (a) $|\varphi(t)| < 1 \;\; \forall\, t \neq 0$.

  (b) $|\varphi(t)| = 1 \;\; \forall\, t \in \mathcal{R}$ y en ese caso $F = \delta_{\{b\}}$ para algún $b \in \mathcal{R}$.

  (c) $\exists\, \theta > 0$ tal que $|\varphi(\theta)| = 1$ y $|\varphi(t)| < 1\ \forall\, t \in (0,\theta)$ y en ese caso, φ es periódica de período θ y $\exists$ un número real b tal que $F(x+b)$ es la distribución de una medida concentrada en los puntos de la forma $\left\{\frac{2\pi}{\theta}m : m \text{ entero}\right\}$.(estas medidas se llaman "aritméticas", están concentradas en los puntos de una progresión aritmética).

Para probar esto, supongamos que $|\varphi(t_0)| = 1$ para algún $t_0 > 0$. Sea $\varphi(t_0) = \exp(ib)$, a real. Entonces $\widetilde{\varphi}(t) = \exp(-ib).\varphi(t)$ es también la transformada de Fourier de una distribución de probabilidad (observar que si φ es la transformada de Fourier de la variable aleatoria ξ, entonces $\widetilde{\varphi}$ lo es de $\xi - b$). Entonces:

$$0 = 1 - \widetilde{\varphi}(t_0) = \mathrm{Re}\,[1 - \widetilde{\varphi}(t_0)] = \int_{-\infty}^{+\infty} [1 - \cos(t_0 x)]\, \widetilde{F}(dx)$$

lo cual sólo es posible si $\widetilde{F}$ está concentrada en los puntos de la forma $\left\{\frac{2\pi}{t_0}m : m \text{ entero}\right\}$. Notar que $\widetilde{F}(x) = F(x+b)$. Pero entonces $\widetilde{\varphi}$ es periódica de período t_0 y lo mismo vale para φ.

Si el conjunto $C = \{t : |\varphi(t)| = 1, t > 0\}$ no es vacío, sea $\theta = \inf(C)$. Si $\theta = 0$, φ resulta periódica con períodos positivos arbitrariamente pequeños, y como es continua, es constante e igual a 1.

Si $\theta > 0$ estamos en el caso (c).

2.3 Teorema del Límite Central 1.

Teorema. Sea $\{\xi_n\}_{n=1,2,...}$ una sucesión de variables aleatorias independientes con igual distribución F.

Suponemos que $E(\xi_n) = \mu,\ Var(\xi_n) = \sigma^2$.

Se define: $S_0 = 0,\ \ S_n = \xi_1 + ... + \xi_n \ \ (n \geq 1)$.

Entonces,

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \Longrightarrow N(0,1) \quad \text{cuando} \ \ n \to +\infty$$

(es decir que la distribución de las sumas parciales normalizadas $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ converge débilmente a la distribución normal típica).

Demostración. Sea F_n la distribución de probabilidad de $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ y φ (resp. φ_n) la transformada de Fourier de $\eta = \frac{\xi_1 - \mu}{\sigma}$ (resp. F_n). Notar que $E(\eta) = 0$ y $E(\eta^2) = 1$ y, por lo tanto, $\varphi(\zeta) = 1 - \frac{1}{2}\zeta^2 + o(\zeta^2)$ cuando $\zeta \to 0$.

Aplicando el teorema de Lévy-Cramér bastará probar que

$$\varphi_n(t) \to \exp(-\frac{t^2}{2})$$

cuando $n \to +\infty$, ya que $\exp(-\frac{t^2}{2})$ es la transformada de Fourier de la distribución normal típica.

Usando las propiedades vistas de la transformada de Fourier, se verifica que:

$$\varphi_n(t) = \left[\varphi\left(\frac{t}{\sqrt{n}}\right)\right]^n = \left[1 - \frac{t^2}{2n} + o\left(\frac{t^2}{2n}\right)\right]^n \to \exp(-\frac{t^2}{2}) \ \ \text{cuando } n \to +\infty$$

$\square$.

2.4 Teorema del Límite Central 2.

Teorema de Lindeberg.

Sea $\{\xi_k\}_{k=1,2,...}$ una sucesión de variables aleatorias independientes. Denotamos por F_k (resp. φ_k) la función de distribución (resp. la transformada de Fourier) de ξ_k. Suponemos que $E(\xi_k) = 0,\ \ Var(\xi_k) = \sigma_k^2 < \infty$.

Se define: $S_0 = 0, \quad S_n = \xi_1 + ... + \xi_n \quad (n \geq 1)$ y denotamos $V_n = [Var(S_n)]^{\frac{1}{2}} = \left[\sum_{k=1}^n \sigma_k^2\right]^{\frac{1}{2}}$.

Entonces, si se cumple la condición (llamada "de Lindeberg"):

$$\forall \delta > 0, \quad \theta_n(\delta) = \frac{1}{V_n^2} \sum_{k=1}^{n} E\left(\xi_k^2 \, 1_{\{|\xi_k| \geq \delta V_n\}}\right) \to 0 \quad \text{cuando} \quad n \to +\infty \tag{33}$$

entonces

$$\frac{S_n}{V_n} \Longrightarrow N(0,1) \quad \text{cuando} \quad n \to +\infty$$

Demostración.

La demostración consiste en aplicar el teorema de Lévy Cramér, probando que la transformada de Fourier de $\frac{S_n}{V_n}$ converge puntualmente a $\exp(-\frac{t^2}{2})$, que es la trasformada de Fourier de la distribución normal típica.

Fijemos entonces $t \in \mathcal{R}$ para el resto de la demostración.

Usando las hipótesis sobre F_k, se tiene, usando (32) y la definición de $\theta_n(\delta)$

$$\begin{aligned}
\left|\varphi_k\left(\frac{t}{V_n}\right) - 1\right| &= \left|\int_{-\infty}^{+\infty} \left[\exp\left(\frac{itx}{V_n}\right) - 1 - \frac{itx}{V_n}\right] F_k(dx)\right| \\
&\leq \int_{-\infty}^{+\infty} \left|\exp\left(\frac{itx}{V_n}\right) - 1 - \frac{itx}{V_n}\right| F_k(dx) \\
&\leq \int_{-\infty}^{+\infty} \frac{1}{2}\left(\frac{tx}{V_n}\right)^2 F_k(dx) = \frac{t^2\sigma_k^2}{2V_n^2} \\
&\leq \frac{t^2}{2}\left(\delta^2 + \theta_n(\delta)\right)
\end{aligned}$$

Como $\delta > 0$ se puede elegir de manera arbitraria, la condición de Lindeberg implica que

$$\sup_{1 \leq k \leq n} \left|\varphi_k\left(\frac{t}{V_n}\right) - 1\right| \to 0 \quad \text{cuando} \quad n \to +\infty. \tag{34}$$

Por lo tanto, si n es suficientemente grande, está bien definida la rama principal de $\log\left[\varphi_k\left(\frac{t}{V_n}\right)\right]$ $\forall k = 1, ..., n$ y podemos escribir la transformada de Fourier de $\frac{S_n}{V_n}$ mediante:

$$\varphi_{\frac{S_n}{V_n}}(t) = \varphi_{S_n}(\frac{t}{V_n}) = \prod_{k=1}^{n} \varphi_k(\frac{t}{V_n}) = \exp\left[\sum_{k=1}^{n} \log\left(\varphi_k(\frac{t}{V_n})\right)\right]$$

En lo que sigue, probamos que el exponente tiende a $-\frac{t^2}{2}$ cuando $n \to +\infty$.

Primero, de aquí en adelante, elegimos n suficientemente grande de modo que

$$\sup_{1 \leq k \leq n} \left| \varphi_k \left(\frac{t}{V_n} \right) - 1 \right| < \frac{1}{2}$$

lo cual es posible en virtud de (34). Por lo tanto, $\forall k = 1, ..., n$ podemos escribir el desarrollo de Taylor

$$\log \left(\varphi_k(\frac{t}{V_n}) \right) = - \left[1 - \varphi_k(\frac{t}{V_n}) \right] + A(k,n,t) \ \left[1 - \varphi_k(\frac{t}{V_n}) \right]^2$$

donde $A(k,n,t)$ está acotado por una constante universal. Entonces

$$\sum_{k=1}^{n} \log \left(\varphi_k(\frac{t}{V_n}) \right) = \sum_{k=1}^{n} \left\{ - \left[1 - \varphi_k(\frac{t}{V_n}) \right] + A(k,n,t) \ \left[1 - \varphi_k(\frac{t}{V_n}) \right]^2 \right\}$$

Usando las cotas anteriores,

$$\sum_{k=1}^{n} \left[1 - \varphi_k(\frac{t}{V_n}) \right]^2 \leq \frac{1}{2} t^2 \left[\sup_{1 \leq k \leq n} \frac{\sigma_k^2}{V_n^2} \right] \sum_{k=1}^{n} \left| 1 - \varphi_k(\frac{t}{V_n}) \right|$$

Asimismo, desarrollando por Taylor y usando la hipótesis sobre la distribución F_k :

$$1 - \varphi_k(\frac{t}{V_n}) = \frac{t^2 \sigma_k^2}{2 V_n^2} - R_{k,n}$$

y reemplazando:

$$\sum_{k=1}^{n} - \left[1 - \varphi_k(\frac{t}{V_n}) \right] = -\frac{1}{2} t^2 + \sum_{k=1}^{n} R_{k,n}$$

$$\sum_{k=1}^{n} \left| 1 - \varphi_k(\frac{t}{V_n}) \right| \leq \frac{1}{2} t^2 + \sum_{k=1}^{n} |R_{k,n}|$$

Si probamos que

$$\sum_{k=1}^{n} |R_{k,n}| \to 0 \quad \text{cuando} \quad n \to +\infty \tag{35}$$

resulta entonces la tesis con un cálculo elemental de límites. Para esto, fijemos $\varepsilon > 0$. Tenemos:

$$|R_{k,n}| = \int_{-\infty}^{+\infty} \left[\exp\left(\frac{itx}{V_n}\right) - 1 - \frac{itx}{V_n} - \frac{1}{2}\left(\frac{itx}{V_n}\right)^2 \right] F_k(dx) \leq |S_{k,n}| + |T_{k,n}|$$

donde $S_{k,n} = \int_{|x| \geq \varepsilon V_n}$ y $T_{k,n} = \int_{|x| < \varepsilon V_n}$.

Para el primer término:

$$\begin{aligned} |S_{k,n}| &\leq \int_{|x| \geq \varepsilon V_n} \left[\left| \exp\left(\frac{itx}{V_n}\right) - 1 - \frac{itx}{V_n} \right| + \frac{1}{2}\left(\frac{tx}{V_n}\right)^2 \right] F_k(dx) \\ &\leq \frac{t^2}{V_n^2} \int_{|x| \geq \varepsilon V_n} x^2 F_k(dx) \end{aligned}$$

aplicando la desigualdad (32) para $m = 2$.

Para el segundo término, aplicamos la misma desigualdad para $m = 3$:

$$\begin{aligned} |T_{k,n}| &\leq \int_{|x| < \varepsilon V_n} \left| \frac{1}{3!} \frac{t^3 x^3}{V_n^3} \right| F_k(dx) \\ &\leq \frac{1}{3!} |t|^3 \varepsilon \int_{|x| < \varepsilon V_n} \frac{x^2}{V_n^2} F_k(dx) \leq \frac{1}{3!} |t|^3 \varepsilon \frac{\sigma_k^2}{V_n^2} \end{aligned}$$

Resumiendo,

$$\sum_{k=1}^{n} |R_{k,n}| \leq \frac{t^2}{V_n^2} \sum_{k=1}^{n} \int_{|x| \geq \varepsilon V_n} x^2 F_k(dx) + \frac{1}{3!} |t|^3 \varepsilon = t^2 \ \theta_n(\delta) + \frac{1}{3!} |t|^3 \varepsilon$$

lo que implica, en razón de la condición de Lindeberg:

$$\overline{\lim}_{n \to +\infty} \sum_{k=1}^{n} |R_{k,n}| \leq \frac{1}{3!} |t|^3 \varepsilon$$

Como t es fijo y $\varepsilon > 0$ es arbitrario, se sigue el resultado.□

2.4.1 Recíproco parcial del Teorema de Lindeberg.

Teorema.

Sea $\{\xi_k\}_{k=1,2,\dots}$ una sucesión de variables aleatorias independientes. Suponemos que $E(\xi_k) = 0$, $Var(\xi_k) = \sigma_k^2 < \infty$ y usamos la misma notación que en el teorema anterior.

Si se verifica que

$$\frac{S_n}{V_n} \Longrightarrow N(0,1) \quad \text{cuando} \;\; n \to +\infty$$

y ademas $\frac{\sigma_n}{V_n} \to 0$ y $V_n \to +\infty$, entonces se cumple la condición de Lindeberg (33).

Demostración.

Primero, probar que

$$\sup_{1 \le k \le n} \frac{\sigma_k^2}{V_n^2} \to 0 \qquad \text{cuando} \;\; n \to +\infty$$

y deducir de aquí (mediante cálculos iguales a los de la demostración del teorema de Lindeberg) que

$$\sum_{k=1}^{n} \log\left(\varphi_k(\frac{t}{V_n})\right) + \sum_{k=1}^{n}\left[1 - \varphi_k(\frac{t}{V_n})\right] \to 0 \quad \text{cuando} \;\; n \to +\infty$$

El primer término tiende a $-\frac{t^2}{2}$, de modo que tomando partes reales, se tiene, para $\delta > 0$ y t fijos, que

$$\begin{aligned} &\frac{t^2}{2} - \sum_{k=1}^{n} \int_{|x|<\delta V_n} \left(1 - \cos(\frac{tx}{V_n})\right) F_k(dx) \\ = \quad &\sum_{k=1}^{n} \int_{|x|\ge\delta V_n} \left(1 - \cos(\frac{tx}{V_n})\right) F_k(dx) + o(1) \end{aligned}$$

cuando $n \to +\infty$. En el segundo miembro, el integrando está acotado por $\frac{2x^2}{\delta^2 V_n^2}$ y en el primer miembro, el integrando está acotado por $\frac{t^2 x^2}{2V_n^2}$. Dividiendo por $\frac{t^2}{2}$ resulta que

$$\theta_n(\delta) \le \frac{4}{\delta^2 t^2} + o(1)$$

Como se puede elegir t arbitrariamente grande, esto prueba que $\theta_n(\delta) \to 0$.□

2.4.2 Condición de Liapunov.

Sea

$$m_{\alpha,k} = E\left(|\xi_k|^{\alpha}\right)$$

El momento de orden α de la k-ésima variable de la sucesión de sumandos dada.

La "condición de Liapunov" es que

$$\exists \alpha > 2 \text{ tal que } m_{\alpha,k} < \infty \forall k \text{ y además } \sum_{k=1}^{n} m_{\alpha,k} = o(V_n^{\alpha}) \quad \text{cuando} \quad n \to +\infty \tag{36}$$

El lector podrá comprobar que la condición de Liapunov implica la condición de Lindeberg y, por lo tanto, que se cumple el Teorema del Límite Central.

En muchas situaciones, es más cómodo verificar la condición de Liapunov que la de Lindeberg.

2.5 Teorema del Límite Central 3. Martingalas.

En esta sección, demostraremos un Teorema del Límite Central de tipo Lindeberg, pero para sumas parciales de variables que no tienen porqué ser independientes, con tal que esas sumas parciales sean una martingala. Esto abre un extenso capítulo relativo a la convergencia débil de las distribuciones de las sumas parciales de variables dependientes, del que solamente daremos el resultado contenido en esta sección.

Teorema (del Límite Central para Martingalas).

Sea $\{M_n\}_{n=0,1,2,...}$ una martingala de cuadrado integrable con respecto a la filtración $\{\mathfrak{F}_n\}_{n=0,1,2,...}$.

Para cada $n = 0, 1, 2, ...$ ponemos $\Delta_n = M_{n+1} - M_n$ y $\sigma_n^2 = E\left(\Delta_n^2/\mathfrak{F}_n\right)$.

Suponemos además que existe una constante positiva K tal que $E\left(|\Delta_n|^3/\mathfrak{F}_n\right) \leq K \ \forall n$.

Denotamos por $\langle M \rangle_n = \sum_{k=0}^{n-1} \sigma_k^2$ el compensador de la submartingala $\{M_n^2\}_{n=0,1,2,...}$.

Entonces, si la sucesión de variables aleatorias $\left\{\frac{1}{n}\langle M \rangle_n\right\}$ converge en probabilidad a la constante $\sigma^2 > 0$, se cumple que

$$\frac{1}{\sqrt{n}} M_n \implies N(0, \sigma^2)$$

Demostración. Es inmediato que no hay inconveniente en suponer que $M_0 = 0$.

Del mismo modo, alcanza con considerar el caso $K = 1$ (si no es así, multiplicar la martingala por una constante apropiada).

Observemos además que, usando la desigualdad de Jensen para esperanzas condicionales, es

$$\sigma_n^3 = \left[E\left(\Delta_n^2/\mathfrak{F}_n\right)\right]^{\frac{3}{2}} \leq E\left(|\Delta_n|^3/\mathfrak{F}_n\right) \leq 1$$

de modo que $0 \leq \sigma_n \leq 1 \ \forall n$ y por lo tanto, también se cumple que $0 \leq \frac{1}{n} \langle M \rangle_n \leq 1 \ \forall n$.

Aplicamos el teorema de Lévy-Cramér y por lo tanto, basta probar que para cada $t \in \mathcal{R}$ fijo, se verifica:

$$\lim_{n \to +\infty} E \left(\exp \left[it \frac{M_n}{\sqrt{n}} + \frac{t^2 \sigma^2}{2} \right] \right) = 1. \tag{37}$$

Para probar (37), veamos que existe una constante universal K_1 tal que $\forall t \in [-1, 1]$ se verifica

$$E \left(\exp \left[itM_n + \frac{t^2 \langle M \rangle_n}{2} \right] \right) = 1 + v_n \quad \text{con} \quad |v_n| \leq nK_1 t^3 \exp(nt^2) \tag{38}$$

Reemplazando en (38) t por $\frac{t}{\sqrt{n}}$ resulta que para cada $t \in \mathcal{R}$:

$$\lim_{n \to +\infty} E \left(\exp \left[it \frac{M_n}{\sqrt{n}} + \frac{t^2 \langle M \rangle_n}{2n} \right] \right) = 1$$

y (37) resulta de

$$E \left(\left| \exp \left[\frac{t^2 \sigma^2}{2} \right] - \exp \left[\frac{t^2 \langle M \rangle_n}{2n} \right] \right| \right) \to 0 \quad \text{cuando} \quad n \to +\infty.$$

Esto último se deduce, a su vez, de que para t fijo,

$$Z_n(t) = \left| \exp \left[\frac{t^2 \sigma^2}{2} \right] - \exp \left[\frac{t^2 \langle M \rangle_n}{2n} \right] \right|$$

tiende a cero en probabilidad y que la sucesión de variables aleatorias $\{Z_n(t)\}$ está acotada uniformemente por la constante $\exp(\frac{t^2}{2})$, de modo que es posible aplicar el teorema de convergencia dominada.

De modo que se trata de probar (38).

Para ello, usamos la fórmula de Taylor para probar que si $x, t \in R$, $y \in [0, 1]$, entonces

$$f(t) = \exp\left(itx + t^2 y\right) = 1 + itx - \frac{t^2 x^2}{2} + t^2 y + R(t, x, y) \tag{39}$$

donde $|R(t, x, y)| \leq K_2 (1 + |x|^3) t^3 \exp(t^2)$.

[Éste es un cálculo inmediato, ya que $f'(t) = (ix + 2ty) f(t), f''(t) = \left[(ix + 2ty)^2 + 2y\right] f(t)$, $f'''(t) = \left[(ix + 2ty)^3 + 6y(ix + 2ty)\right] f(t)$]

Fijemos $t \in \mathcal{R}$ en lo que sigue.

Sean:

$$\begin{aligned} Y_n &= \exp\left[itM_n + \frac{t^2 \langle M \rangle_n}{2}\right] \\ Z_n &= \exp\left[it\Delta_n + \frac{t^2\sigma_n^2}{2}\right] \end{aligned}$$

de modo que $Y_{n+1} = Y_n Z_n$. Usando (39) y teniendo en cuenta las hipótesis:

$$\begin{aligned} E\left(Y_{n+1}/\mathfrak{F}_n\right) &= Y_n E\left(Z_n/\mathfrak{F}_n\right) \\ &= Y_n \left[E\left(1 + it\Delta_n - \frac{t^2}{2}\left(\Delta_n^2 - \sigma_n^2 +\right) + R(t, \Delta_n, \sigma_n^2)/\mathfrak{F}_n\right)\right] \\ &= Y_n \left[1 + E\left(R(t, \Delta_n, \sigma_n^2)/\mathfrak{F}_n\right)\right] = Y_n \left[1 + \alpha_n\right] \end{aligned}$$

donde $|\alpha_n| \leq 2K_2 t^3 \exp(t^2)$.

Ponemos ahora $K_1 = 2K_2$ y probamos (38) por inducción completa en n.

Observar que

$$1 + v_{n+1} = E\left(Y_n \left[1 + \alpha_n\right]\right) = 1 + v_n + E(Y_n \alpha_n)$$

y que $|Y_n| \leq \exp\left(\frac{t^2 n}{2}\right)$. Por lo tanto, admitiendo la desigualdad para n :

$$|v_{n+1}| \leq nK_1 t^3 \exp(nt^2) + \exp\left(\frac{t^2 n}{2}\right) K_1 t^3 \exp(t^2) \leq (n+1)K_1 t^3 \exp(nt^2)$$

Esto prueba (38) y, por lo tanto, concluye la demostración.□

Ejemplo. Sea $\{\xi_k\}_{k=1,2,\ldots}$ una sucesión de variables aleatorias independientes, que supondremos uniformemente acotadas por una cierta constante C y $S_n = \xi_1 + \ldots + \xi_n \quad (n \geq 1)$.

Suponemos que $E(\xi_k) = \mu \neq 0$, $Var(\xi_k) = 1$.

Consideremos la sucesión $\{M_n\}$, definida por:

$$M_1 = S_1, \quad M_n = M_{n-1} + \frac{1}{n-1} S_{n-1}(\xi_n - \mu) \quad \text{para} \;\; n \geq 2 \qquad (40)$$

Es inmediato verificar que $\{M_n\}_{n=1,2,\ldots}$ es una martingala con respecto a la filtración engendrada por la sucesión $\{\xi_k\}$. Dado que la sucesión $\left\{\frac{S_n}{n}\right\}$ es acotada por la constante C, se deduce que $E\left(|\Delta_n|^3 /\mathfrak{F}_n\right) \leq C^4$.

Por otra parte,

$$\sigma_n^2 = E\left(\Delta_n^2/\mathfrak{F}_n\right) = E\left(\left(\frac{S_n}{n}\right)^2 (\xi_n - \mu)^2/\mathfrak{F}_n\right) = \left(\frac{S_n}{n}\right)^2$$

En virtud de la ley fuerte de los grandes números, casi seguramente $\frac{S_n}{n} \to \mu$ cuando $n \to +\infty$. Por lo tanto $\frac{1}{n}\langle M\rangle_n = \frac{1}{n}\sum_{k=1}^n \sigma_k^2 \to \mu^2 > 0$ y el teorema anterior implica que $\frac{1}{\sqrt{n}}M_n \implies N(0,\mu^2)$.

No parece obvio reducir este tipo de resultado a las sumas de variables aleatorias independientes.

Obsérvese además que este ejemplo se puede extender de varias maneras, a modelos más generales, en los que se obtiene un resultado similar.

Por ejemplo, si en lugar de (40) se pone

$$M_1 = S_1, \quad M_n = M_{n-1} + H\left(\frac{S_{n-1}}{n-1}\right)(\xi_n - \mu) \quad \text{para} \;\; n \geq 2 \qquad (41)$$

y en lugar de la condición $\mu \neq 0$ ponemos $H(\mu) \neq 0$, se obtiene el mismo resultado, cambiando μ^2 por $[H(\mu)]^2$.

Las ecuaciones de recurrencia (40) o (41) son casos particulares de "modelos autorregresivos" no lineales, en los que el estado del sistema en el instante n se obtiene a partir del estado en el instante $n-1$ más una perturbación aleatoria. Ésta última contiene al "ruido" o "innovación" $(\xi_n - \mu)$ - independiente de lo ocurrido hasta el instante anterior - amplificado por una función del pasado (que en algunos contextos se denomina la "volatilidad"), que para el ejemplo es de la forma $H\left(\frac{S_{n-1}}{n-1}\right)$, pero que podría ser distinta.

2.6 Velocidad de convergencia. Teorema de Berry-Essen.

Existe una gran variedad de teoremas que permiten estimar la velocidad de convergencia en el Teorema del Límite Central y otros resultados relacionados. Nos limitamos aquí a un sólo resultado. Por mayores detalles, se puede consultar, por ejemplo, el libro de V. Petrov.

Teorema (Berry-Essen). Sea $\{\xi_n\}_{n=1,2,...}$ una sucesión de variables aleatorias independientes con igual distribución F, $E(\xi_n) = 0$, $Var(\xi_n) = \sigma^2 < \infty$.

Como antes, $S_0 = 0$, $S_n = \xi_1 + ... + \xi_n$ $(n \geq 1)$.

Suponemos además que $\mu_3 = E(|\xi_n|^3) < \infty$

Sea F_n la función de distribución de probabilidad de $\frac{S_n}{\sigma\sqrt{n}}$ y Φ la función de distribución normal típica, i.e.

$$F_n(x) = P(\frac{S_n}{\sigma\sqrt{n}} \leq x) \quad \text{y} \quad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-\frac{y^2}{2})\ dy$$

Entonces,

$$\sup_{x\in\mathcal{R}} |F_n(x) - \Phi(x)| \leq \frac{33}{4} \frac{\mu_3}{\sigma^3\sqrt{n}}$$

La demostración está basada en el Lema siguiente.

Lema. Sea F la distribución de probabilidad de una variable aleatoria ξ a valores en $\mathcal{R}$ y φ su transformada de Fourier.

Sea $G : \mathcal{R} \to \mathcal{R}$ una función de clase C^1, $G(-\infty) = 0, G(+\infty) = 1, G' = g, \|g\|_\infty \leq m$. Denotamos por γ la transformada de Fourier de g, i.e., $\gamma(t) = \int_{-\infty}^{+\infty} \exp(itx)g(x)dx$ que suponemos bien definida y, además, $\gamma \in C^1, \gamma(0) = 1, \gamma'(0) = 0$.

Entonces, para todo $T > 0$:

$$\sup_{x\in\mathcal{R}} |F(x) - G(x)| \leq \frac{1}{\pi} \int_{-T}^{T} \left| \frac{\varphi(t) - \gamma(t)}{t} \right| \ dt + \frac{24.m}{\pi T} \tag{42}$$

Demostración.

Paso 1. Sean w_1 la función (”de Slepian”)

$$w_1(t) = (1 - |t|) \vee 0$$

y para $T > 0$,

$$w_T(t) = w_1\left(\frac{t}{T}\right)$$

w_T es la transformada de Fourier de la medida de probabilidad μ_T cuya densidad es

$$v_T(x) = \frac{1 - \cos(Tx)}{\pi T x^2}.$$

Sean f_T, g_T las densidades de las convoluciones $F_T = \mu_T * F$ y $G_T = \mu_T * G$ respectivamente y φ_T, γ_T sus transformadas de Fourier. Se tiene

$$\varphi_T = w_T.\varphi \quad \gamma_T = w_T.\gamma$$

y usando la fórmula de inversión:

$$f_T(x) - g_T(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(-ixt)\,[\varphi(t) - \gamma(t)]\;\; w_T(t)\; dt \tag{43}$$

De aquí se deduce la igualdad:

$$F_T(x) - G_T(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(-ixt)\, \frac{\varphi(t) - \gamma(t)}{-it}\; w_T(t)\; dt \tag{44}$$

ya que derivando ambos miembros de (44) se obtiene ambos miembros de (43) respectivamente, y además, ambos miembros de (44) se anulan en $-\infty$: el primero por definición de las funciones de distribución y el segundo en virtud del Lema de Riemann-Lebesgue.

Se deduce que

$$|F_T(x) - G_T(x)| \leq \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left| \frac{\varphi(t) - \gamma(t)}{t} \right| \;\; dt \tag{45}$$

Paso 2. Usamos las siguientes notaciones:

$$\Delta = F - G,\ \Delta_T = F_T - G_T,\ \eta = \|\Delta\|_\infty,\ \eta_T = \|\Delta_T\|_\infty$$

Se tiene:

$$\begin{aligned} \Delta_T(t) &= \int_{-\infty}^{+\infty} \Delta(t-x) v_T(x) dx \\ &\geq -\eta \int_{|x|>h} v_T(x) dx + \int_{-h}^{+h} \Delta(t-x) v_T(x) dx \end{aligned}$$

donde elegiremos luego $h > 0$.

Por un lado:

$$\int_{|x|>h} v_T(x) dx \leq \frac{4}{\pi T} \int_h^{+\infty} \frac{dx}{x^2} = \frac{4}{\pi T h} \implies \int_{|x|\leq h} v_T(x) dx \ \geq 1 - \frac{4}{\pi T h}.$$

Como $\Delta(+\infty) = \Delta(-\infty) = 0$ y Δ posee límites laterales, se deduce que $\exists x_0$ tal que $\eta = \Delta(x_0^+)$ ó $\eta = \Delta(x_0^-)$. Entonces, si $s > 0$:

$$\Delta(x_0 + s) - \eta = F(x_0 + s) - F(x_0) + G(x_0 + s) - G(x_0) \geq -ms$$

(aquí, en cada caso, cuando corresponda habrá que poner o bien x_0, o bien x_0^+ o bien x_0^-).

Sea $t = x_0 + h$. Si $|x| < h \implies t - x > t - h = x_0$ y si $s = t - x - x_0$:

$$\begin{aligned} \eta_T &\geq \Delta_T(t) \geq -\eta\frac{4}{\pi T h} + \int_{-h}^{+h} [\eta - m(t - x - x_0)]\, v_T(x)dx \\ &= -\eta\frac{4}{\pi T h} + [\eta - mh] \int_{-h}^{+h} v_T(x)dx \\ &\geq -\eta\frac{4}{\pi T h} + [\eta - mh]\left(1 - \frac{4}{\pi T h}\right) \end{aligned}$$

Eligiendo $h = \frac{\eta}{2m}$ resulta $\eta_T \geq \frac{\eta}{2} - \frac{12.m}{\pi T}$ y usando (45):

$$\eta \leq 2\eta_T + \frac{24.m}{\pi T} \leq \frac{1}{\pi}\int_{-\infty}^{+\infty} \left|\frac{\varphi(t) - \gamma(t)}{t}\right| \; dt + \frac{24.m}{\pi T}$$

lo que prueba el Lema.□

Demostración del Teorema de Berry-Essen.

Aplicamos el Lemma, reemplazando F por F_n y G por la distribución normal Φ. Entonces

$$\sup_{x\in\mathcal{R}} |F_n(x) - \Phi(x)| \leq \frac{1}{\pi}\int_{-T}^{T} \left|\varphi^n\left(\frac{t}{\sigma\sqrt{n}}\right) - \exp\left(-\frac{t^2}{2}\right)\right| \frac{1}{|t|}dt + \frac{24}{T\pi\sqrt{\pi}} \tag{46}$$

Usando la desigualdad (verficación sencilla a cargo del lector) $\sigma^6 < \mu_3^2$, resulta que $|1 - \varphi(t)| \leq \frac{1}{2}\sigma^2 t^2$. Por lo tanto, $|t| < \frac{\sigma^2}{\mu_3}$ implica que $|1 - \varphi(t)| \leq \frac{1}{2}$ y también la validez del desarrollo en serie:

$$-\log\varphi(t) = \sum_{k=1}^{\infty} \frac{1}{k}\,[1 - \varphi(t)]^k$$

donde log representa la rama principal del logaritmo.

Se deduce que si $|t| < \frac{\sigma^2}{\mu_3}$, entonces

$$\left|\log\varphi(t) + \frac{1}{2}\sigma^2 t^2\right| \leq \frac{1}{6}\mu_3\,|t|^3 + \frac{1}{4}\sigma^4 t^4 \leq \frac{5}{12}\mu_3\,|t|^3$$

Para terminar, poner en (46) $T = \frac{\sigma^3}{\mu_3}\sqrt{n}$ y usar la acotación simple $\left|e^t - 1\right| \leq |t|\,.e^{|t|}$ para obtener la conclusión del teorema (estos cálculos finales quedan a cargo del lector).□

2.7 Convergencia débil de medidas de probabilidad en $\mathcal{R}^d, d > 1$.

Toda la teoría anterior se extiende sin mayores dificultades a $\mathcal{R}^d$ para $d > 1$. (Convergencia débil de medidas de probabilidad, transformada de Fourier, teoremas del límite central). No habremos de dar detalles aquí, sino indicar solamente algunos puntos, sin demostraciones, que por otra parte son enteramente similares a las que hemos expuesto en dimensión 1.

1. Para la definición de convergencia débil (todas las medidas son medidas de Borel en $\mathcal{R}^d$) podemos adoptar la siguiente definición.

   Una sucesión $\{F_n\}_{n=1,2,\ldots}$ de medidas de probabilidad en $\mathcal{R}^d$ converge débilmente a F (medida de probabilidad propia) si para todo conjunto de Borel B tal que $F(\partial B) = 0$, se tiene que $F_n(B) \to F(B)$ cuando $n \to +\infty$. Se denota $F_n \Longrightarrow F$.

   Se verifica que $F_n \Longrightarrow F$ (propia) si y sólo si $\int_{\mathcal{R}^d} f\ dF_n \to \int_{\mathcal{R}^d} f\ dF$ $\forall f$ continua y acotada.

2. Se define la transformada de Fourier de F (o de una variable aleatoria X a valores en $\mathcal{R}^d$ que tiene la distribución F - es decir que para todo boreliano B es $P(X \in B) = F(B)$ - mediante:

   $$\varphi_X(t) = \varphi(t) = \int_{\mathcal{R}^d} \exp\left[i\left\langle t, x\right\rangle\right]\ F(dx), \quad t \in \mathcal{R}^d$$

   Aquí, $\langle , \rangle$ denota el producto escalar usual en $\mathcal{R}^d$.

   Un caso particular importante es el de la distribución normal en $\mathcal{R}^d$. Se dice que X es una variable aleatoria a valores en $\mathcal{R}^d$ con distribución normal, si su transformada de Fourier φ_X tiene la forma:

   $$\varphi_X(t) = \exp\left[i\left\langle t, \mu\right\rangle - \frac{1}{2}\left\langle \Sigma t, t\right\rangle\right]$$

   donde $\mu \in \mathcal{R}^d$ y Σ es una matriz simétrica, definida no negativa con d filas y d columnas.

   Se verifica sin mucha dificultad que, en este caso

   $$E(X) = \mu, \quad Var(X) = \Sigma$$

   donde las definiciones de la esperanza y la varianza para variables aleatorias vectoriales es la siguiente, siempre y cuando existan cada una de las esperanzas que figuran:

Si en una base ortonormal de $\mathcal{R}^d$ el vector X se escribe $X = (X_1, ..., X_d)^T$, entonces

$$E(X) = (E(X_1), ..., E(X_d))^T$$

$$Var(X) = ((\Sigma_{jk}))_{j,k=1,\ldots,d} \quad \text{con} \quad \Sigma_{jk} = E\left[(X_j - E(X_j))(X_k - E(X_k))\right]$$

o lo que es lo mismo

$$Var(X) = E\left[(X - E(X))(X - E(X)^T)\right].$$

3. Se verifican propiedades análogas a las que vimos en dimensión 1, incluyendo los teoremas de unicidad, fórmula de inversión y pasaje al límite (teorema de Lévy-Cramér). En particular, la fórmula de inversión más cómoda adquiere la forma siguiente: si $\varphi \in L^1(\mathcal{R}^d, \lambda_d)$ (donde λ_d denota la medida de Lebesgue), entonces F tiene una densidad continua f que está dada por la fórmula

$$f(x) = \frac{1}{(2\pi)^d} \int_{\mathcal{R}^d} \exp\left[-i \langle t, x \rangle\right] \ \varphi(t) \ dt$$

En el caso particular de la distribución normal con parámetros μ, Σ, se verifica fácilmente que si Σ es definida positiva, entonces vale esta fórmula de inversión y

$$f(x) = \frac{1}{(2\pi)^{\frac{d}{2}} \left[\det(\Sigma)\right]^{\frac{1}{2}}} \exp\left[(x - \mu)^T \Sigma^{-1} (x - \mu)\right]$$

y si Σ, en cambio, es singular, entonces la distribución no tiene densidad y está concentrada en una variedad afín de dimensión menor que d.

4. Con los mismos métodos, se prueba el teorema del límite central. Por ejemplo, en el caso de sumas de variables aleatorias independientes con igual distribución, se tiene el enunciado siguiente: Sea $\xi_1, \xi_2, ...$ una sucesión de variables aleatorias independientes a valores en $\mathcal{R}^d$, con igual distribución, $E(\xi_n) = 0$, $Var(\xi_n) = \Sigma$. Definimos la sucesión de sumas parciales $S_n = \xi_1 + ... + \xi_n$, para $n \geq 1$. Entonces, $\frac{S_n}{\sqrt{n}}$ converge débilmente a la distribución normal en $\mathcal{R}^d$, con media $\mu = 0$ y matriz de varianza Σ.

2.8 Grandes desviaciones.

Sea $\{\xi_k\}_{k=1,2,...}$ una sucesión de variables aleatorias independientes y supongamos que $E(\xi_k) = \mu, Var(\xi_k) = \sigma^2$ y pongamos como antes $S_n = \xi_1 + ... + \xi_n$ $(n \geq 1)$. Entonces, bajo ciertas condiciones, hemos probado que la distribución de las sumas parciales normalizadas $\frac{S_n - \mu}{\sigma\sqrt{n}}$ converge débilmente a la distribución normal típica.

Como sabemos esto significa que si $a, b \in \mathcal{R}, a < b$, son números reales **fijos** entonces

$$P\left(a < \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq b\right) \to \frac{1}{\sqrt{2\pi}} \int_a^b \exp\left(-\frac{x^2}{2}\right)\ dx \tag{47}$$

cuando $n \to +\infty$.

Este tipo de resultado es insuficiente para muchas aplicaciones relevantes. Por ejemplo, sabemos que en virtud de la ley débil de los grandes números $\forall \varepsilon > 0$,

$$\alpha_n(\varepsilon) = P\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \to 0$$

cuando $n \to +\infty$. Se suele decir que $\alpha_n(\varepsilon)$ es la probabilidad del "intervalo de confianza de radio ε para la media μ".

Dicho de otro modo, en esta aplicación estadística sencilla de la ley de los grandes números, estaremos interesados en saber, para cada $\varepsilon > 0$, el valor de $\alpha_n(\varepsilon)$, o por lo menos, con qué velocidad $\alpha_n(\varepsilon)$ tiende a cero cuando $n \to +\infty$. Nos interesa saber de que tamaño deberíamos elegir n (es decir, de que tamaño debe ser la muestra que habremos de observar) para que la probabilidad de cometer un error mayor que ε, cuando reemplazamos la media teórica μ por la media empírica $\frac{S_n}{n}$, no supere un valor prestablecido.

Si intentamos aplicar el Teorema del Límite Central para calcular - o aproximar - el valor de $\alpha_n(\varepsilon)$, vemos que

$$\alpha_n(\varepsilon) = 1 - P\left(-\frac{\varepsilon}{\sigma}\sqrt{n} < \frac{S_n - n\mu}{\sigma\sqrt{n}} < \frac{\varepsilon}{\sigma}\sqrt{n}\right)$$

y deberíamos poder calcular el equivalente de esta expresión cuando $n \to +\infty$, con la complicación de que $a = -\frac{\varepsilon}{\sigma}\sqrt{n}$, $b = \frac{\varepsilon}{\sigma}\sqrt{n}$ no son fijos, sino que varían con n. En estas condiciones, no es más cierto de que los complementos a 1 de ambos miembros de (47) sean infinitésimos equivalentes y debemos introducir otros métodos para estudiar el problema.

En esta sección habremos de limitarnos a enunciar y probar un teorema sobre esta cuestión, que es un tema vasto y con una gran diversidad de aplicaciones.

Preliminares. Transformada de Cramér (también llamada de Fenchel-Legendre).

Sea F la distribución de probabilidad de una variable aleatoria a valores reales X. En lo que sigue, supondremos que $E(X) = m$ y que la distribución F no se reduce a la delta de Dirac en el punto m.

Se define la transformada de Laplace de la distribución F mediante

$$\Phi_F(t) = E(\exp(tX)) = \int_{-\infty}^{+\infty} \exp(tx)\ F(dx)$$

y denotamos por $J = \{t \in \mathcal{R} : \Phi_F(t) < \infty\}$. Es claro que $\Phi_F(t) > 0\ \forall t \in \mathcal{R}$.

Es claro que J es un intervalo, que en principio puede reducirse al punto $\{0\}$.

Habremos de suponer de aquí en adelante, a los efectos de simplificar la exposición, que si denotamos $\alpha_F = \inf(J)$, $\beta_F = \sup(J)$, entonces el intervalo abierto $(\alpha_F,\ \beta_F)$ no es vacío y contiene al origen. Puede ocurrir que $\alpha_F = -\infty$ o $\beta_F = +\infty$.

El lector verificará sin dificultad que esta hipótesis implica que todos los momentos de F son finitos.

También definimos $\Psi_F : (\alpha_F,\ \beta_F) \to \mathcal{R}$ mediante $\Psi_F(t) = \log[\Phi_F(t)]$. Entonces la transformada de Cramér de F es la función $h_F : \mathcal{R} \to \overline{\mathcal{R}}$

$$h_F(a) = \sup\{at - \Psi_F(t) : t \in (\alpha_F,\ \beta_F)\}$$

Obsérvese que $h_F(a)$ puede valer $+\infty$.

Veamos algunas propiedades básicas de las funciones así definidas.

- Ψ es estrictamente convexa en $(\alpha_F,\ \beta_F)$. Es claro que Ψ_F es infinitamente derivable (verificación a cargo del lector) y que

$$\Psi_F''(t) = \frac{\Phi_F''(t)\Phi_F(t) - [\Phi_F'(t)]^2}{[\Phi_F(t)]^2}$$

Además,

$$\begin{aligned}&\Phi_F''(t)\Phi_F(t) - \left[\Phi_F'(t)\right]^2 \qquad (48)\\ = &\int_{-\infty}^{+\infty} x^2 \exp(tx)\ F(dx).\int_{-\infty}^{+\infty} \exp(tx)\ F(dx) - \left[\int_{-\infty}^{+\infty} x.\exp(tx)\ F(dx)\right]^2 \geq 0\end{aligned}$$

mediante aplicación de la desigualdad de Cauchy-Schwartz.

En la última desigualdad, la igualdad ocurre si y sólo si las funciones $f(x) = 1$ y $g(x) = x \ \forall x \in \mathcal{R}$ son linealmente dependientes en el espacio $L^2(\mathcal{R}, \mathfrak{B}, F_t)$, donde F_t es la medida $F_t(dx) = \exp(tx)\ F(dx)$, lo cual quiere decir que $\exists$ constantes C_1, C_2, no ambas nulas, tales que

$$\int_{-\infty}^{+\infty} (C_1 + C_2.x)^2 \exp(tx)\ F(dx) = 0 \tag{49}$$

Es claro que no puede ser $C_2 = 0$ en (49). Además, esta igualdad implica que F esté concentrada en el punto $-C_1/C_2$, lo cual no es posible por la hipótesis que hicimos sobre F. En consecuencia no puede ocurrir la igualdad en (48) y por lo tanto $\Psi''_F(t) > 0 \ \forall t \in (\alpha_F,\ \beta_F)$. Esto prueba que Ψ_F es estrictamente convexa.

- Observar que si parametrizamos con la esperanza m la distribución - usamos la notación F_m para la distribución de X y por lo tanto F_0 es la distribución de $X - m$ - entonces $\Psi_{F_m}(t) = mt + \Psi_{F_0}(t) \implies h_{F_m}(a) = h_{F_0}(a - m)$.

- Bajo las hipótesis que hemos hecho, podemos calcular la transformada de Cramér mediante las fórmulas siguientes:

  (a) Si $\Psi'_F(\alpha_F) < a < \Psi'_F(\beta_F)$, entonces el máximo de la función $t \rightsquigarrow at - \Psi_F(t)$ ocurre en el interior del intervalo $(\alpha_F,\ \beta_F)$, en el punto $t = (\Psi'_F)^{-1}(a)$ y vale

$$h_F(a) = a.\left(\Psi'_F\right)^{-1}(a) - \Psi_F\left[\left(\Psi'_F\right)^{-1}(a)\right]$$

  (b) Si $a \geq \Psi'_F(\beta_F)$, entonces $h_F(a) = a.\beta_F - \Psi_F(\beta_F)$.

  (c) Si $a \leq \Psi'_F(\alpha_F)$, entonces $h_F(a) = a.\alpha_F - \Psi_F(\alpha_F)$.

- El lector verificará fácilmente que si $\Psi'_F(\alpha_F) < a < \Psi'_F(\beta_F)$, entonces

$$h'_F(a) = \left(\Psi'_F\right)^{-1}(a) \quad \text{y} \quad h''_F(a) = \frac{1}{\Psi''_F\left[\left(\Psi'_F\right)^{-1}(a)\right]}$$

  Por lo tanto, h_F es estrictamente convexa en el intervalo $(\Psi'_F(\alpha_F), \Psi'_F(\beta_F))$.

- También dejamos a cargo del lector la verificación de que

$$\Psi'_F(0) = m \quad \text{y} \quad h_F(m) = 0.$$

- Finalmente, observemos que si $a > m$, entonces

$$h_F(a) = \sup\{at - \Psi_F(t) : t \in (0,\ \beta_F)\}$$

En efecto, $a > m \implies (\Psi'_F)^{-1}(a) > (\Psi'_F)^{-1}(m) = 0$. O bien el punto en el que ocurre el máximo de la función $t \rightsquigarrow at - \Psi_F(t)$ está en el interior de $(\alpha_F,\ \beta_F)$ y es $(\Psi'_F)^{-1}(a)$,o bien el supremo se alcanza para $t \uparrow \beta_F$.

Teorema (Chernof).

Sea $\{\xi_k\}_{k=1,2,...}$ una sucesión de variables aleatorias independientes a valores reales con igual distribución F y supongamos que $E(\xi_k) = m$ y $S_n = \xi_1 + ... + \xi_n \quad (n \geq 1)$. Con la misma notación que antes, suponemos que $0 \in J^0$ (interior del intervalo J).

Entonces:

(a) $a > m \implies P\left(\frac{S_n}{n} \geq a\right) \leq \exp\left[-n.h_F(a)\right]$

$a < m \implies P\left(\frac{S_n}{n} \leq a\right) \leq \exp\left[-n.h_F(a)\right]$

En ambos casos es $h_F(a) > 0$.

(b) $m < a < \Psi'_F(\beta_F) \implies \lim_{n\to+\infty} \frac{1}{n}\log\left[P\left(\frac{S_n}{n} \geq a\right)\right] = -h_F(a)$

$\Psi'_F(\alpha_F) < a < m \implies \lim_{n\to+\infty} \frac{1}{n}\log\left[P\left(\frac{S_n}{n} \leq a\right)\right] = -h_F(a)$.

Demostración. Basta considerar el caso $a > m$. [Si $a < m$, reemplazar ξ_k por $-\xi_k$].

Para la parte (a):

$$P\left(S_n \geq n.a\right) \leq E\left(\exp\left[u(S_n - n.a)\right]\right) \quad \forall u \in (0, \beta_F)$$

que implica

$$\begin{aligned} P\left(S_n \geq n.a\right) &\leq \inf_{u\in(0,\beta_F)} E\left(\exp\left[u(S_n - n.a)\right]\right) \leq \inf_{u\in(0,\beta_F)} \exp\left[-n(u.a - \Psi_F(u))\right] \\ &= \exp\left[-n \sup_{u\in(0,\beta_F)} (u.a - \Psi_F(u))\right] = \exp\left[-n.h_F(a)\right] \end{aligned}$$

Para la parte (b), usando (a) alcanza con probar que

$$\underline{\lim}_{n\to+\infty} \frac{1}{n}\log\left[P\left(\frac{S_n}{n} \geq a\right)\right] \geq -h_F(a) \tag{50}$$

Para probar (50), poniendo $s_n = x_1 + ... + x_n$:

$$\begin{aligned} P\left(S_n \geq n.a\right) &= \int_{\mathcal{R}^n} 1_{\{s_n \geq n.a\}} F(dx_1)....F(dx_n) \\ &= \int_{\mathcal{R}^n} 1_{\{s_n \geq n.a\}} \left[\Phi_F(t)\right]^n \ \exp\left[-t.s_n\right] \ G(dx_1)....G(dx_n) \end{aligned}$$

donde G es la medida de probabilidad en la recta: $G(dx) = \left[\Phi_F(t)\right]^{-1} \exp\left[tx\right] \ F(dx)$.

Sea $Y_1, ..., Y_n$ una sucesión de variables aleatorias independientes con igual distribución G y $T_n = Y_1 + ... + Y_n \quad (n \geq 1)$. La igualdad anterior se rescribe como

$$P\left(S_n \geq n.a\right) = \left[\Phi_F(t)\right]^n \exp\left[-nat\right] \ E\left(1_{\{T_n \geq n.a\}} \exp\left[-nt\left(\frac{T_n}{n} - a\right)\right]\right) \tag{51}$$

Sea $m < a < \Psi_F'(\beta_F)$ y $\varepsilon > 0$ tal que $a + \varepsilon < \Psi_F'(\beta_F)$. Elegimos t tal que:

$a < \Psi_F'(t) < a{+}\varepsilon < \Psi_F'(\beta_F) \quad \Longrightarrow \quad \left(\Psi_F'\right)^{-1}(a) < t < \left(\Psi_F'\right)^{-1}(a{+}\varepsilon) < \beta_F.$

Observar que

$$E\left(Y_1\right) = \left[\Phi_F(t)\right]^{-1} \int_{-\infty}^{+\infty} x \exp(tx) \ F(dx) = \frac{\Phi_F'(t)}{\Phi_F(t)} = \Psi_F'(t)$$

Por lo tanto, en virtud de la ley fuerte de los grandes números, casi seguramente, $\frac{T_n}{n} \to \Psi_F'(t)$ cuando $n \to +\infty$. Tenemos, a partir de (51):

$$\begin{aligned} & \frac{1}{n}\log\left[P\left(\frac{S_n}{n} \geq a\right)\right] \\ = \quad & -at + \Psi_F(t) + \frac{1}{n}\log\left[E\left(1_{\{T_n \geq n.a\}} \exp\left[-nt\left(\frac{T_n}{n} - a\right)\right]\right)\right] \\ \geq \quad & -at + \Psi_F(t) + \frac{1}{n}\log\left[E\left(1_{\{na \leq T_n \leq n(a+\varepsilon)\}} \exp\left[-nt\left(\frac{T_n}{n} - a\right)\right]\right)\right] \\ \geq \quad & -at + \Psi_F(t) + \frac{1}{n}\log\left[\exp(-nt\varepsilon).P(na \leq T_n \leq n(a+\varepsilon))\right] \\ = \quad & -at + \Psi_F(t) - \varepsilon t + \frac{1}{n}\log\left[P(a \leq \frac{T_n}{n} \leq a + \varepsilon)\right] \end{aligned}$$

Esto prueba que

$$\underline{\lim}_{n\to+\infty}\frac{1}{n}\log\left[P\left(\frac{S_n}{n} \geq a\right)\right] \geq -at + \Psi_F(t) - \varepsilon t$$

ya que $P(a \leq \frac{T_n}{n} \leq a{+}\varepsilon) \to 1$. Como $\varepsilon > 0$, es arbitrario, dada la definición de $h_F(a)$, esto termina la demostración.

2.8.1 Ejemplos.

1. Consideremos, como en el teorema anterior, una sucesión de variables aleatorias independientes a valores reales con igual distribución $\{\xi_k\}_{k=1,2,\ldots}$, y que la distribución común F está dada por $P(\xi_k = 1) = P(\xi_k = -1) = 1/2$.

   Es claro que $E(\xi_k) = 0$, $Var(\xi_k) = 1$.

   Supongamos que queremos estudiar el comportamiento de $\alpha_n(\varepsilon) = P(\left|\frac{S_n}{n}\right| \geq \varepsilon)$, para $0 < \varepsilon < 1$.

   En virtud del teorema de Chernof, sabemos que $\frac{1}{n}\log\left[\alpha_n(\varepsilon)\right] \to -h_F(\varepsilon)$. Los cálculos en este caso son inmediatos y se obtiene:

   - $\Phi(t) = \cosh t$

   - $\Psi(t) = \log\left[\cosh t\right]$

   - $h_F(\varepsilon) = \varepsilon \tanh^{-1}\varepsilon - \log\left[\cosh\left(\tanh^{-1}\varepsilon\right)\right] = \varepsilon\tanh^{-1}\varepsilon + \frac{1}{2}\log\left[\left(1-\varepsilon^2\right)\right]$ (usar que $\cosh^2 x = \frac{1}{1-\tanh^2 x}$).

   Si en lugar del teorema de grandes desviaciones se pone el equivalente que viene del Teorema del Límite Central, lo que se obtiene como equivalente es

   $$2\frac{1}{\sqrt{2\pi}}\int_{\varepsilon\sqrt{n}}^{+\infty}\exp\left(-\frac{x^2}{2}\right)dx \approx \sqrt{\frac{2}{\pi}}\frac{1}{\varepsilon\sqrt{n}}\exp\left(-\frac{\varepsilon^2 n}{2}\right) \quad \text{cuando } n \to +\infty$$

   o sea

   $$\frac{1}{n}\log\left[2\frac{1}{\sqrt{2\pi}}\int_{\varepsilon\sqrt{n}}^{+\infty}\exp\left(-\frac{x^2}{2}\right)dx\right] \to -\frac{\varepsilon^2}{2}$$

   El comportamiento preciso, dado por $h_F(\varepsilon)$ se ve que difiere del que resulta de utilizar (inadecuadamente) las colas que da el TLC. Sin embargo, si se hace el desarrollo de Taylor de $h_F(\varepsilon)$ en $\varepsilon = 0$, se ve que al primer orden significativo, se obtiene $-\frac{\varepsilon^2}{2}$. Si ε es suficientemente pequeño, ambos exponentes difieren poco, pero ello deja de ser cierto para valores moderados de ε.

2. Veamos una variante del ejemplo anterior. Supongamos que la distribución F está dada por $P(\xi_k = 1) = P(\xi_k = -1) = p$, $P(\xi_k = 0) = r$, $2p + r = 1$. Es claro que $E(\xi_k) = 0$ y $Var(\xi_k) = 2p$. Se obtiene $\Psi(t) = \log\left[2p\cosh t + r\right]$ y

   $$h_F(a) = a\tau - \Psi(\tau)$$

donde τ es la solución de la ecuación

$$a - \Psi'(\tau) = a - \frac{2p\sinh\tau}{2p\cosh\tau + r} = 0$$

Con estos ingredientes, se puede escribir los primeros términos del desarrollo de Taylor de $h_F(a)$ para a cerca de cero.

3. Si F es la distribución normal típica, se verifica inmediatamente que $\Phi(t) = \exp\left(\frac{t^2}{2}\right)$ y por lo tanto $\Psi(t) = \frac{t^2}{2}$, $h_F(a) = \frac{a^2}{2}$. El comportamiento asintótico del logaritmo de $P(\left|\frac{S_n}{n}\right| \geq \varepsilon)$ coincide con el de las colas en el TLC.

4. Si F es la distribución exponencial con parámetro igual a 1 (es decir que para $x > 0$, es $1 - F(x) = \exp(-x)$), entonces

$$\Psi(t) = -\log(1-t) \text{ si } \ t < 1 \quad \text{y} \quad \Psi(t) = +\infty \ \text{ si } \ t \geq 1.$$

Un cálculo elemental muestra que

$$h_F(a) = a - 1 - \log a \text{ siempre que } a > 0$$

y si $0 < \varepsilon$:

$$\frac{1}{n}\log\left[P(\frac{S_n}{n} - 1 \geq \varepsilon)\right] \to \varepsilon - \log(1+\varepsilon) = \frac{\varepsilon^2}{2} - \frac{\varepsilon^3}{3} + \frac{\varepsilon^4}{4} -$$

donde el último desarrollo vale si además $\varepsilon \leq 1$.

Del otro lado, si $0 < \varepsilon < 1$:

$$\frac{1}{n}\log\left[P(\frac{S_n}{n} - 1 \leq -\varepsilon)\right] \to \varepsilon - \log(1+\varepsilon)$$

2.9 Distribuciones infinitamente divisibles

Preliminares, medidas canónicas.

Definición. Se dice que la medida M definida sobre los Borelianos de $\mathcal{R}$ es "canónica" si verifica las condiciones:

1. $M(I) < \infty$ para todo intervalo acotado I

2. $\forall x > 0$ se cumple que

$$\begin{aligned} M^{+}(x) &= \int_{x^{-}}^{+\infty} \frac{M(dy)}{y^2} < \infty \\ M^{-}(-x) &= \int_{-\infty}^{-x^{+}} \frac{M(dy)}{y^2} < \infty \end{aligned}$$

Ejemplo básico. Si F es una medida de probabilidad en la recta y C es una constante positiva, entonces $M(dx) = C.x^2.F(dx)$ es una medida canónica.

Definición. Sea $\{F_n\}$ una sucesión de medidas de probabilidad en $\mathcal{R}$ y $\{c_n\}$ una sucesión de números reales positivos. Diremos que $c_n.x^2.F_n(dx)$ converge propiamente a la medida canónica $M(dx)$ si

(a) $c_n.\int_I x^2.F_n(dx) \quad \rightarrow \quad M(I) \quad \forall$ intervalo I cuyos extremos son de continuidad de M

(b) $c_n \int_{x^{-}}^{+\infty} \frac{F_n(dy)}{y^2} \quad \rightarrow \quad M^{+}(x), \qquad c_n \int_{-\infty}^{-x^{+}} \frac{M(dy)}{y^2} \rightarrow M^{-}(-x)$ si $x, -x$ son de cont de M

Se verifica fácilmente que la pareja de condiciones (a),(b) es equivalente a (a),(b'), donde

$$(b') \ \forall \varepsilon > 0 \ \exists a > 0 \text{ tal que } c_n \left[F_n(-a) + (1 - F_n(a))\right] < \varepsilon \ \forall n$$

Proposición técnica auxiliar.

Sea $\{F_n\}$ una sucesión de medidas de probabilidad en $\mathcal{R}$ y $\{\varphi_n\}$ la sucesión de sus transformadas de Fourier. Consideramos la sucesión de funciones

$$\psi_n(t) = c_n \left[\varphi_n(t) - 1 - i\beta_n t\right]$$

donde $c_n > 0, \beta_n \in \mathcal{R}$.

Entonces, para cada $t \in \mathcal{R}$:

$$\psi_n(t) \rightarrow \rho(t)$$

donde ρ es una función continua, si y sólo si, existen una medida canónica M y una constante real b, tales que:

1. $c_n.x^2.F_n(dx)$ converge a M propiamente
2. $c_n \left(b_n - \beta_n\right) \rightarrow b$, donde $b_n = \int_{-\infty}^{+\infty} senx \ F_n(dx)$.

En caso de que haya convergencia, el límite es de la forma:

$$
\begin{aligned}
\rho(t) &= \psi(t) + ibt \quad \text{con} \\
\psi(t) &= \int_{-\infty}^{+\infty} \frac{\exp(itx) - 1 - i.t.senx}{x^2} M(dx)
\end{aligned}
\tag{52}
$$

La medida M está unívocamente determinada por (52).

Demostración. En realidad, no daremos una demostración completa de esta proposición auxiliar, sino que indicaremos solamente los pasos a seguir, cuyos detalles quedan a cargo del lector.

Antes que nada, reescribimos $\psi_n(t)$:

$$
\psi_n(t) = \int_{-\infty}^{+\infty} [\exp(itx) - 1 - i.t.senx] \ c_n F_n(dx) + i \ c_n \left(b_n - \beta_n\right) \ t \tag{53}
$$

*Paso 1.*Mostrar que si $c_n.x^2.F_n(dx)$ converge a M propiamente, entonces, para toda función continua y acotada $f : \mathcal{R} \to \mathcal{R}$, tal que $\frac{f(x)}{x^2}$ es continua en $x = 0$, se tiene:

$$
\int_{-\infty}^{+\infty} f(x) \ c_n F_n(dx) \to \int_{-\infty}^{+\infty} \frac{f(x)}{x^2} M(dx)
$$

[Usar un método análogo al empleado inicialmente para la convergencia débil de medidas de probabilidad].

Paso 2. Si $\psi_n(t) \to \rho(t)$ uniformemente en $|t| \le t_0$ y $0 < h \le t_0$, entonces:

$$
\int_{-\infty}^{+\infty} \left[1 - \frac{sen(xh)}{xh}\right] c_n F_n(dx) \to -\frac{1}{2h} \int_{-h}^{h} \rho(t) \ dt
$$

Paso 3. Si $\psi_n(t) \to \rho(t)$ uniformemente en $|t| \le t_0$, entonces existe una subsucesión $\{n_k\}$ tal que $c_{n_k}.x^2.F_{n_k}(dx)$ converge a M propiamente, donde M es una medida canónica.

Para probar esto, usar el mismo tipo de argumento diagonal que para la compacidad débil de las medidas de probabilidad y también, el resultado del paso 2.

Paso 4. Si M es canónica, la función ψ definida por (52) es continua y la correspondencia $M \rightsquigarrow \psi$ es inyectiva.

La continuidad es inmediata. Para ver la inyectividad, observar que, de la definición de ψ resulta que $\forall h > 0$:

$$\psi(t) - \frac{\psi(t+h) + \psi(t-h)}{2} = \int_{-\infty}^{+\infty} \exp(itx) \frac{1 - \cos(xh)}{x^2} M(dx)$$

Dicho de otro modo, la función $\psi(t) - \frac{\psi(t+h)+\psi(t-h)}{2}$ es la transformada de Fourier de la medida finita $\mu_h(dx) = \frac{1-\cos(xh)}{x^2} M(dx)$. Por la unicidad ya conocida, esto dice que ψ determina μ_h $\forall h > 0$. Esto "casi" determina M, salvo por eventuales átomos en los puntos x tales que $\cos(xh) = 1$. Queda a cargo del lector mostrar que entonces M está unívocamente determinada.

Paso 5. Con los elementos anteriores, el lector puede completar la demostración de la Proposición.□

Definición. Se dice que la distribución de probabilidad propia en la recta F es "infinitamente divisible" si para cada entero positivo n existe una distribución de probabilidad F_n tal que, si $\xi_1, ..., \xi_n$ son n variables aleatorias independientes con igual distribución F_n, entonces, la suma $\xi_1 + ... + \xi_n$ tiene distribución F.

De manera equivalente, podemos escribir esta condición en términos de transformadas de Fourier. Si denotamos por φ la transformada de Fourier de F, que ésta es infinitamente divisible significa que para cada entero positivo n podemos encontrar una transformada de Fourier de una medida de probabilidad propia, φ_n tal que

$$[\varphi_n(t)]^n = \varphi(t) \quad \forall t \in \mathcal{R}$$

El resultado fundamental de esta sección es el teorema siguiente:

Teorema (Lévy, Kinchín). φ es la transformada de Fourier de una distribución de probabilidad infinitamente divisible, si y sólo si, existen una medida canónica M y una constante real b tales que:

$$\begin{aligned} \varphi(t) &= \exp[\rho(t)] \quad \text{donde} \\ \rho(t) &= ibt + \int_{-\infty}^{+\infty} \frac{\exp(itx) - 1 - i.t.senx}{x^2} M(dx) \end{aligned} \tag{54}$$

M y b son únicas.

Antes que la demostración veamos algunos (primeros) ejemplos.

Ejemplo 1. Si la medida M está concentrada en el origen, $M(\{0\}) = \sigma^2$, entonces la fórmula (54) da $\varphi(t) = \exp\left[ibt - \frac{1}{2}\sigma^2 t^2\right]$ que es la transformada de Fourier de la distribución normal con media b y varianza σ^2.

Ejemplo 2. Si la variable aleatoria ξ tiene la distribución de Poisson con parámetro λ, entonces su transformada de Fourier es

$$E\left(\exp(it\xi)\right) = \sum_{n=0}^{\infty} \exp(itn)\frac{\lambda^n}{n!}e^{-\lambda} = \exp\left[\lambda\left(e^{it} - 1\right)\right] \tag{55}$$

Obsérvese que en la fórmula (54), si la medida M está concentrada en el punto $x = 1$ y $M(\{1\}) = \lambda$, entonces $\rho(t) = \lambda\left(e^{it} - 1\right)$ si se hace una elección adecuada de la constante b. Es decir que la distribución de Poisson es infinitamente divisible, cosa que podemos verificar directamente a partir de (55), ya que es claro de aquí que si ξ tiene distribución de Poisson con parámetro λ, es suma de n variables aleatorias independientes, cada una de ellas con distribución de Poisson con parámetro $\frac{\lambda}{n}$.

Recíprocamente, si M está concentrada en un punto $x_0 \neq 0$ y $M(\{x_0\}) = c$, entonces $\rho(t) = ibt + \frac{\exp(itx_0) - 1 - itsenx_0}{x_0^2}c = i\bar{b}t + \lambda\left(e^{itx_0} - 1\right)$

lo que muestra que la distribución de la variable aleatoria ξ cuya transformada de Fourier es $\exp\left[\rho(t)\right]$ se obtiene de la distribución de Poisson mediante un cambio de posición y escala, más precisamente, que $\frac{\xi - \bar{b}}{x_0}$ tiene distribución de Poisson con parámetro λ.

Ejemplo 3 (Distribución de Poisson compuesta).- Sea $\xi_1, \xi_2, ...$ una sucesión de variables aleatorias independientes con igual distribución F. Definimos, como antes, la sucesión de sumas parciales $S_0 = 0$, $S_n = \xi_1 + ... + \xi_n$ $(n \geq 1)$. Denotamos por φ la transformada de Fourier de F.

Sea además υ una variable aleatoria con distribución de Poisson de parámetro λ, independiente de la sucesión $\{\xi_n\}$.

Definimos $X = S_\nu$, es decir, son sumas con un número aleatorio de términos. La distribución de X se denomina "de Poisson compuesta".

Calculamos la transformada de Fourier de la distribución de X, es decir:

$$\begin{aligned}
\varphi_X(t) &= E\left(\exp(itX)\right) = E\left[E\left(\exp(itX)/N=k\right)\right] \\
&= \sum_{k=0}^{\infty} E\left(\exp(itS_k)/N=k\right) P(N=k) \\
&= \sum_{k=0}^{\infty} E\left(\exp(itS_k)\right) P(N=k) = \sum_{k=0}^{\infty} [\varphi(t)]^k \frac{\lambda^k}{k!} \exp(-\lambda) \\
&= \exp\left[\lambda\left(\varphi(t)-1\right)\right] = \exp\left[i\lambda t\bar{b} + \lambda \int_{-\infty}^{+\infty} \frac{\exp(itx)-1-i.t.senx}{x^2} x^2 F(dx)\right]
\end{aligned}$$

con $\bar{b} = \int_{-\infty}^{+\infty} senx\ F(dx)$.

Es claro entonces que φ_X tiene la forma indicada en (52) con $b = \lambda.\bar{b}$ y $M(dx) = \lambda.x^2 F(dx)$.

El lector observará que es posible dar una demostración directa simple de que la distribución de X es infinitamente divisible, que no depende de (52).

Observe el lector que en la proposición técnica auxiliar de esta sección, $\exp\left[\psi_n(t)\right]$ son transformadas de Fourier de distribuciones de Poisson compuestas, a menos de una traslación y un cambio de escala.

Ejemplo 4.- El lector verificará que si $M(dx) = x.e^{-x}.1_{\{x>0\}}\ dx$ entonces F es la distribución exponencial de parámetro igual a 1. Para ver esto, observar que la transformada de Fourier de la exponencial con parámetro 1 es $\omega(t) = 1/(1-it)$ y que $\psi(t) = \log\left[\omega(t)\right] = \int_0^{+\infty} \frac{\exp(itx)-1}{x} \exp(-x)\ dx$.

Ejemplo 5.- El lector probará que F es una distribución infinitamente divisible concentrada en $[0,+\infty)$ si, y sólo si, su transformada de Fourier φ se escribe como $\varphi = \exp(\rho)$ donde

$$\rho(t) = ibt + \int_0^{+\infty} \frac{\exp(itx)-1}{x} \mu(dx)$$

donde μ es una medida de Borel concentrada en la semirrecta $(0,+\infty)$, $\int_0^{+\infty} \frac{1}{1+x}\mu(dx) < \infty$.

Observación sobre el enunciado del Teorema de Lévy-Kinchín.

En la representación dada por (54), se puede escribir el exponente de diversas formas. Por ejemplo, se suele usar la forma alternativa

$$\rho(t) = ib't + \int_{-\infty}^{+\infty} \frac{\exp(itx)-1-i.t1_{\{|x|<1\}}}{x^2} M(dx) \tag{56}$$

donde $b' = b + \int_{-\infty}^{+\infty} \frac{1_{\{|x|<1\}} - senx}{x^2} M(dx)$. Véase que esta última integral está bien definida, en virtud de que M es una medida canónica.

La forma que es quizá la más utilizada de la representación de Lévy-Kinchín, es la siguiente

$$\rho(t) = ib't - \frac{1}{2}\sigma^2 t^2 + \int_{-\infty}^{+\infty} \left[\exp(itx) - 1 - i.t1_{\{|x|<1\}}\right] \, N(dx) \tag{57}$$

donde se ha reescrito la medida canónica M como $M(dx) = \sigma^2 \delta_0(dx) + x^2 N(dx)$. Aquí, δ_0 es la medida de Dirac en el punto $x = 0$, es decir que el primer sumando es una masa de tamaño σ^2, concentrada en el origen, y N es una medida de Borel en $R \setminus \{0\}$ tal que

$$\int_{|x|<1} x^2 N(dx) < \infty, \qquad \int_{|x|\geq 1} N(dx) < \infty \tag{58}$$

Es inmediata la deducción de estas condiciones de las que verifica M por ser canónica.

Demostración del teorema de Lévy-Kinchín.

Paso 1. En este primer paso, probamos que si $\{\varphi_n\}$ es una sucesión de transformadas de Fourier de distribuciones (propias) de probabilidad, entonces:

$$\varphi_n^n \to \varphi \text{ puntualmente, } \varphi \text{ continua} \iff n(\varphi_n - 1) \to \rho \text{ puntualmente, } \rho \text{ continua} \tag{59}$$

En ese caso, $\varphi(t) = \exp\left[\rho(t)\right]$.

Probemos primero la implicación de derecha a izquierda. La hipótesis implica que $\varphi_n \to 1$ puntualmente, y por lo tanto, uniformemente en cada compacto, dado que tanto φ_n como 1 son transformadas de Fourier de medidas de probabilidad. Por lo tanto, si t varía en un compacto dado, cuando n es suficientemente grande está bien definido log (la rama principal del logaritmo) de $\varphi_n(t)$ y la hipótesis implica, desarrollo de Taylor mediante, que $n \log\left[\varphi_n(t)\right] \to \rho(t)$. Se sigue que $\left[\varphi_n(t)\right]^n \to \exp\left[\rho(t)\right]$. Es claro que esto vale entonces para cada t fijo.

Veamos ahora (59) de izquierda a derecha. Como $\{\varphi_n^n\}$ es también una sucesión de transformadas de Fourier de medidas de probabilidad, la hipótesis implica, a raíz del teorema de Lévy-Cramér, que la convergencia es uniforme en cada compacto. Dado que φ es continua, elijamos $t_1 > 0$ tal que $|t| \leq t_1 \Longrightarrow |\varphi(t) - 1| < 1/2$. Por lo tanto, si n es suficientemente grande, está definido log $\left[\varphi_n(t)\right]^n$ (rama principal) $\forall t$ tal que $|t| \leq t_1$. Se deduce entonces que $n\log\left[\varphi_n(t)\right]\to\log\left[\varphi(t)\right]\Longrightarrow n\left[\varphi_n(t)-1\right]\to\log\left[\varphi(t)\right]$ cuando $|t|\leq t_1$. Queremos ver que hay convergencia para todo t.

De la convergencia en el intervalo $[-t_1,t_1]$ se deduce que existe una subsucesión $\left\{\varphi_{n_k}\right\}$ tal que $n_k\left[\varphi_{n_k}(t)-1\right]$ converge puntualmente en toda la recta, a una función continua ρ (Usar el Paso 3 de la demostración de la proposición auxiliar). Pero entonces, podemos aplicar a esta subsucesión la implicación de derecha a izquierda de (59), que ya fue probada, y resulta que $\left[\varphi_{n_k}(t)\right]^{n_k}\to\exp\left[\rho(t)\right]$ puntualmente, en toda la recta. Entonces este límite es continuo y no depende de la subsucesión. El lector aprovechará esto para terminar la argumentación.

Paso 2. Veamos ahora el directo en el teorema de Lévy-Kinchín.

Supongamos que φ es infinitamente divisible. Entonces, $\forall n=1,2,...$ $\exists$ una transformada de Fourier φ_n tal que $\varphi_n^n=\varphi$ y por lo tanto estamos en las condiciones de aplicar la implicación de izquierda a derecha en (59) del Paso 1, y resulta que $n(\varphi_n-1)\to\rho$ puntualmente, ρ continua y que $\varphi=\exp(\rho)$. Aplicando ahora la proposición auxiliar, resulta que ρ tiene la forma requerida en (52).

Paso 3. Para el recíproco, supongamos que φ tiene la forma de (54).

Consideremos primero el caso en que la medida canónica M está concentrada en $\{x:|x|>\delta\}$ con $\delta>0$. Entonces, si definimos la medida de probabilidad

$$G(dx)=\frac{K}{x^2}M(dx)$$

con $K=\left[\int_{-\infty}^{+\infty}\frac{M(dx)}{x^2}\right]^{-1}$ y sea γ la transformada de Fourier de G. El lector verificará que entonces, con una elección apropiada de b, resulta $\varphi(t)=\exp\left[i\bar{b}t+K\left(\gamma(t)-1\right)\right]$ que es infinitamente divisible (es una la transformada de Fourier de una Poisson compuesta, trasladada y con un cambio de escala).

Veamos ahora el caso general, en que no necesariamente M está concentrada en un conjunto de la forma $\{x:|x|>\delta\}$ con $\delta>0$.

Consideramos entonces, para cada $\delta>0$, la medida canónica $M_\delta(dx)=1_{\{|x|>\delta\}}M(dx)$ y le aplicamos lo anterior. Denotamos por γ_δ la transformada de Fourier de $\frac{K_\delta}{x^2}M(dx)$ (mismas notaciones). Aplicando nuevamente la proposición auxiliar, resulta que exite el límite puntual

$$\rho(t)=\lim_{\delta\downarrow 0}\left[\gamma_\delta(t)-\frac{1}{2}t^2M\left(\{0\}\right)\right]$$

y como $\exp\left[\gamma_\delta(t)-\frac{1}{2}t^2M\left(\{0\}\right)\right]$ es la transformada de Fourier de una distribución de probabilidad (convolución de una Poisson compuesta y una normal) también $\exp[\rho(t)]$, en virtud del teorema de Lévy-Cramér.

Esto prueba que si φ tiene la forma (54), entonces es la transformada de Fourier de una distribución de probabilidad propia. Pero lo mismo se aplica si uno reemplaza, para cada $n \geq 1$, M por $\frac{1}{n}M$ y b por $\frac{1}{n}b$. Si φ_n es la transformada de Fourier con esta medida canónica y esta constante, entonces es obvio que se verifica que $\varphi_n^n = \varphi$.□

2.10 Distribuciones estables. Dominios de atracción.

Definición. Sea F una distribución de probabilidad (propia), $F \neq \delta_0$. Diremos que F es "estable en sentido amplio" si para todo n, entero positivo, si $X_1, ..., X_n$ son variables aleatorias independientes con distribución F y $S_n = X_1 + ... + X_n$ entonces, existen números reales c_n, d_n, $c_n > 0$ tales que

$$S_n \stackrel{(d)}{=} c_n X_1 + d_n$$

donde $\xi \stackrel{(d)}{=} \eta$ significa que las variables aleatorias ξ y η tienen la misma distribución de probabilidad.

Diremos que F es "estrictamente estable", si una relación análoga se verifica con $d_n = 0$.

Es obvio que si F es estable, entonces es infinitamente divisible, ya que con esa notación, para todo $n \geq 1$, X_1 tiene la misma distribución que la suma de las n variables aleatorias independientes con igual distribución $\left(X_k - \frac{d_n}{n}\right)\frac{1}{c_n}$ $(k = 1, ..., n)$. Entonces, su transformada de Fourier se puede representar mediante (54) y el primer objetivo de esta sección es calcular las medidas canónicas M y las constantes b de las distribuciones estables. El segundo objetivo es relacionar las distribuciones estables con una extensión del teorema del límite central para sumas de variables aleatorias independientes con igual distribución, cuando no se cumplen las hipótesis para que la distribución límite sea normal, pero sin embargo hay convergencia débil.

Definición. Sea F una distribución de probabilidad (propia). Se define la "simetrizada" de F, que denotamos F^s, como la distribución de probabilidad de $X - Y$, donde X e Y son variables aleatorias independientes, cada una de ellas con distribución F.

Es inmediato que si F es estable, entonces F^s es estrictamente estable, con el mismo c_n.

Teorema. En la definición de distribución estable, necesariamente debe ser $c_n = n^{\frac{1}{\alpha}}$ para algún real α, $0 < \alpha \leq 2$.

Demostración. El lector verificará que basta probarlo cuando F es estrictamente estable (usar F^s).

En ese caso, con la misma notación anterior, aplicando la definición, resulta que si m, n, k son enteros positivos:

$$\begin{aligned} c_{m+n}X_1 &\overset{(d)}{=} c_m X_1 + c_n X_2 \\ c_{m.n} &= c_m.c_n \quad \Longrightarrow c_{n^k} = (c_n)^k \end{aligned}$$

Veamos que la sucesión $\{c_n\}_{n=1,2,..}$ es monótona creciente. En efecto, si

$a > 0$:

$$\begin{aligned} P(c_{m+n}X_1 &\geq c_m.a) \geq P(c_n X_1 \geq 0).P(c_m X_1 \geq c_m.a) \\ &= P(X_1 \geq 0).P(X_1 \geq a) \end{aligned} \tag{60}$$

Esto implica que $\left\{\frac{c_m}{c_{m+n}}\right\}_{m,n=1,2,...}$ es una sucesión (doble) acotada, puesto que existe $a > 0$ de modo que el último miembro de (60), que no depende de m, n, sea estrictamente positivo. Sea entonces C una constante positiva tal que

$$\frac{c_m}{c_{m+n}} \leq C \quad \forall m, n$$

Se sigue que

$$\left(\frac{c_m}{c_{m+1}}\right)^k = \frac{c_{m^k}}{c_{(m+1)^k}} \leq C \quad \forall m, k$$

lo que implica que $\frac{c_m}{c_{m+1}} \leq 1 \; \forall m$. Esto prueba que $\{c_n\}$ es monótona creciente en sentido amplio.

Veamos ahora que

$$\frac{\log c_m}{\log m} = const \tag{61}$$

Fijemos m entero, $m \geq 2$ y para cada n entero positivo, sea $k(n)$ el único entero tal que

$$\begin{aligned} 2^{k(n)} &\leq m^n < 2^{k(n)+1} \quad \text{es decir que} \\ k(n)\ log2 &\leq n\ \log m < (k(n) + 1)\ log2 \end{aligned}$$

Además

$$\begin{aligned}
(c_2)^{k(n)} &= c_{2^{k(n)}} \leq c_{m^n} = (c_m)^n < c_{2^{k(n)+1}} = (c_2)^{k(n)+1} \implies \\
k(n)\log c_2 &\leq n \log c_m < [k(n)+1]\log c_2 \implies \\
\frac{k(n)\log c_2}{n}.\frac{n}{[k(n)+1]\log 2} &\leq \frac{\log c_m}{\log m} < \frac{[k(n)+1]\log c_2}{n}\frac{n}{k(n)\log 2}
\end{aligned}$$

y de aquí se sigue fácilmente, pasando al límite cuando $n \to +\infty$ que $\frac{\log c_m}{\log m} = \frac{\log c_2}{\log 2}$. Como $\{c_m\}$ no es constante y es monótona creciente, esto prueba que $c_m = m^{\frac{1}{\alpha}}$ con $\alpha > 0$.

Lo que falta es probar que $\alpha \leq 2$. Para probar esto, usamos la representación de Lévy-Kinchín, cosa que podemos hacer, puesto que F es infinitamente divisible, de modo que su transformada de Fourier φ es escribe mediante la fórmula (54) y del hecho que es estable se deduce inmediatamente que

$$\begin{aligned}
n\rho(t) &= \rho(c_n t) + id_n t \\
&= \int_{-\infty}^{+\infty} \frac{\exp(itx) - 1 - i.c_n t.sen(c_n^{-1}x)}{x^2}\frac{1}{c_n^{-2}}M(c_n^{-1}dx) + id_n t \\
&= \int_{-\infty}^{+\infty} \frac{\exp(itx) - 1 - i.t.senx}{x^2}c_n^2 M(c_n^{-1}dx) + i\bar{d}_n t
\end{aligned}$$

En virtud de la unicidad de la representación de Lévy-Kinchín, esto implica que para cada $n = 1, 2, ...$

$$nM(dx) = c_n^2 M(c_n^{-1}dx) \tag{62}$$

donde, por otra parte, ya sabemos que $c_n = n^{\frac{1}{\alpha}}$.
Tenemos dos casos posibles:

1. $\alpha = 2$. En este caso, (62) se traduce en que $M(dx) = M(n^{-\frac{1}{2}}dx)$ $\forall n = 1, 2, ...$ Por lo tanto $\forall a > 0$:

   $$M([-a, a]) = M\left(\left[-n^{-\frac{1}{2}}a, n^{-\frac{1}{2}}a\right]\right) \to M(\{0\}) \text{ cuando } n \to +\infty$$

   i.e. $M([-a, a]) = M(\{0\})$ $\forall a > 0$, es decir que la medida canónica M está concentrada en el origen y la distribución F es normal.

2. $\alpha \neq 2$. Poniendo, para $x > 0$, $G(x) = M(-x, x)$, (62) se traduce en que

   $$G(x) = n^{\frac{2}{\alpha}-1}G(n^{-\frac{1}{\alpha}}x) \quad \forall x > 0, \forall n = 1, 2, ...$$

y por lo tanto, para $m, n = 1, 2, ...$

$$G\left(\left[\frac{m}{n}\right]^{\frac{1}{\alpha}}\right) = n^{-\left(\frac{2}{\alpha}-1\right)}G(m^{\frac{1}{\alpha}}) = \left[\frac{m}{n}\right]^{\frac{2}{\alpha}-1}G(1)$$

Esto prueba que $G(x) = G(1)x^{2-\alpha}$ para los x que pertenecen al conjunto $E = \left\{\left[\frac{m}{n}\right]^{\frac{1}{\alpha}} : m, n = 1, 2, ...\right\}$. Dado que G es monótona creciente, esto prueba ya que $\alpha < 2$.

Asimismo, como el conjunto E es denso en $\mathcal{R}^+$ esto implica que

$$G(x) = G(1)x^{2-\alpha} \quad \forall x \geq 0.$$

Un argumento enteramente análogo - a cargo del lector - permite probar que

$$M\left((0, x)\right) = pG(1)x^{2-\alpha}, M\left((-x, 0)\right) = qG(1)x^{2-\alpha} \quad \forall x \geq 0,$$

donde p, q son constantes no negativas, $p + q = 1.\square$

Teorema (caracterización de las distribuciones estables)

Con las notaciones anteriores, una distribución es estable, si y sólo si, $\rho(t)$ se puede escribir de alguna de las dos formas siguientes, que corresponden respectivamente a $\alpha = 1$ y a $\alpha \neq 1, 0 < \alpha \leq 2$ en el teorema anterior:

$$(i) \quad \rho(t) = -C_1 \left|t\right| \left(\frac{\pi}{2} + i.sg(t) \log\left|t\right|\right) \tag{63}$$

$$(ii) \quad \rho(t) = -C_\alpha \left|t\right|^\alpha \left(1 - i.sg(t)(p-q)tg(\frac{\pi\alpha}{2}\right) \tag{64}$$

donde $C_\alpha > 0$.

Demostración. En el caso $\alpha = 2$ ya hemos visto el resultado, que coincide con el dado por (64) cuando $\alpha = 2$ y que corresponde a la distribución normal. De aquí en adelante, suponemos $0 < \alpha < 2$.

Usando los mismos cálculos de la demostración del teorema anterior y, en especial, la forma de la medida de Lévy-Kinchín de una distribución estable, resulta que

$$\begin{aligned} \rho(t) \quad = \quad & C.p(2-\alpha)\int_0^{+\infty} \frac{\exp(itx) - 1 - i.t.senx}{x^{\alpha+1}}dx \qquad (65) \\ & +C.q(2-\alpha)\int_{-\infty}^{0} \frac{\exp(itx) - 1 - i.t.senx}{x^{\alpha+1}}dx \end{aligned}$$

Para facilitar el cálculo, no usamos en el centramiento la función $senx$ en todos los casos, sino que usamos la función $\tau_\alpha(x)$, dependiente de x.

- si $0 < \alpha < 1 \implies \tau_\alpha(x) = 0$. Entonces, si $t > 0$

$$
\begin{aligned}
J_\alpha(t) &= \int_0^{+\infty} \frac{\exp(itx) - 1}{x^{\alpha+1}} dx = \lim_{\lambda \downarrow 0} \int_0^{+\infty} \frac{\exp(itx - \lambda x) - 1}{x^{\alpha+1}} dx \\
&= \lim_{\lambda \downarrow 0} \left[-\frac{1}{\alpha x^\alpha} \left(\exp(itx - \lambda x) - 1 \right) |_0^{+\infty} + \frac{1}{\alpha}(it - \lambda) \int_0^{+\infty} \frac{\exp(itx - \lambda x)}{x^\alpha} dx \right] \\
&= \lim_{\lambda \downarrow 0} \left[\frac{1}{\alpha}(it - \lambda) \int_0^{+\infty} \frac{\exp(itx - \lambda x)}{x^\alpha} dx \right]
\end{aligned}
$$

  Si denotamos por $H(t)$ a la última integral, es fácil verificar - integrando por partes - que

$$H'(t) = i \frac{1 - \alpha}{\lambda - it} H(t)$$

  y que

$$H(0) = \lambda^{\alpha - 1} \Gamma(1 - \alpha)$$

  de modo que

$$H(t) = \Gamma(1 - \alpha) \, (\lambda - it)^{\alpha - 1}$$

  donde la potencia de base compleja debe interpretarse como rama principal, que está bien definida ya que $\lambda - it$ está siempre en el cuarto cuadrante del plano (λ, t). Por lo tanto

$$
\begin{aligned}
J_\alpha(t) &= -\lim_{\lambda \downarrow 0} \frac{1}{\alpha} \Gamma(1 - \alpha) \, (\lambda - it)^\alpha \\
&= -\lim_{\lambda \downarrow 0} \frac{1}{\alpha} \Gamma(1 - \alpha) \exp \left[\alpha \left(\log \sqrt{\lambda^2 + t^2} + i\theta \right) \right]
\end{aligned}
$$

  donde log denota el logaritmo real y θ es el argumento del complejo $\lambda - it$, $-\pi/2 < \theta < 0$. El pasaje al límite es ahora inmediato y se obtiene:

$$J_\alpha(t) = -\frac{1}{\alpha} \Gamma(1 - \alpha) \; t^\alpha \exp \left(i\alpha \frac{\pi}{2} \right)$$

  La segunda integral en (65) se calcula de manera enteramente análoga, y eso permite obtener el resultado.

- si $\alpha > 1 \implies \tau_\alpha(x) = x$. El cálculo sigue los lineamientos del anterior, después de haber reemplazado $J_\alpha(t)$ por la nueva integral

$$\widetilde{J}_\alpha(t) = \int_0^{+\infty} \frac{\exp(itx) - 1 - itx}{x^{\alpha+1}} dx$$

- si $\alpha = 1 \implies \tau_1(x) = senx.$

Para este caso, el cálculo es similar, sólo que el punto clave es calcular la integral

$$J_1(t) = \int_0^{+\infty} \frac{\cos(tx) - 1}{x^2} dx + i \int_0^{+\infty} \frac{sen(tx) - t.senx}{x^2} dx$$

El cálculo de la primera integral es sencillo (hacer el cambio de variables $tx = y$ y luego integrar por partes). En cuanto a la segunda, observar que, si $t > 0$:

$$\begin{aligned}
\int_0^{+\infty} \frac{sen(tx) - t.senx}{x^2} dx &= \lim_{\delta \downarrow 0} \int_\delta^{+\infty} \frac{sen(tx) - t.senx}{x^2} dx \\
&= \lim_{\delta \downarrow 0} t \int_{\delta t}^{+\infty} \frac{senw}{w^2} dw - t \int_\delta^{+\infty} \frac{senw}{w^2} dw \\
&= t \lim_{\delta \downarrow 0} \int_{\delta t}^{\delta} \frac{senw}{w^2} dw = t \lim_{\delta \downarrow 0} \int_t^1 \frac{sen(\delta y)}{\delta y^2} dy \\
&= t \int_t^1 \frac{dy}{y} = -t \log t.
\end{aligned}$$

Reemplazando, se obtiene el resultado anunciado.$\square$

Definición. Sea G una distribución de probabilidad (propia) en la recta, que no está concentrada en un sólo punto. Se dice que la distribución F pertenece al dominio de atracción de G si para cada $n = 1, 2, ...$ existen constantes reales $a_n, b_n, a_n > 0$ tales que si $X_1, ..., X_n$ son variables aleatorias indendientes con igual distribución F y $S_n = X_1 + ... + X_n$, entonces, cuando $n \to +\infty$

$$\frac{S_n - b_n}{a_n} \Longrightarrow G \tag{66}$$

Teorema (TLC estable)
G tiene un dominio de atracción si y sólo si G es estable.

Demostración. Si G es una distribución de probabilidad estable, es obvio que pertenece a su propio dominio de atracción, en virtud de la definición de distribución estable. Lo interesante es la afirmación recíproca.

Supongamos entonces que existe una distribución de probabilidad F que pertenece al dominio de atracción de G, es decir que se verifica (66) Probaremos que G es estable.

Observemos primero que para cada pareja de enteros positivos m, n las variables aleatorias
$$\frac{S_{mn} - b_{mn}}{a_{mn}}, \quad \frac{S_{mn} - nb_m}{a_m}$$
difieren solamente en un cambio de posición y escala (es decir, una se obtiene de la otra mediante una afinidad).

Para cada n fijo, la primera tiende débilmente a G cuando $m \to +\infty$ mientras que la segunda tiende débilmente a $G^{*(n)} = G * G * ... * G$. La conclusión será probada, entonces, si demostramos que G y $G^{*(n)}$ difieren solamente en una afinidad (es decir que $\exists$ constantes reales $a, b, a \neq 0$, tales que $G(x) = G^{*(n)}(ax + b)$ $\forall x \in \mathcal{R}$. Esto está contenido en la proposición auxiliar siguiente:

Proposición auxiliar. Sea $\{X_m\}_{m=1,2,..}$ una sucesión de variables aleatorias con distribuciones respectivas $\{F_m\}_{m=1,2,..}$ que converge débilmente, cuando $m \to +\infty$ a la distribución (propia) F, no concentrada en un sólo punto. Sea además $\left\{\widetilde{X}_m\right\}_{m=1,2,..}$ la sucesión de variables aleatorias $\widetilde{X}_m = c_m X_m + d_m$, donde c_m, d_m $(m = 1, 2, ...)$ son números reales, $c_m > 0$. Supongamos además que $\widetilde{X}_m \Longrightarrow \widetilde{F}$ (propia) y que $\widetilde{F}$ tampoco está concentrada en un sólo punto.

Entonces, necesariamente las sucesiones $\{c_m\}$ y $\{d_m\}$ convergen, $c_m \to c, d_m \to d$, con $c > 0$ y $\widetilde{F}(cx + d) = F(x)$ $\forall x \in \mathcal{R}$.

Demostración. Denotemos por $\varphi_m, \widetilde{\varphi}_m, \varphi, \widetilde{\varphi}$ respectivamente, las transformadas de Fourier de $F_m, \widetilde{F}_m, F, \widetilde{F}$.

Observemos que alcanza con demostrar que las sucesiones $\{c_m\}$,$\left\{\frac{1}{c_m}\right\}$, $\{d_m\}$ están acotadas. En efecto, si así fuera, y si $\{c_{m_k}\}$ y $\{d_{m_k}\}$ son subsucesiones convergentes a c^* y d^* respectivamente, entonces:

- $c^* > 0$
- $\varphi_{m_k}(t) \to \varphi(t)$ $\forall t \in \mathcal{R}$
- $\widetilde{\varphi}_{m_k}(t) = \exp[i.d_{m_k} t]\ \ \varphi_{m_k}(c_{m_k} t) \to \exp[i.d^* t]\ \ \varphi(c^* t) = \widetilde{\varphi}(t)$ $\forall t \in \mathcal{R}$ (para verificar esta convergencia, usar el hecho de que la convergencia de las transformadas de Fourier es uniforme en cada compacto).

En consecuencia, $\widetilde{F}(x) = F\left(\frac{x - d^*}{c^*}\right)$ $\forall x \in \mathcal{R}$, o sea que $\widetilde{F}(c^* x + d^*) = F(x)$ $\forall x \in \mathcal{R}$.

Esto identifica los límites c^*, d^* que no dependen de la subsucesión. En consecuencia, las sucesiones enteras $\{c_m\}$,$\{d_m\}$ convergen y los límites verifican la propiedad enunciada.

Veamos a continuación la acotación de $\{c_m\}$,$\left\{\frac{1}{c_m}\right\}$, $\{d_m\}$.

Como $\widetilde{F}$ no está concentrada en un punto, podemos encontrar x', x'', puntos de continuidad de $\widetilde{F}$, tales que $x' < x''$ y

$$0 < \widetilde{F}(x') \leq \widetilde{F}(x'') < 1$$

y también puntos de continuidad de F, y', y'', tales que

$$F(y') < \widetilde{F}(x') \leq \widetilde{F}(x'') < F(y'') \tag{67}$$

esto último, simplemente porque $F(-\infty) = 0, F(+\infty) = 1$. Es obvio además que

$$\widetilde{F}_m(c_m x + d_m) = P(c_m X_m + d_m \leq c_m x + d_m) = F_m(x)$$

y como $F_m(y') \to F(y')$, y $\widetilde{F}_m(x') \to \widetilde{F}(x')$, para m suficientemente grande es $F_m(y') < \widetilde{F}_m(x') \implies \widetilde{F}_m(c_m y' + d_m) < \widetilde{F}_m(x') \implies c_m y' + d_m < x'$.

Del mismo modo, usando la desigualdad de la derecha en (67), resulta que $c_m y'' + d_m > x''$ y podemos concluir que $c_m(y'' - y') > x'' - x' \implies \frac{1}{c_m} \leq \frac{y''-y'}{x''-x'}$ si m es suficientemente grande. Esto prueba que $\left\{\frac{1}{c_m}\right\}$ es acotada.

Intercambiando los roles de F_m y $\widetilde{F}_m$ resulta que $\{c_m\}$ también es acotada. Finalmente, de las mismas desigualdades se tiene que $x'' - c_m y'' < d_m < x' - c_m y'$ para m suficientemente grande, lo que prueba que $\{d_m\}$ es acotada.$\square$

Es también posible caracterizar las distribuciones que pertenecen al dominio de atracción de una cierta ley estable. Ése es el contenido del teorema siguiente, que enunciamos sin demostración (para la prueba y otra serie de aspectos del mismo problema, ver Feller, Vol. 2, Cap. XVII).

Usamos la notación siguiente:

Si F es una distribución de probabilidad en la recta,

$$s_F(x) = \int_{[-x,x]} y^2 F(dy).$$

Además, diremos que la función $L : \mathcal{R}^+ \to \mathcal{R}$ es "de variación lenta en $+\infty$" si para cada $x > 0$ fijo, se tiene

$$\frac{L(tx)}{L(t)} \to 1 \text{ cuando } t \to +\infty$$

(ejemplos típicos de funciones de variación lenta en $+\infty$ son: (a) el logaritmo, (b) las funciones que tienen límite finito no nulo en $+\infty$).

Teorema (caracterización de los dominios de atracción)

(a) F pertenece al dominio de atracción de la distribución normal, si y sólo si, la función s_F es de variación lenta en $+\infty$.

(b) F pertenece al dominio de atracción de una distribución estable no normal, si y sólo si se cumplen las dos condiciones siguientes con $0 < \alpha < 2$:

$$s_F(x) \sim x^{2-\alpha} L(x) \quad \text{para} \;\; x \to +\infty \text{ con } L \text{ de variación lenta en} +\infty \tag{68}$$

$$\exists \lim_{x \to +\infty} \frac{1 - F(x)}{1 - F(x) + F(-x)} = p \tag{69}$$

En las condiciones del teorema, (68) es equivalente a

$$1 - F(x) + F(-x) \sim \frac{2-\alpha}{\alpha} x^{-\alpha} L(x)$$

para $x \to +\infty$ con L de variación lenta en $+\infty$. Asimismo, de cumplirse las condiciones de (b), la medida canónica de la ley estable está dada por

$$M\left((0,x)\right) = Cpx^{2-\alpha}, M\left((-x,0)\right) = C(1-p)x^{2-\alpha} \quad \forall x \geq 0,$$

donde C es una constante positiva.