Capítulo 3

Convergencia Débil y Funciones Características

3.1. Convergencia en Distribución

Definición 3.1 Si F es una f.d. definimos el *conjunto de puntos de continuidad* o *conjunto de continuidad* de F por

$$\mathcal{C}(F) = \{x \in \mathbb{R} : F \text{ es continua en } x\}$$

Un intervalo finito I con extremos $a < b$ es un *intervalo de continuidad* para F si $a, b \in \mathcal{C}(F)$.

Definición 3.2 Sea $(X_n,\ n \geq 1)$ una sucesión de v.a. con f.d. $F_{X_n},\ n \geq 1$. X_n *converge en distribución* a la v.a. X cuando $n \to \infty$ si

$$F_{X_n}(x) \to F_X(x) \quad \text{cuando } n \to \infty,\ \forall x \in \mathcal{C}(F_X).$$

Notación: $X_n \xrightarrow{d} X$ o $X_n \xrightarrow{w} X$ cuando $n \to \infty$.

Observación 3.1 Como en esta definición sólo intervienen las funciones de distribución, las variables no necesariamente están definidas en un mismo espacio de probabilidad.

Abusando la notación, escribiremos $X_n \xrightarrow{w} \mathcal{N}(0,1)$ en lugar de $X_n \xrightarrow{w} X$ cuando $n \to \infty$ donde $X \sim \mathcal{N}(0,1)$.

Ejemplo 3.1
Sea $X_n \sim \delta_{1/n}$, es decir, la delta de Dirac concentrada en el punto $1/n$. De acuerdo a la definición 3.2, $X_n \xrightarrow{d} \delta_0$ cuando $n \to \infty$. Tenemos

$$F_{X_n}(x) = \begin{cases} 0, & \text{si } x < 1/n, \\ 1, & \text{si } x \geq 1/n, \end{cases} \quad \to \quad \begin{cases} 0, & \text{si } x \leq 0, \\ 1, & \text{si } x > 0. \end{cases}$$

Por lo tanto $F_{X_n} \to F_{\delta_0}(x)$ cuando $n \to \infty$ para todo $x \in \mathcal{C}(F_{\delta_0})$ pero no para todo x. Si en cambio tuviéramos $Y_n \sim \delta_{-1/n}$ entonces

$$F_{Y_n}(x) = \begin{cases} 0, & \text{si } x < -1/n, \\ 1, & \text{si } x \geq -1/n, \end{cases} \quad \to \quad \begin{cases} 0, & \text{si } x < 0, \\ 1, & \text{si } x \geq 1/n. \end{cases}$$

y en este caso si tenemos convergencia para todo x.

Teorema 3.1 *Sean X y X_n, $n \geq 1$ v.a. Si $X_n \xrightarrow{P} X$ entonces $X_n \xrightarrow{d} X$.*

Demostración. Sea $\varepsilon > 0$, entonces

$$\begin{aligned} F_{X_n}(x) = P(X_n \leq x) &= P(X_n \leq x,\ |X_n - X| \leq \varepsilon) + P(X_n \leq x,\ |X_n - X| > \varepsilon) \\ &\leq P(X \leq x + \varepsilon,\ |X_n - X| \leq \varepsilon) + P(|X_n - X| > \varepsilon) \\ &\leq P(X \leq x + \varepsilon) + P(|X_n - X| > \varepsilon), \end{aligned}$$

y por la convergencia en probabilidad

$$\limsup_{n\to\infty} F_{X_n}(x) \leq F_X(x+\varepsilon).$$

De manera similar, cambiando X_n por X, x por $x - \varepsilon$, X por X_n y $x + \varepsilon$ por x, sigue que

$$\liminf_{n\to\infty} F_{X_n}(x) \geq F_X(x-\varepsilon).$$

Estas dos relaciones valen para todo x y todo $\varepsilon > 0$. Suponiendo ahora que $x \in \mathcal{C}(F_X)$ y haciendo $\varepsilon \to 0$ obtenemos que

$$F_X(x) = F_X(x^-) \leq \liminf_{n\to\infty} F_{X_n}(x) \leq \limsup_{n\to\infty} F_{X_n}(x) \leq F_X(x).$$

■

Observación 3.2 Si F_X tiene un salto en x, sólo podemos concluir que

$$F_X(x^-) \leq \liminf_{n\to\infty} F_{X_n}(x) \leq \limsup_{n\to\infty} F_{X_n}(x) \leq F_X(x).$$

Como $F_X(x) - F_X(x^-) > 0$, no es posible obtener convergencia en el punto x.

Ejemplo 3.2
Sea $N \sim \mathcal{N}(0,1)$, de modo que la distribución es simétrica. Definimos $X_n = (-1)^n N$ para $n \geq 1$, entonces $X_n \stackrel{d}{=} N$ de modo que $X_n \xrightarrow{d} N$, pero (X_n) no converge ni c.s. ni en probabilidad.

Teorema 3.2 *Sea $(X_n,\ n \geq 1)$ una sucesión de v.a.i. y c una constante, entonces*

$$X_n \xrightarrow{d} \delta_c \quad (n \to \infty) \quad \Leftrightarrow \quad X_n \xrightarrow{P} c \quad (n \to \infty).$$

Demostración. Supongamos que $X_n \xrightarrow{d} \delta_c$ cuando $n \to \infty$ y sea $\varepsilon > 0$. Si $c - \varepsilon,\ c + \varepsilon \in \mathcal{C}(F_X)$ entonces

$$\begin{aligned} P(|X_n - c| > \varepsilon) &= 1 - P(c - \varepsilon \leq X_n \leq c + \varepsilon) \\ &= 1 - F_{X_n}(c+\varepsilon) + F_{X_n}(c-\varepsilon) - P(X_n = c - \varepsilon) \\ &\leq 1 - F_{X_n}(c+\varepsilon) + F_{X_n}(c-\varepsilon) \\ &\to 0 \quad \text{cuando } n \to \infty, \end{aligned}$$

ya que $F_{X_n}(c+\varepsilon) \to F_X(c+\varepsilon) = 1,\ F_{X_n}(c-\varepsilon) \to F_X(c-\varepsilon) = 0,.$ ■

3.1.1. Caracterización de la Convergencia en Distribución

Denotamos por $\mathcal{C}_0(\mathbb{R})$ a la clase de funciones reales continuas tales que $|f(x)| \to 0$ cuando $x \to \pm\infty$. Una función cualquiera en un espacio arbitrario tiene *soporte* en un subconjunto S del espacio si f se anula fuera del conjunto S. Llamamos $\mathcal{C}_K(\mathbb{R})$ a la clase de funciones continuas a soporte compacto. Finalmente, llamamos $\mathcal{C}_B(\mathbb{R})$ a las funciones continuas y acotadas. Tenemos

$$\mathcal{C}_K \subset \mathcal{C}_0 \subset \mathcal{C}_B \subset \mathcal{C}.$$

En consecuencia, si $f \in \mathcal{C}_K(\mathbb{R})$, f tiene soporte en algún conjunto compacto y por lo tanto también en un intervalo compacto. Una *función escalera* en el intervalo (finito o infinito) (a, b) es una función con soporte en este intervalo tal que para alguna partición finita $a = a_1 < \cdots < a_k = b$ se tiene que $f(x) = c_j$ para $x \in (a_j, a_{j+1})$ y $1 \leq j \leq k$, donde los c_j son números reales. Decimos que la función es a valores en D si $c_j \in D$ para $1 \leq j \leq k$. Observamos que hemos dejado sin especificar los valores de la función f en los puntos de la partición para tener flexibilidad.

Lema 3.1 (Lema de Aproximación) *Sea f una función a valores reales tal que se cumple alguna de las siguientes condiciones:*

- $f \in \mathcal{C}_0$.
- $f \in \mathcal{C}[a, b]$ *para* $a, b \in \mathbb{R}$.

Sea A un subconjunto denso de $\mathbb{R}$ entonces, para cada $\varepsilon > 0$ existe una función escalera g en (a, b) definida usando una partición con valores en A tal que

$$\sup_{x \in A} |f(x) - g(x)| < \varepsilon. \tag{3.1}$$

Demostración. El resultado es consecuencia del teorema de Stone-Weierstrass. ■

Lema 3.2 *Sea $\{X_n,\ n \geq 1\}$ una sucesión de v.a. y supongamos que $X_n \xrightarrow{d} X$ cuando $n \to \infty$. Si h es una función continua a valores reales o complejos definida en el intervalo acotado $[a, b]$, donde $a, b \in \mathcal{C}(F_X)$, entonces*

$$\mathrm{E}\, h(X_n) \to \mathrm{E}\, h(X) \quad \textit{cuando } n \to \infty. \tag{3.2}$$

Demostración. El caso complejo sigue del caso real. Usaremos el lema de aproximación. Sea $A \subset \mathcal{C}(F_X) \subset \mathbb{R}$ un subconjunto denso numerable. Si $h(x) = \mathbf{1}_{(c,d]}$ para $c, d \in A$, $a \leq c < d \leq b$, (3.2) se reduce a demostrar que $P(c < X_n \leq d) \to P(c < X \leq d)$, lo cual es cierto por hipótesis. Algo similar ocurre para intervalos de otro tipo cuyo extremos pertenezcan a A. Por linealidad la conclusión vale para funciones escalera cuyos 'escalones' tengan extremos en A. Sea ahora $h \in \mathcal{C}[a, b]$, y sea g una función simple aproximante, entonces usando la desigualdad de Jensen,

$$\begin{aligned} |\mathrm{E}\, h(X_n) - \mathrm{E}(h(X)| &\leq |\mathrm{E}\, h(X_n) - \mathrm{E}\, g(X_n)| + |\mathrm{E}\, g(X_n) - \mathrm{E}\, g(X)| + |\mathrm{E}\, g(X) - \mathrm{E}\, h(X)| \\ &\leq \mathrm{E}\, |h(X_n) - g(X_n)| + |\mathrm{E}\, g(X_n) - \mathrm{E}\, g(X)| + \mathrm{E}\, |g(X) - h(X)| \\ &\leq \varepsilon + |\mathrm{E}\, g(X_n) - \mathrm{E}\, g(X)| + \varepsilon. \end{aligned}$$

Como ya hemos visto que para funciones simples el término central tiende a cero cuando $n \to \infty$ obtenemos

$$\limsup_{n \to \infty} |\mathrm{E}\, h(X_n) - \mathrm{E}\, h(X)| < 2\varepsilon,$$

lo cual demuestra el teorema. ■

Teorema 3.3 *Sea $\{X_n,\ n \geq 1\}$ una sucesión de v.a. y supongamos que $X_n \xrightarrow{d} X$ cuando $n \to \infty$. Si h es una función a valores reales o complejos, continua y acotada, entonces*

$$\mathrm{E}\, h(X_n) \to \mathrm{E}\, h(X) \quad \textit{cuando } n \to \infty. \tag{3.3}$$

Demostración. De nuevo, el caso complejo sigue del caso real. Supongamos que $|h(x)| \leq M$ para todo x y sea $K > 0$, entonces

$$\begin{aligned} |\mathrm{E}\, h(X_n) - \mathrm{E}\, h(X)| &\leq |\mathrm{E}\, h(X_n)\mathbf{1}_{\{|X_n| \leq K\}} - \mathrm{E}\, h(X)\mathbf{1}_{\{|X| \leq K\}}| + |\mathrm{E}\, h(X_n)\mathbf{1}_{\{|X_n| > K\}}| + |\mathrm{E}\, h(X)\mathbf{1}_{\{|X| > K\}}| \\ &\leq |\mathrm{E}\, h(X_n)\mathbf{1}_{\{|X_n| \leq K\}} - \mathrm{E}\, h(X)\mathbf{1}_{\{|X| \leq K\}}| + \mathrm{E}\,|h(X_n)|\mathbf{1}_{\{|X_n| > K\}} + \mathrm{E}\,|h(X)|\mathbf{1}_{\{|X| > K\}} \\ &\leq |\mathrm{E}\, h(X_n)\mathbf{1}_{\{|X_n| \leq K\}} - \mathrm{E}\, h(X)\mathbf{1}_{\{|X| \leq K\}}| + MP(|X_n| > K) + MP(|X| > K) \end{aligned}$$

Sea $\varepsilon > 0$ y escojamos $K \in \mathcal{C}(F_X)$ suficientemente grande como para que $2MP(|X| > K) < \varepsilon$. Usando el lema 3.2 para el primer término y la convergencia en distribución para el segundo, obtenemos que

$$\limsup_{n \to \infty} |\mathrm{E}\, h(X_n) - \mathrm{E}\, h(X)| < 2MP(|X| > K) < \varepsilon.$$

■

Teorema 3.4 *Sea $\{X_n,\ n \geq 1\}$ una sucesión de v.a. Entonces $X_n \xrightarrow{d} X$ cuando $n \to \infty$ si y sólo si*

$$\mathrm{E}\, h(X_n) \to \mathrm{E}\, h(X) \quad \textit{cuando } n \to \infty. \tag{3.4}$$

para toda función real continua y acotada.

Demostración. Por el teorema 3.3 basta demostrar la suficiencia. Sean $a, b \in \mathcal{C}(F)$, $-\infty < a < b < \infty$. Ponemos

$$g_k(x) = \begin{cases} 0 & \text{para } x < a - \delta_k, \\ \frac{x-(a-\delta_k)}{\delta_k} & \text{para } x \in [a - \delta_k, a], \\ 1 & \text{para } x \in [a, b], \\ \frac{b+\delta_k - x}{\delta_k} & \text{para } x \in [b, b + \delta_k], \\ 0 & \text{para } x > b + \delta_k \end{cases}$$

con $\delta_k \downarrow 0$ cuando $k \to \infty$ y observamos que $\mathbf{1}_{(a,b]}(x) \leq g_k(x)$. Supongamos que (3.4) vale. Por la monotonía de la función de distribución y el teorema 3.3 obtenemos

$$F_n(a, b] = \int_a^b dF_n(x) \leq \int g_k(x) dF_n(x) \leq \mathrm{E}\, g_k(X_n) \to \mathrm{E}\, g_k(X) \quad \text{cuando } n \to \infty.$$

Haciendo $n \to \infty$

$$\limsup_{n \to \infty} F_n(a, b] \leq \mathrm{E}\, g_k(X) \leq F([a - \delta_k, b + \delta_k]).$$

Haciendo ahora $k \to \infty$ tenemos $\delta_k \downarrow 0$ y $g_k(x) \downarrow \mathbf{1}_{[a,b]}$, y como a y b son puntos de continuidad concluimos que

$$\limsup_{n \to \infty} F_n(a, b] \leq F(a, b]. \tag{3.5}$$

Si, en cambio,

$$h_k(x) = \begin{cases} 0 & \text{para } x < a, \\ \frac{x-a}{\delta_k} & \text{para } x \in [a, a + \delta_k], \\ 1 & \text{para } x \in [a + \delta_k, b - \delta_k], \\ \frac{b-x}{\delta_k} & \text{para } x \in [b - \delta_k, b], \\ 0 & \text{para } x > b \end{cases}$$

los mismos argumentos nos dan

$$\liminf_{n\to\infty} F_n(a,b] \geq F([a+\delta_k, b-\delta_k]),$$

y como ahora $h_k(x) \uparrow \mathbf{1}_{(a,b)}$, tenemos finalmente

$$\liminf_{n\to\infty} F_n(a,b] \geq F(a,b]. \tag{3.6}$$

Las ecuaciones (3.5) y (3.6) demuestran que $X_n \xrightarrow{d} X$ cuando $n \to \infty$. ■

Como corolario obtenemos los siguientes resultados:

Corolario 3.1 *Sea $\{F_n,\ n \geq 0\}$ una familia de funciones de distribución. Las siguientes afirmaciones son equivalentes*

(1) $F_n \xrightarrow{d} F_0$.

(2) Para toda función real f acotada y uniformemente continua,

$$\int f\, dF_n \to \int f\, dF_0,$$

cuando $n \to \infty$.

Demostración. Basta observar que las funciones g_k y h_k que usamos en la demostración del teorema anterior son continuas con soporte compacto y por lo tanto son uniformemente continuas. ■

Teorema 3.5 *Sean X e Y v.a.*

$$X \stackrel{d}{=} Y \quad \Leftrightarrow \quad \mathrm{E}\, h(X) = \mathrm{E}\, h(Y)$$

para toda función h continua y acotada.

Teorema 3.6 *Sean X y $\{X_n,\ n \geq 1\}$ v.a. y supongamos que $X_n \xrightarrow{d} X$ cuando $n \to \infty$. Entonces*

$$\mathrm{E}(|X|) \leq \liminf_{n\to\infty} \mathrm{E}(|X_n|).$$

Demostración. Sea $K \in \mathcal{C}(F_X)$ un número positivo. Por el teorema 3.2 tenemos

$$\liminf_{n\to\infty} \mathrm{E}(|X_n|) \geq \liminf_{n\to\infty} \mathrm{E}(|X_n| \mathbf{1}_{\{|X_n| \leq K\}}) = \mathrm{E}(|X| \mathbf{1}_{\{|X| \leq K\}}).$$

La conclusión sigue haciendo K tender a infinito a través de una sucesión de puntos de continuidad de F_X. ■

3.2. El Lema de Scheffé

La pregunta que nos hacemos ahora es cual es la relación entre convergencia débil y convergencia de densidades.

Ejemplo 3.3
Sea $X_n,\ n \geq 1$ una sucesión de v.a. con densidades

$$f_n(x) = \begin{cases} 1 - \cos(2\pi n x), & \text{para } 0 \leq x \leq 1, \\ 0, & \text{en otro caso.} \end{cases}$$

Entonces

$$F_n(x) = \begin{cases} 0, & \text{para } x \leq 0, \\ x - \frac{\text{sen}(2\pi n x)}{2\pi n} \to x, & \text{para } 0 \leq x \leq 1, \\ 1, & \text{para } x \geq 1. \end{cases}$$

lo cual implica que $X_n \xrightarrow{d} \mathcal{U}(0,1)$ cuando $n \to \infty$. Sin embargo, las densidades oscilan y por lo tanto no pueden converger.

Definición 3.3 La *distancia en variación total* entre las f.d. F y G se define como

$$d(F,G) = \sup_{A \in \mathcal{B}} |F(A) - G(A)|.$$

Si las variables X e Y tienen f.d. F y G respectivamente, la definición es equivalente a

$$d(X,Y) = \sup_{A \in \mathcal{B}} |P(X \in A) - P(Y \in A)|.$$

Si X y X_n, $n \geq 1$ son v.a. tales que

$$d(X_n, X) \to 0 \quad (n \to \infty),$$

decimos que X_n *converge a* X *en variación total* cuando $n \to \infty$.

Como los conjuntos $(-\infty, x]$ son de Borel para todo x, tenemos que

$$|P(X_n \leq x) - P(X \leq x)| \leq \sup_{A \in \mathcal{B}} |P(X_n \in A) - P(X \in A)|,$$

lo cual implica el siguiente resultado.

Lema 3.3 *Sea $\{X_n,\ n \geq 1\}$ una sucesión de v.a. Si $X_n \to X$ en variación total cuando $n \to \infty$ entonces $X_n \xrightarrow{d} X$ cuando $n \to \infty$.*

Teorema 3.7 (Lema de Scheffé) *Supongamos que X y $X_n,\ n \geq 1$ son v.a. absolutamente continuas. Entonces*

$$\sup_{A \in \mathcal{B}} |P(X_n \in A) - P(X \in A)| \leq \int_{\mathbb{R}} |f_{X_n}(x) - f_X(x)|\, dx, \tag{3.7}$$

y si $f_{X_n}(x) \to f_X(x)$ c.s. cuando $n \to \infty$, entonces $d(X_n, X) \to 0$ cuando $n \to \infty$. En particular, $X_n \xrightarrow{d} X$ cuando $n \to \infty$.

Demostración. Sea $A \in \mathcal{B}$, tenemos

$$\begin{aligned} |P(X_n \in A) - P(X \in A)| &= \Big| \int_A f_{X_n}(x) - \int_A f_X(x)\, dx \Big| \\ &\leq \int_A |f_{X_n}(x) - f_X(x)|\, dx \\ &\leq \int_{\mathbb{R}} |f_{X_n}(x) - f_X(x)|\, dx, \end{aligned}$$

lo cual demuestra (3.7).

Para demostrar la convergencia observamos que

$$\int_{\mathbb{R}} (f_{X_n}(x) - f_X(x))\,dx = 1 - 1 = 0,$$

de modo que

$$\int_{\mathbb{R}} (f_{X_n}(x) - f_X(x))^+\,dx = \int_{\mathbb{R}} (f_{X_n}(x) - f_X(x))^-\,dx.$$

Además, $(f_X(x) - f_{X_n}(x))^+ \to 0$ c.s. cuando $n \to \infty$, y $0 \leq (f_X(x) - f_{X_n}(x))^+ \leq f_X(x)$, y usando el teorema de convergencia dominada

$$\int_{\mathbb{R}} |f_X(x) - f_{X_n}(x)| dx = 2\int_{\mathbb{R}} (f_X(x) - f_{X_n}(x))^+ dx \to 0 \quad (n \to 0).$$

Esto demuestra la convergencia en variación total. ■

Observación 3.3 Es posible demostrar que

$$d(X_n, X) = \frac{1}{2}\int_{\mathbb{R}} |f_{X_n}(x) - f_X(x)|\,dx.$$

3.3. Convergencia Vaga

En nuestra definición de convergencia en distribución hemos supuesto que la función de distribución límite F es propia, es decir que $F(\mathbb{R}) = 1$ pero este no es siempre el caso. Esta es la diferencia fundamental entre la convergencia en distribución y un nuevo tipo de convergencia que definimos a continuación.

Definición 3.4 Una sucesión $\{F_n,\ n \geq 1\}$ de f.d. converge vagamente a la f.d (posiblemente impropia) H si, para todo intervalo finito $I = (a, b]$ con $a, b \in \mathcal{C}(H)$,

$$F_n(I) \to H(I) \quad (n \to \infty)$$

Notación: $F \xrightarrow{v} H$ cuando $n \to \infty$.

Ejemplo 3.4
Sea $P(X_n = n) = P(X_n = -n) = 1/2$, entonces

$$F_n(x) = \begin{cases} 0, & \text{para } x < -n, \\ \frac{1}{2}, & \text{para } -n \leq x < n, \\ 1, & \text{para } x \geq n. \end{cases}$$

De modo que para todo $x \in \mathbb{R}$

$$F_n(x) \to H(x) = \frac{1}{2} \quad (n \to \infty).$$

H tiene todas las propiedades de una función de distribución salvo que $F(\mathbb{R}) = 0$.

Ejemplo 3.5
Sea $X_n \sim \mathcal{U}(-n, n),\ n \geq 1$. La función de distribución es

$$F_{X_n}(x) = \begin{cases} 0, & \text{para } x < -n, \\ \frac{x+n}{2n}, & \text{para } -n \leq x < n, \\ 1, & \text{para } x \geq n. \end{cases}$$

De nuevo, esta sucesión converge cuando $n \to \infty$ a $H(x) = 1/2$ para todo $x \in \mathbb{R}$.

Si $X_n \xrightarrow{d} X$ cuando $n \to \infty$ entonces, por definición, $F_{X_n}(x) \to F_X(x)$ cuando $n \to \infty$ para $x \in \mathcal{C}(F_X)$, de donde se obtiene de inmediato que para cualquier intervalo $I = (a, b]$ acotado tal que $a, b \in \mathcal{C}(F_X)$,

$$F_{X_n}(I) = F_{X_n}(b) - F_{X_n}(a) \to F_X(b) - F_X(a) = F_X(I)$$

cuando $n \to \infty$, de modo que convergencia en distribución siempre implica convergencia vaga.

Teorema 3.8 (Principio de Selección de Helly) *Sea $\{F_n,\ n \geq 1\}$ una sucesión de f.d. Entonces existe una subsucesión no-decreciente $\{n_k,\ k \geq 1\}$ tal que*

$$F_{n_k} \xrightarrow{v} H \quad \textit{cuando } k \to \infty,$$

para alguna función de distribución H, posiblemente impropia.

Demostración. La demostración usa el método diagonal. Sea $\{r_k,\ k \geq 1\}$ una enumeración de los racionales. Como la sucesión $\{F_n(r_1), n \geq 1\}$ está acotada, por el teorema de Bolzano-Weierstrass existe un punto de acumulación j_1 y una subsucesión $\{F_{1,k}(r_1),\ k \geq 1\}$ tal que

$$F_{1,k}(r_1) \to j_1 \quad \text{cuando } k \to \infty,$$

donde $0 \leq j_1 \leq 1$. El mismo argumento aplicado a la subsucesión $\{F_{2,k}(r_2), k \geq 1\}$ produce un punto de acumulación j_2 y una subsubsucesión $\{F_{2,k}(r_2), k \geq 1\}$ tal que

$$F_{2,k}(r_2) \to j_2 \quad \text{cuando } k \to \infty,$$

donde $0 \leq j_2 \leq 1$. Además, esta subsucesión converge también para r_1, ya que es una subsucesión de la primera subsucesión. Continuando este procedimiento obtenemos, cuando $k \to \infty$,

$$\begin{aligned}
F_{1,k}(r_1) &\to j_1, & & \\
F_{2,k}(r_i) &\to j_i, & &\text{para } i = 1, 2, \\
F_{3,k}(r_i) &\to j_i, & &\text{para } i = 1, 2, 3 \\
&\ \vdots & &\ \vdots \\
F_{m,k}(r_i) &\to j_i, & &\text{para } i = 1, 2, \dots, m \\
&\ \vdots & &\ \vdots
\end{aligned}$$

Ahora consideramos la diagonal de este arreglo, $\{F_{k,k},\ k \geq 1\}$ que, salvo por un número finito de términos al comienzo, es una subsucesión de todas las subsucesiones horizontales, y por lo tanto converge en todo $\mathbb{Q}$. Tenemos, por lo tanto, un conjunto denso numerable en el cual la sucesión converge.

Definimos ahora la función H: Ponemos

$$H(r_i) = j_i, \quad \text{para } r_i \in \mathbb{Q}.$$

Como $0 \leq j_i \leq 1$ para todo i tenemos que $0 \leq H(r_i) \leq 1$ para $r_i \in \mathbb{Q}$. Además, si $r < s,\ r, s \in \mathbb{Q}$, entonces

$$0 \leq F_{k,k}(s) - F_{k,k}(r) \to H(s) - H(r)$$

cuando $k \to \infty$, lo cual muestra que H es no-decreciente en $\mathbb{Q}$.

Sea ahora $x \in \mathbb{R}$ arbitrario y escogemos $r, s \in \mathbb{Q}$ tal que $r < x < s$. Como $F_{k,k}(r) \leq F_{k,k}(x) \leq F_{k,k}(s)$, tomando límites obtenemos que

$$H(r) \leq \liminf_{k\to\infty} F_{k,k}(x) \leq \limsup_{k\to\infty} F_{k,k}(x) \leq H(s),$$

y como podemos escoger r y s tan cerca como queramos de x,

$$H(x^-) \leq \liminf_{k\to\infty} F_{k,k}(x) \leq \limsup_{k\to\infty} F_{k,k}(x) \leq H(x^+)$$

Esto prueba que siempre existe una subsucesión que converge a una función no-decreciente H que toma valores en $[0,1]$ en sus puntos de continuidad. Para que H sea continua por la derecha, completamos su definición poniendo

$$H(x) = \lim_{r_i \downarrow x} H(r_i) \quad \text{para } x \in \mathbb{R},$$

es decir, escogemos la versión de H que es continua por la derecha. ■

Observación 3.4 Es importante notar que el teorema de Helly no dice que $F_{k,k}(\infty) \to H(\infty)$.

Si llamamos $\mathcal{M}$ a la clase de las funciones de distribución posiblemente degeneradas, es decir, a las funciones $H : \mathbb{R} \to \mathbb{R}$ que son no-decrecientes, continuas por la derecha y satisfacen $G(-\infty) \geq 0$, $G(\infty) \leq 1$, entonces el teorema de Helly dice que $\mathcal{M}$ es *relativamente o secuencialmente compacto*. Ya hemos visto en los ejemplos 3.4 y 3.5 que la clase de las funciones de distribución no es compacta.

Corolario 3.2 *Sea F_n una sucesión en $\mathcal{M}$. Si existe $H \in \mathcal{M}$ tal que para toda subsucesión vagamente convergente (F_{n_k}) se tiene que $F_{n_k} \xrightarrow{v} H$, entonces la sucesión también converge vagamente a H: $F_n \xrightarrow{v} H$.*

Demostración. Supongamos que F_n no converge vagamente a H, entonces existe $x_0 \in \mathcal{C}(H)$ tal que $F_n(x_0) \nrightarrow H(x_0)$. Pero toda subsucesión de las F_n tiene una subsucesión convergente F_{n_k} y $F_{n_k}(x_0) \to H(x_0)$. Esto implica que $F_n(x_0) \to H(x_0)$, lo cual es una contradicción. ■

Hemos visto en el teorema 3.4 que las integrales de funciones en $\mathcal{C}_B$ caracterizan la convergencia en distribución: $F_n \xrightarrow{d} F$ si y sólo si $\int f\,dF_n \to \int f\,dF$ para toda $f \in \mathcal{C}_B(\mathbb{R})$. El siguiente teorema nos da una caracterización similar para la convergencia vaga. Primero unas definiciones y un lema de aproximación.

Teorema 3.9 *Sea $\{H_n,\ n \geq 1\}$ una sucesión de funciones en $\mathcal{M}$. Entonces $H_n \xrightarrow{v} H$ si y sólo si para toda $f \in \mathcal{C}_K$,*

$$\int_{\mathbb{R}} f\,dH_n \to \int f\,dH. \tag{3.8}$$

Demostración. Supongamos que $H_n \xrightarrow{v} H$. (3.8) es cierta si f es la función indicadora de $(a,b]$ para $a,b \in A$ donde A es un subconjunto denso de $\mathcal{C}(H)$. Por la linealidad de las integrales también es cierto si f es una función escalera definida a partir de una partición con valores en A. Si $f \in \mathcal{C}_K$ y $\varepsilon > 0$, por el lema de aproximación existe una función escalera f_ε que satisface (3.1). Tenemos

$$\left|\int f\,dH_n - \int f\,dH\right| \leq \left|\int (f-f_\varepsilon)\,dH_n\right| + \left|\int f_\varepsilon\,dH_n - \int f_\varepsilon\,dH\right| + \left|\int (f_\varepsilon - f)\,dH\right|. \tag{3.9}$$

El primer término del lado derecho está acotado por

$$\int |f - f_\varepsilon|\,dH_n \leq \varepsilon \int dH_n \leq \varepsilon,$$

y una acotación similar sirve para el tercer término. El segundo término tiende a cero porque f_ε es una función escalera cuyos intervalos tienen extremos en A. En consecuencia, el lado derecho de (3.9) está acotado por 2ε cuando $n \to \infty$ y por lo tanto converge a 0.

Para demostrar el recíproco basta observar que las funciones g_k y h_k que definimos para la demostración del teorema 3.4 son funciones continuas a soporte compacto y por lo tanto el mismo argumento demuestra la convergencia

$$H_n(a,b] \to H(a,b] \quad (n \to \infty)$$

para $a,b \in \mathcal{C}(H)$. ■

3.3.1. Convergencia Vaga y Tensión.

Ahora consideraremos la relación entre convergencia vaga y convergencia en distribución.

Teorema 3.10 (Prohorov) *Sea $\{F_n,\ n \geq 1\}$ una sucesión de f.d. Para que toda subsucesión que converge vagamente converja en distribución es necesario y suficiente que*

$$\int_{|x|>a} dF_n(x) \to 0 \quad \textit{cuando } a \to \infty \tag{3.10}$$

uniformemente en n.

Demostración.

Supongamos que $F_{n_k} \xrightarrow{v} H$ para alguna $H \in \mathcal{M}$ cuando $k \to \infty$ y que (3.10) vale. Como la masa total de H es finita tenemos, para cualquier $\varepsilon > 0$, existe a_H tal que

$$\int_{|x|>a} dH(x) < \varepsilon \quad \text{para } a > a_H.$$

Además, por hipótesis tenemos que

$$\sup_n \int_{|x|>a} dF_n(x) < \varepsilon \quad \text{para } a > a_0.$$

Entonces, para $a > \text{máx}\{a_h, a_0\}$, tenemos que para todo k,

$$\begin{aligned}
1 - \int_{-\infty}^{\infty} dH(x) &= \int_{-\infty}^{\infty} dF_{n_k}(x) - \int_{-\infty}^{\infty} dH(x) \\
&\leq \int_{|x|>a} dF_{n_k}(x) + \Big| \int_{-a}^{a} dF_{n_k}(x) - \int_{-a}^{a} dH(x) \Big| + \int_{|x|>a} dH(x) \\
&\leq \varepsilon + \Big| \int_{-a}^{a} dF_{n_k}(x) - \int_{-a}^{a} dH(x) \Big| + \varepsilon
\end{aligned}$$

de modo que, por el teorema 3.9,

$$1 - \int_{-\infty}^{\infty} dH(x) < 2\varepsilon,$$

lo cual implica que la masa total de H es 1.

Veamos ahora la demostración de la necesidad. Sabemos que $F_n \xrightarrow{d} H$ cuando $n \to \infty$, de modo que H es propia, y queremos verificar que (3.10) vale. Supongamos lo contrario, entonces, dado $\varepsilon > 0$ existen sucesiones crecientes y no acotadas $\{n_k,\ k \geq 1\}$ y $\{a_k,\ k \geq 1\}$ tales que

$$\int_{|x|>a_k} dF_{n_k}(x) > 2\varepsilon.$$

Por otro lado, como antes,

$$\int_{|x|>a} dH(x) < \varepsilon \quad \text{para } a > a_H,$$

Además, como sólo estamos considerando una cantidad numerable de distribuciones y una f.d. tiene, a lo sumo, una cantidad numerable de discontinuidades, podemos suponer sin restricción que a y $(a_k)_{k\geq 1}$ son puntos de continuidad de todas las distribuciones que consideramos.

Dado cualquier $a > a_H$ existe un k_0 tal que $a_k > a$ para $k > k_0$. Sean a y $k > k_0$ dados, combinando los resultados anteriores obtenemos

$$\begin{aligned}1 - \int_{-\infty}^{\infty} dH(x) &= \int_{-\infty}^{\infty} dF_{n_k}(x) - \int_{-\infty}^{\infty} dH(x) \\ &= \int_{|x|>a} dF_{n_k}(x) - \int_{|x|>a} dH(x) + \Big(\int_{-a}^{a} dF_{n_k}(x) - \int_{-a}^{a} dH(x) \Big) \\ &> 2\varepsilon - \varepsilon + \Big(\int_{-a}^{a} dF_{n_k}(x) - \int_{-a}^{a} dH(x) \Big)\end{aligned}$$

y por el teorema 3.9,

$$1 - \int_{-\infty}^{\infty} dH(x) > \varepsilon,$$

de modo que H no es propia. Esta contradicción concluye la demostración del teorema. ■

Una sucesión de distribuciones que satisface (3.10) no permite que la masa escape a infinito por la uniformidad de la condición en n: Todas las distribuciones tienen masas uniformemente pequeñas fuera de intervalos suficientemente grandes. Decimos que una sucesión de distribuciones que satisfacen esta propiedad es *tensa*. Si $\{X_n,\ n \geq 1\}$ es una sucesión de v.a. cuyas distribuciones asociadas forman una sucesión tensa decimos que la sucesión $\{X_n,\ n \geq 1\}$ es *tensa*.

El teorema de Prohorov dice que una familia de f.d. es tensa si y sólo si es relativamente compacta. Combinando el teorema de Prohorov y el corolario 3.2 obtenemos el siguiente resultado.

Corolario 3.3 *Sea $\{F_n,\ n \geq 1\}$ una sucesión tensa de f.d. Si todas las subsucesiones convergentes de F_n convergen al mismo límite F, entonces $F_n \xrightarrow{d} F$.*

Observamos que para cualquier $p > 0$

$$P(|X_n| > a) \leq \frac{1}{a^p} \mathrm{E}\, |X_n|^p \mathbf{1}_{\{|X_n|>a\}} \leq \frac{1}{a^p} \sup_n \mathrm{E}\, |X_n|^p,$$

de donde obtenemos la primera parte del siguiente corolario.

Corolario 3.4 *(i) Sea $p > 0$. Una sucesión acotada en L^p es tensa. En particular, sucesiones uniformemente integrables son tensas.*

(ii) Sea $g \uparrow \infty$ una función no negativa y supongamos que $\sup_n \mathrm{E}\, g(X_n) < \infty$. Entonces $\{X_n,\ n \geq 1\}$ es tensa.

Usando el teorema 3.10 podemos demostrar el siguiente resultado.

Teorema 3.11 *Sea $\{X_n,\ n \geq 1\}$ una sucesión tensa. Entonces $X_n \xrightarrow{d} X$ cuando $n \to \infty$ si y sólo si toda subsucesión contiene una subsucesión que converge vagamente a X.*

3.3.2. Clases Separantes de Funciones

Definición 3.5 Una clase $\mathcal{E}$ de funciones continuas y acotadas en $\mathbb{R}$ es *separante* si para cualquier par F, G de f.d. se tiene que

$$\int f\, dF = \int f\, dG, \quad \forall f \in \mathcal{E} \quad \Rightarrow \quad F = G.$$

Proposición 3.1 *Sea $\mathcal{E}$ una clase separante y $\{F_n,\ n \geq 1\}$ una sucesión tensa de f.d. Entonces existe una f.d. F tal que $F_n \xrightarrow{d} F$ si y sólo si*

$$\lim_n \int f\, dF_n \text{ existe para toda } f \in \mathcal{E}.$$

Si esto es cierto, entonces $\lim_n \int f\, dF_n = \int f\, dF$, para toda $f \in \mathcal{E}$.

Demostración. Si $F_n \xrightarrow{d} F$ entonces $\lim_n \int f dF_n = \int f\, dF$ porque las funciones en $\mathcal{E}$ son continuas y acotadas. Para ver el recíproco sea F_{n_k} una subsucesión convergente. Por la tensión $F_{n_k} \xrightarrow{d} F$ donde F es una f.d. propia. Si tomamos cualquier otra subsucesión débilmente convergente $F_{n'_k} \xrightarrow{d} G$, para $f \in \mathcal{E}$ se tiene que

$$\lim_k \int f\, dF_{n_k} = \int f\, dF, \quad \lim_k \int f\, dF_{n'_k} = \int f\, dG,$$

de modo que $\int f\, dF = \int f\, dG$ para toda $f \in \mathcal{E}$. Esto implica que $F = G$ y en consecuencia todas las subsucesiones convergentes de F_n tienen el mismo límite F. Por el corolario 3.3 tenemos $F_n \xrightarrow{d} F$. ■

Corolario 3.5 *Sea $\mathcal{E}$ una clase separante y $\{F_n,\ n \geq 1\}$ una colección tensa de f.d. Si F es una f.d. tal que $\int f\, dF_n \to \int f\, dF$ para toda $f \in \mathcal{E}$, entonces $F_n \xrightarrow{d} F$.*

3.4. El Teorema de Skorohod

Sabemos que convergencia casi segura implica convergencia débil y que el recíproco es falso. Sin embargo, el teorema de Skorohod nos da un recíproco parcial.

Teorema 3.12 (Skorohod) *Sea $\{X_n,\ n \geq 0\}$ una sucesión de v.a. tales que $X_n \xrightarrow{d} X_0$. Entonces existen v.a. $\{Y_n,\ n \geq 0\}$ definidas en el espacio de Lebesgue $([0,1], \mathcal{B}[0,1], m)$ tales que para cada $n \geq 0$ fijo, $X_n \stackrel{d}{=} Y_n$ y $Y_n \xrightarrow{c.s.} Y_0$ donde la convergencia casi segura es respecto a la medida de Lebesgue.*

Para la demostración del teorema necesitamos el siguiente resultado

Lema 3.4 *Sea F_n la f.d. de X_n, de modo que $F_n \xrightarrow{d} F_0$. Si $t \in (0,1) \cap \mathcal{C}(F_0^{\leftarrow})$, entonces $F_n^{\leftarrow}(t) \to F_0^{\leftarrow}(t)$.*

Demostración. Como $\mathcal{C}(F_0)^c$ es a lo sumo numerable, dado $\varepsilon > 0$ existe $x \in \mathcal{C}(F_0)$ tal que

$$F_0^{\leftarrow}(t) - \varepsilon < x < F_0^{\leftarrow}(t).$$

A partir de la definición de la función inversa, $x < F_0^{\leftarrow}(t)$ implica que $F_0(x) < t$. Además, $x \in \mathcal{C}(F_0)$ implica $F_n(x) \to F_0(x)$, de modo que para n grande tenemos $F_n(x) < t$. De nuevo, usando la definición de la función inversa, tenemos $x \leq F_n^{\leftarrow}(t)$. Por lo tanto,

$$F_0^{\leftarrow}(t) - \varepsilon < x \leq F_n^{\leftarrow}(t).$$

para todo n grande, y como $\varepsilon > 0$ es arbitrario, obtenemos que

$$F_0^{\leftarrow}(t) \leq \liminf_{n\to\infty} F_n^{\leftarrow}(t). \tag{3.11}$$

Para cualquier $t' > t$ podemos hallar $y \in \mathcal{C}(F_0)$ tal que

$$F_0^{\leftarrow}(t') < y < F_0^{\leftarrow}(t') + \varepsilon.$$

y esto implica $F_0(y) \geq t' > t$. Como $y \in \mathcal{C}(F_0)$, $F_n(y) \to F_0(y)$ y para n grande, $F_n(y) \geq t$, de modo que $y \geq F_n^{\leftarrow}(t)$ y en consecuencia

$$F_n^{\leftarrow}(t) \leq y < F_0^{\leftarrow}(t') + \varepsilon.$$

para todo n grande. Como $\varepsilon > 0$ es arbitrario obtenemos

$$\limsup_{n\to\infty} F_n^{\leftarrow}(t) \leq F_0^{\leftarrow}(t').$$

Hacemos ahora $t' \downarrow t$ y usamos la continuidad de $F_0^{\leftarrow}$ en t para obtener

$$\limsup_{n\to\infty} F_n^{\leftarrow}(t) \leq F_0^{\leftarrow}(t). \tag{3.12}$$

Combinando (3.11) y (3.12) obtenemos el resultado. ■

Demostración del teorema de Skorohod. En el espacio $[0,1]$ definimos la variable aleatoria $U(t) = t$ de modo que U tiene distribución uniforme, ya que para $0 \leq x \leq 1$

$$m(U \leq x) = m(t \in [0,1] : U(t) \leq x) = m([0,x]) = x.$$

Para $n \geq 0$ definimos Y_n en $[0,1]$ por $Y_n = F_n^{\leftarrow}(U)$. Entonces para $y \in \mathbb{R}$,

$$m(Y_n \leq y) = m(t \in [0,1] : F_n^{\leftarrow}(t) \leq y) = m(t \in [0,1] : t \leq F_n(y)) = F_n(y),$$

y concluimos que $Y_n \stackrel{d}{=} X_n$ para todo $n \geq 0$. Ahora escribimos

$$m(t \in [0,1] : Y_n(t) \nrightarrow Y_0(t)) = m(t \in [0,1] : F_n^{\leftarrow}(t) \nrightarrow F_0^{\leftarrow}(t)),$$

y usando el lema 3.4, esto esta acotado por

$$m(t \in [0,1] : F_0^{\leftarrow} \text{ no es continua en } t) = 0$$

porque el conjunto de discontinuidades es numerable. ■

3.4.1. El Método Delta

Supongamos que tenemos una familia de modelos $\{(\Omega, \mathcal{F}, P_\theta),\ \theta \in \Theta\}$, y queremos estimar el parámetro θ con base en una muestra de tamaño n usando el estadístico $T_n = T_n(X_1, \dots, X_n)$. El estimador T_n es consistente si

$$T_n \xrightarrow{P_\theta} \theta,$$

para cada θ. El estimador es consistente y asintóticamente normal (CAN) si para todo $\theta \in \Theta$

$$\lim_{n\to\infty} P_\theta(\sigma_n(T_n - \theta) \leq x) = \mathcal{N}(x; 0, 1),$$

para alguna sucesión $\sigma_n \to \infty$.

Supongamos que tenemos un estimador CAN de θ pero queremos estimar una función suave $g(\theta)$. Por ejemplo, en una familia de densidades exponenciales, θ puede representar la media pero estamos interesados en la varianza θ^2. Usando el método delta vemos que $g(T_n)$ también es CAN para $g(\theta)$.

Ilustramos el método usando el TCL que demostraremos más adelante. Sea $\{X_j,\ j \geq 1\}$ una sucesión i.i.d. con $\mathrm{E}(X_n) = \mu$ y $\mathrm{Var}(X_n) = \sigma^2$. El TCL dice que

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} \mathcal{N}(0,1).$$

Equivalentemente, podemos expresar este resultado en términos de $\bar{X}_n = \sum_1^n X_i/n$ como

$$\sqrt{n}\Big(\frac{\bar{X}_n - \mu}{\sigma}\Big) \xrightarrow{d} \mathcal{N}(0,1),$$

de modo que $\bar{X}$ es un estimador CAN de μ. El método delta dice que si $g(x)$ tiene derivada no nula $g'(\mu)$ en μ entonces

$$\sqrt{n}\Big(\frac{g(\bar{X}_n) - g(\mu)}{\sigma g'(\mu)}\Big) \xrightarrow{d} \mathcal{N}(0,1). \tag{3.13}$$

Vemos que $g(\bar{X})$ es CAN para $g(\mu)$.

Observación 3.5 La demostración no depende de que la variable límite tenga distribución $\mathcal{N}(0,1)$.

Demostración de (3.13). Por el teorema de Skorohod existen variables Z'_n y N' en el espacio de Lebesgue tales que

$$Z'_n \stackrel{d}{=} \sqrt{n}\Big(\frac{\bar{X}_n - \mu}{\sigma}\Big), \qquad N' \sim \mathcal{N}(0,1)$$

y

$$Z'_n \to N' \quad c.s.\ (m).$$

Definimos

$$\bar{X}'_n = \mu + \sigma Z'_n/\sqrt{n},$$

de modo que $\bar{X}_n \stackrel{d}{=} \bar{X}'_n$. Usando la definición de derivada

$$\begin{aligned}\sqrt{n}\Big(\frac{g(\bar{X}_n) - g(\mu)}{\sigma g'(\mu)}\Big) &\stackrel{d}{=} \sqrt{n}\Big(\frac{g(\mu + \sigma Z'_n/\sqrt{n}) - g(\mu)}{\sigma g'(\mu)}\Big)\\ &= \frac{g(\mu + \sigma Z'_n/\sqrt{n}) - g(\mu)}{\sigma Z'_n/\sqrt{n}}\,\frac{Z'_n}{g'(\mu)}\\ &\xrightarrow{c.s.} g'(\mu)\frac{N'}{g'(\mu)} = N' \sim \mathcal{N}(0,1),\end{aligned}$$

ya que $\sigma Z'_n/\sqrt{n} \to 0$ c.s. ■

3.5. Convergencia en Distribución II

Teorema 3.13 (Slutsky) *Sean $\{X, X_n, Y_n, \xi_n, n \geq 1\}$ variables aleatorias. Si $X_n \xrightarrow{d} X$ y $X_n - Y_n \xrightarrow{P} 0$ entonces $Y_n \xrightarrow{d} X$. Equivalentemente, si $X_n \xrightarrow{d} X$ y $\xi_n \xrightarrow{P} 0$ entonces $X_n + \xi_n \xrightarrow{d} X$.*

Demostración. Basta probar la segunda afirmación. Sea f una función real, acotada y uniformemente continua. Definimos el módulo de continuidad por

$$\omega_\delta(f) = \sup_{|x-y|\leq\delta} |f(x) - f(y)|.$$

Como f es uniformemente continua tenemos que

$$\omega_\delta(f) \to 0, \qquad \text{cuando } \delta \to 0. \tag{3.14}$$

Por el corolario 3.1 basta ver que $\mathrm{E}(f(X_n + \xi_n)) \to \mathrm{E}(f(X))$. Para esto observemos que

$$\begin{aligned}|\,\mathrm{E}(f(X_n+\xi_n)) - \mathrm{E}(f(X))| &\leq |\,\mathrm{E}(f(X_n+\xi_n)) - \mathrm{E}(f(X_n))| + |\,\mathrm{E}(f(X_n)) - \mathrm{E}(f(X))|\\ &\leq \mathrm{E}\left(|f(X_n+\xi_n) - f(X_n)|\mathbf{1}_{\{|\xi_n|\leq\delta\}}\right) + 2\sup_x |f(x)| P(|\xi_n| > \delta) + o(1)\\ &\leq \omega_\delta(f) + \mathrm{Const}\, P(|\xi_n| > \delta) + o(1).\end{aligned}$$

Haciendo $n \to \infty$ y luego $\delta \to 0$ y usando (3.14) obtenemos el resultado. ∎

A continuación presentamos una generalización del teorema de Slutsky. La demostración de este resultado puede verse en el libro de Resnick.

Teorema 3.14 *Sean $\{X_{un}, X_u, Y_n, X, n \geq 1\}$ variables aleatorias tales que para cada n, Y_n, X_{un}, $u \geq 1$ están definidas en un mismo espacio de probabilidad. Supongamos que para cada u, cuando $n \to \infty$, $X_{un} \xrightarrow{d} X_u$ y cuando $u \to \infty$, $X_u \xrightarrow{d} X$. Supongamos además que para todo $\varepsilon > 0$,*

$$\lim_{u\to\infty} \limsup_{n\to\infty} P(|X_{un} - Y_n| > \varepsilon) = 0.$$

Entonces tenemos $Y_n \xrightarrow{d} X$ cuando $n \to \infty$.

3.6. Teorema de Convergencia a Familias

Muchos resultados de convergencia de variables aleatorias son del siguiente tipo: Para una sucesión de v.a. $\xi_n, n \geq 1$ y constantes $a_n > 0$ y $b_n \in \mathbb{R}$, se demuestra que

$$\frac{\xi_n - b_n}{a_n} \xrightarrow{d} Y,$$

donde Y es una v. a. no-degenerada. Usando esto tenemos

$$P(\frac{\xi_n - b_n}{a_n} \leq x) \approx P(Y \leq x) = G(x),$$

o poniendo $y = a_n x + b_n$,

$$P(\xi_n \leq y) \approx G(\frac{y - b_n}{a_n}).$$

Esto permite aproximar la distribución de ξ_n por una familia de distribuciones con parámetros de ubicación y escala.

La pregunta ahora es ¿Hasta qué punto son únicas estas constantes de normalización a_n y b_n? La respuesta la da el teorema de convergencia a familias de distribuciones: Las constantes están determinadas salvo por equivalencias asintóticas y la distribución límite está determinada salvo por parámetros de ubicación y escala.

Definición 3.6 Dos distribuciones F y G son *del mismo tipo* o *pertenecen a la misma familia* si para algunas constantes $a > 0$, $b \in \mathbb{R}$,

$$G(x) = F(ax + b), \qquad x \in \mathbb{R}.$$

En términos de variables aleatorias, si $X \sim F$ y $Y \sim G$ entonces

$$Y \stackrel{d}{=} \frac{X - b}{a}.$$

Por ejemplo, podemos considerar la familia gaussiana. Si $X_{\mu,\sigma}$ tiene distribución $N(\mu, \sigma^2)$, entonces $X_{\mu,\sigma} \stackrel{d}{=} \sigma X_{0,1} + \mu$.

Teorema 3.15 (Convergencia a familias, Gnedenko & Khinchin) *Sean $G(x)$ y $H(x)$ dos funciones de distribución propias, ninguna de las cuales está concentrada en un punto. Supongamos que para $n \geq 1$, X_n son v.a. con funciones de distribución F_n y U y V son v.a. con f.d. G y H, respectivamente. Sean, además, constantes $a_n > 0$, $\alpha_n > 0$, $b_n \in \mathbb{R}, \beta_n \in \mathbb{R}$.*

a) Si

$$F_n(a_n x + b_n) \to G(x), \quad F_n(\alpha_n x + \beta_n) \to H(x) \tag{3.15}$$

o equivalentemente

$$\frac{X_n - b_n}{a_n} \xrightarrow{d} U, \quad \frac{X_n - \beta_n}{\alpha_n} \xrightarrow{d} V,$$

entonces existen constantes $A > 0$ y $B \in \mathbb{R}$ tales que, cuando $n \to \infty$,

$$\frac{\alpha_n}{a_n} \to A > 0, \quad \frac{\beta_n - b_n}{a_n} \to B, \tag{3.16}$$

y

$$H(x) = G(Ax + B), \quad V \stackrel{d}{=} \frac{U - B}{A}. \tag{3.17}$$

b) Recíprocamente, si (3.16) vale, entonces cualquiera de las relaciones en (3.15) implica la otra y (3.17) vale.

Demostración

Veamos primero la demostración de (b). Supongamos que

$$G_n(x) := F_n(a_n x + b_n) \to G(x)$$

y

$$\frac{\alpha_n}{a_n} \to A > 0, \quad \frac{\beta_n - b_n}{a_n} \to B,$$

entonces

$$F_n(\alpha_n x + \beta_n) = G_n(\frac{\alpha_n}{a_n} x + \frac{\beta_n - b_n}{a_n}).$$

Escogemos x de modo que $Ax + B \in \mathcal{C}(G)$. Supongamos que $x > 0$, un argumento similar sirve si $x \leq 0$. Dado $\varepsilon > 0$ para n grande tenemos

$$(A - \varepsilon)x + B - \varepsilon \leq \frac{\alpha_n}{a_n} x + \frac{\beta_n - b_n}{a_n} \leq (A + \varepsilon)x + B + \varepsilon$$

y en consecuencia

$$\limsup_{n\to\infty} F_n(\alpha_n x + \beta_n) \leq \limsup_{n\to\infty} G_n((A + \varepsilon)x + B + \varepsilon).$$

Esto implica que para cualquier $z \in \mathcal{C}(G)$ con $z > (A + \varepsilon)x + B + \varepsilon$ se tiene

$$\limsup_{n\to\infty} F_n(\alpha_n x + \beta_n) \leq \limsup_{n\to\infty} G_n(z) = G(z).$$

Por lo tanto,

$$\limsup_{n\to\infty} F_n(\alpha_n x + \beta_n) \leq \inf\{G(z) : z > (A + \varepsilon)x + B + \varepsilon, z \in \mathcal{C}(G)\}.$$

y como $\varepsilon > 0$ es arbitrario,

$$\limsup_{n\to\infty} F_n(\alpha_n x + \beta_n) \leq \inf\{G(z) : z > Ax + B\} = G(Ax + B)$$

por la continuidad por la derecha de G. De manera similar,

$$\begin{aligned}\liminf_{n\to\infty} F_n(\alpha_n x + \beta_n) &\geq \liminf_{n\to\infty} G_n((A - \varepsilon)x + B - \varepsilon) \\ &\geq \liminf_{n\to\infty} G_n(z) = G(z)\end{aligned}$$

para cualquier $z < (A-\varepsilon)x + B - \varepsilon$, $z \in \mathcal{C}(G)$. Como estas desigualdades valen para todo $\varepsilon > 0$,

$$\liminf_{n\to\infty} F_n(\alpha_n x + \beta_n) \geq \sup\{G(z) : z < Ax + B,\ z \in \mathcal{C}(G)\} = G(Ax + B)$$

porque $Ax + B \in \mathcal{C}(G)$.

Veamos ahora la demostración de la parte (a). Supongamos que

$$G_n(x) := F_n(a_n x + b_n) \to G(x), \qquad H_n(x) := F_n(\alpha_n x + \beta_n) \to H(x).$$

Usando el lema 3.4 tenemos que $G_n^{\leftarrow}(y) \to G^{\leftarrow}(y)$ para $y \in \mathcal{C}(G^{\leftarrow})$ y $H_n^{\leftarrow}(y) \to H^{\leftarrow}(y)$ para $y \in \mathcal{C}(H^{\leftarrow})$, y es fácil verificar que

$$G_n^{\leftarrow}(y) = \frac{F_n^{\leftarrow}(y) - b_n}{a_n}, \qquad H_n^{\leftarrow}(y) = \frac{F_n^{\leftarrow}(y) - \beta_n}{\alpha_n}.$$

Por lo tanto

$$\begin{aligned} \frac{F_n^{\leftarrow}(y) - b_n}{a_n} &\to G^{\leftarrow}(y), \quad y \in \mathcal{C}(G^{\leftarrow}), \\ \frac{F_n^{\leftarrow}(y) - \beta_n}{\alpha_n} &\to H^{\leftarrow}(y), \quad y \in \mathcal{C}(H^{\leftarrow}). \end{aligned}$$

Como las f.d. $G(x)$ y $H(x)$ no están concentradas en un punto, podemos hallar $y_1 < y_2$ con $y_i \in \mathcal{C}(G^{\leftarrow}) \cap \mathcal{C}(H^{\leftarrow})$ para $i = 1, 2$, tales que

$$-\infty < G^{\leftarrow}(y_1) < G^{\leftarrow}(y_2) < \infty, \qquad -\infty < H^{\leftarrow}(y_1) < H^{\leftarrow}(y_2) < \infty.$$

Por lo tanto, para $i = 1, 2$ tenemos

$$\frac{F_n^{\leftarrow}(y_i) - b_n}{a_n} \to G^{\leftarrow}(y_i), \quad \frac{F_n^{\leftarrow}(y_i) - \beta_n}{\alpha_n} \to H^{\leftarrow}(y_i). \tag{3.18}$$

En las expresiones anteriores restamos las ecuaciones con $i = 1$ de las expresiones con $i = 2$ para obtener

$$\frac{F_n^{\leftarrow}(y_2) - F_n^{\leftarrow}(y_1)}{a_n} \to G^{\leftarrow}(y_2) - G^{\leftarrow}(y_1), \tag{3.19}$$

$$\frac{F_n^{\leftarrow}(y_2) - F_n^{\leftarrow}(y_1)}{\alpha_n} \to H^{\leftarrow}(y_2) - H^{\leftarrow}(y_1). \tag{3.20}$$

Ahora dividimos (3.19) entre (3.20) y obtenemos

$$\frac{\alpha_n}{a_n} \to \frac{G^{\leftarrow}(y_2) - G^{\leftarrow}(y_1)}{H^{\leftarrow}(y_2) - H^{\leftarrow}(y_1)} := A > 0.$$

Además, de (3.18) vemos que

$$\begin{aligned} \frac{F_n^{\leftarrow}(y_1) - b_n}{a_n} &\to G^{\leftarrow}(y_1), \\ \frac{F_n^{\leftarrow}(y_1) - \beta_n}{a_n} = \frac{F_n^{\leftarrow}(y_1) - \beta_n}{\alpha_n}\frac{\alpha_n}{a_n} &\to H^{\leftarrow}(y_1)A, \end{aligned}$$

y restando obtenemos

$$\frac{\beta_n - b_n}{a_n} \to H^{\leftarrow}(y_1)A - G^{\leftarrow}(y_1) := B,$$

como queríamos ver, de modo que (3.16) vale. Por la parte (b) obtenemos (3.17).

■

Observación 3.6 Una consecuencia de la demostración es que, a partir de (3.18), (3.19) y (3.20), una posible selección de las constantes de normalización es

$$a_n = F_n^{\leftarrow}(y_2) - F_n^{\leftarrow}(y_1), \qquad b_n = F_n^{\leftarrow}(y_1).$$

El siguiente ejemplo muestra la importancia de la hipótesis de que las distribuciones límite no estén concentradas en un punto.

Ejemplo 3.6
Sea

$$G(x) = \begin{cases} 0, & \text{si } t < c, \\ 1, & \text{si } t \geq c. \end{cases}$$

Entonces,

$$G^{\leftarrow}(t) = \inf\{y : G(y) \geq t\} = \begin{cases} -\infty, & \text{si } t = 0 \\ c, & \text{si } 0 < t \leq 1, \\ \infty, & \text{si } t > 1. \end{cases}$$

El siguiente corolario es consecuencia inmediata del teorema de convergencia a familias.

Corolario 3.6 *Sea F_n una sucesión de f. d. y $a_n > 0$ y b_n sucesiones de constantes tales que*

$$F_n(a_n x + b_n) \to G(x) \tag{3.21}$$

en todo punto de continuidad de G, que es una f.d. propia y no está concentrada en un punto. Sean $c_n > 0$ y d_n sucesiones de constantes tales que

$$\frac{a_n}{c_n} \to 1, \qquad \frac{d_n - b_n}{a_n} \to 0.$$

Entonces (3.21) vale con c_n y d_n en lugar de a_n y b_n.

3.7. Variables Aleatorias Complejas

Una variable aleatoria compleja X es una función medible $X : \Omega \to \mathbb{C}$ de la forma $X(\omega) = U(\omega) + iV(\omega)$ donde U, V son funciones medibles de Ω a valores reales, que se conocen como la parte real e imaginaria de X, respectivamente. Si U, V son integrables entonces X es integrable y tenemos que

$$\mathrm{E}(X) = \mathrm{E}(U) + i\,\mathrm{E}(V).$$

Teniendo en cuenta las desigualdades

$$\max(|U|, |V|) \leq |X| \leq |U| + |V|$$

observamos que al igual que en el caso real, X es integrable si y sólo si $|X|$ es integrable. También es posible demostrar la siguiente desigualdad, que queda como ejercicio:

$$|\,\mathrm{E}(X)| \leq \mathrm{E}(|X|).$$

3.8. Funciones Características

Definición 3.7 La función característica (f.c.) de una variable aleatoria X con f.d. F es la función a valores complejos definida por

$$\varphi_X(t) = \mathrm{E}(e^{itX}) = \int_{\mathbb{R}} e^{itx}\, dF(x).$$

Propiedades.
(a) Para cualquier v.a. X la función característica siempre existe:

$$|\,\mathrm{E}\, e^{itX}| \leq \mathrm{E}\,|e^{itX}| = 1.$$

(b) Toda función característica satisface $|\varphi(t)| \leq 1$ y $\varphi(0) = 1$.
(c) Toda f.c. es uniformemente continua:

$$\begin{aligned}|\varphi(t+h) - \varphi(t)| &= |\,\mathrm{E}\, e^{i(t+h)X} - \mathrm{E}\, e^{itX}| \\ &= |\,\mathrm{E}\, e^{itX}(e^{ihX} - 1)| \\ &\leq \mathrm{E}\,|e^{ihX} - 1| \to 0\end{aligned}$$

cuando $h \downarrow 0$ por el TCD. Como la cota es independiente de t, la convergencia es uniforme.
(d) Si X es una v.a. con f.c. φ_X, $Y = aX + b$ tiene f.c.

$$\varphi_Y(t) = \mathrm{E}(e^{itY}) = \mathrm{E}(e^{it(aX+b)}) = \varphi_X(at)e^{ibt}.$$

(e) Si $\bar{z}$ denota el conjugado complejo del número z, $\overline{\varphi_X(t)} = \varphi_X(-t) = \varphi_{-X}(t)$:

$$\overline{\varphi_X(t)} = \mathrm{E}(\overline{e^{itX}}) = \mathrm{E}(e^{-itX}) = \mathrm{E}(e^{i(-t)X}) = \mathrm{E}(e^{it(-X)}) = \varphi_{-X}(t).$$

(f) Tenemos

$$Re(\varphi(t)) = \int \cos(tx)\, dF(x); \qquad Im(\varphi(t)) = \int \mathrm{sen}(tx)\, dF(x)$$

donde la primera función es par mientras que la segunda es impar.
(g) La f.c. φ_X es real si y sólo si $X \stackrel{d}{=} -X$ si y sólo si la medida asociada a F_X es simétrica. Esto es cierto porque φ_X es real si y sólo si $\varphi_X = \bar{\varphi}_X$ si y sólo si X y $-X$ tienen la misma función característica. Por el teorema de unicidad que veremos más adelante, esto implica que $X \stackrel{d}{=} -X$.
(h) Si X e Y son independientes se tiene $\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$:

$$\varphi_{X+Y}(t) = \mathrm{E}(e^{it(X+Y)}) = \mathrm{E}(e^{itX}e^{itY}) = \mathrm{E}(e^{itX})\,\mathrm{E}(e^{itY}) = \varphi_X(t)\varphi_Y(t)$$

(i) Podemos generalizar la propiedad anterior de la siguiente manera: Si $X_1, \dots, X_n$ son i.i.d. con f.c. común φ, entonces

$$\mathrm{E}(e^{it(a_n^{-1}S_n - nb_n)}) = e^{-itnb_n}(\varphi(t/a_n))^n.$$

Distribución	Notación	Función Característica
Delta de Dirac	δ_a	e^{ita}
Bernoulli	Be(p)	$q + pe^{it}$
Binomial	Bin(n,p)	$(q + pe^{it})^n$
Geométrica	Ge(p)	$p/(1 - qe^{it})$
Poisson	Po(λ)	$e^{\lambda(e^{it}-1)}$

Tabla 1. Funciones características de algunas distribuciones discretas.

Distribución	Notación	Función Característica
Uniforme	$U(a,b)$	$\frac{e^{itb}-e^{ita}}{it(b-a)}$
	$U(0,1)$	$\frac{e^{it}-1}{it}$
	$U(-1,1)$	$\frac{\sin t}{t}$
Exponencial	$\text{Exp}(\lambda)$	$1/(1-\lambda it)$
Gamma	$\Gamma(p,\theta)$	$\left(\frac{1}{1-\theta it}\right)^p$
Normal	$\mathcal{N}(\mu,\sigma^2)$	$e^{it\mu-\frac{1}{2}t^2\sigma^2}$
	$\mathcal{N}(0,1)$	$e^{-\frac{1}{2}t^2}$
Cauchy	$C(0,1)$	$e^{-\lvert t\rvert}$
Estables Simétricas		$e^{-c\lvert t\rvert^\alpha}$

Tabla 2. Funciones características de algunas distribuciones continuas.

Proposición 3.2 (Riemann-Lebesgue) *Sea X una v.a. con f.d. F y f.c. φ_X. Si F tiene densidad entonces*

$$\lim_{|t|\to\infty} \big|\varphi_X(t)\big| = 0.$$

Demostración Sea f la densidad de X, entonces, para $t \in \mathbb{R}$

$$\varphi_X(t) = \int e^{itx} f(x)\,dx.$$

Sea $\varepsilon > 0$. Como f es integrable, existe una función escalera $f_\varepsilon(x) = \sum_{k=1}^n c_k \mathbf{1}_{(a_k,b_k)}$ con $k \in \mathbb{N}$ y $a_k, b_k, c_k \in \mathbb{R}$ para $1 \le k \le n$ tal que

$$\int |f - f_\varepsilon|\,dm < \frac{\varepsilon}{2} \tag{3.22}$$

Observamos que para $t \neq 0$,

$$\Big|\int e^{itx} f_\varepsilon(x)\,dx\Big| = \Big|\sum_{k=1}^n c_k \int_{a_k}^{b_k} e^{itx}\,dx\Big| \le \sum_{k=1}^n |c_k| \frac{2}{|t|} \tag{3.23}$$

Usando ahora (3.22) y (3.23) obtenemos

$$\begin{aligned} |\varphi_X(t)| &= \Big|\int e^{itx} f(x)\,dx\Big| \\ &\le \int |f - f_\varepsilon| dx + \Big|\int e^{itx} f_\varepsilon(x)\,dx\Big| < \varepsilon. \end{aligned}$$

para todo t con $|t|$ suficientemente grande. ■

Definición 3.8 Una función $\phi : \mathbb{R} \to \mathbb{C}$ es *definida no-negativa* si para cualesquiera $n \in \mathbb{N}, t_1, t_2, \ldots, t_n \in \mathbb{R}, \alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{C}$ se tiene que

$$\sum_{i=1}^n \sum_{j=1}^n \alpha_i \bar{\alpha}_j \phi(t_i - t_j) \ge 0.$$

Proposición 3.3 *Sea φ la f.c. de una v.a. X, entonces φ es definida no-negativa.*

Demostración Tomamos $n, t_i, \alpha_i, 1 \leq i \leq n$ como en la definición 3.8

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \bar{\alpha}_j \varphi(t_i - t_j) &= \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \bar{\alpha}_j \operatorname{E}(e^{iXi(t_i - t_j)}) \\
&= \operatorname{E}\Big(\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \bar{\alpha}_j e^{iXi(t_i - t_j)}\Big) \\
&= \operatorname{E}\Big(\Big|\sum_{i=1}^{n} \alpha_i e^{iXi(t_i)}\Big|^2\Big) \geq 0
\end{aligned}$$

El recíproco también es cierto y se conoce como el teorema de Bochner-Khinchin: Si φ es definida no-negativa, continua y $\varphi(0) = 1$ entonces es la función característica de una variable aleatoria X.

3.9. Desarrollos

3.9.1. Desarrollos de e^{ix}

Comenzamos con una integración por partes. Para $n \geq 0$ tenemos

$$\int_0^x (x-s)^n e^{is}\, ds = \frac{x^{n+1}}{n+1} + \frac{i}{n+1}\int_0^x (x-s)^{n+1} e^{is} ds. \tag{3.24}$$

Para $n = 0$ esta relación es

$$\int_0^x e^{is} ds = \frac{e^{ix} - 1}{i} = x + i\int_0^x (x-s) e^{is} ds,$$

de donde obtenemos

$$\begin{aligned}
e^{ix} &= 1 + ix + i^2 \int_0^x (x-s) e^{is}\, ds \\
&= 1 + ix + i^2 \big[\frac{x^2}{2} + \frac{i}{2}\int_0^x (x-s)^2 e^{is} ds\big] \\
&= 1 + ix + \frac{(ix)^2}{2} + \frac{i^3}{2}\big[\frac{x^3}{3} + \frac{i}{3}\int_0^x (x-s)^3 e^{is} ds\big] \\
&\ \ \vdots
\end{aligned}$$

En general obtenemos para $n \geq 0$ y $x \in \mathbb{R}$

$$e^{ix} = \sum_{k=0}^{n} \frac{(ix)^k}{k!} + \frac{i^{n+1}}{n!}\int_0^x (x-s)^n e^{is}\, ds. \tag{3.25}$$

Por lo tanto

$$\Big|e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!}\Big| \leq \frac{|x|^{n+1}}{(n+1)!}. \tag{3.26}$$

Concluimos que truncar el desarrollo de e^{ix} luego de un número finito de términos da una aproximación cuyo error está acotado por el módulo del primer término no tomado en cuenta.

Escribiendo ahora (3.24) con $n-1$ en lugar de n obtenemos

$$\int_0^x (x-s)^{n-1} e^{is} ds - \frac{x^n}{n} = \frac{i}{n}\int_0^x (x-s)^n e^{is} ds.$$

Si multiplicamos por $i^n/(n-1)!$ e intercambiamos ambos lados de la ecuación obtenemos

$$\frac{i^{n+1}}{n!}\int_0^x (x-s)^n e^{is}ds = \frac{i^n}{(n-1)!}\int_0^x (x-s)^{n-1}e^{is}ds - \frac{(ix)^n}{n!}.$$

Sustituyendo esta expresión en el lado derecho de (3.25) obtenemos

$$e^{ix} - \sum_{k=0}^{n}\frac{(ix)^k}{k!} = \frac{i^n}{(n-1)!}\int_0^x (x-s)^{n-1}e^{is}ds - \frac{(ix)^n}{n!},$$

y por lo tanto

$$\Big|e^{ix} - \sum_{k=0}^{n}\frac{(ix)^k}{k!}\Big| \leq \frac{|x|^n}{n!} + \frac{|x|^n}{n!} = \frac{2|x|^n}{n!}. \tag{3.27}$$

Combinando (3.26) y (3.27) obtenemos

$$\Big|e^{ix} - \sum_{k=0}^{n}\frac{(ix)^k}{k!}\Big| \leq \frac{|x|^{n+1}}{(n+1)!} \wedge \frac{2|x|^n}{n!}. \tag{3.28}$$

Observamos que el primer término del lado derecho da una mejor estimación para x pequeño mientras que el segundo es mejor para x grande.

3.9.2. Desarrollos de la Función Característica

Supongamos que X es una v.a. cuyos n primeros momentos absolutos son finitos: $\mathrm{E}\,|X|^k < \infty$ para $1 \leq k \leq n$, entonces

$$\begin{aligned}\Big|\varphi(t) - \sum_{k=0}^{n}\frac{(it)^k}{k!}\,\mathrm{E}(X^k)\Big| &= \Big|\mathrm{E}(e^{itX}) - \mathrm{E}\,\Big(\sum_{k=0}^{n}\frac{(it)^k}{k!}X^k\Big)\Big| \\ &\leq \mathrm{E}\,\Big|e^{itX} - \sum_{k=0}^{n}\frac{(itX)^k}{k!}\Big|\end{aligned}$$

y usando (3.28) con tX en lugar de x obtenemos

$$\Big|\varphi(t) - \sum_{k=0}^{n}\frac{(it)^k\,\mathrm{E}(X^k)}{k!}\Big| \leq \mathrm{E}\,\Big(\frac{|tX|^{n+1}}{(n+1)!} \wedge \frac{2|tX|^n}{n!}\Big). \tag{3.29}$$

Supongamos ahora que existen momentos de todos los órdenes y que para todo $t \in \mathbb{R}$ se tiene

$$\lim_{n\to\infty}\frac{|t|^n\,\mathrm{E}(|X|^n)}{n!} = 0. \tag{3.30}$$

Si ahora hacemos $n \to \infty$ en (3.29) obtenemos

$$\varphi(t) = \sum_{k=0}^{\infty}\frac{(it)^k}{k!}\,\mathrm{E}(X^k).$$

Una condición suficiente para (3.30) es

$$\sum_{k=0}^{\infty}\frac{|t|^k}{k!}\,\mathrm{E}(|X|^k) = \mathrm{E}(e^{|t||X|}) < \infty$$

la cual vale si $\Psi(t) = \mathrm{E}(e^{tX}) < \infty$ para todo $t \in \mathbb{R}$, es decir, si la función generadora de momentos existe en todo $\mathbb{R}$.

Ejemplo 3.7
Sea X una v.a. con distribución $\mathcal{N}(0,1)$. Para cualquier $t \in \mathbb{R}$

$$\begin{aligned}
\mathrm{E}(e^{tX}) &= \int_{-\infty}^{\infty} e^{tu} \frac{e^{-u^2/2}}{\sqrt{2\pi}} du \\
&= e^{t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\Big\{-\frac{1}{2}(u^2 - 2tu + t^2)\Big\} du \\
&= e^{t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(u-t)^2} du \\
&= e^{t^2/2}
\end{aligned}$$

y concluimos que, para todo $t \in \mathbb{R}$, $\mathrm{E}(e^{tX}) < \infty$. Por lo tanto podemos desarrollar tanto la fgm como $e^{t^2/2}$ para obtener

$$\mathrm{E}(e^{tX}) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \mathrm{E}(X^k), \qquad e^{t^2/2} = \sum_{l=0}^{\infty} \frac{1}{l!}\Big(\frac{t^2}{2}\Big)^l,$$

es decir,

$$\sum_{k=0}^{\infty} \frac{\mathrm{E}(X^k)}{k!} t^k = \sum_{l=0}^{\infty} \frac{2^{-l}}{l!} t^{2l}.$$

Igualando coeficientes obtenemos que

$$\frac{\mathrm{E}(X^{2n})}{(2n)!} = \frac{2^{-n}}{n!}; \qquad \frac{\mathrm{E}(X^{2n+1})}{(2n+1)!} = 0$$

y concluimos que

$$\mathrm{E}(X^{2n}) = \frac{(2n)!}{n!} 2^{-n}, \qquad \mathrm{E}(X^{2n+1}) = 0$$

Ahora bien, como

$$\varphi(t) = \mathrm{E}\, e^{itX} = \sum_{k=0}^{\infty} \frac{(it)^k}{k!} \mathrm{E}(X^k),$$

obtenemos

$$\begin{aligned}
\varphi(t) &= \sum_{k=0}^{\infty} \frac{(it)^{2k}}{(2k)!} \mathrm{E}(X^{2k}) = \sum_{k=0}^{\infty} \frac{(it)^{2k}}{(2k)!} \frac{(2k)!}{k!} 2^{-k} \\
&= \sum \frac{(i^2 t^2 2^{-1})^k}{k!} = \sum_{k=0}^{\infty} \frac{1}{k!}\Big(\frac{-t^2}{2}\Big)^k \\
&= e^{-t^2/2}
\end{aligned}$$

de modo que, salvo por constantes multiplicativas, la densidad normal típica y su f.c. son iguales.

3.10. Momentos y Derivadas

Si el k-ésimo momento absoluto de X existe, podemos calcularlo usando la $k$ésima derivada de la f.c.

Teorema 3.16 *Sea X una v.a. con f.d. F y f.c. φ. Si $\mathrm{E}\,|X|^n < \infty$ para algún $n \geq 1$, entonces $\varphi^{(k)}$ existe para $k = 1, 2, \ldots, n$, es uniformemente continua y*

$$\varphi^{(k)}(t) = \int_{-\infty}^{\infty} (ix)^k e^{itx} dF(x), \tag{3.31}$$

$$\varphi^{(k)}(0) = i^k\,\mathrm{E}(X^k) \tag{3.32}$$

$$\varphi(t) = 1 + \sum_{k=1}^{n} \frac{(it)^k}{k!}\,\mathrm{E}(X^k) + o(|t|^n) \quad (t \to 0). \tag{3.33}$$

Demostración. Supongamos que $\mathrm{E}(|X|) < \infty$, entonces

$$\begin{aligned}\frac{\varphi(t+h) - \varphi(t)}{h} - \mathrm{E}(iXe^{itX}) &= \mathrm{E}\,\Big(\frac{e^{i(t+h)X} - e^{itX} - ihXe^{itX}}{h}\Big)\\ &= \mathrm{E}\,\Big(e^{itX}\,\frac{e^{ihX} - 1 - ihX}{h}\Big).\end{aligned}$$

Usando (3.27) con $n = 1$ obtenemos

$$\big|e^{itX}\big|\Big|\frac{e^{ihX} - 1 - ihX}{h}\Big| \leq 2|X| \in L^1.$$

Por otro lado usando (3.26) obtenemos

$$\Big|\frac{e^{ihX} - 1 - ihX}{h}\Big| \leq \frac{h^2X^2}{2h} = h\frac{X^2}{2} \to 0$$

cuando $h \downarrow 0$ y por el TCD vemos que

$$\lim_{h\downarrow 0}\Big(\frac{\varphi(t+h) - \varphi(t)}{h} - \mathrm{E}(iXe^{itX})\Big) = \mathrm{E}\,\Big(\lim_{h\downarrow 0} e^{itX}\Big(\frac{e^{ihX} - 1 - ihX}{h}\Big)\Big) = 0.$$

En consecuencia

$$\varphi'(t) = \int_{-\infty}^{\infty} ixe^{itx}\,dF(x).$$

El caso general de la fórmula sigue por inducción: Usando el mismo procedimiento

$$\begin{aligned}\frac{\varphi^{(k)}(t+h) - \varphi^{(k)}(t)}{h} &= \int_{-\infty}^{\infty} e^{itx}(ix)^k\,\frac{e^{ihx} - 1}{h}\,dF(x)\\ &\to \int_{-\infty}^{\infty} (ix)^{k+1} e^{itx} dF(x) \qquad (h \to 0),\end{aligned}$$

por el TCD, ya que el integrando converge a $(ix)^{k+1}e^{itx}$ cuando $h \to 0$, está acotado por $|x|^{k+1}$ para $|h| < 1$ y $\mathrm{E}\,|X|^{k+1} < \infty$. Además, el lado izquierdo converge a la derivada de orden $k+1$ cuando $h \to 0$, lo que demuestra la fórmula (3.31). (3.32) se obtiene poniendo $t = 0$.

Veamos la continuidad uniforme. Para $k = 1$ tenemos

$$\begin{aligned}|\varphi'(t+h) - \varphi'(t)| = \Big|\int_{-\infty}^{\infty} ixe^{itx}(e^{ihx} - 1)dF(x)\Big| &\leq \int_{-\infty}^{\infty} |x||e^{ihx} - 1|dF(x)\\ &\leq \int_{|x|\leq A} hx^2\,dF(x) + 2\int_{|x|>A} |x|\,dF(x)\\ &\leq hA^2 + 2\int_{|x|>A} |x|\,dF(x)\end{aligned}$$

Dado $\varepsilon > 0$ escogemos A de modo que la integral sea menor que $\varepsilon/2$ y luego escogemos h de modo que $hA^2 < \varepsilon/2$. Para el caso general usamos

$$|\varphi^{(k+1)}(t+h) - \varphi^{(k+1)}(t)| = \Big| \int_{-\infty}^{\infty} (ix)^{k+1} e^{itx}(e^{ihx} - 1) dF(x) \Big| \leq hA^{k+2} + 2 \int_{|x|>A} |x|^{k+1}\, dF(x),$$

y completamos la demostración como antes.

Para ver la última fórmula del teorema recordamos (3.29):

$$\Big| \varphi(t) - \sum_{k=0}^{n} \frac{(it)^k \operatorname{E}(X^k)}{k!} \Big| \leq |t|^n \operatorname{E}\Big(\frac{|t||X|^{n+1}}{(n+1)|} \wedge \frac{2|X|^n}{n!} \Big). \tag{3.34}$$

Pero

$$\text{mín}\{ \frac{|t||X|^{n+1}}{(n+1)|},\ \frac{2|X|^n}{n!} \} \to 0 \qquad (t \to 0)$$

y está acotado por $2|X|^n/n!$ que es integrable por hipótesis, así que por el TCD

$$\operatorname{E}\big[\text{mín}\{ \frac{|t||X|^{n+1}}{(n+1)|},\ \frac{2|X|^n}{n!} \} \big] \to 0 \qquad (t \to 0)$$

y en consecuencia la cota superior en (3.34) es $o(|t|^n)$ cuando $t \to 0$. ■

3.11. Unicidad y Continuidad

Teorema 3.17 (Unicidad) *La función característica de una distribución de probabilidad determina de manera única la distribución de probabilidad.*

Demostración. Sea X una v.a. con f.d. F y f.c. φ. Veamos que φ determina F. Para cualquier f.d. G con f.c. γ y cualquier real θ tenemos la siguiente versión de la relación de Parseval, que se obtiene aplicando el teorema de Fubini:

$$\begin{aligned} \int_{\mathbb{R}} e^{-i\theta y} \varphi(y)\, dG(y) &= \int_{\mathbb{R}} e^{-i\theta y} \Big[\int_{\mathbb{R}} e^{iyx}\, dF(x) \Big] dG(y) \\ &= \iint_{\mathbb{R}^2} e^{iy(x-\theta)}\, dF(x)\, dG(y) \\ &= \int_{\mathbb{R}} \Big[\int_{\mathbb{R}} e^{iy(x-\theta)}\, dG(y) \Big] dF(x) \\ &= \int_{\mathbb{R}} \gamma(x-\theta) dF(x). \end{aligned} \tag{3.35}$$

Sea ahora $N \sim \mathcal{N}(0,1)$ con densidad $\phi(x)$, de modo que σN tiene densidad normal con varianza σ^2. Reemplazamos $dG(y)$ por $\frac{1}{\sigma}\phi(y/\sigma)dy$. Luego de hacer un cambio de variables en el lado izquierdo y usando la forma de la f.c. Gaussiana γ en el lado derecho obtenemos

$$\int_{\mathbb{R}} e^{-i\theta\sigma y} \varphi(\sigma y) \phi(y)\, dy = \int_{\mathbb{R}} e^{-\sigma^2(z-\theta)^2/2} dF(z). \tag{3.36}$$

Ahora integramos ambos lados de (3.36) sobre θ de $-\infty$ a x y obtenemos

$$\int_{-\infty}^{x} \int_{\mathbb{R}} e^{-i\theta\sigma y} \varphi(\sigma y) \phi(y) dy\, d\theta = \int_{-\infty}^{x} \int_{\mathbb{R}} e^{-\sigma^2(z-\theta)^2/2} dF(z)\, d\theta,$$

y usando el teorema de Fubini para invertir el orden de integración en el lado derecho obtenemos que la expresión anterior es

$$= \int_{\mathbb{R}} \sqrt{2\pi}[\int_{-\infty}^{x} \frac{e^{-\sigma^2(z-\theta)^2/2}}{\sqrt{2\pi}} d\theta]dF(z).$$

En la integral interior hacemos el cambio de variables $s = \theta - z$ para obtener

$$\begin{aligned}\int_{-\infty}^{x}\int_{\mathbb{R}} e^{-i\theta\sigma y}\varphi(\sigma y)\phi(y)dy\, d\theta &= \frac{1}{\sigma}\int_{\mathbb{R}} \sqrt{2\pi}[\int_{-\infty}^{x-z} \frac{e^{-\sigma^2 s^2/2}}{\sqrt{2\pi}\sigma^{-1}} ds]dF(z) \\ &= \frac{\sqrt{2\pi}}{\sigma}\int_{\mathbb{R}}[\int_{-\infty}^{x-z} \phi(s; 0, \sigma^{-2})\, ds]\, dF(z) \\ &= \frac{\sqrt{2\pi}}{\sigma} P(\sigma^{-1}N + X \leq x).\end{aligned}$$

Dividimos por $\sqrt{2\pi}\sigma^{-1}$ y hacemos $\sigma \to \infty$. Dada la f.c. φ obtenemos por el Teorema de Slutsky

$$\lim_{\sigma\to\infty} \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{x}\int_{\mathbb{R}} e^{-i\theta\sigma y}\varphi(\sigma y)\phi(y)dy\, d\theta = \lim_{\sigma\to\infty} P(\sigma^{-1}N + X \leq x) = F(x) \tag{3.37}$$

para todo $x \in \mathcal{C}(F)$, de modo que φ determina el valor de F en sus puntos de continuidad y eso es suficiente. ■

Poniendo $\theta = 0$ en (3.35) obtenemos la siguiente identidad, que se conoce como la Relación de Parseval:

$$\int_{\mathbb{R}} \varphi(y)\, dG(y) = \int_{\mathbb{R}} \gamma(x)\, dF(x).$$

El siguiente corolario nos da una fórmula de inversión

Corolario 3.7 *Sea F una f.d. con f.c. φ integrable:*

$$\int_{\mathbb{R}} |\varphi(t)|\, dt < \infty.$$

Entonces F tiene una densidad f continua y acotada dada por

$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-iyx}\varphi(y)\, dy.$$

Demostración. A partir de (3.37) definimos

$$F_\sigma(x) := P(\sigma^{-1}N + X \leq x)$$

y observamos que esta f.d. tiene densidad, que llamamos f_σ, ya que $\sigma^{-1}N$ tiene densidad. A partir del lado izquierdo de (3.37) obtenemos

$$f_\sigma(\theta) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-i\theta y}\varphi(y)\phi(y/\sigma)\, dy = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-i\theta y}\varphi(y)e^{-y^2/2\sigma^2} dy.$$

Observamos que

$$\left|e^{-i\theta y}\varphi(y)e^{-y^2/2\sigma^2}\right| \leq |\varphi(y)| \in L^1$$

y cuando $\sigma \to \infty$

$$e^{-i\theta y}\varphi(y)e^{-y^2/2\sigma^2} \to e^{-i\theta y}\varphi(y).$$

Por el TCD, $f_\sigma(\theta) \to f(\theta)$. Además, para todo intervalo finito I

$$\sup_{\theta \in I} f_\sigma(\theta) \leq \frac{1}{2\pi} \int_{\mathbb{R}} |\varphi(y)| e^{-y^2/2\sigma^2} \, dy \leq \frac{1}{2\pi} \int_{\mathbb{R}} |\varphi(y)| \, dy < \infty.$$

Así, cuando $\sigma \to \infty$, tenemos por los teoremas de Slutsky y convergencia acotada

$$F(I) = \lim_{\sigma \to \infty} P(\sigma^{-1} N + X \in I) = \lim_{\sigma \to \infty} \int_I f_\sigma(\theta) \, d\theta = \int_I f(\theta) \, d\theta.$$

Por lo tanto f es la densidad de F. ∎

Antes de demostrar el teorema de continuidad presentamos un lema preliminar.

Lema 3.5 *Si F es una f.d. con f.c. φ, entonces existe $\alpha \in (0, \infty)$ tal que para todo $x > 0$*

$$F([-x, x]^c) \leq \alpha x \int_0^{1/x} (1 - \operatorname{Re} \varphi(t)) \, dt.$$

Demostración. Como

$$\operatorname{Re} \varphi(t) = \int_{-\infty}^{\infty} \cos ty \, dF(y),$$

tenemos, usando Fubini

$$\begin{aligned} x \int_0^{1/x} (1 - \operatorname{Re} \varphi(t)) \, dt &= x \int_0^{1/x} \int_{-\infty}^{\infty} (1 - \cos ty) \, dF(y) \, dt \\ &= x \int_{-\infty}^{\infty} \Big[\int_0^{1/x} (1 - \cos ty) \, dt \Big] dF(y) \\ &= x \int_{-\infty}^{\infty} \Big(\frac{1}{x} - \frac{\operatorname{sen}(y/x)}{y} \Big) \, dF(y) \\ &= \int_{-\infty}^{\infty} \Big(1 - \frac{\operatorname{sen}(y/x)}{y/x} \Big) \, dF(y). \end{aligned}$$

Como el integrando es positivo, esto es mayor que

$$\int_{|y| > x} \Big(1 - \frac{\operatorname{sen}(y/x)}{y/x} \Big) \, dF(y) \geq \alpha^{-1} F([-x, x]^c),$$

donde

$$\alpha^{-1} = \inf_{|y/x| \geq 1} \Big(1 - \frac{\operatorname{sen}(y/x)}{y/x} \Big).$$

∎

Teorema 3.18 (Continuidad, Lévy) *(i) Sea $\{X_n, \ n \geq 1\}$ una sucesión de v.a. tal que X_n tiene f.d. F_n y f.c. φ_n. Si $X_n \xrightarrow{d} X_0$ entonces*

$$\varphi_n(t) \to \varphi_0(t), \qquad \forall t \in \mathbb{R}.$$

(ii) Si

a) $\lim_{n \to \infty} \varphi_n(t)$ existe para todo t. Llamemos $\varphi_0(t)$ al límite.

b) $\varphi_0(t)$ es continua en 0.

Entonces para alguna f.d. F_0
$$F_n \xrightarrow{d} F_0$$
y φ_0 *es la f.c. de* F_0*. Si* $\varphi_0(0) = 1$ *entonces* F_0 *es propia.*

Demostración. (i) Si $X_n \xrightarrow{d} X_0$, como consecuencia del teorema de Skorohod tenemos que $e^{itX_n} \xrightarrow{d} e^{itX_0}$ y como $|e^{itX_n}| \leq 1$, tenemos por convergencia dominada que
$$\varphi_n(t) = \mathrm{E}\, e^{itX_n} \to Ee^{itX_0} = \varphi_0(t).$$
(ii) Supongamos ahora que para todo $t \in \mathbb{R}$, se tiene que $\varphi_n(t) \to \varphi_0(t)$, con φ_0 continua en 0, entonces demostramos primero que $\{F_n\}$ es tensa. Sea $M > 0$ y usemos el lema 3.5 tenemos
$$\limsup_{n\to\infty} F_n([-M, M]^c) \leq \limsup_{n\to\infty} \alpha M \int_0^{1/M} (1 - \mathrm{Re}\, \varphi_n(t))\, dt.$$
Pero $\varphi_n(t) \to \varphi_0(t)$ implica que $\mathrm{Re}\, \varphi_n(t) \to \mathrm{Re}\, \varphi_0(t)$ y $1 - \mathrm{Re}\, \varphi_n(t) \to 1 - \mathrm{Re}\, \varphi_0(t)$. Como $1 - \varphi_n$ está acotada, también lo está $\mathrm{Re}(1 - \varphi_n) = 1 - \mathrm{Re}\, \varphi_n$. Por el TCD
$$\limsup_{n\to\infty} F_n([-M, M]^c) \leq \alpha M \int_0^{1/M} (1 - \mathrm{Re}\, \varphi_0(t))\, dt.$$
Como φ_0 es continua en 0,
$$\lim_{t\to 0} \varphi_0(t) = \varphi_0(0) = \lim_{n\to\infty} \varphi_n(0) = 1.$$
En consecuencia, $1 - \mathrm{Re}\, \varphi_0(t) \to 0$ cuando $t \to 0$ y por lo tanto, dados $\varepsilon > 0$ y M suficientemente grande tenemos
$$\alpha M \int_0^{1/M} (1 - \mathrm{Re}\, \varphi_0(t))\, dt \leq \alpha M \int_0^{1/M} \varepsilon\, dt = \alpha\varepsilon.$$
Por lo tanto $\{F_n\}$ es tensa y en consecuencia cualquier par de subsucesiones convergentes de $\{F_n\}$ tienen que converger al mismo límite, porque si
$$F_{n'} \xrightarrow{d} F, \quad \text{y} \quad F_{n''} \xrightarrow{d} G,$$
entonces F y G son propias. Por la parte (i) del teorema que ya hemos probado
$$\varphi_{n'} \to \varphi_F = \varphi_0, \qquad \text{y} \qquad \varphi_{n''} \to \varphi_G = \varphi_0,$$
y por lo tanto $\varphi_F = \varphi_G$. Por el teorema de unicidad $F = G$. Por lo tanto cualquier par de subsucesiones convergentes convergen al mismo límite y por lo tanto $\{F_n\}$ converge a un límite cuya f.c. es φ_0. ■

Ejemplo 3.8 (Aproximación Poisson a la Binomial)
Sea S_n una v.a. con distribución binomial de parámetros n y p:
$$P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, \dots, n.$$
Si $p = p(n) \to 0$ cuando $n \to \infty$ de modo que $np \to \lambda > 0$ entonces
$$S_n \xrightarrow{d} Po(\lambda).$$
Para verificar este resultado calculamos inicialmente la f.c. de $Y \sim Po(\lambda)$. Tenemos
$$\mathrm{E}\, e^{itY} = \sum_{k=0}^{\infty} e^{itk} \frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^{it})^k}{k!} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it}-1)}.$$

Recordemos que podemos representar una variable binomial como suma de variables iid Bernoulli $\xi_1, \dots, \xi_n$ con $P(\xi_i = 1) = p = 1 - P(\xi_i = 0)$. Por lo tanto,

$$\begin{aligned} \mathrm{E}\, e^{itS_n} &= (\mathrm{E}\, e^{it\xi_1})^n = (1 - p + e^{it}p)^n \\ &= (1 + p(e^{it} - 1))^n = \big(1 + \frac{np(e^{it} - 1)}{n}\big)^n \\ &\to e^{\lambda(e^{it}-1)}. \end{aligned}$$

Ejemplo 3.9 (Aproximación Normal a la Binomial)
Sea S_n una v.a. con distribución binomial de parámetros n y $p(n) = p_n$. Supongamos que $s_n^2 = \mathrm{Var}(S_n) = np_n(1 - p_n) \to \infty$ cuando $n \to \infty$, entonces

$$T_n = \frac{S_n - np_n}{s_n} \xrightarrow{d} \mathcal{N}(0, 1).$$

Para demostrar este resultado comenzamos por la f.c. φ_n de T_n. Observamos que T_n se puede escribir como la suma de variables $(X_k - p_n)/s_n$ con $X_k \sim Be(p_n)$, $1 \le k \le n$. Por lo tanto, la f.c. está dada por

$$\begin{aligned} \varphi_n(t) &= \big(p_n e^{it(1-p_n)/s_n} + (1 - p_n)e^{itp_n/s_n}\big)^n \\ &= \Big(1 + \frac{u_n(t)}{n}\Big)^n \end{aligned}$$

donde

$$u_n(t) = n\big(p_n e^{it(1-p_n)/s_n} + (1 - p_n)e^{itp_n/s_n} - 1\big)$$

Basta verificar que para todo $t \in \mathbb{R}$ se tiene que $u_n(t) \to -t^2/2$ cuando $n \to \infty$. Llamemos $a_n(t) = tp_n/s_n$ y $b_n(t) = t(1 - p_n)/s_n$ y recordemos que para cualquier $x \in \mathbb{R}$,

$$\Big|e^{ix} - 1 - ix - \frac{(ix)^2}{2}\Big| \le \frac{|x|^3}{3!}.$$

Como $s_n \to \infty$,

$$\begin{aligned} u_n(t) &= n\big(p_n e^{ib_n(t)} + (1 - p_n)e^{ia_n(t)} - 1\big) \\ &= n\big(p_n(e^{ib_n(t)} - 1 - ib_n(t)) + (1 - p_n)(e^{ia_n(t)} - 1 - ia_n(t))\big) \\ &= n\Big(\frac{1}{2}p_n(ib_n(t))^2 + \frac{1}{2}(1 - p_n)(ia_n(t))^2 + \mathcal{O}\big(\frac{p_n(1 - p_n)|t|^3}{s_n^3}\big)\Big) \\ &= \frac{-t^2}{2} + o(1) \end{aligned}$$

cuando $n \to \infty$.