Capítulo 6

Funciones características

Una[1] función característica es una función que toma valores complejos y tiene argumento real. Definida a partir de una variable aleatoria X, *caracteriza* la distribución $F(x)$ de esta variable aleatoria. Las funciones características son especialemente adecuadas para el estudio de la convergencia débil de variables aleatorias independientes, y serán utilizadas a lo largo del capítulo 7. Dadas dos variables aleatorias X e Y, definimos la *variable aletoria compleja* $Z = X + iY$, donde $i = \sqrt{-1}$ es la unidad imaginaria. Si la variables aleatorias X, Y tienen esperanzas respectivas $\mathbf{E}\, X, \mathbf{E}\, Y$, definimos la esperanza matemática de la variable aleatoria compleja Z mediante $\mathbf{E}\, Z = \mathbf{E}\, X + i\, \mathbf{E}\, Y$. No es difícil verificar (tomando partes real e imaginaria), que si a, b son dos números complejos, se tiene $\mathbf{E}(aZ + b) = a\, \mathbf{E}\, Z + b$; y que si Z_1, Z_2 son dos variables aleatorias complejas, se tiene $\mathbf{E}(Z_1 + Z_2) = \mathbf{E}\, Z_1 + \mathbf{E}\, Z_2$. Si $z = a + ib$ es un número complejo, designamos $\overline{z} = a - ib$ el *complejo conjugado* de z.

6.1. Definiciones y primeras propiedades

Consideremos una variable aleatoria X definida en un espacio de probabilidad $(\Omega, \mathcal{A}, \mathbf{P})$. Llamamos *función característica* de la variable aleatoria X a la función $f(t)$, definida para todo t real, mediante la igualdad

$$f(t) = \mathbf{E}\, e^{itX}. \tag{6.1}$$

[1]Capítulo 6 del libro "Teoría de la Probabilidad" por Valentín Petrov y Ernesto Mordecki

La fórmula (6.1) es equivalente a

$$f(t) = \mathbf{E} \cos tX + i\, \mathbf{E} \operatorname{sen} tX. \tag{6.2}$$

Como las variables aleatorias $\cos tX$ y $\operatorname{sen} tX$ están acotadas para todo t real, sus esperanzas matemáticas existen. Por ésto, la función característica de una variable aleatoria arbitraria X está correctamente definida para todo t real.

De acuerdo a la definición de esperanza matemática, tenemos

$$f(t) = \int_\Omega e^{itX} d\, \mathbf{P}\,. \tag{6.3}$$

Si la variable aleatoria X tiene función de distribución $F(x)$, aplicando la identidad (**??**) obtenemos

$$f(t) = \int_{-\infty}^{\infty} e^{itx} dF(x). \tag{6.4}$$

Consideremos ahora los dos tipos más importantes de distribuciones. Si la variable aleatoria X tiene distribución discreta, y toma los valores $x_1, x_2, \dots$ con probabilidades $p_1, p_2, \dots$ respectivamente, entonces

$$f(t) = \sum_k e^{itx_k} p_k, \tag{6.5}$$

como se obtiene de aplicar cualquiera de las fórmulas (6.1), (6.2), ó (6.3).

Si X tiene distribución absolutamente continua, con densidad dada por $p(x)$, entonces

$$f(t) = \int_{-\infty}^{\infty} e^{itx} p(x) dx, \tag{6.6}$$

como se obtiene de aplicar (6.4).

Calculemos las funciones características de variables aleatorias con distribuciones de ambos tipos, en los casos mas importantes.

Ejemplo 6.1. Supongamos que la variable aleatoria X tiene distribución degenerada, es decir, existe una constante c tal que $\mathbf{P}(X = c) = 1$. Aplicando la fórmula (6.5), tenemos

$$f(t) = \mathbf{E}\, e^{itX} = e^{itc}.$$

En particular, si $\mathbf{P}(X = 0) = 1$, tenemos $f(t) = 1$ para todo t real.

Ejemplo 6.2. Sea X una variable aleatoria con distribución binomial con parámetros (n, p). Aplicando (6.5), obtenemos

$$f(t) = \sum_{m=0}^{n} e^{itm} \binom{n}{m} p^m q^{n-m} = \sum_{m=0}^{n} \binom{n}{m} (pe^{it})^m q^{n-m} = (pe^{it} + q)^n,$$

donde $q = 1 - p$.

Ejemplo 6.3. Si X tiene distribución de Poisson con parámetro $\lambda > 0$, aplicando la fórmula (6.5), obtenemos

$$f(t) = \mathbf{E}\, e^{itX} = \sum_{m=0}^{\infty} e^{itm} \frac{\lambda^m}{m!} e^{-\lambda} = e^{-\lambda} \sum_{m=0}^{\infty} \frac{(e^{it}\lambda)^m}{m!} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it}-1)}.$$

Ejemplo 6.4. Consideremos una variable aleatoria X con distribución normal estándar, con densidad dada por $p(x) = e^{-x^2/2}/\sqrt{2\pi}$. Aplicando la fórmula (6.6), tenemos

$$f(t) = \mathbf{E}\, e^{itX} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx - x^2/2} dx.$$

Para calcular la integral anterior derivamos con respecto de t, y obtenemos

$$f'(t) = \frac{i}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{itx - x^2/2} dx = \frac{i}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} d\Big(- e^{-x^2/2} \Big).$$

Luego de integrar por partes, resulta

$$f'(t) = \frac{-1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t e^{itx - x^2/2} dx = -t f(t).$$

En consecuencia, $(\ln f(t))' = -t$, $\ln f(t) = -t^2/2 + C$. Como $f(0) = 1$, obtenemos que $C = 0$, y en conclusión

$$f(t) = e^{-t^2/2}.$$

Si X es una variable aleatoria con distribución normal con parámetros (a, σ), entonces, como veremos en el ejemplo 6.6, su función característica está dada por

$$f(t) = e^{iat - \sigma^2 t^2/2}.$$

Ejemplo 6.5. Sea X una variable aleatoria con distribución exponencial con parámetro $\alpha > 0$. La densidad de esta variable aleatoria está dada por $p(x) = \alpha e^{-\alpha x}$ si $x \geq 0$, $p(x) = 0$ si $x < 0$. Por ésto, aplicando la fórmula (6.6), tenemos

$$f(t) = \mathbf{E}\, e^{itX} = \alpha \int_0^\infty e^{(it-\alpha)x} dx = \frac{\alpha}{\alpha - it}.$$

Estudiemos ahora las propiedades que verifica la función característica $f(t)$ de una variable aleatoria X con función de distribución $F(x)$. Las dos primeras propiedades son evidentes.

Propiedad 1. *Se tiene* $f(0) = 1$.

Propiedad 2. *Se tiene* $|f(t)| \leq 1$ *para cualquier* $t \in \mathbb{R}$.

Propiedad 3. *La función característica* $f(t)$ *es uniformemente continua en la recta real.*

Demostración. La demostración, basada en la fórmula (6.3), se adapta sin dificultad si tomamos como definición cualquiera de las fórmulas (6.4), (6.5) ó (6.6).

Sean t, h reales arbitrarios. De (6.3), tenemos

$$f(t+h) - f(t) = \int_\Omega e^{itX}(e^{ihX} - 1) d\,\mathbf{P}\,. \tag{6.7}$$

Utilizamos la acotación $|e^{iu} - 1| \leq |u|$, válida para todo u real, porque

$$|e^{iu} - 1| = \Big| \int_0^u e^{ix} dx \Big| \leq \int_0^{|u|} |e^{ix}| dx = |u|.$$

Sea $\varepsilon > 0$, arbitrario. Existe un real A que verifica: A y $-A$ son puntos de continuidad de $F(x)$; $1 - F(A) < \varepsilon/8$; $F(-A) < \varepsilon/8$. Tomando valor absoluto a ambos lados en (6.7), y designando $\mathbf{B} = \{\omega \in \Omega \colon |X(\omega)| \geq A\}$, obtenemos

$$\begin{aligned}
|f(t+h) - f(t)| &\leq \int_\Omega |e^{ihX} - 1| d\,\mathbf{P} \leq 2 \int_{\mathbf{B}} d\,\mathbf{P} + \int_{\bar{\mathbf{B}}} |e^{ihX} - 1| d\,\mathbf{P}, \\
&\leq 2\,\mathbf{P}(\mathbf{B}) + \int_{\bar{\mathbf{B}}} |hX| d\,\mathbf{P} \leq 2\,\mathbf{P}(|X| \geq A) + A|h| \\
&\leq 2\big(1 - F(A) + F(-A)\big) + Ah \leq \frac{\varepsilon}{2} + Ah < \varepsilon,
\end{aligned}$$

si tomamos $h < \varepsilon/(2A)$. Como la acotación es independiente de t, esto concluye la demostración. □

Propiedad 4. *Consideremos la variable aleatoria $Y = aX + b$, donde a, b son constantes. Entonces, la función característica $g(t)$ de la variable aleatoria Y verifica $g(t) = e^{ibt}f(at)$, donde $f(t)$ es la función característica de la variable aleatoria X.*

Demostración. Utilizando la propiedad $e^{i(\alpha+\beta)} = e^{i\alpha}e^{i\beta}$ y aplicando la definición (6.2), tenemos

$$g(t) = \mathbf{E}\, e^{itY} = \mathbf{E}\, e^{it(aX+b)} = \mathbf{E}\big(e^{ibt}e^{iatX}\big) = e^{ibt}\,\mathbf{E}\, e^{iatX} = e^{ibt}f(at),$$

lo que concluye la demostración. □

Ejemplo 6.6. Calculemos la función característica $g(t)$ de una variable aleatoria Y con función de distribución normal con parámetros (a, σ).

Es claro que la variable aleatoria $X = (Y - a)/\sigma$ tiene distribución normal con parámetros $(0, 1)$, y por lo tanto función característica $f(t) = e^{-t^2/2}$, como vimos en el ejemplo 6.4. Aplicando la propiedad anterior, obtenemos

$$g(t) = e^{iat}f(\sigma t) = e^{iat-\sigma^2t^2/2}.$$

Propiedad 5. *Consideremos una variable aleatoria X para la cual existe $\alpha_k = \mathbf{E}(X^k)$, el momento de orden k, para algun natural $k \geq 1$. Entonces, su función característica $f(t)$ tiene derivadas continuas, para todo t real, hasta el orden k inclusive. Además*

$$f^{(m)}(0) = i^m\alpha_m \quad (1 \leq m \leq k), \tag{6.8}$$

donde $\alpha_m = \mathbf{E}(X^m)$.

Demostración. Derivando formalmente m veces, con respecto de t, bajo el signo de integración en la fórmula (6.4), obtenemos

$$f^{(m)}(t) = \int_{-\infty}^{\infty} (ix)^m e^{itx} dF(x). \tag{6.9}$$

Es claro que como existe el momento de orden $m \leq k$ (proposición **??**), tenemos

$$\Big|\int_{-\infty}^{\infty} (ix)^m e^{itx} dF(x)\Big| \leq \int_{-\infty}^{\infty} |x|^m dF(x) < \infty.$$

Esto permite demostrar que la derivación bajo el signo de integral es válida, y obtener la fórmula (6.9). Sustituyendo $t = 0$ en (6.9) se obtiene (6.8).
□

Dada una variable aleatoria X, si para algún natural $k \geq 1$ existe $\alpha_k = \mathbf{E}(X^k)$, el momento de orden k de la variable aleatoria, aplicando la propiedad 5, el desarrollo de Taylor, y la igualdad $f(0) = 1$, se obtiene, que

$$f(t) = 1 + \sum_{m=1}^{k} \frac{\alpha_m}{m!}(it)^m + o(|t|^k) \ (t \to 0). \tag{6.10}$$

En la demostración de la próxima propiedad, utilizamos el resultado siguiente.

Lema 6.1. *Consideremos dos variables aleatorias independientes X, Y, dos funciones reales $u(x), v(x)$, definidas en la recta real, y supongamos que existen $\mathbf{E}\,u(X)$ y $\mathbf{E}\,v(Y)$. Entonces, tiene lugar la identidad*

$$\mathbf{E}\big(u(X)v(Y)\big) = \mathbf{E}\,u(X)\,\mathbf{E}\,v(Y).$$

Demostración. Veremos la demostración en el caso en que ambas variables tienen distribución discreta, y en el caso en que tienen distribución absolutamente continua.

Supongamos primero que X toma los valores $x_1, x_2, \dots$, con probabilidades $p_1, p_2, \dots$, respectivamente; Y toma los valores $y_1, y_2, \dots$, con probabilidades $q_1, q_2, \dots$, respectivamente. Aplicando la proposición **??**, obtenemos que $\mathbf{P}(X = x_k, Y = y_j) = \mathbf{P}(X = x_k)\,\mathbf{P}(Y = y_j) = p_k q_j$ $(k, j = 1, 2, \dots)$. Por ésto, tenemos

$$\begin{aligned}\mathbf{E}\big(u(X)v(Y)\big) &= \sum_{k,j} u(x_k)v(y_j)\,\mathbf{P}(X = x_k, Y = y_j)\\ &= \sum_{k} u(x_k)p_k \sum_{j} v(y_j)p_j = \mathbf{E}\,u(X)\,\mathbf{E}\,v(Y).\end{aligned}$$

Si X e Y tienen distribución absolutamente continua y $r(x, y)$ designa la densidad del vector (X, Y), aplicando la proposición **??**, obtenemos que $r(x, y) = p(x)q(y)$, donde $p(x)$ es la densidad de la variable aleatoria X, $q(y)$ la densidad de la variable aleatoria Y. Por ésto, tenemos

$$\begin{aligned}\mathbf{E}\big(u(X)v(Y)\big) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(x)v(y)r(x, y)dxdy\\ &= \int_{-\infty}^{\infty} u(x)p(x)dx \int_{-\infty}^{\infty} v(y)q(y)dy = \mathbf{E}\,u(X)\,\mathbf{E}\,v(Y),\end{aligned}$$

concluyendo la demostración. $\square$

Observación. El lema recién formulado es válido también en el caso en que las funciones de $u(x)$ y $v(x)$ de argumento real, tomen valores complejos, es decir, si tenemos

$$u(x) = u_1(x) + iu_2(x), \qquad v(x) = v_1(x) + iv_2(x),$$

donde $u_k(x)$ y $v_k(x)$ son funciones de argumento real, que toman valores reales ($k = 1, 2$). Esto es sencillo de demostrar, aplicando el lema anterior a las partes real e imaginaria del producto $u(X)v(Y)$.

Propiedad 6. *Consideremos dos variables aleatorias independientes X e Y, con funciones características $f(t)$ y $g(t)$ respectivamente. Sea $h(t)$ la función característica de la suma $X + Y$. Entonces, se verifica $h(t) = f(t)g(t)$.*

Demostración. Tenemos

$$h(t) = \mathbf{E}\, e^{it(X+Y)} = \mathbf{E}\left(e^{itX}e^{itY}\right) = \mathbf{E}\, e^{itX}\,\mathbf{E}\, e^{itY} = f(t)g(t),$$

en vista de la observación posterior al lema 6.1. □

Es válida la siguiente generalización de la propiedad recién demostrada: si $X_1, X_2, \dots, X_n$ es un conjunto de variables aleatorias mutuamente independientes, con funciones características $f_1(t), f_2(t), \dots, f_n(t)$ respectivamente, entonces, la función característica $h(t)$ de la suma $X_1+X_2+\cdots+X_n$ es igual al producto de las funciones características de los sumandos:

$$h(t) = f_1(t)f_2(t)\cdots f_n(t).$$

Propiedad 7. *Para todo t real, se verifica $f(-t) = \overline{f(t)}$.*

La propiedad anterior se obtiene de la igualdad

$$f(-t) = \mathbf{E}\, e^{-itX} = \mathbf{E}\,\overline{e^{itX}} = \overline{\mathbf{E}\, e^{itX}} = \overline{f(t)}.$$

Definición 6.1. *Una variable aleatoria X y su distribución $F(x)$ se dicen* simétricas *cuando las funciones de distribución de las variables aleatorias X y $-X$ son idénticas.*

Propiedad 8. *Si la variable aleatoria X es simétrica, su función característica $f(t)$ es una función real.*

A la conclusión de la propiedad anterior conducen las igualdades

$$f(t) = \mathbf{E}\, e^{itX} = \mathbf{E}\, e^{it(-X)} = f(-t) = \overline{f(t)}.$$

En la sección 2 demostraremos el recíproco de la propiedad anterior: si la función característica de una variable aleatoria X es real, la variable aleatoria X es simétrica.

6.2. Fórmula de inversión. Teorema de unicidad

Teorema 6.1. *Consideremos una variable aleatoria X con función de distribución $F(x)$, y función característica $f(t)$. Sean x_1, x_2 dos puntos de continuidad de $F(x)$. Entonces, tiene lugar la igualdad*

$$F(x_2) - F(x_1) = \frac{1}{2\pi} \lim_{T\to\infty} \int_{-T}^{T} \frac{e^{-itx_2} - e^{-itx_1}}{-it} f(t)dt. \tag{6.11}$$

La igualdad (6.11) *se denomina* fórmula de inversión*.*

Demostración. Comenzamos introduciendo la función auxiliar

$$R(h,T) = \int_0^T \frac{\operatorname{sen} ht}{t} dt = \int_0^{hT} \frac{\operatorname{sen} u}{u} du.$$

Del cálculo integral, son conocidas los siguientes afirmaciones:

$$\int_0^\infty \frac{\operatorname{sen} u}{u} du = \frac{\pi}{2}, \qquad \Big| \int_0^x \frac{\operatorname{sen} u}{u} du \Big| \leq C \ (x \geq 0),$$

de donde obtenemos, que

$$\lim_{T\to\infty} R(h,T) = \begin{cases} -\pi/2, & \text{si } h < 0, \\ \pi/2, & \text{si } h > 0. \end{cases} \tag{6.12}$$

Es importante observar, que esta convergencia es uniforme en los intervalos de la forma $(-\infty, \delta]$, y $[\delta, \infty)$, para todo $\delta > 0$.

La segunda etapa de la demostración, consiste en representar a la integral en (6.11), que designamos I, en términos de la función $R(h,T)$.

Aplicando la definición (6.4), e intercambiando el orden de integración (dado que el integrando es una función continua y acotada), tenemos

$$
\begin{aligned}
I &= \int_{-T}^{T} \frac{e^{-itx_2} - e^{-itx_1}}{-it} f(t)dt = \int_{-T}^{T} \frac{e^{-itx_2} - e^{-itx_1}}{-it} \Big(\int_{-\infty}^{\infty} e^{ity} dF(y) \Big) dt \\
&= \int_{-\infty}^{\infty} \Big(\int_{-T}^{T} \frac{e^{it(y-x_2)} - e^{it(y-x_1)}}{-it} dt \Big) dF(y) \\
&= \int_{-\infty}^{\infty} \Big(\int_{-T}^{T} \frac{\operatorname{sen}\big(t(y-x_2)\big) - \operatorname{sen}\big(t(y-x_1)\big)}{t} dt \Big) dF(y) \\
&= 2 \int_{-\infty}^{\infty} \big(R(y-x_2,T) - R(y-x_1,T) \big) dF(y).
\end{aligned}
$$

donde utilizamos que $\int_{-T}^{T} \big(\cos(\alpha t)/t\big) dt = 0$ para todo α real, para obtener la última igualdad. Respecto del comportamiento asintótico del último integrando, tomando por ejemplo $x_1 < x_2$, en vista de (6.12), tenemos

$$
\lim_{T\to\infty} R(y-x_2,T) - R(y-x_1,T) = \begin{cases} 0, & \text{si } y < x_1, \\ \pi, & \text{si } x_1 < y < x_2, \\ 0, & \text{si } x_2 < y. \end{cases} \tag{6.13}
$$

La última etapa de la demostración consiste en verificar que el límite de I cuando $T \to \infty$ es la integral del límite obtenido en (6.13). Para ésto elegimos $\delta > 0$, de forma que $x_1 + \delta < x_2 - \delta$, y consideramos la integral I como la suma de cinco integrales, designadas I_k $(i = 1, \dots, 5)$, en los intervalos de integración $(-\infty, x_1 - \delta]$, $(x_1 - \delta, x_1 + \delta]$, $(x_1 + \delta, x_2 - \delta]$, $(x_2 - \delta, x_2 + \delta]$, y $(x_2 + \delta, \infty)$. Tenemos entonces $I = \sum_{i=1}^{5} I_k$.

Sea $\varepsilon > 0$ arbitrario. Tenemos

$$
I_1 = 2 \int_{(-\infty, x_1-\delta]} \big(R(y-x_2,T) - R(y-x_1,T) \big) dF(y).
$$

Como, en vista de (6.13), el integrando converge uniformemente a cero, obtenemos que $|I_1| < \varepsilon$ si T es suficientemente grande. Una situación análoga ocurre con T_5, por lo que, $|I_5| < \varepsilon$ si T es suficientemente grande.

Para la segunda integral, tenemos

$$
\begin{aligned}
|I_2| &\le 2 \int_{(x_1-\delta, x_1+\delta]} \big(R(y-x_2,T) - R(y-x_1,T) \big) dF(y) \\
&\le 4C\big(F(x_1+\delta) - F(x_1-\delta) \big) < \varepsilon,
\end{aligned}
$$

si δ es suficientemente pequeño (independientemente de T), dado que x_1 es un punto de continuidad de $F(x)$. La situación con I_4 es análoga, y por ésto $|I_4| < \varepsilon$ si δ es suficientemente pequeño. Por último, como la convergencia en (6.13) es uniforme, para la tercer integral tenemos

$$\begin{aligned} I_3 &= 2\int_{(x_1+\delta,x_2-\delta]} \big(R(y-x_2,T) - R(y-x_1,T)\big)dF(y) \\ &\to 2\pi\big(F(x_2-\delta) - F(x_1+\delta)\big), \end{aligned}$$

si $T \to \infty$. En conclusión, para todo δ suficientemente pequeño, tenemos

$$\begin{aligned} \Big|I - 2\pi\big(F(x_2-\delta) - F(x_1+\delta)\big)\Big| &\le |I_1| + |I_2| + |I_4| + |I_5| \\ &\quad + \Big|I - 2\pi\big(F(x_2-\delta) - F(x_1+\delta)\big)\Big| < 5\varepsilon, \end{aligned}$$

si T es suficientemente grande. Como x_1 y x_2 son puntos de continuidad de la función $F(x)$, esto concluye la demostración. □

Teorema 6.2 (Unicidad)**.** *Sean $f(t), g(t)$ las funciones características correspondientes a dos funciones de distribución $F(x), G(x)$. Supongamos que $f(t) = g(t)$ para todo t real. Entonces, se verifica $F(x) = G(x)$ para todo x real.*

Demostración. Sea $\mathbf{C}$ el conjunto de los puntos en el que ambas funciones $F(x)$ y $G(x)$ son continuas. Como $F(x)$ y $G(x)$ son funciones de distribución, el complemento del conjunto $\mathbf{C}$ es, a lo sumo, numerable. Sean entonces $x, y_1, y_2, \dots$ puntos de $\mathbf{C}$, tales que $y_n \to -\infty$ $(n \to \infty)$. Aplicando el teorema 6.1, obtenemos que $F(x) - F(y_n) = G(x) - G(y_n)$, y tomando límite si $n \to \infty$ en la igualdad anterior, resulta

$$F(x) = G(x) \text{ para todo } x \text{ en } \mathbf{C}. \tag{6.14}$$

Sea ahora z un real arbitrario. Consideremos $x_1 > x_2 > \cdots$ puntos de $\mathbf{C}$, tales que $x_n \to z$ $(n \to \infty)$. En vista de (6.14), tenemos $F(x_n) = G(x_n)$. Como ambas funciones de distribución son continuas por la derecha, al tomar límite si $n \to \infty$ en la igualdad anterior, obtenemos la igualdad $F(z) = G(z)$. Como z es arbitrario, esto concluye la demostración. □

De acuerdo al teorema 6.2, denominado *teorema de unicidad*, la función característica de una variable aleatoria define unívocamente (es decir "caracteriza") su función de distribución. Veamos algunas aplicaciones de este teorema.

Ejemplo 6.7. Consideremos dos variables aleatorias independientes X_1 y X_2, cada una de las cuales tiene distribución normal con parámetros (a_1, σ_1) y (a_2, σ_2) respectivamente. Con la ayuda del teorema de unicidad es sencillo demostrar que la suma X_1+X_2 tiene distribución normal (como vimos en la sección 3.4).

La función característica de X_k es $f_k(t) = e^{ia_k t - \sigma_k^2 t^2/2}$ $(k = 1, 2)$. Como las variables aleatorias son independientes, la función característica de la suma $X_1 + X_2$ es

$$\begin{aligned} g(t) &= \mathbf{E}\, e^{it(X_1+X_2)} = \mathbf{E}\, e^{itX_1}\, \mathbf{E}\, e^{itX_2} \\ &= f_1(t) f_2(t) = e^{i(a_1+a_2)t - (\sigma_1^2+\sigma_2^2)t^2/2}. \end{aligned}$$

Es claro que la función $g(t)$ coincide con la función característica de una variable aleatoria con distribución normal, con parámetros $\big(a_1 + a_2, (\sigma_1^2 + \sigma_2^2)^{1/2}\big)$. Aplicando el teorema 6.2 se deduce que la suma $X_1 + X_2$ es una variable aleatoria con distribución normal con éstos parámetros.

Ejemplo 6.8. Consideremos dos variables aleatorias independientes X e Y, cada una de las cuales tiene distribución exponencial con parámetros $\alpha > 0$ y $\beta > 0$ respectivamente. Veamos que la variable aleatoria $Z = X - Y$ tiene densidad, dada por

$$p(x) = \begin{cases} \frac{\alpha\beta}{\alpha+\beta} e^{-\alpha x} & \text{si } x > 0, \\ \frac{\alpha\beta}{\alpha+\beta} e^{\beta x} & \text{si } x \leq 0. \end{cases}$$

Por una parte, utilizando el ejemplo 6.5, tenemos

$$\mathbf{E}\, e^{itZ} = \mathbf{E}\, e^{itX}\, \mathbf{E}\, e^{-itY} = \frac{\alpha}{\alpha - it} \frac{\beta}{\beta + it}. \tag{6.15}$$

Por otra parte,

$$\begin{aligned} \int_{-\infty}^{\infty} e^{itx} p(x) dx &= \frac{\alpha\beta}{\alpha+\beta} \Big(\int_0^\infty e^{(it-\alpha)x} dx + \int_{-\infty}^0 e^{(it+\beta)x} dx \Big) \\ &= \frac{\alpha\beta}{\alpha+\beta} \Big(\frac{1}{\alpha - it} + \frac{1}{\beta + it} \Big) = \frac{\alpha}{\alpha - it} \frac{\beta}{\beta + it}. \end{aligned} \tag{6.16}$$

Como los resultados en (6.15) y (6.16) coinciden, del teorema de unicidad se deduce que $p(x)$ es la densidad de Z.

Veamos un corolario más del teorema de unicidad: si la función característica $f(t)$ de una variable aleatoria dada es real, entonces la variable aleatoria es simétrica. En efecto, como $f(t)$ es real, aplicando la propiedad 7, tenemos

$$f(t) = \overline{f(t)} = f(-t).$$

Por otra parte, $f(-t) = \mathbf{E}\, e^{-itX}$ es la función característica de la variable aleatoria $-X$ en el punto t. De la coincidencia de las funciones características se obtiene la igualdad de las distribuciones de las variables aleatorias X y $-X$, aplicando el teorema 6.2. Concluímos que X es simétrica.

En vista de lo recién demostrado y de la propiedad 8, llegamos a la siguiente conclusión: una variable aleatoria es simétrica si y solo si su función característica es real.

6.3. Teoremas de Helly

Los teoremas de Helly juegan un importante rol en la demostración de los teoremas de convergencia de funciones características, que estudiamos en la sección 4. Dado que no se incluyen en los cursos habituales de cálculo, los presentamos aquí con sus correspondientes demostraciones.

Definición 6.2. *Consideremos funciones $F(x), F_1(x), F_2(x), \dots,$ acotadas y no decrecientes.*
(a) *Decimos que la sucesión $\{F_n(x)\}$* converge débilmente *a $F(x)$ y escribimos $F_n \to F$, si $F_n(x) \to F(x)$ $(n \to \infty)$ para todo punto x de continuidad de la función $F(x)$.*
(b) *Decimos que la sucesión $\{F_n(x)\}$* converge completamente *a $F(x)$ y escribimos $F_n \Rightarrow F$, si $\{F_n(x)\}$ converge débilmente a $F(x)$ (es decir, si $F_n \to F$) y además*[2] *$F_n(-\infty) \to F(-\infty)$, $F_n(\infty) \to F(\infty)$ si $n \to \infty$.*

Observemos que si $F(x), F_1(x), F_2(x), \dots$ son funciones de distribución, la convergencia débil definida en (a) coincide con la convergencia débil de variables aleatorias definida en la sección 5.1. Además, en este caso, las definiciones (a) y (b) coinciden. Sin embargo, en el caso general esta equivalencia no es cierta, como se ve en el siguiente ejemplo.

[2]Designamos $G(-\infty) = \lim_{x\to-\infty} G(x)$, cuando existe este límite para una cierta función $G(x)$. Análogamente, designamos $G(\infty) = \lim_{x\to\infty} G(x)$.

Consideremos, para cada $n = 1, 2, \dots$, la función

$$F_n(x) = \begin{cases} 0, & \text{si } x \le -n, \\ 1/2, & \text{si } -n < x \le n, \\ 1, & \text{si } n \le x. \end{cases}$$

Es claro que $F_n(x) \to 1/2$ $(n \to \infty)$ para todo x real, y en consecuencia la sucesión $\{F_n(x)\}$ considerada converge débilmente a la función $F(x) = 1/2$. Sin embargo, la convergencia completa no tiene lugar, dado que $F_n(-\infty) = 0$, $F_n(+\infty) = 1$ para todo n, y tenemos $F(-\infty) = F(\infty) = 1/2$.

Teorema 6.3 (Helly). *Consideremos una sucesión $\{F_n(x)\}$ de funciones no decrecientes. Supongamos que existen constantes A y B tales que se verifica $A \le F_n(x) \le B$ para todo x real y para todo n. Entonces, la sucesión dada contiene una subsucesión $\{F_{n_k}(x)\}$ que converge débilmente a una cierta función $F(x)$, no decreciente y acotada.*

En la demostración de este teorema utilizamos el siguiente resultado.

Lema 6.2. *Si $F_n(x) \to F(x)$ para todo $x \in \mathbf{D}$, donde $\mathbf{D}$ es un conjunto denso en la recta real, entonces $F_n \to F$.*

Demostración del lema. Sea x un punto de continuidad de $F(x)$. Como $\mathbf{D}$ es denso, existen dos sucesiones $\{x'_k\}$ y $\{x''_k\}$ que verifican $x'_k < x < x''_k$ para todo k y $\lim_{k\to\infty} x'_n = \lim_{k\to\infty} x''_k = x$. Para cada k y cada n, tenemos

$$F_n(x'_k) \le F_n(x) \le F_n(x''_k).$$

Como $\lim_{n\to\infty} F_n(x'_k) = F(x'_k)$ y $\lim_{n\to\infty} F_n(x''_k) = F(x''_k)$, de las desigualdades anteriores, si $n \to \infty$, obtenemos

$$F(x'_k) \le \liminf_{n\to\infty} F_n(x) \le \limsup_{n\to\infty} F_n(x) \le F(x''_k).$$

Hagamos ahora tender k a infinito. Dado que x es un punto de continuidad de $F(x)$, se verifica $\lim_{k\to\infty} F(x'_k) = \lim_{k\to\infty} F(x''_k) = F(x)$, por lo que

$$F(x) \le \liminf_{n\to\infty} F_n(x) \le \limsup_{n\to\infty} F_n(x) \le F(x).$$

Entonces, el $\lim_{n\to\infty} F_n(x)$ existe, y vale $F(x)$. Como el punto de continuidad es arbitrario, esto concluye la demostración. $\square$

Demostración del teorema 6.3 de Helly. Sea $\mathbf{D}$ un conjunto denso numerable de números reales $x'_1, x'_2, \dots$. La sucesión numérica $\{F_n(x'_1)\}$ es acotada, por lo que contiene una subsucesión $\{F_{1n}(x'_1)\}$ que converge a un cierto límite, que designamos $F(x'_1)$.

La sucesión $\{F_{1n}(x'_2)\}$ también contiene una subsucesión $\{F_{2n}(x'_2)\}$, convergente a un cierto límite, que designamos $F(x'_2)$. Además, se verifica $\lim_{n\to\infty} F_{2n}(x'_1) = F(x'_1)$.

Continuando este proceso, obtenemos, que para cualquier natural k, existen k sucesiones $\{F_{kn}(x'_i)\}$ $(i = 1, \dots, k)$ para las cuales se verifica $\lim_{n\to\infty} F_{kn}(x'_i) = F(x'_i)$ $(i = 1, \dots, k)$.

Consideremos ahora la sucesión diagonal compuesta por las funciones $\{F_{nn}(x)\}$. Sea $x'_k \in \mathbf{D}$. Es claro que $\lim_{n\to\infty} F_{nn}(x'_k) = F(x'_k)$, dado que $\{F_{nn}(x'_k)\}$ es una subsucesión de la sucesión numérica $\{F_{kn}(x'_k)\}$, si $n \geq k$. Hemos así definido una función $F(x)$ en el conjunto $\mathbf{D}$. Si $x < y$ son dos puntos de $\mathbf{D}$, entonces $F(x) = \lim_{n\to\infty} F_{nn}(x) \leq \lim_{n\to\infty} F_{nn}(y) = F(y)$, y la función $F(x)$ es no decreciente en $\mathbf{D}$. Es claro también que $A \leq F(x) \leq B$. Estas propiedades permiten extender la función $F(x)$ a toda la recta real, conservando las propiedades mencionadas. Estamos entonces en condiciones de aplicar el lema 6.2, para concluir la demostración del teorema. □

Observación. Se puede ver que la función límite $F(x)$ puede elegirse continua por la derecha, si definimos $F(x) = \lim_{n\to\infty} F(x_n)$, donde $\{x_n\} \in \mathbf{D}$, $x_n \to x$ $(n \to \infty)$, y $x_n \geq x$ para todo n.

Teorema 6.4 (Helly)**.** *Consideremos funciones no decrecientes y acotadas $F(x), F_1(x), F_2(x), \dots$ tales que $F_n \Rightarrow F$, y una función $g(x)$ continua y acotada. Entonces*

$$\int_{-\infty}^{\infty} g(x)dF_n(x) \to \int_{-\infty}^{\infty} g(x)dF(x) \quad (n \to \infty).$$

Demostración. Sea $\varepsilon > 0$ arbitrario. Designemos

$$G = \sup_{x\in\mathbb{R}} |g(x)|, \quad C = F(\infty) - F(-\infty).$$

Como $\lim_{x\to-\infty} F(x) = F(-\infty)$, $\lim_{x\to\infty} F(x) = F(\infty)$, existen $a < b$, puntos de continuidad de $F(x)$, tales que se verifica

$$F(\infty) - F(b) < \varepsilon/(3G), \qquad F(a) - F(-\infty) < \varepsilon/(3G). \tag{6.17}$$

Como $g(x)$ es una función continua, existe una partición del intervalo $[a,b]$, designada $a = x_0 < x_1 < \cdots < x_N = b$, formada por puntos de continuidad de $F(x)$, y tales que se verifica $|g(x) - g(x_k)| < \delta$ si $x \in (x_{k-1}, x_k)$ $(k = 1, \dots, N)$.

Consideremos la función auxiliar

$$g_0(x) = \begin{cases} g(x_k), & \text{si } x \in (x_{k-1}, x_k] \ (k = 1, \dots, N), \\ 0, & \text{en otro caso,} \end{cases}$$

que podemos también definir como $g_0(x) = \sum_{k=1}^{N} g(x_k)\mathbf{1}_{(x_{k-1},x_k]}(x)$. Sean

$$I = \int_{-\infty}^{\infty} g(x)dF(x), \quad I_n = \int_{-\infty}^{\infty} g(x)dF_n(x).$$

Sumando y restando, obtenemos

$$\begin{aligned} I_n - I &= \int_{-\infty}^{\infty} \big(g(x) - g_0(x)\big)dF_n(x) && (6.18) \\ &\quad + \int_{-\infty}^{\infty} g_0(x)dF_n(x) - \int_{-\infty}^{\infty} g_0(x)dF(x) && (6.19) \\ &\quad + \int_{-\infty}^{\infty} \big(g(x) - g_0(x)\big)dF(x) && (6.20) \\ &= S_1 + S_2 + S_3. \end{aligned}$$

Acotemos cada uno de los tres sumandos anteriores. Para S_3 en (6.20), tenemos

$$\begin{aligned} |S_3| &\le \int_{-\infty}^{a} \big|g(x)\big|dF(x) + \int_{a}^{b} \big|g(x) - g_0(x)\big|dF(x) + \int_{b}^{\infty} \big|g(x)\big|dF(x) \\ &\le G\big(F(a) - F(-\infty)\big) + \delta\big(F(b) - F(a)\big) + G\big(F(\infty) - F(b)\big) \quad (6.21) \\ &\le \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon, \end{aligned}$$

en vista de (6.17) y la desigualdad $F(b) - F(a) \le C$.

Para S_1 en (6.18), cambiando F_n por F, obtenemos

$$\begin{aligned} |S_1| &\le G\big(F_n(a) - F_n(-\infty)\big) \\ &\quad + \delta\big(F_n(b) - F_n(a)\big) + G\big(F_n(\infty) - F_n(b)\big) < \varepsilon, \quad (6.22) \end{aligned}$$

si n es suficientemente grande, dado que, por la convergencia completa $F_n \Rightarrow F$, la cota obtenida en (6.22) converge a la cota en (6.21).

Finalmente, también utilizando la convergencia completa $F_n \Rightarrow F$, obtenemos $S_2 \to 0$ $(n \to \infty)$ en (6.19), porque

$$\int_{-\infty}^{\infty} g_0(x)dF_n(x) = \sum_{k=1}^{N} g(x_k)\big(F_n(x_k) - F_n(x_{k-1})\big)$$
$$\to \sum_{k=1}^{N} g(x_k)\big(F(x_k) - F(x_{k-1})\big) = \int_{-\infty}^{\infty} g_0(x)dF(x),$$

si $n \to \infty$. En conclusión, para n suficientemente grande, tenemos

$$\Big|\int_{-\infty}^{\infty} g(x)dF_n(x) - \int_{-\infty}^{\infty} g(x)dF(x)\Big| < \varepsilon + \varepsilon + \varepsilon = 3\varepsilon,$$

lo que concluye la demostración. □

6.4. Relación entre la convergencia de distribuciones y de funciones características

El objetivo de esta sección es la demostración del siguiente resultado.

Teorema 6.5. *Consideremos las funciones de distribución*

$$F(x), F_1(x), F_2(x), \dots$$

con sus correspondientes funciones características

$$f(t), f_1(t), f_2(t), \dots$$

Entonces, la sucesión $\{F_n(x)\}$ converge débilmente a $F(x)$ (es decir $F_n \to F$), si y solo si se verifica

$$f_n(t) \to f(t) \ (n \to \infty) \ \textit{para todo } t \textit{ real.} \tag{6.23}$$

Demostración. Supongamos primero que $F_n \to F$. Tenemos

$$f_n(t) = \int_{-\infty}^{\infty} e^{itx}dF_n(x), \quad f(t) = \int_{-\infty}^{\infty} e^{itx}dF(x).$$

Como $e^{itx} = \cos tx + i \operatorname{sen} tx$, donde las funciones $\operatorname{sen} tx, \cos tx$ son continuas y acotadas; y tenemos $F_n \Rightarrow F(x)$ (porque se trata de funciones de distribución), obtenemos (6.23) aplicando el teorema 6.4 de Helly.

Supongamos ahora que se verifica la condición (6.23). En virtud del teorema 6.3 de Helly, la sucesión $\{F_n(x)\}$ contiene una cierta subsucesión $\{F_{n_k}(x)\}$, que converge débilmente a una cierta función $F(x)$ no decreciente, que verifica $0 \leq F(x) \leq 1$, y que elegimos continua por la derecha. Demostremos que $F(x)$ es una función de distribución. Para ésto hay que demostrar que $F(-\infty) = 0$ y $F(+\infty) = 1$, lo que equivale a demostrar que $\delta = F(+\infty) - F(-\infty) = 1$.

Supongamos que $\delta < 1$ y sea ε tal que $0 < \varepsilon < 1 - \delta$. Como $f(0) = 1$ y $f(t)$ es continua, para γ suficientemente pequeño es válida la desigualdad

$$\frac{1}{2\gamma} \int_{-\gamma}^{\gamma} f(t)dt > 1 - \frac{\varepsilon}{2} > \delta + \frac{\varepsilon}{2}. \tag{6.24}$$

Como $F_{n_k} \to F$, podemos elegir un real $A > \gamma/(4\varepsilon)$ que verifique: $F(x)$ es continua en A y en $-A$; para todo k suficientemente grande $\delta_k(A) = F_{n_k}(A) - F_{n_k}(-A) < \delta + \varepsilon/4$.

Introducimos ahora la función

$$B(x) = \int_{-\gamma}^{\gamma} e^{itx} dt = \frac{2}{x} \operatorname{sen}(\gamma x)$$

que, como $|e^{itx}| \leq 1$, verifica $|B(x)| \leq 2\gamma$. Además, como $|\operatorname{sen}(\gamma x)| \leq 1$, tenemos $|B(x)| \leq 2/A$, si $|x| \geq A$.

Cambiando el orden de integración y utilizando la función $B(x)$ recién introducida, tenemos

$$\int_{-\gamma}^{\gamma} f_{n_k}(t)dt = \int_{-\gamma}^{\gamma} \Big(\int_{-\infty}^{\infty} e^{itx} dF_{n_k}(x) \Big) dt = \int_{-\infty}^{\infty} B(x) dF_{n_k}(x).$$

Partiendo el intervalo de integración, obtenemos la siguiente acotación:

$$\begin{aligned} \Big| \int_{-\gamma}^{\gamma} f_{n_k}(t)dt \Big| &\leq \Big| \int_{\{|x| \leq A\}} B(x) dF_{n_k}(x) \Big| \\ &\quad + \Big| \int_{\{|x| > A\}} B(x) dF_{n_k}(x) \Big| \leq 2\gamma\delta_k + \frac{2}{A}. \end{aligned}$$

En vista de la elección de A, dividendo por 2γ, obtenemos

$$\frac{1}{2\gamma} \Big| \int_{-\gamma}^{\gamma} f_{n_k}(t)dt \Big| \leq \delta + \frac{\varepsilon}{2}.$$

Como $f_{n_k}(t) \to f(t)$ $(n \to \infty)$ con $|f_{n_k}(t)| \leq 1$, tomando límite en la desigualdad anterior, obtenemos

$$\frac{1}{2\gamma}\Big|\int_{-\gamma}^{\gamma} f(t)dt\Big| \leq \delta + \frac{\varepsilon}{2},$$

lo que contradice la desigualdad (6.24). Entonces $\delta = 1$ y $F(x)$ es una función de distribución.

Por último observemos, que como $F_{n_k} \to F$, aplicando la primera parte del teorema, obtenemos que $f(t)$ es la función característica que corresponde a esta distribución $F(x)$.

Para terminar, resta demostrar que toda la sucesión $\{F_n(x)\}$ converge débilmente a $F(x)$. Supongamos que esto no es cierto. Existe entonces una subsucesión $\{F_{n'}(x)\}$, que converge débilmente a una cierta función $G(x)$, que es distinta de $F(x)$, en por lo menos un punto de continuidad. Una argumentación similar a la aplicada a $\{F_{n_k}(x)\}$ nos permite obtener, que $G(x)$ es una función de distribución, y que $f(t)$ es su función característica. Aplicando el teorema 6.2 de unicidad de funciones características obtenemos que $F(x) = G(x)$ para todo x real, contradiciendo nuestro supuesto. Entonces $F_n \to F$, lo que concluye la demostración. $\square$

En el capítulo 7 estudiaremos distintas variantes del *teorema central del límite*, cuyas demostraciones se basan el teorema recién demostrado.

6.5. Ejercicios

1. Hallar la función característica de una variable aleatoria: (a) con distribución uniforme en el intervalo $(-\ell, \ell)$; (b) con densidad dada por $p(x) = (1 - \cos x)/(\pi x^2)$.

2. Dada una variable aleatoria X, la variable aleatoria simetrizada, designada X^s, se define mediante la igualdad $X^s = X - Y$, donde Y es una variable aleatoria independiente de X, y con su misma distribución. Demostrar que si X tiene función característica $f(t)$, entonces, X^s tiene función característica $|f(t)|^2$.

3. Consideremos una función característica $f(t)$. Demostrar la desigualdad

$$1 - |f(2t)|^2 \leq 4(1 - |f(t)|^2),$$

válida para todo t real.

4. Consideremos funciones características $f_1(t), \dots, f_n(t)$, y constantes positivas $b_1, \dots, b_n$, que verifican $b_1 + \cdots + b_n = 1$. Demostrar que $b_1 f_1(t) + \cdots + b_n f_n(t)$ es una función característica.

5. Determinar si las siguientes son funciones características: (a) $\operatorname{sen} t$; (b) $\cos t$; (c) $\cos^2 t$; (d) $\operatorname{sen} t + \cos t$; (e) $\big(e^{it} + e^{2it}\big)^3/8$; (f) $\operatorname{Re} f(t)$, donde $f(t)$ es una función característica; (g) $\operatorname{Im} f(t)$ donde $f(t)$ es una función característica.

6. Calcular la varianza $\mathbf{var}\, X$, de una variable aleatoria X, con función característica $f(t) = \big(1 + e^{3it}\big)^2/4$.

7. Sea X una variable aleatoria con distribución discreta. Esta distribución se denomina *látice*, si existen dos reales $h > 0$ y a, tales que se verifica $\sum_{k=-\infty}^{\infty} \mathbf{P}(X = a + hk) = 1$. (a) Encontrar una distribución discreta que no sea látice. (b) Demostrar que una distribución con función característica $f(t)$ es látice si y solo si existe $t_0 \neq 0$ tal que $|f(t_0)| = 1$.

8. Sea X una variable aleatoria con distribución látice, que toma los valores $a+hk$ $(k = 0, \pm 1, \pm 2, \dots)$, con probabilidades $p_k = \mathbf{P}(X = a+kh)$. Demostrar que se verifica

$$p_k = \frac{h}{2\pi} \int_{\{|t|<\frac{\pi}{h}\}} e^{-it(a+kh)} f(t) dt$$

para todo k entero, donde $f(t)$ es la función característica de X.

9. Utilizando el teorema 6.2 de unicidad, demostrar: si X e Y son variables aleatorias independientes, con distribución de Poisson con parámetros λ_1 y λ_2 respectivamente, entonces $X+Y$ tiene distribución de Poisson con parámetro $\lambda_1 + \lambda_2$.

10. Consideremos una función característica $f(t)$ y dos constantes b, c, que verifican $0 < c < 1$, $b > 0$. Demostrar que si $|f(t)| \leq c$, cuando $|t| \geq b$, entonces $|f(t)| \leq 1 - (1-c^2)t^2/(8b^2)$, si $|t| < b$.

11. Demostrar que si una función característica verifica la condición

$$\limsup_{|t|\to\infty} |f(t)| < 1$$

(denominada *condición* (C) *de Cramér*), entonces, para todo $\varepsilon > 0$, existe un real positivo $c < 1$, tal que $|f(t)| \leq c$, cuando $|t| \geq \varepsilon$. Sugerencia: Utilizar el ejercicio 10.

12. Consideremos una variable aleatoria con función característica $f(t)$. Demostrar que si la variable aleatoria es absolutamente continua, entonces $f(t) \to 0$, si $|t| \to \infty$. (Sugerencia: utlizar el teorema de Riemann-Lebesgue).

13. Sea $F(x)$ una función de distribución con función característica $f(t)$. Demostrar la igualdad

$$\int_{-\infty}^{\infty} \frac{x^2}{1+x^2} dF(x) = \int_0^{\infty} e^{-t}\big(1 - \operatorname{Re} f(t)\big) dt.$$

14. Una función característica se denomina *infinitamente divisible* si para cada natural $n \geq 1$, existe una función característica $f_n(t)$, tal que $f(t) = \big(f_n(t)\big)^n$. Demostrar las siguientes proposiciones: (a) Si $f(t)$ y $g(t)$ son funciones características infinitamente divisibles, entonces, $f(t)g(t)$ es una función característica infinitamente divisible. (b) Si $f(t)$ es función característica infinitamente divisible, entonces, $f(t) \neq 0$ para todo t real. (Sugerencia: utilizar la desigualdad del ejercicio 3.)

15. Si una función característica verifica $f(t) = 1+o(t^2)$ $(t \to 0)$, entonces $f(t) = 1$ para todo t real.

16. Sea $f(t)$ la función característica de una variable aleatoria con distribución no degenerada. Demostrar que existen reales positivos δ y ε, tales que $|f(t)| \leq 1 - \varepsilon t^2$ para $|t| \leq \delta$.

17. Sea X una variable aleatoria con función característica $f(t)$, con distribución látice, que toma los valores $a + hk$ $(k = 0, \pm 1, \pm 2, \dots)$, donde $h > 0$ y a son números reales fijos. El número h se denomina el *paso* de la distribución. El paso h se denomina *maximal*, si no existe un par h_1, a_1, con $h_1 > h$, tal que $\sum_{k=-\infty}^{\infty} \mathbf{P}(X = a_1 + h_1 k) = 1$. Demostrar que el paso h es maximal si y solo si se verifican las condiciones $\big|f(2\pi/h)\big| = 1$, y $|f(t)| < 1$ para $0 < |t| < 2\pi/h$.

18. Una función de distribución, o su función característica $f(t)$, se denomina *estable*, cuando para todo par a_1 y a_2 de números reales positivos, existen $a > 0$ y b reales, tales que $f(a_1t)f(a_2t) = e^{ibt}f(at)$. (a) Demostrar que esta definición es equivalente a la siguiente: la función de distribución $F(x)$ es estable, si para todo $a_1 > 0$, $a_2 > 0$, b_1 y b_2, existen $a > 0$ y b reales tales que $F(a_1x + b_1) * F(a_2x + b_2) = F(ax + b)$, donde $*$ designa la convolución. (b) Determinar si son estables las siguientes distribuciones: (i) degenerada, (ii) normal, (iii) uniforme, (iv) binomial, (v) de Poisson.

19. (a) Hallar la función característica de una variable aleatoria con distribución Gama, con parámetros (α, β). (b) Demostrar que si $T_1, \dots, T_n$ son variables aleatorias independientes con distribución común exponencial de parámetro α, entonces, su suma $T_1 + \cdots + T_n$, tiene densidad dada por $p(x) = \alpha^n x^{n-1} e^{-\alpha x}/(n-1)!$.

Capítulo 7

Teorema central del límite

Se[1] denomina *teorema central del límite* a cualquier proposición que establece, bajo determinadas condiciones, que la función de distribución de la suma de una cantidad creciente a infinito de variables aleatorias, converge a la función de distribución normal. Aplicando el teorema central del límite podemos aproximar la distribución de la suma de un gran número de variables aleatorias, mediante la distribución normal. En este capítulo, dedicado a estudiar diversas variantes del teorema central del límite, comenzamos por el teorema de Lindeberg–Lévy que considera sumas de variables independientes e idénticamente distribuídas. En la sección 2 se demuestra un resultado más general: el teorema de Lindeberg, en el cual no se supone que las variables aleatorias consideradas tienen la misma distribución.

7.1. Teorema de Lindeberg–Lévy

Decimos que $X_1, X_2, \dots$ es una *sucesión de variables aleatorias independientes*, cuando las variables aleatorias $X_1, \dots, X_n$ son mutuamente independientes para cada $n = 1, 2, \dots$. Recordemos, que $X_1, X_2, \dots$ es una *sucesión de variables aleatorias idénticamente distribuidas* cuando todas las variables aleatorias consideradas tiene la misma distribución.

Teorema 7.1 (Lindeberg–Lévy).
Consideremos una sucesión $X_1, X_2, \dots$ de variables aleatorias independi-

[1]Capítulo 7 del libro "Teoría de la Probabilidad" por Valentín Petrov y Ernesto Mordecki

entes e idénticamente distribuidas, con esperanza matemática $\mathbf{E}\, X_1 = a$ *y varianza* $\mathbf{var}\, X_1 = \sigma^2 > 0$. *Designemos*

$$F_n(x) = \mathbf{P}\Big(\frac{X_1 + \cdots + X_n - na}{\sigma\sqrt{n}} \leq x\Big).$$

Entonces,

$$F_n(x) \to \Phi(x) \ (n \to \infty) \text{ para todo } x \text{ real}, \tag{7.1}$$

donde $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$ *es la distribución normal estándar.*

Demostración. La primer etapa consiste en demostrar el teorema en el caso particular en el que $\mathbf{E}\, X_1 = a = 0$ y $\mathbf{var}\, X_1 = \sigma^2 = 1$.

Consideremos entonces para cada $n = 1, 2, \dots$ la variable aleatoria

$$Z_n = \frac{1}{\sqrt{n}}\big(X_1 + \cdots + X_n\big).$$

Como $a = 0$ y $\sigma^2 = 1$, tenemos $\mathbf{P}(Z_n \leq x) = F_n(x)$. La demostración se basa en la aplicación del teorema 6.5. Calculemos $f_n(t)$, la función característica de Z_n, en términos de $v(t)$, la función característica de X_1:

$$\begin{aligned} f_n(t) &= \mathbf{E}\, e^{itZ_n} = \mathbf{E}\, e^{i(t/\sqrt{n}) \sum_{k=1}^n X_k} = \mathbf{E} \prod_{k=1}^{n} e^{i(t/\sqrt{n})X_k} \\ &= \prod_{k=1}^{n} \mathbf{E}\, e^{i(t/\sqrt{n})X_k} = \Big[v\Big(\frac{t}{\sqrt{n}}\Big)\Big]^n, \end{aligned} \tag{7.2}$$

donde utilizamos que las variables aleatorias son idénticamente distribuidas en la última igualdad, y que son independientes en la ante última. Como $\alpha_2 = \mathbf{var}\, X_1 = 1 < \infty$, aplicando el desarrollo de Taylor (6.10) de orden $k = 2$ para la función característica de X_1, tenemos

$$v(u) = 1 - \frac{u^2}{2} + o(u^2) \ (u \to 0), \tag{7.3}$$

dado que $\alpha_1 = \mathbf{E}\, X_1 = 0$. Consideremos un real t arbitrario y fijo. Queremos calcular $\lim_{n\to\infty} f_n(t)$. Si en (7.3) ponemos $u = t/\sqrt{n}$, tenemos

$$v\Big(\frac{t}{\sqrt{n}}\Big) = 1 - \frac{t^2}{2n} + o\Big(\frac{1}{n}\Big) \quad (n \to \infty),$$

dado que $1/n \to 0$ si y solo si $u \to 0$.

Verifiquemos ahora la validez de la identidad

$$\ln(1+z) = z + r(z), \tag{7.4}$$

donde $|r(z)| \leq 2|z|^2$, si z es un número complejo que verifica $|z| < 1/2$. En efecto, considerando el desarrollo de Taylor de la función logaritmo, tenemos

$$r(z) = \ln(1+z) - z = \sum_{m=2}^{\infty} (-1)^{m-1} z^m/m.$$

Acotando y sumando la serie geométrica que se obtiene, tenemos

$$|r(z)| \leq \sum_{m=2}^{\infty} |z|^m = |z|^2 \sum_{m=0}^{\infty} |z|^m = \frac{|z|^2}{1-|z|} \leq 2|z|^2,$$

porque $(1-|z|)^{-1} \leq 2$, dado que $|z| \leq 1/2$. Esto prueba (7.4).

Si $z = v(t/\sqrt{n}) - 1 = -t^2/(2n) + o(1/n)$, para n suficientemente grande podemos aplicar la fórmula (7.4), para obtener

$$\ln v\Big(\frac{t}{\sqrt{n}}\Big) = -\frac{t^2}{2n} + o\Big(\frac{1}{n}\Big),$$

porque, en este caso $|r(z)| \leq 2|z|^2 = t^4/(2n^2) + o(1/n^2)$.

Estamos en condiciones de tomar logaritmo en la fórmula (7.2):

$$\ln f_n(t) = n \ln v\Big(\frac{t}{\sqrt{n}}\Big) = n\Big[-\frac{t^2}{2n} + o\Big(\frac{1}{n}\Big)\Big] = -\frac{t^2}{2} + o(1).$$

En otras palabras, $f_n(t) \to e^{-t^2/2}$ si $n \to \infty$. Como $f(t) = e^{-t^2/2}$ es la función característica de la distribución normal $\Phi(x)$, aplicando el teorema 6.5 obtenemos que $F_n(x) \to \Phi(x)$ para todo x real (dado que $\Phi(x)$ es continua en $\mathbb{R}$), concluyendo la primer etapa de la demostración ($a = 0, \sigma^2 = 1$).

Supongamos ahora que $\mathbf{E}\, X_1 = a, \mathbf{var}\, X_1 = \sigma^2$, con $a, \sigma > 0$ arbitrarios. Consideremos las variables aleatorias auxiliares $Y_n = (X_n - a)/\sigma$ ($n = 1, 2, \dots$). Es fácil de ver que $\{Y_n\}$ es una sucesión de variables aleatorias independientes e idénticamente distribuidas, con $\mathbf{E}\, Y_1 = 0$, y $\mathbf{var}\, Y_1 = 1$. Entonces

$$\begin{aligned} F_n(x) = \mathbf{P}\Big(\frac{X_1 + \cdots + X_n - na}{\sigma\sqrt{n}} \leq x\Big) \\ = \mathbf{P}\Big(\frac{Y_1 + \cdots + Y_n}{\sqrt{n}} \leq x\Big) \to \Phi(x), \end{aligned}$$

para todo x real si $n \to \infty$, dado que $\{Y_n\}$ verifica las condiciones de la primer etapa de la demostración. Esto es la tesis del teorema. $\square$

Observación. Es posible demostrar que la convergencia en (7.1) es uniforme en el conjunto de los x reales. No es difícil verificar esta afirmación directamente; es consecuencia del *teorema de Pólya*: si una sucesión de funciones de distribución $\{G_n(x)\}$ converge a una función de distribución continua $G(x)$ para todo x, entonces, esta convergencia es uniforme en la recta real (ver ejercicio **??**, capítulo 5).

Veamos que el teorema límite integral de De Moivre–Laplace de la sección 2.3, es un corolario del teorema de Lindeberg–Lévy recién demostrado.

Consideremos entonces una serie de n experimentos independientes, con dos resultados posibles cada uno (éxito y fracaso), y probabilidad de éxito igual a p en cada experimento ($0 < p < 1$). Sea μ la cantidad de éxitos en n experimentos. Veamos como se formula el teorema límite integral de De Moivre–Laplace en términos de variables aleatorias. Para ésto consideremos una sucesión de variables aleatorias $X_1, X_2, \dots$, cada una de las cuales toma el valor 1 con probabilidad p (si ocurre un éxito) y el valor 0 con probabilidad $q = 1-p$ (si ocurre un fracaso). Tenemos $\mathbf{E}\, X_k = p$, $\mathbf{var}\, X_k = pq > 0$ para cada $i = 1, 2, \dots$. Además, $\mu = \sum_{k=1}^{n} X_k$, porque la suma contiene tantos sumandos iguales a uno como éxitos ocurren en los primeros n experimentos, siendo nulos los sumandos restantes. La sucesión $\{X_n\}$ es una sucesión de variables aleatorias independientes e idénticamente distribuidas, por lo que es aplicable el teorema de Lindeberg–Lévy. De (7.1) obtenemos, que

$$\mathbf{P}\left(\frac{\mu - np}{\sqrt{npq}} \leq x\right) - \Phi(x) \to 0 \quad (n \to \infty), \tag{7.5}$$

uniformemente en el conjunto de los x reales, en vista de la última observación. Poniendo entonces en (7.5) primero $x = b$, luego $x = a$, y restando, obtenemos

$$\mathbf{P}\left(a < \frac{\mu - np}{\sqrt{npq}} \leq b\right) - \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx \to 0 \quad (n \to \infty),$$

uniformemente, en el conjunto de los reales $a < b$, que es el contenido del teorema límite integral de De Moivre–Laplace.

7.2. Teorema de Lindeberg

Teorema 7.2 (Lindeberg)**.**
Consideremos una sucesión $X_1, X_2, \dots$ de variables aleatorias independientes, con esperanzas $a_1, a_2, \dots$ y varianzas $\sigma_1^2, \sigma_2^2, \dots$, no todas nulas. Designemos

$$V_n(x) = \mathbf{P}(X_n \leq x), \quad B_n = \sum_{k=1}^{n} \sigma_k^2,$$

$$F_n(x) = \mathbf{P}\Big(\frac{1}{\sqrt{B_n}} \sum_{k=1}^{n} (X_k - a_k) \leq x\Big).$$

Supongamos que se verifica la condición de Lindeberg: *Para todo $\varepsilon > 0$*

$$\Lambda_n(\varepsilon) = \frac{1}{B_n} \sum_{k=1}^{n} \int_{\{|x-a_k| \geq \varepsilon\sqrt{B_n}\}} (x - a_k)^2 dV_k(x) \to 0 \quad (n \to \infty). \tag{7.6}$$

Entonces

$$F_n(x) \to \Phi(x) \ (n \to \infty) \textit{ para todo } x \textit{ real.}$$

La demostración del teorema de Lindeberg utiliza el siguiente resultado de cálculo, que incluimos por conveniencia del lector.

Lema 7.1. *Vale la desigualdad*

$$\Big| e^{ix} - \sum_{\nu=0}^{k-1} \frac{(ix)^\nu}{\nu!} \Big| \leq \frac{1}{k!} |x|^k, \tag{7.7}$$

para todo real x, y todo natural $k \geq 1$.

Demostración del lema 7.1. Como

$$\int_0^x e^{it} dt = \frac{1}{i}(e^{ix} - 1)$$

obtenemos que $|e^{ix} - 1| \leq |x|$, demostrando la fórmula (7.7) para $k = 1$.

Para demostrar la validez de (7.7) para $k + 1$, a partir de su validez para k, escribimos

$$I = \int_0^x \Big(e^{it} - \sum_{\nu=0}^{k-1} \frac{(it)^\nu}{\nu!}\Big) dt = \frac{1}{i}\Big(e^{ix} - \sum_{\nu=0}^{k} \frac{(ix)^\nu}{\nu!}\Big).$$

Entonces

$$\left|e^{ix} - \sum_{\nu=0}^{k} \frac{(ix)^\nu}{\nu!}\right| = |I| \leq \int_0^{|x|} \frac{t^k}{k!} dt = \frac{|x|^{k+1}}{(k+1)!},$$

concluyendo la demostración. □

Demostración del teorema 7.2 de Lindeberg. Observemos en primer lugar que podemos suponer $a_n = 0$ $(n = 1, 2, \dots)$. El caso general en el que las variables aleatorias tienen esperanzas arbitrarias, se reduce a este caso particular mediante la consideración de las variables aleatorias $Y_n = X_n - a_n$ $(n = 1, 2, \dots)$ que verifican $\mathbf{E}\,Y_n = 0$, $\mathbf{var}\,Y_n = \sigma_n^2$.

Supongamos entonces que $\mathbf{E}\,X_n = 0$ $(n = 1, 2, \dots)$. En primer lugar, demostremos que la condición (7.6) de Lindeberg, que en nuestro caso es: Para todo $\varepsilon > 0$

$$\Lambda_n(\varepsilon) = \frac{1}{B_n} \sum_{k=1}^{n} \int_{\{|x| \geq \varepsilon\sqrt{B_n}\}} x^2 dV_k(x) \to 0 \quad (n \to \infty),$$

implica la condición:

$$\frac{1}{B_n} \max_{1 \leq k \leq n} \sigma_k^2 \to 0 \quad (n \to \infty). \tag{7.8}$$

En efecto, para $\varepsilon > 0$ arbitrario y para cada $k = 1, 2, \dots, n$, tenemos

$$\begin{aligned}\sigma_k^2 &= \int_{\{|x| < \varepsilon\sqrt{B_n}\}} x^2 dV_k(x) + \int_{\{|x| \geq \varepsilon\sqrt{B_n}\}} x^2 dV_k(x) \\ &\leq \varepsilon^2 B_n + \sum_{k=1}^{n} \int_{\{|x| \geq \varepsilon\sqrt{B_n}\}} x^2 dV_k(x).\end{aligned}$$

Entonces, como la cota obtenida no depende de k, dividiendo por B_n tenemos

$$\frac{1}{B_n} \max_{1 \leq k \leq n} \sigma_k^2 \leq \varepsilon^2 + \Lambda_n(\varepsilon) < \varepsilon^2 + \varepsilon,$$

para n suficientemente grande. Como $\varepsilon > 0$ es arbitrario, hemos demostrado (7.8).

Consideremos ahora $Z_n = \sum_{k=1}^{n} X_k/\sqrt{B_n}$ y calculemos su función característica $f_n(t)$ en términos de $v_k(t)$, la función característica de X_k $(k =$

$1, \ldots, n)$. Tenemos

$$f_n(t) = \mathbf{E}\, e^{itZ_n} = \mathbf{E}\, e^{i(t/\sqrt{B_n})\sum_{k=1}^n X_k} = \mathbf{E} \prod_{k=1}^n e^{i(t/\sqrt{B_n})X_k}$$

$$= \prod_{k=1}^n \mathbf{E}\, e^{i(t/\sqrt{B_n})X_k} = \prod_{k=1}^n v_k\Big(\frac{t}{\sqrt{B_n}}\Big), \tag{7.9}$$

donde utilizamos que las variables aleatorias son independientes.

Para la demostración del teorema es suficiente verificar que $f_n(t) \to e^{-t^2/2}$ $(n \to \infty)$ y aplicar el teorema 6.5. Para demostrar entonces la convergencia de las funciones características, tomamos logaritmo en (7.9) y utilizamos el siguiente resultado.

Lema 7.2. *Para cada t real, tiene lugar la igualdad*

$$\ln f_n(t) = \sum_{k=1}^n \ln v_k\Big(\frac{t}{\sqrt{B_n}}\Big) = \sum_{k=1}^n \Big[v_k\Big(\frac{t}{\sqrt{B_n}}\Big) - 1\Big] + R_n(t),$$

donde $R_n(t) \to 0$ $(n \to \infty)$.

Demostración del lema 7.2. Consideremos

$$r_k(t) = \ln v_k\Big(\frac{t}{\sqrt{B_n}}\Big) - \Big[v_k\Big(\frac{t}{\sqrt{B_n}}\Big) - 1\Big] \text{ para } k = 1, \ldots, n,$$

$$R_n(t) = \sum_{k=1}^n r_k(t).$$

Como $\mathbf{E}\, X_k = \int_{-\infty}^{\infty} x dV_k(x) = 0$, aplicando el lema 7.1 con $k = 2$, tenemos

$$\Big|v_k\Big(\frac{t}{\sqrt{B_n}}\Big) - 1\Big| = \Big|\int_{-\infty}^{+\infty} \Big(e^{\frac{itx}{\sqrt{B_n}}} - 1 - \frac{itx}{\sqrt{B_n}}\Big) dV_k(x)\Big|$$

$$\leq \frac{t^2}{2B_n} \int_{-\infty}^{+\infty} x^2 dV_k(x) = \frac{t^2 \sigma_k^2}{2B_n} \tag{7.10}$$

$$\leq \frac{t^2}{2} \times \frac{1}{B_n} \max_{1 \leq k \leq n} \sigma_k^2 \to 0 \quad (n \to \infty), \tag{7.11}$$

según vimos en la fórmula (7.8). Luego, si n es suficientemente grande, designando $z_k = v_k(t/\sqrt{B_n}) - 1$, se verifica $|z_k| < 1/2$ para todo $k = 1, \ldots, n$, y podemos utilizar el desarrollo del logaritmo (7.4), para obtener

$$|r_k(t)| \leq 2\Big|v_k\Big(\frac{t}{\sqrt{B_n}}\Big) - 1\Big|^2 \leq \frac{t^4}{2B_n^2}\sigma_k^4 \leq \frac{t^4 \sigma_k^2}{2B_n} \times \frac{1}{B_n} \max_{1 \leq k \leq n} \sigma_k^2,$$

donde utilizamos (7.10). Por ésto,

$$|R_n(t)| = \sum_{k=1}^{n} |r_k(t)| \le \frac{t^4}{2} \sum_{k=1}^{n} \frac{\sigma_k^2}{B_n} \times \frac{1}{B_n} \max_{1\le k\le n} \sigma_k^2 \to 0 \quad (n \to \infty),$$

en vista de (7.8), concluyendo la demostración del lema. □

Resta la etapa final, que consiste en demostrar que

$$\ln f_n(t) + t^2/2 \to 0 \quad (n \to \infty). \tag{7.12}$$

Con este fin, introducimos

$$I_k = v_k\Big(\frac{t}{\sqrt{B_n}}\Big) - 1 + \frac{t^2\sigma_k^2}{2B_n} = \int_{-\infty}^{\infty} \Big(e^{itx/\sqrt{B_n}} - 1 - \frac{itx}{\sqrt{B_n}} - \frac{(itx)^2}{2B_n}\Big) dV_k(x)$$

para cada $k = 1, 2, \dots$. Acotamos I_k partiendo el dominio de integración en en las regiones $\{|x| < \varepsilon\sqrt{B_n}\}$ y $\{|x| \ge \varepsilon\sqrt{B_n}\}$, y utilizando el lema 7.1 con $k = 3$ en la primer región, y con $k = 2$ en la segunda. Las acotaciones que obtenemos de dicho lema, son

$$\Big|e^{iy} - 1 - iy - \frac{(iy)^2}{2}\Big| \le \frac{1}{6}|y|^3,$$
$$\Big|e^{iy} - 1 - iy - \frac{(iy)^2}{2}\Big| \le |e^{iy} - 1 - iy| + \frac{1}{2}|y|^2 \le |y|^2.$$

Por eso, si $y = ix/\sqrt{B_n}$, tenemos

$$\begin{aligned} |I_k| &\le \int_{\{|x|<\varepsilon\sqrt{B_n}\}} \frac{|tx|^3}{6B_n^{3/2}} dV_k(x) + \int_{\{|x|\ge\varepsilon\sqrt{B_n}\}} \frac{|tx|^2}{B_n} dV_k(x) \\ &\le \frac{\varepsilon\sigma_k^2|t|^3}{6B_n} + \frac{t^2}{B_n} \int_{\{|x|\ge\varepsilon\sqrt{B_n}\}} |x|^2 dV_k(x). \end{aligned} \tag{7.13}$$

Estamos en condiciones de concluir la demostración. Utilizando el lema 7.2, tenemos

$$\begin{aligned} \ln f_n(t) + \frac{t^2}{2} &= \sum_{k=1} \Big[\ln v_k\Big(\frac{t}{\sqrt{B_n}}\Big) + \frac{t^2\sigma_k^2}{2B_n}\Big] \\ &= \sum_{k=1} \Big[v_k\Big(\frac{t}{\sqrt{B_n}}\Big) - 1 - \frac{(it)^2\sigma_k^2}{2B_n}\Big] + R_n(t) \\ &= \sum_{k=1}^{n} I_k + R_n(t). \end{aligned}$$

Aplicando ahora la acotación (7.13), tenemos

$$\left| \ln f_n(t) + \frac{t^2}{2} \right| \leq \sum_{k=1}^{n} |I_k| + R_n(t) \leq \frac{\varepsilon |t|^3}{6} + t^2 \Lambda_n(\varepsilon) + R_n(t).$$

Como el primer sumando es arbitrariamente pequeño, el segundo converge a cero (aplicando la condición de Lindeberg), y el tercero también tiende a cero (según demostramos en el lema 7.2), obtuvimos (7.12). Con la aplicación del teorema 6.5 concluimos la demostración. □

El teorema 7.1 de Lindeberg–Lévy, resulta ser un corolario del teorema de Lindeberg, recién demostrado. En efecto, en el caso de variables aleatorias con distribución común $V(x)$, esperanza matemática a y variancia $\sigma^2 > 0$, obtenemos que $B_n = n\sigma^2$, y, para $\varepsilon > 0$ arbitrario, se verifica

$$\Lambda_n(\varepsilon) = \frac{1}{\sigma^2} \int_{\{|x-a| \geq \varepsilon \sigma \sqrt{n}\}} (x-a)^2 dV(x) \to 0 \quad (n \to \infty)$$

porque $\sigma^2 = \int_{-\infty}^{\infty} (x-a)^2 V(x) < \infty$. En conclusión, si se verifican las hipótesis del teorema de Lindeberg–Lévy, también se verifican las del teorema de Lindeberg; mientras que las tesis de estos dos teoremas, en el caso particular considerado, coinciden.

Volvamos ahora al caso general, en el que las distribuciones no necesariamente son idénticas. Sea $\{X_n\}$ una sucesión de variables aleatorias que verifica las condiciones del teorema 7.2 de Lindeberg. Consideremos

$$X_{nk} = \frac{X_k - a_k}{\sqrt{B_n}} \quad (k = 1, \dots, n).$$

Poniendo $Z_n = \sum_{k=1}^{n} (X_k - a_k)/\sqrt{B_n}$, tenemos $Z_n = \sum_{k=1}^{n} X_{nk}$. Demostremos que la condición de Lindeberg implica la condición: Para todo $\varepsilon > 0$

$$\mathbf{P}\left(\max_{1 \leq k \leq n} |X_{nk}| \geq \varepsilon \right) \to 0 \quad (n \to \infty). \tag{7.14}$$

La fórmula (7.14) significa que las variables aleatorias son uniformemente "pequeñas". Veamos su demostración. Dado $\varepsilon > 0$, tenemos

$$\begin{aligned} \mathbf{P}(|X_{nk}| \geq \varepsilon) = \mathbf{P}(|X_k - a_k| \geq \varepsilon \sqrt{B_n}) &= \int_{\{|x-a_k| \geq \varepsilon \sqrt{B_n}\}} dV_k(x) \\ &\leq \frac{1}{\varepsilon^2 B_n} \int_{\{|x-a_k| \geq \varepsilon \sqrt{B_n}\}} (x-a_k)^2 dV_k(x). \end{aligned}$$

Por ésto,

$$\begin{aligned}\mathbf{P}\Big(\max_{1\leq k\leq n}|X_{nk}|\geq\varepsilon\Big)&\leq\mathbf{P}\Big(\bigcup_{k=1}^{n}\big\{|X_{nk}|\geq\varepsilon\big\}\Big)\leq\sum_{k=1}^{n}\mathbf{P}\left(|X_{nk}|\geq\varepsilon\right)\\&\leq\frac{1}{\varepsilon^2 B_n}\sum_{k=1}^{n}\int_{\{|x-a_k|\geq\varepsilon\sqrt{B_n}\}}(x-a_k)^2dV_k(x)=\frac{1}{\varepsilon^2}\Lambda_n(\varepsilon),\end{aligned}$$

obteniendo la condición (7.14).

7.3. Teorema de Lyapunov

Teorema 7.3 (Lyapunov)**.**
Consideremos una sucesión $X_1, X_2, \dots$ de variables aleatorias independientes, con esperanzas $a_1, a_2, \dots$ y varianzas $\sigma_1^2, \sigma_2^2, \dots$, no todas nulas. Designemos

$$B_n=\sum_{k=1}^{n}\sigma_k^2,\quad F_n(x)=\mathbf{P}\Big(\frac{1}{\sqrt{B_n}}\sum_{k=1}^{n}(X_k-a_k)\leq x\Big).$$

Supongamos que se verifica la condición de Lyapunov: *Existe $\delta>0$ tal que*

$$L_n(\delta)=\frac{1}{B_n^{1+\delta/2}}\sum_{k=1}^{n}\mathbf{E}\,|X_n-a_n|^{2+\delta}\to 0\quad(n\to\infty).\tag{7.15}$$

Entonces

$$F_n(x)\to\Phi(x)\ (n\to\infty)\ \textit{para todo } x \textit{ real.}$$

Observemos que la condición de Lyapunov implica que existen los momentos de orden $2+\delta$ de las variables aleatorias X_k $(k=1,2,\dots)$.

Demostración. Como la tesis del teorema 7.2 y la del teorema 7.3 coinciden, es suficiente demostrar que la condición (7.15) de Lyapunov implica la condición (7.6) de Lindeberg. En efecto, supongamos que se verifica la condición (7.15) para un cierto $\delta > 0$. Dado $\varepsilon > 0$ arbitrario, tenemos

$$\begin{aligned}\Lambda_n(\varepsilon) &= \frac{1}{B_n}\sum_{k=1}^{n}\int_{\{|x-a_k|\geq\varepsilon\sqrt{B_n}\}}(x-a_k)^2 dV_k(x)\\ &\leq \frac{1}{\varepsilon^\delta B_n^{1+\delta/2}}\sum_{k=1}^{n}\int_{\{|x-a_k|\geq\varepsilon\sqrt{B_n}\}}|x-a_k|^{2+\delta}dV_k(x)\\ &\leq \frac{1}{\varepsilon^\delta B_n^{1+\delta/2}}\sum_{k=1}^{n}\int_{-\infty}^{\infty}|x-a_k|^{2+\delta}dV_k(x) = \frac{1}{\varepsilon^\delta}L_n(\delta)\to 0 \quad (n\to\infty).\end{aligned}$$

De aquí se obtiene la demostración, resultando el teorema de Lyapunov un caso particular del teorema de Lindeberg. $\square$

Como conclusión de este capítulo haremos algunas consideraciones relativas a la *velocidad de convergencia* en el teorema central del límite.

Lyapunov demostró que si se verifican las condiciones del teorema 7.3 con $0 < \delta < 1$, entonces, existe una constante C tal que, para n suficientemente grande, se verifica

$$|F_n(x) - \Phi(x)| \leq C L_n(\delta),$$

uniformemente en el conjunto de los x reales.

En el caso $\delta = 1$ Esseen obtuvo la siguiente desigualdad, válida para *todo* natural $n = 1, 2, \dots$. Sean $X_1, \dots, X_n$ variables aleatorias independientes, con esperanzas $a_1, \dots, a_n$ y varianzas $\sigma_1^2, \dots, \sigma_n^2$, no todas nulas. Supongamos que $\mathbf{E}\,|X_k - a_k|^3 < \infty$ $(k = 1, \dots, n)$. Designemos

$$B_n = \sum_{k=1}^{n}\sigma_k^2, \quad F_n(x) = \mathbf{P}\Big(\frac{1}{\sqrt{B_n}}\sum_{k=1}^{n}(X_k - a_k) \leq x\Big),$$

$$L_n = L_n(1) = \frac{1}{B_n^{3/2}}\sum_{k=1}^{n}\mathbf{E}\,|X_n - a_n|^3.$$

Entonces

$$|F_n(x) - \Phi(x)| \leq A L_n \quad \text{para todo } x \text{ real y todo } n = 1, 2, \dots, \tag{7.16}$$

donde $A > 0$ es una constante absoluta. (Está demostrado que la acotación (7.16) es válida con $A = 0{,}8$). De (7.16) se obtiene que si $L_n \to 0$ entonces $F_n(x) \to \Phi(x)$ para todo x real.

La desigualdad (7.16) es válida también si $\mathbf{E}\,|X_n - a_n|^{2+\delta} < \infty$ $(k = 1, \dots, n)$ para algún $0 < \delta \leq 1$, definiendo $L_n(\delta)$ como en (7.15).

En el caso particular en el que $X_1, \dots, X_n$ son variables aleatorias idénticamente distribuidas, con $a = \mathbf{E}\,X_1$, $\sigma^2 = \mathbf{var}\,X_1$, y $\delta = 1$, los resultados anteriores son los siguientes. Si designamos

$$L_n = \frac{\rho}{\sqrt{n}}, \text{ con } \rho = \frac{\mathbf{E}\,|X_1 - a|^3}{\sigma^3},$$

aplicando la desigualdad (7.16) de Esseen, obtenemos

$$|F_n(x) - \Phi(x)| \leq \frac{A\rho}{\sqrt{n}} \tag{7.17}$$

para todo x real y todo $n = 1, 2, \dots$ Es posible demostrar (ver por ejemplo §5.2 en [**?**]), que sin la introducción de condiciones adicionales, esta acotación es óptima en el siguiente sentido: el término $\sqrt{n}$ en el denominador, a la derecha en (7.17), no se puede sustituir por una función $g(n)$, que verifique $g(n)/\sqrt{n} \to \infty$ $(n \to \infty)$.

7.4. Ejercicios

1. Al disparar a un blanco, se obtienen 10 puntos con probabilidad 0,3; 9 puntos con probabilidad 0,3; 8 con probabilidad 0,2; 7 con probabilidad 0,1 y 6 con probabilidad 0,1. Utilizando el teorema central del límite, estimar la probabilidad de que, al realizar 100 disparos, se obtengan más de 870 puntos.

2. Se tira un dado 500 veces. Hallar un valor aproximado para la probabilidad de que la suma de puntos obtenidos sea mayor que 1800.

3. Sea $\{X_n\}$ una sucesión de variables aleatorias independientes, e idénticamente distribuidas, con $\mathbf{E}\,X_1 = a$, y $\mathbf{var}\,X_1 = \sigma^2 > 0$. Verificar que $\{Y_n\}$, con $Y_n = (X_n - a)/\sigma$, es una sucesión de variables aleatorias independientes, idénticamente distribuidas, con esperanza nula, y varianza igual a uno.

4. Sea $\{X_n\}$ una sucesión de variables aleatorias independientes, que verifican $\mathbf{P}(X_n = n^a) = \mathbf{P}(X_n = -n^a) = 1/2$ para cada $n = 1, 2, \dots$, donde $a > -1/2$. ¿Es aplicable el teorema central del límite a esta sucesión?

5. Sea $\{X_n\}$ una sucesión de variables aleatorias independientes, con esperanza nula. Supongamos que $|X_n| \leq C$ para todo n, donde C es una cierta constante. Sea $B_n = \sum_{k=1}^{n} \mathbf{var}\, X_k^2$. Demostrar que si $B_n \to \infty$ ($n \to \infty$), entonces, es aplicable el teorema central del límite a esta sucesión, es decir $\mathbf{P}\left(\frac{1}{\sqrt{B_n}} \sum_{k=1}^{n} X_k \leq x\right) \to \Phi(x)$.

6. Sea $\{X_n\}$ una sucesión de variables aleatorias independientes, que verifican $\mathbf{P}(X_n = -1/\sqrt{n}) = \mathbf{P}(X_n = 1/\sqrt{n}) = p$, $\mathbf{P}(X_n = 0) = 1 - 2p$ para todo n, y algún p en el intervalo $0 < p \leq 1/2$. ¿Es aplicable el teorema central del límite a esta sucesión?

7. Sea $\{X_n\}$ una sucesión de variables aleatorias independientes, tales que, cada variable aleatoria X_n, tiene distribución uniforme, en el intervalo $(-\sqrt{n}, \sqrt{n})$. Demostrar que para esta sucesión, es aplicable el teorema central del límite.

8. Sea μ la cantidad de éxitos, en una serie n experimentos independientes, con dos resultados posibles cada uno (éxito y fracaso). Sea p_k la probabilidad de que ocurra un éxito en el k-ésimo experimento, y $q_k = 1 - p_k$, la probabilidad de que ocurra un fracaso ($k = 1, 2, \dots$). Demostrar que si $\sum_{k=1}^{\infty} p_k q_k = \infty$, entonces, la función de distribución de la variable aleatoria $(\mu - \sum_{k=1}^{n} p_k)(\sum_{k=1}^{n} p_k q_k)^{-1/2}$ (cantidad de éxitos, normalizados), converge a la distribución normal estándar, si $n \to \infty$.

9. Sea $\{X_n\}$ una sucesión de variables aleatorias independientes, que verifica la condición de Lindeberg. Demostrar que $B_n \to \infty$. (Sugerencia: Utilizar, que la condición de Lindeberg implica la condición (7.8).)

10. Sea $\{X_n\}$ una sucesión de variables aleatorias independientes, con esperanza matemática nula. Supongamos que se verifica

$$\sum_{k=1}^{n} \mathbf{E}\,|X_k|^3 \leq Bn, \qquad \sum_{k=1}^{n} \mathbf{E}\,|X_k|^2 \geq An$$

para todo n, donde A y B son constantes positivas. ¿Es aplicable el teorema central del límite a esta sucesión?

11. Sea $\{X_n\}$ una sucesión de variables aleatorias independientes, con esperanza matemática nula. Supongamos que se verifica

$$\sum_{k=1}^{n} \mathbf{E}\,|X_k|^2 \Big| \ln |X_k| \Big|^{1+\delta} \leq Bn, \qquad \sum_{k=1}^{n} \mathbf{E}\,|X_k|^2 \geq An,$$

para todo n y algún $\delta > 0$, donde A y B son constantes positivas. Demostrar que para esta sucesión es aplicable el teorema central del límite. (Sugerencia: verificar la condición de Lindeberg).

12. Consideremos una sucesión $\{X_n\}$ de variables aleatorias independientes, idénticamente distribuidas, con distribución de Poisson con parámetro 1. Hallar $\lim_{n\to\infty} \mathbf{P}\left(\frac{1}{\sqrt{n}}\left(\sum_{k=1}^n X_k - n\right) \le x\right)$.

13. Sea $\{X_n\}$ una sucesión de variables aleatorias independientes, tal que que para cada $n = 1, 2, \dots$, la variable aleatoria X_n tiene distribución normal, con $\mathbf{E}\, X_n = 0$, y $\mathbf{var}\, X_n = 2^{2n}$. (a)¿Es aplicable el teorema central del límite a esta sucesión? (b)¿Se cumple la condición de Lindeberg para esta sucesión de variables aleatorias?

14. Sea $\Phi(x)$ la distribución normal estándar. Demostrar la desigualdad

$$1 - \Phi(x) < \frac{1}{x\sqrt{2\pi}} e^{-x^2/2},$$

para todo $x > 0$.

15. Demostrar la fórmula

$$1 - \Phi(x) = \frac{1}{x\sqrt{2\pi}} e^{-x^2/2}\left(1 + O\left(\frac{1}{x^2}\right)\right) \ (x \to \infty). \tag{7.18}$$

(Sugerencia: Utilizar la identidad $\int_x^\infty e^{-t^2/2} dt = \int_x^\infty \frac{1}{t} d\left(e^{-t^2/2}\right)$ e integrar por partes.)

16. Sea $\{X_n\}$ una sucesión de variables aleatorias independientes, idénticamente distribuidas, con esperanza nula y varianza $\sigma^2 > 0$. Designemos $F_n(x) = \mathbf{P}\left(\frac{1}{\sigma\sqrt{n}} \sum_{k=1}^n X_k \le x\right)$. Demostrar que si se verifica

$$|F_n(x) - \Phi(x)| \le \frac{C}{\sqrt{n}}$$

para todo x real, y todo n natural, donde C es una constante positiva, entonces, para $0 < \varepsilon < 1$, se verifica

$$\frac{1 - F_n(x)}{1 - \Phi(x)} \to 1 \quad (n \to \infty),$$

uniformemente, en el conjunto $0 \le x \le (1 - \varepsilon)\sqrt{\ln n}$. (Sugerencia: utilizar la fórmula (7.18).)