Teorema Central del Límite

Aronna María Soledad,

del Barco Viviana Jorgelina,

Tolomei Paola Beatriz

y

Vittone Francisco

Departamento de Matemática, Escuela de Ciencias Exactas y Naturales

Facultad de Ciencias Exactas, Ingeniería y Agrimensura, Universidad Nacional de Rosario

Disciplina: Probabilidad y Estadística

31 de Julio de 2004

Secretaría de Ciencia y Técnica

Facultad de Ciencias Exactas, Ingeniería y Agrimensura

Universidad Nacional de Rosario

Av. Pellegrini 250 - 2000 Rosario - Argentina

http://www.fceia.unr.edu.ar/secyt

Este documento es publicado por la FCEIA para su consulta externa. El mismo se publica como Reporte de Investigación para divulgación de las tareas científicas que se desarrollan en la FCEIA, Universidad Nacional de Rosario. Los autores conservan los derechos de autoría y copia de la totalidad de su trabajo aquí publicado. Luego de su posterior eventual publicación externa a la FCEIA, los requerimientos deberán dirigirse a los autores respectivos. El contenido de este reporte refleja la visión de los autores, quienes se responsabilizan por los datos presentados, los cuales no necesariamente reflejan la visión de la SeCyT-FCEIA. Tanto la SeCyT-FCEIA como los autores del presente reporte no se responsabilizan por el uso que pudiera hacerse de la información y/o metodologías publicadas.

Cualquier sugerencia dirigirla a: rtsecyt@fceia.unr.edu.ar

Teorema Central del Límite.

María Soledad Aronna [1],

Viviana Jorgelina del Barco[2],

Paola Beatriz Tolomei [3],

Francisco Vittone[4].

Departamento de Matemática, Escuela de Ciencias Exactas y Naturales

Facultad de Ciencias Exactas, Ingeniería y Agrimensura,

Universidad Nacional de Rosario

[1]aronna@fceia.unr.edu.ar
[2]delbarc@fceia.unr.edu.ar
[3]ptolomei@fceia.unr.edu.ar
[4]vittone@fceia.unr.edu.ar

Resumen. Este trabajo fue presentado en el concurso de monografías de la Unión Matemática Argentina, en homenaje al Dr. Jorge D. Samur en el 2004. El mismo resultó premiado. El tema propuesto fue el Teorema Central del Límite. En el escrito comenzamos con un pequeño resumen de la historia que lleva este teorema. Exponemos la demostración que dio William Feller del mismo y finalizamos el trabajo mostrando algunas aplicaciones conocidas y otras no tanto.

Palabras claves: Suma de variables aleatorias, convergencia en distribución.

Índice general

1. Introducción

En esta monografía nos dedicaremos al estudio del Teorema Central del Límite. Comenzamos en el capítulo 2 presentando una breve reseña histórica. En el capítulo 3 incluimos algunos resultados preliminares que consideramos necesarios para una mejor comprensión de la demostración del Teorema. Se desarrollan algunos aspectos de la Integral de Lebesgue - Stieltjes y de las funciones características. La demostración completa se encuentra en el Capítulo 4. Vamos a exhibir la realizada por William Feller en [**1**] que se basa en parte en la realizada por Paul Levy. Por último, en el Capítulo 5, mostraremos algunas de las aplicaciones que se le dieron al Teorema.

En el desarrollo de este trabajo vamos a considerar conocidas las definiciones y propiedades básicas de la probabilidad. También obviaremos algunos pasos en las demostraciones que surgen exclusivamente del trabajo algebraico de ciertas expresiones ya que nos alejan del interés central de esta monografía.

2. Breve reseña histórica

La Teoría de la Probabilidad surgió en el siglo XVII en Francia, cuando los reconocidos matemáticos Pierre de Fermat y Blaise Pascal comenzaron a interesarse en los juegos de azar, a raíz de lo cual muchos matemáticos de la época empezaron a introducirse en el tema, entre ellos Abraham De Moivre, Chistiaan Huygens y Jacob Bernoulli. Así fue como se fundó definitivamente lo que hoy se conoce como Teoría de la Probabilidad.

Como en toda área de la matemática, era necesario establecer una estructura adecuada para la investigación de esas "experiencias aleatorias". Para tal fin se crearon las llamadas variables aleatorias, objetos fundamentales dentro de la Teoría de la Probabilidad. Éstas representan resultados numéricos obtenidos a partir de la realización de un experimento; y matemáticamente se modelizan como una función a valores reales definida sobre el conjunto de posibles resultados del experimento (espacio muestral). Mientras se fue avanzando en el estudio de estos sucesos azarosos, resultó cada vez más frecuente la necesidad de relacionar muchas variables para poder analizarlos. Surgió naturalmente de esta manera el problema de estudiar la distribución de la suma de una infinidad de variables aleatorias.

Los dos resultados fundamentales vinculados a este problema son la Ley de los Grandes Números y el Teorema Central del Límite.

Este último establece que la distribución de la suma de una gran cantidad de variables aleatorias independientes, bajo ciertas condiciones adicionales, se aproxima a una distribución Normal. Más adelante estableceremos cuáles son estas condiciones. Por ahora, nos dedicaremos a hacer un repaso de cómo evolucionó históricamente este teorema.

Empecemos por aclarar que la denominación Teorema Central del Límite, es relativamente reciente. Fue utilizada por primera vez en 1920 por George Polya en uno de sus artículos [**2**]. El término "central" significa "fundamental" o de "importancia central".

Su primera aparición, en su versión más sencilla, se produjo en 1718, cuando De Moivre publicó y demostró en su libro "The Doctrine of Chances" [**3**] la aproximación de la distribución binomial $B(n, p)$ simétrica ($p = 1/2$) por lo que hoy conocemos como la distribución normal $N(np, \sqrt{np(1-p)})$ para valores grandes de n. En la forma general actual del Teorema Central del Límite, el resultado de De Moivre indica que es posible aproximar por una distribución normal, la distribución de la suma de n variables aleatorias de Bernoulli independientes e idénticamente distribuídas con $p = 1/2$ (ya que, en efecto, esta suma tiene la distribución de la binomial antes mencionada). En realidad, De Moivre notó que cuando el número de repeticiones de una experiencia (que consistía en lanzar una moneda) aumentaba, la forma de la correspondiente distribución binomial se aproximaba a una curva suave. Entonces razonó que si podía encontrar una expresión matemática para esta curva, podría resolver problemas (como por ejemplo, calcular la probabilidad de obtener 90 caras en 150 lanzamientos) de manera mucho más fácil. Esto es lo que hizo y esta curva es la que hoy llamamos curva normal. Su importancia, y por ende, la de la distribución normal, recae principalmente en que muchos fenómenos naturales tienen, al menos, una distribución aproximadamente normal. En 1809 Gauss desarrolló la fórmula de la distribución normal y mostró que la distribución de los errores cometidos en las observaciones astronómicas, se adaptaba perfectamente a ella.

Más tarde, alrededor de 1810, Pierre Simon, marqués de Laplace, mostró que bajo condiciones casi siempre alcanzadas en la práctica cualquier suma de un número considerable de variables aleatorias mutuamente independientes e idénticamente distribuidas puede aproximarse por una normal. De esta manera, Laplace generalizó el resultado de De Moivre. En su "Théorie analytique des probabilités" [**4**]. que publicó en 1812, dejó enunciadas las bases del teorema, ya que, aunque nunca lo formuló en su forma más general, se deduce del mismo método utilizado en cada una de las aplicaciones. Además, logró demostrarlo para distribuciones discretas cualesquiera y para ciertas distribuciones continuas, pero no realizó ninguna prueba rigurosa que lo hiciera válido para distribuciones arbitrarias.

Los pasos siguientes hacia la formalización del teorema los dio Siméon Denis Poisson en 1824, quien mejoró y generalizó la demostración de Laplace. Poisson realizó la prueba para variables independientes e idénticamente distribuidas, primero para una suma y luego para una combinación lineal. Lo siguieron Johann Dirichlet, Friedrich Bessel y Louis Cauchy, quienes dieron distintas pruebas para los resultados ya establecidos. Por su parte, Cauchy lo conectó con una forma más abstracta y con una perspectiva más moderna de la teoría de errores, dándole así más relevancia en el campo matemático.

Sin embargo, las demostraciones realizadas por todos estos matemáticos, si bien importantes, eran insatisfactorias en tres aspectos. En primer lugar, faltaba demostrar el teorema para distribuciones arbitrarias, o sea, el caso más general. Segundo, se deseaba determinar condiciones necesarias y suficientes generales para las cuales fuera posible aproximar por una normal. Y tercero, determinar el orden de convergencia de la suma, es decir, cuántas variables debían sumarse para obtener una aproximación considerablemente buena.

Estos interrogantes fueron resueltos por matemáticos rusos entre 1870 y 1910. Los más importantes fueron Pafnuty Chebyshev, Andrei Markov y Aleksandr Mikhailovich Lyapunov. Chebyshev y Markov intentaron demostrarlo utilizando el denominado método de los momentos. Sin embargo, fue necesario esperar hasta 1901 cuando Lyapunov, alumno de Chebyshev y compañero de Markov, dio la primera demostración completa mediante funciones características, que sólo establecía condiciones suficientes. Éstas se conocen como Condiciones de Lyapunov.

Este matemático, además de la demostración del teorema, en 1923 propuso una cota superior del error cometido al substituir una determinada distribución por una distribución normal (el tercero de los problemas por resolver).

Finalmente, en 1922, J. W. Lindeberg estableció una condición suficiente para la cual es válida la aproximación de la suma de variables aleatorias independientes e idénticamente distribuidas por una variable aleatoria normal. La condición de Lindeberg es más general que las condiciones de Lyapunov, y la enunciaremos en capítulos posteriores (ver [**5, 6, 1**] para profundizar en la relación entre ambas condiciones). De cualquier manera, la demostración de Lindeberg del teorema, a pesar de ser directa, es muy engorrosa. Se simplifica muchísimo mediante el uso de funciones características, idea que debemos a Paul Levy.

Nuestro recorrido por la historia del Teorema Central del Límite termina al llegar al año 1937 cuando William Feller demostró las condiciones necesarias para su cumplimiento. Así se resolvieron los tres problemas que habían quedado pendientes hacia fines del siglo XIX.

3. Preliminares

3.1. Observaciones a la Integral de Lebesgue - Stieltjes. Durante el desarrollo del trabajo utilizaremos la integral conocida como integral de Lebesgue - Stieltjes, la cual se nota $\int_M g(x)F\{dx\}$. Esta integral es una extensión de la integral de Riemann que permite integrar funciones más generales y sobre dominios más generales que las estudiadas en el cálculo elemental.

No desarrollaremos este concepto pues nos alejaríamos de nuestro objetivo central, simplemente aclararemos que esta integral verifica las propiedades de la integral de Riemann. Para buscar más información acerca de este tema ver [**7, 6**].

En nuestro caso F será una función de distribución y vale realizar las siguientes observaciones:

Observación 1. *Cuando F es la distribución de una variable discreta X, la integral se reduce a una suma.*

Si $P(X = x_i) = p(x_i) > 0 \ \forall i \in I$ entonces

$$\int_a^b g(x)F\{dx\} = \sum_{i/a \leq x_i \leq b} g(x_i)p(x_i).$$

Una explicación intuitiva de esta propiedad resulta de interpretar al diferencial $F\{dx\}$ como $p(x_i)$ en los puntos x_i del recorrido de X y cero en el resto de $\mathbb{R}$.

Observación 2. *Cuando F es la función de distribución de una variable aleatoria continua con función densidad f, entonces f es la derivada de F en todos los puntos donde F sea derivable, esto es en casi todo punto de $\mathbb{R}$. Luego la integral $\int_a^b g(x)F\{dx\}$ se convierte en la usual integral de Riemann. Es decir,*

$$\int_a^b g(x)F\{dx\} = \int_a^b g(x)f(x)dx.$$

Observación 3. *Sea F una función de distribución cualquiera. Es sabido que F se descompone en partes discreta, absolutamente continua y singular en el siguiente sentido*

$$F = F_d + F_{ac} + F_s.$$

Luego, tenemos por linealidad que

$$\int_a^b g(x)F\{dx\} = \int_a^b g(x)F_d\{dx\} + \int_a^b g(x)F_{ac}\{dx\} + \int_a^b g(x)F_s\{dx\}.$$

3.2. Funciones características. A continuación desarrollaremos el concepto de función característica y algunas de sus propiedades que serán de utilidad para llevar a cabo la demostración del Teorema.

Definición 4. *Sea X una variable aleatoria con distribución de probabilidad F. La función característica de F (o de X) es la función φ definida para ζ real por*

$$\varphi(\zeta) = \int_{-\infty}^{+\infty} e^{i\zeta x} F\{dx\} = u(\zeta) + iv(\zeta),$$

donde

$$u(\zeta) = \int_{-\infty}^{+\infty} \cos(\zeta x) F\{dx\} \quad y \quad v(\zeta) = \int_{-\infty}^{+\infty} \sin(\zeta x) F\{dx\}$$

Desde luego, si F es una distribución continua con una densidad f, tendremos

$$\varphi(\zeta) = \int_{-\infty}^{+\infty} e^{i\zeta x} f(x) dx$$

Es claro que dada una variable aleatoria X, su función característica φ representa la esperanza de la variable $e^{i\zeta X}$, es decir, $\varphi(\zeta) = E(e^{i\zeta X})$.

De aquí vemos que si X_1 y X_2 son variables aleatorias independientes con funciones caraterísticas φ_1 y φ_2, la función caraterística φ de $X_1 + X_2$ resulta

$$\varphi(\zeta) = E(e^{i\zeta(X_1+X_2)}) = E(e^{i\zeta X_1} e^{i\zeta X_2}) = E(e^{i\zeta X_1})E(e^{i\zeta X_2}) = \varphi_1(\zeta)\varphi_2(\zeta),$$

ya que $e^{i\zeta X_1}$ y $e^{i\zeta X_2}$ son también independientes.

Observación 5. *Dado que la función densidad de una variable con distribución normal estandarizada es $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$, tenemos que su función característica es:*

$$\varphi(\zeta) = e^{-\frac{1}{2}\zeta^2}.$$

El siguiente teorema (ver [**8**]) asegura que cada función característica corresponde a una única función de distribución, hecho que juega un papel principal en la demostración del Teorema Central del Límite que aquí daremos.

Teorema 6. *Dada la función característica φ de una variable aleatoria X, la función de distribución F queda determinada de la siguiente manera:*

$$F(b)-F(a)=\lim_{U\to\infty}\frac{1}{2\pi}\int_{-U}^{U}\frac{e^{-iua}-e^{-iub}}{iu}\varphi(u)du,$$

con $a<b$ puntos de continuidad de F.

Esta expresión se conoce como Fórmula de Inversión.

Definición 7. *Decimos que F es normalizada si los valores de F en sus puntos de discontinuidad x se toman como $\frac{F(x-0)+F(x+0)}{2}$, donde $F(x-0)=\lim_{t\to x^-}F(t)$ y $F(x+0)=\lim_{t\to x^+}F(t)$*

Observación 8. *La normalización destruye la continuidad por izquierda de F en sus puntos de discontinuidad. Sin embargo la función de distribución normalizada determina a la original.*

Observación 9. *Si F es normalizada la fórmula de inversión se cumple para todos los $a<b$ $\in\mathbb{R}$.*

Observación 10. *En la integral que figura en el miembro de la derecha de la fórmula de inversión, el integrando es continuo en todo $\mathbb{R}$. En efecto, el mismo puede definirse en $u=0$ por continuidad, teniendo en cuenta que $\lim_{u\to 0}\frac{e^{-iua}-e^{-iub}}{iu}=b-a$. El integrando también está acotado sobre $\mathbb{R}$ por su valor $(b-a)\,\varphi(0)$. Consecuentemente, para cada U real finita esta integral es una integral de Riemann ordinaria, y al demostrar la fórmula de inversión hallaremos que existe el límite de esta integral cuando $U\to\infty$.*

Demostración. Sea

$$I_U=\frac{1}{2\pi}\int_{-U}^{U}\frac{e^{-iua}-e^{-iub}}{iu}\varphi(u)du,\ \text{con } a<b\ \in\mathbb{R}.$$

Sustituyendo en I_U a $\varphi(u)=\int_{-\infty}^{\infty}e^{iux}F\{dx\}$ y realizando adecuados cambios de variables, combinadas con operaciones elementales del cálculo obtenemos:

$$I_U=\int_{-\infty}^{+\infty}J_U(x)F\{dx\},\qquad \text{donde}\qquad J_U(x)=\frac{1}{\pi}\int_{U(x-b)}^{U(x-a)}\frac{sen\,(v)}{v}dv$$

Como J_U está acotada uniformemente en U, podemos intercambiar los extremos de integración y el límite con $U\to\infty$, en

$$\lim_{U\to+\infty}I_U=\lim_{U\to+\infty}\int_{-\infty}^{+\infty}J_U(x)F\{dx\}=\int_{-\infty}^{+\infty}J(x)F\{dx\}$$

donde para calcular $\lim_{U\to\infty} J_U(x) = J(x)$ utilizamos la clásica fórmula de Dirichlet $\int_{-\infty}^{\infty} \frac{sen(v)}{v} dv = \pi$. Este límite dependerá de dónde esté ubicado x con respecto a a y b de la siguiente forma:

$$J(x) = \begin{cases} 1 & \text{si } a < x < b \\ \frac{1}{2} & \text{si } x = a,\ x = b \\ 0 & \text{si } x < a,\ x > b, \end{cases}$$

y luego obtenemos así

$$\begin{aligned} \lim_{U\to\infty} I_U(x) &= \frac{1}{2}(F(a+0) - F(a-0)) \\ &+ (F(b-0) - F(a+0)) + \frac{1}{2}(F(b+0) - F(b-0)) \\ &= \frac{F(b-0) + F(b+0)}{2} - \frac{F(a-0) + F(a+0)}{2} \end{aligned}$$

Por consiguiente, si F es normalizada o si $a < b$ son puntos de continuidad de F, $\lim_{U\to\infty} I_u = F(b) - F(a)$, y queda demostrada la fórmula de inversión. □

Definición 11. *Sean F una función de distribución, definimos $m_n = \int_{-\infty}^{+\infty} x^n F\{dx\}$ y $M_n = \int_{-\infty}^{+\infty} |x|^n F\{dx\}$. Los valores de estas integrales son denominados momentos y momentos absolutos de F respectivamente.*

Lema 12. *Si $M_n < \infty$ entonces la n-ésima derivada de la función característica φ de una variable X es una función continua dada por*

$$\varphi^{(n)}(\zeta) = i^n \int_{-\infty}^{+\infty} e^{i\zeta x} x^n F\{dx\}.$$

Lema 13. *Si $m_n < \infty$, entonces*

$$\varphi'(0) = im_1 \quad y \quad \varphi''(0) = -m_2.$$

4. El Teorema

Dada una sucesión de variables aleatorias independientes $\{X_k\}$ con $k = 0, 1, 2, ..$ tales que

$$E(X_k) = 0 \quad E(X_k^2) = V(X_k) = \sigma_k^2,$$

denotamos mediante F_k la función de distribución de X_k y mediante φ_k su función característica. Llamamos $S_n = X_1 + X_2 + \ldots + X_n$ a la n-ésima suma parcial y $s_n^2 = V(S_n)$. Por la independencia de las variables aleatorias consideradas se tiene que $s_n^2 = \sigma_1^2 + \ldots + \sigma_n^2$.

Definición 14. *Se dice que una sucesión de variables aleatorias satisface la condición de Lindeberg cuando*

$$\frac{1}{s_n^2} \sum_{k=1}^{n} \int_{|x|>ts_n} x^2 F_k\{dx\} \to 0, \quad \text{si } n \to \infty \text{ para cada } t > 0 \text{ fijo.}$$

Si observamos que

$$\begin{aligned}
\frac{\sigma_k^2}{s_n^2} &= \frac{1}{s_n^2} \int_{|x|\leq\ ts_n} x^2 F_k\{dx\} + \frac{1}{s_n^2} \int_{|x|>ts_n} x^2 F_k\{dx\} \leq \\
&\leq \frac{1}{s_n^2} \int_{|x|\leq\ ts_n} t^2 s_n^2 F_k\{dx\} + \frac{1}{s_n^2} \sum_{j=1}^{n} \int_{|x|>ts_n} x^2 F_j\{dx\} \leq \\
&\leq \frac{1}{s_n^2} \int_{-\infty}^{\infty} t^2 s_n^2 F_k\{dx\} + \frac{1}{s_n^2} \sum_{j=1}^{n} \int_{|x|>ts_n} x^2 F_j\{dx\} = \\
&= t^2 + \frac{1}{s_n^2} \sum_{j=1}^{n} \int_{|x|>ts_n} x^2 F_j\{dx\},
\end{aligned}$$

donde t es independiente de k. Deducimos entonces que si la condición de Lindeberg se verifica en la sucesión $\{X_k\}$, la última suma del segundo miembro tiende a cero y por lo tanto se cumple que

$$\text{(A)} \qquad \frac{\sigma_k^2}{s_n^2} \leq \varepsilon \quad \text{si } n \text{ es suficientemente grande.}$$

También podemos observar que $s_n \to \infty$ si $n \to \infty$.

Teorema 15. *Suficiencia de la condición de Lindeberg: Si la condición de Lindeberg se satisface, la distribución de las sumas normalizadas S_n/s_n tiende a la distribución normal $N(0,1)$.*

Demostración. Vamos a probar que la función característica de S_n/s_n tiende a la función característica de $N(0,1)$ cuando $n \to \infty$. Teniendo esto y mediante el teorema 6, resulta que las funciones de distribución correspondientes a S_n/s_n convergen a la función de distribución de $N(0,1)$.

Para llegar a esto tomaremos $\zeta > 0$ arbitraria pero fija y probaremos que

$$\varphi_1(\zeta/s_n)\dots\varphi_n(\zeta/s_n) \to e^{-\frac{1}{2}\zeta^2}.$$

De los lemas 12 y 13 anterior surge que $\varphi_k'(0) = 0$ y $|\varphi_k''(x)| \leq \sigma_k^2$ para toda x donde φ sea derivable. Tomando el desarrollo de Taylor hasta el segundo término

$$\begin{aligned}\varphi_k(\zeta/s_n) &= \varphi_k(0) + \varphi_k'(0)\frac{\zeta}{s_n} + \varphi_k''(t)\frac{\zeta^2}{2!s_n^2} \\ &= \varphi_k(0) + \varphi_k''(t)\frac{\zeta^2}{2!s_n^2}, \quad 0 < t < \zeta.\end{aligned}$$

Si n es suficientemente grande y por la propiedad (A), resulta

$$|\varphi_k(\zeta/s_n) - 1| \leq \frac{\sigma_k^2\zeta^2}{2s_n^2} \leq \varepsilon\zeta^2.$$

Dado que $|\varphi_k(\zeta/s_n) - 1| \leq \varepsilon\zeta^2$, se cumple que

$$\left| e^{\left(\sum_{k=1}^n (\varphi_k - 1)\right)} - \varphi_1 \dots \varphi_n \right| \to 0. \tag{1}$$

En efecto,

$$\begin{aligned}|e^{\sum_{k=1}^n(\varphi_k-1)} - \varphi_1 ... \varphi_n| &\leq \sum_{k=1}^n |e^{\varphi_k(\zeta/s_n)-1} - \varphi_k(\zeta/s_n)| \leq \\ &\leq \delta \sum_{k=1}^n |\varphi_k(\zeta/s_n) - 1| \leq \delta\frac{\zeta^2}{s_n^2}\sum_{k=1}^n \sigma_k^2 = \delta\zeta^2,\end{aligned}$$

donde la primera desigualdad se mantiene por la propiedad de los números complejos que dice que si $|a_k| \leq 1$ y $|b_k| \leq 1$ entonces $|a_1 \dots a_k - b_1 \dots b_k| \leq \sum_{k=1}^n |a_k - b_k|$; mientras que la segunda desigualdad se debe a que para cualquier $\delta > 0$ se tiene $|e^z - 1 - z| < \delta|z|$ si $|z|$ es suficientemente pequeño. Dado que δ es arbitraria concluimos que el primer miembro tiende a cero, con lo cual queda probada la validez del límite (1).

Además, por la continuidad de la función exponencial tenemos que $|e^{\sum_{k=1}^{n}(\varphi_k-1)} - e^{-\frac{1}{2}\zeta^2}| \to 0$ si y sólo si $|\sum_{k=1}^{n}(\varphi_k - 1) + \frac{1}{2}\zeta^2| \to 0$.

Por otro lado vamos a probar que $|\sum_{k=1}^{n}(\varphi_k - 1) + \frac{1}{2}\zeta^2|$ es quivalente a

$$\sum_{k=1}^{n}\int_{-\infty}^{\infty}\left[e^{ix\zeta/s_n} - 1 - \frac{ix\zeta}{s_n} + \frac{x^2\zeta^2}{2s_n^2}\right]F_k\{dx\}. \tag{2}$$

En efecto,

$$\begin{aligned}
&\sum_{k=1}^{n}\int_{-\infty}^{\infty}\left[-1 - \frac{ix\zeta}{s_n} + \frac{x^2\zeta^2}{2s_n^2}\right]F_k\{dx\} = \\
&\quad = \sum_{k=1}^{n}\left[-\int_{-\infty}^{\infty}F_k\{dx\} - \frac{i\zeta}{s_n}\int_{-\infty}^{\infty}xF_k\{dx\} + \frac{\zeta^2}{2s_n^2}\int_{-\infty}^{\infty}x^2F_k\{dx\}\right] \\
&\quad = \sum_{k=1}^{n}\left[-1 - \frac{i\zeta}{s_n}E(X_k) + \frac{\zeta^2}{2s_n^2}\sigma_k^2\right] \\
&\quad = \sum_{k=1}^{n} -1 + \frac{\zeta^2}{2s_n^2}\sum_{k=1}^{n}\sigma_k^2 \\
&\quad = \sum_{k=1}^{n}\left[-1 + \frac{\zeta^2}{2}\right].
\end{aligned}$$

Finalmente, sumando a ambos miembros $\sum_{k=1}^{n}\varphi_k(\zeta/s_n) = \sum_{k=1}^{n}\int_{-\infty}^{\infty}e^{ix\zeta/s_n}F_k\{dx\}$ nos queda probada la equivalencia.

Tomemos ahora el desarrollo de Taylor de la función $e^{ix\zeta/s_n}$ según el siguiente criterio:

$|e^{ix\zeta/s_n} - 1 - \frac{i\zeta x}{s_n}| = |\theta_1(\zeta/s_n)\frac{\zeta^2x^2}{2s_n^2}\ | \leq |\frac{\zeta^2x^2}{2s_n^2}|$ con $|\theta_1(\zeta/s_n)| \leq 1$, si $|x| > ts_n$,

$|e^{ix\zeta/s_n} - 1 - \ \frac{i\zeta x}{s_n} + \frac{\zeta^2x^2}{2s_n^2}| = |\theta_2(\zeta/s_n)\frac{\zeta^3x^3}{6s_n^3}| \leq |\frac{\zeta^3x^3}{s_n^3}|$ con $|\theta_2(\zeta/s_n)| \leq 1$ si $|x| \leq ts_n$,

y utilizamos las acotaciones anteriores en la ecuación (2):

$$\left|\sum_{k=1}^{n}\int_{-\infty}^{\infty}\left[e^{ix\zeta/s_n}-1-\frac{ix\zeta}{s_n}+\frac{x^2\zeta^2}{2s_n^2}\right]F_k\{dx\}\right|$$

$$\leq\sum_{k=1}^{n}\int_{|x|\leq ts_n}\left|\frac{x^3\zeta^3}{s_n^3}\right|F_k\{dx\}+\sum_{k=1}^{n}\int_{|x|>ts_n}\left|\frac{\zeta^2x^2}{s_n^2}\right|F_k\{dx\}$$

$$\leq\frac{t\zeta^3}{s_n^2}\sum_{k=1}^{n}\int_{|x|\leq ts_n}x^2F_k\{dx\}+\frac{\zeta^2}{s_n^2}\sum_{k=1}^{n}\int_{|x|>ts_n}x^2F_k\{dx\}$$

$$\leq t\zeta^3+\frac{\zeta^2}{s_n^2}\sum_{k=1}^{n}\int_{|x|>ts_n}x^2F_k\{dx\},$$

donde por la condición de Lindeberg el segundo término de

$$t\zeta^3+\frac{\zeta^2}{s_n^2}\sum_{k=1}^{n}\int_{|x|>ts_n}x^2F_k\{dx\}$$

tiende a cero. Ya que t se puede elegir tan pequeña como se desee, (2) converge a cero, con lo que $\sum_{k=1}^{n}(\varphi_k(\zeta/s_n)-1)+\frac{1}{2}\zeta^2$ también lo hace ya que probamos su equivalencia. Por consiguiente, $|e^{\sum_{k=1}^{n}(\varphi_k-1)}-e^{-\frac{1}{2}\zeta^2}|\to 0$. Pero teniendo (1), llegamos a que $|e^{-\frac{1}{2}\zeta^2}-\varphi_1...\varphi_n|\to 0$ cuando n tiene a $+\infty$ que es la tesis del teorema. □

TEOREMA 16. *Necesidad de la condición de Lindeberg: Supongamos que $s_n\to\infty$ y $\sigma_n/s_n\to 0$. Entonces la condición de Lindeberg es necesaria para la convergencia de la distribución de S_n/s_n a la distribución Normal.*

DEMOSTRACIÓN. Puesto que por hipótesis $\sigma_n/s_n\to 0$ cuando n tiende a infinito, dado $\epsilon>0$ existe v_ϵ tal que $\sigma_k/s_k<\epsilon$ para todo $k\geq\ v_\epsilon$. Luego, si k es tal que $v<k\leq n$ para un n fijo, se tiene que $\sigma_k/s_n\leq\sigma_k/s_k<\varepsilon$. Ahora, fijando k resulta que $\frac{\sigma_k}{s_n}\to 0$ pues $s_n\to+\infty$ cuando $n\to+\infty$.

Ahora, utilizando la hipótesis de que la ditribución de S_n/s_n tiende a la de $N(0,1)$ tenemos que (2) tiende a cero ya que habíamos probado su equivalencia en el teorema anterior. Luego

$$\sum_{k=1}^{n}\int_{-\infty}^{\infty}\left[e^{ix\zeta/s_n}-1-\frac{ix\zeta}{s_n}+\frac{x^2\zeta^2}{2s_n^2}\right]F_k\{dx\}\to 0.$$

Usando el hecho de que $\cos z-1+\frac{1}{2}z^2\geq 0$, la parte real del integrando dada por $\cos(\frac{x\zeta}{s_n})-1+\frac{x^2\zeta^2}{2s_n^2}$ es no negativa. De este modo, se ve que

$$\left|\sum_{k=1}^{n}\int_{-\infty}^{\infty}\left[e^{ix\zeta/s_n}-1-\frac{ix\zeta}{s_n}+\frac{x^2\zeta^2}{2s_n^2}\right]F_k\{dx\}\right|$$

$$\geq \operatorname{Re}\left(\sum_{k=1}^{n}\int_{-\infty}^{\infty}\left[e^{ix\zeta/s_n}-1-\frac{ix\zeta}{s_n}+\frac{x^2\zeta^2}{2s_n^2}\right]F_k\{dx\}\right)$$

$$=\sum_{k=1}^{n}\int_{-\infty}^{\infty}\left[cos\left(\frac{x\zeta}{s_n}\right)-1+\frac{x^2\zeta^2}{2s_n}\right]$$

$$\geq\sum_{k=1}^{n}\int_{|x|>ts_n}\left(\frac{x^2\zeta^2}{2s_n^2}-2\right)F_k\{dx\}$$

$$\geq\sum_{k=1}^{n}\frac{\zeta^2}{2s_n^2}\int_{|x|>ts_n}x^2F_k\{dx\}+$$

$$+\sum_{k=1}^{n}\frac{2}{t^2s_n^2}\int_{|x|>ts_n}x^2F_k\{dx\}$$

$$=\left(\frac{1}{2}\zeta^2-2t^{-2}\right)\frac{1}{s_n^2}\sum_{k=1}^{n}\int_{|x|>ts_n}x^2F_k\{dx\}.$$

Para justificar la segunda desigualdad, debemos tener en cuenta que $|\cos z-1+\frac{z^2}{2}|\geq|-1-1+\frac{z^2}{2}|=|\frac{z^2}{2}-2|$ y en la penúltima de ellas teniendo en cuenta la región de integración tenemos que $\frac{x^2}{t^2s_n^2}>1$ y por lo tanto $-2\int_{|x|>ts_n}F_k\{dx\}\geq\frac{-2}{t^2s_n^2}\int_{|x|>ts_n}x^2F_k\{dx\}$.

Observando la expresión

$$\text{(B)}\qquad\left(\frac{1}{2}\zeta^2-2t^{-2}\right)\frac{1}{s_n^2}\sum_{k=1}^{n}\int_{|x|>ts_n}x^2F_k\{dx\},$$

el último miembro de las desigualdades anteriores , y al ser ζ y t arbitrarias, las podemos elegir de manera que dicha expresión sea no negativa. Luego haciendo tender

$$\left|\sum_{k=1}^{n}\int_{-\infty}^{\infty}\left[e^{ix\zeta/s_n}-1-\frac{ix\zeta}{s_n}+\frac{x^2\zeta^2}{2s_n^2}\right]F_k\{dx\}\right|$$

a cero tendremos que (B) converge a cero para todas t y ζ. Por lo tanto, $\frac{1}{s_n^2}\sum_{k=1}^{n}\int_{|x|>ts_n}x^2F_k\{dx\}\to 0$ y en consecuencia la sucesión $X_1,X_2,\ldots,X_n,\ldots$ satisface la condición de Lindeberg. $\square$

5. Aplicaciones

A continuación vamos a mostrar algunas aplicaciones del Teorema que nos parecieron interesantes (ver [**9**]).

5.1. Aproximación normal a la binomial. Consideramos X_i variables aleatorias independientes de Bernoulli, todas con probabilidad de éxito p. Luego $S_n = \sum_{k=1}^{n} X_k$ resulta una $Bi(n,p)$.

El Teorema Central del Límite implica que la distribución de $S_n \sim Bi(n,p)$ se puede aproximar con $N(np, np(1-p))$ para n suficientemente grande.

5.2. Movimiento browniano. Si observáramos al microscopio partículas de algún líquido, veríamos los movimientos caóticos e incesantes que éstas realizan. Este fenómeno fue descubierto por el botánico Robert Brown y explicado, como consecuencia de la agitación de las moléculas del líquido, por Albert Einstein. Daremos a continuación un enfoque simple del problema, pero sólo para el caso unidimensional.

Sea la variable X_t la posición de una partícula en el instante t, suponiendo $X_0 = 0$. Hacemos dos suposiciones adicionales:

a) La partícula no tiene inercia.

b) Las condiciones no cambian en el tiempo.

La primera suposición la representamos postulando que los incrementos de X_t son independientes, esto es

$t_1 < t_2 < \ldots < t_n \Longrightarrow (X_{t_2} - X_{t_1}), ..., (X_{t_n} - X_{t_{n-1}})$ son independientes.

La segunda, postulando que $X_{t+s} - X_s = X_t \ \forall s, t$. Esto significa que los incrementos son estacionarios.

Para aproximar la ditribución de X_t, tomaremos un "paseo al azar" de una partícula cualquiera. Denominaremos δ al tiempo transcurrido entre impactos recibidos por las moléculas del líquido, y ε a la distancia que se desplaza cada partícula con cada impacto, a la derecha o a la izquierda, con probabilidad 1/2 cada una. Consideraremos lo sucedido en un impacto independiente de lo ocurrido en los demás.

Sea $\{Z_i\}$ la sucesión de variables independientes que valen ± 1 dependiendo de si el movimiento es hacia la izquierda o hacia la derecha, con probabilidad 0,5. Entonces podemos expresar $X_t = \varepsilon \sum_{i=1}^{n} Z_i$, donde $n = [t/\delta]$ (donde "[.]" representa la función parte entera). En consecuencia $E(X_t) = 0$ y $V(X_t) = [t/\delta]\varepsilon^2$. Nos es natural considerar que los impactos son muy frecuentes y los desplazamientos muy pequeños, por lo que hacemos $\delta \to 0$ y $\varepsilon \to 0$. Llamando $c = lim(\varepsilon^2/\delta)$, el Teorema Central implica que, en el límite, X_t es aproximadamente $N(0, ct)$.

Una familia de variables aleatorias $\{X_t\}$ con incrementos independientes, estacionarios y con distribución aproximadamente normal con esperanza 0 y desvío ct se denomina un proceso estocástico con tiempo continuo llamado “movimiento browniano” o “proceso de Wiener”.

5.3. Tamaño de las piedras. Es fácil notar que los tamaños de las rocas no son todos iguales. En Mineralogía es de gran utilidad representar esta variación mediante una distribución. Un modelo sencillo postula a la *lognormal* para estos casos.

Consideremos una roca de masa M. En el primer paso partimos la roca en dos, resultando de esa forma dos piedras de masa MU_1 y $M(1-U_1)$ respectivamente, donde $U_1 \in (0,1)$ es una variable aleatoria con función densidad F. Como la numeración es arbitraria, podemos suponer que U_1 y $1-U_1$ tienen la misma función de distribución. En el segundo paso volvemos a partir las piedras en dos, y así, en el n-ésimo paso, tendremos 2^n trozos de roca, cada uno de ellos con masa igual al producto $MW_1W_2\ldots W_n$, donde las W_i tienen distribución F_i (la W_1 puede ser U_1 o $1-U_1$,etc.). Si llamamos $Z_i = logW_i$, y X a la masa de cualquier partícula, es $logX = logM + \sum_{i=1}^{n} Z_i$. Si suponemos que todas las W_i tienen igual distribución (y por lo tanto las Z_i), para n grande, tendremos que $logX$ es aproximadamente normal por el Teorema Central del Límite, y por lo tanto que X es aproximadamente *lognormal* (esta suposición proviene de que si $Z_n \to Z$ y g es una función continua, entoces $g(Z_n) \to g(Z)$). Si bien nada nos garantiza que las suposiciones del modelo se verifiquen en la realidad, el hecho es que la *lognormal* resulta en la práctica una muy buena aproximación para muchas distribuciones empíricas de los tamaños de trozos minerales.

5.4. La actualidad. Hoy en día el Teorema Central del Límite da apoyo al uso de la normal como distribución de los errores aleatorios de medición. El error (aleatorio) de medición en un experimento físico está compuesto de muchos errores pequeños no observables que pueden considerarse aditivos.

También hay muchas situaciones en las que se puede justificar el uso de la normal a través del Teorema Central de Límite, aunque no se trate de casos sujetos al error de medición. Así por ejemplo, la distribución de alturas de hombres adultos de cierta ciudad puede ser considerada aproximadamente normal, ya que la misma puede ser pensada como suma de muchos factores pequeños e independientes.

Otro ejemplo sería la distribución del consumo de electricidad en una ciudad en cualquier hora dada, la cual es la suma de los pedidos de un gran número de consumidores individuales.

6. Comentarios finales

Si bien en la década del '30 Feller demostró la necesidad de la condición de Lindeberg para que una suma de variables aleatorias converja a la distribución normal, en la actualidad, diversos matemáticos siguen planteando nuevos interrogantes sobre el tema.

En 1980, Araújo y Giné demostraron resultados que combinados de manera adecuada nos dan el Teorema Central del Límite en su versión actual. Esta última, probada por Genedenko, describe la ley límite de la suma de variables aleatorias independientes como una convolución de una ley normal y una ley de Poisson generalizada (ver [**10**]).

¿Será ésta la última versión del Teorema Central del Límite? ¿O tendremos en este siglo la oportunidad de conocer nuevas versiones que enriquezcan nuestra matemática?

Alguien dijo una vez: "Quien busca, encuentra..."

FIN DE CONTENIDOS

Agradecimientos

Queremos agradecer a todos los profesores del Departamento de Matemática de nuestra Facultad por el apoyo que nos brindaron, en particular a los Licenciados Pedro Marangunic, Graciela Garguichevich y Marina Fernández de Luco por el material bibliográfico que nos facilitaron.

Damos las gracias muy especialmente a los Licenciados Jorge Flamini, Raúl Katz, Alejandro Kocsard y Elisa Petrone, y a las Dras. Graciela Nasini, Silvia Di Marco y Cristina Sanziel, por las importantes observaciones al trabajo que nos han aportado y por habernos instruido en el uso de LaTeX. Sin la colaboración de todos ellos no nos hubiese sido posible realizar este trabajo.

Bibliografía

Bibliografía

[1] Feller W. *Introducción a la teoría de probabilidades y sus aplicaciones*, Editorial Limusa México, (1978).

[2] Polya G. *Über den zentralen grenzwestatz, der wahrscheinlichkeitsrechnung und über das momentenproblem*, Mathematische Zeitschrift (1920).

[3] De Moivre A. *The Doctrine of Chances*, England, (1718).

[4] Laplace M. le Comte. *Théorie Analytique des Probabilités*, Mme Veuve Courcier París, (1812).

[5] Shiryayer A. *Probability*, Springer Verlag New York, (1984).

[6] James B. *Probabilidade: um curso em nivel intermediario*, IMPA Rio de Janeiro (1996).

[7] Apostol T. *Análisis Matemático*, Editorial Reverté, Barcelona (1996).

[8] Loeve M. *Teoría de la Probabilidad*, Tecnos, Madrid, (1976).

[9] Maronna R. *Probabilidad y Estadísticas Elementales*, Editorial Exacta, la Plata, (1995).

[10] Pardo L. Quesada V. *Curso Superior de Probabilidades*, Editorial PPU, Barcelona, (1987).

[11] García A. Quesada V. *Lecciones de Cálculo de Probabilidades.* Ediciones Díaz de Santos, Madrid, (1988).

[12] Landro. *Acerca de la Probabilidad*, Ediciones C.E.C.E., Universidad Nacional de Buenos Aires, (1999).

[13] Papoulis A. *Probability, random variables and stochastic processes*, Mc. Graw-Hill, New York, (1991).

[14] Boyer C. *Historia de la matemática*, Editorial Alianza, Madrid (1996).

[15] Parzen E. *Teoría moderna de probabilidades y sus aplicaciones*, Editorial Limusa, Editorial Limusa, (1976).

[16] Meyer P. *Probabilidad y aplicaciones estadísticas*, Fondo Educativo Interamericano S.A., Fondo Educativo Interamericano, (1973).