EL SIGNIFICADO DEL TEOREMA CENTRAL DEL LÍMITE: EVOLUCIÓN HISTÓRICA A PARTIR DE SUS CAMPOS DE PROBLEMAS[1]

Hugo Alvarado, Carmen Batanero [2]

[1] Universidad Católica de la Santísima Concepción, Chile

[2] Universidad de Granada

halvarad@ucsc.cl, batanero@ugr.es

RESUMEN

En este trabajo se analiza la evolución histórica del teorema central del límite, desde la perspectiva de la TFS con el fin de identificar los campos de problemas que lo originaron. La finalidad es mostrar la evolución de sus diferentes significados, iniciar la reconstrucción de su significado global y comprender las dificultades encontradas, que pueden posteriormente reproducirse en el aprendizaje de los estudiantes.

Palabras claves: Teorema central del límite, Evolución histórica, Campos de problemas

1. INTRODUCCIÓN

En ocasiones, los profesores asumimos sin cuestionarnos los enunciados matemáticos, de manera que llegan a los estudiantes como un producto acabado. En estos casos, nuestras exposiciones en el aula no muestran los conflictos del proceso histórico creativo ni la forma en que han llegado a construirse los conceptos y proposiciones importantes.

En este trabajo iniciamos un estudio histórico sobre el teorema central del límite (TCL), desde la perspectiva del marco teórico de la TFS (Godino, 2003), en donde se consideran los objetos matemáticos como emergentes de la actividad de resolución de campos de problemas. Es parte de un trabajo más amplio que pensamos completar y está orientado a diseñar y evaluar procesos de estudio del TCL para estudiantes de ingeniería. Se inscribe en una línea de investigación llevada a cabo en la Universidad de Granada sobre significado y comprensión de conceptos estadísticos en entornos tecnológicos el que ya se han realizado otros trabajos, también basados en el mismo marco teórico (por ejemplo, Tauber, 2001).

En el marco teórico considerado, las matemáticas se asumen como una actividad humana implicada en la solución de cierta clase de situaciones problemáticas de la cual emergen y evolucionan progresivamente los objetos matemáticos y se pretende elaborar

[1] En A. Contreras (Ed.) (2005). Investigación en Didáctica de las Matemáticas (pp. 13-36). Jaén: Universidad de Jaén

un modelo de los procesos de comprensión de las matemáticas que tenga en cuenta los factores institucionales y socioculturales implicados en los mismos, además de los psicológicos (Godino y Batanero, 1994; 1998). Dicho marco nos parece especialmente adecuado, en cuanto permite resaltar el significado específico de los conceptos estadísticos inducido por la tecnología, los conflictos semióticos potenciales y los criterios de idoneidad de los procesos de estudio diseñados.

Partiremos principalmente de los trabajos de Fisher (2000), Mether (2003), Xiuyu (2003), que completamos con otros trabajos sobre historia de la estadística y reanalizamos desde el marco teórico. Nos centramos, precisamente en la identificación de los campos de problemas asociados al TCL y de su evolución histórica, desde su primera forma simple cuando la teoría de la probabilidad todavía no había sido considerada parte de la matemática hasta llegar a la etapa actual, resaltando las prácticas matemáticas que llevaron a la solución de los mismos, así como la contribución por diferentes matemáticos.

La finalidad principal es iniciar la reconstrucción del significado global del TCL, que sirva como base para el diseño de procesos de estudio e instrumentos de evaluación. Asimismo, el conocimiento de la historia del TCL puede ser un recurso para los profesores, contribuyendo a enriquecer su actividad docente en el aula, favoreciendo la comprensión de los diversos campos de problemas asociados al teorema y mostrando algunas dificultades que pueden también estar presentes en nuestros estudiantes.

2. EL TCL Y SU IMPORTANCIA EN ESTADÍSTICA

El Teorema Central del Límite es fundamental en la enseñanza universitaria, especialmente en las carreras de ingenierías y ciencias, donde es un instrumento que se utiliza a diario en situaciones cotidianas y profesionales. En la actualidad conocemos como Teorema (o teoremas) central del límite a una serie de resultados acerca del comportamiento de la distribución de la suma (o media) de variables aleatorias, que en forma simplificada establece: "*la suma de un gran número de variables aleatorias independientes tiende a seguir de manera asintótica una distribución normal, siempre que determinadas condiciones queden satisfechas*" (Wisniewski yVelasco, 2001, pp. 211).

Pero en el marco de la teoría de las funciones semióticas (Godino y Batanero, 1994; 1998; Godino, 2003), el significado de un objeto matemático no se considera único, ni estable en el tiempo y se diferencia entre el *significado institucional* de un objeto matemático, que es compartido dentro de una institución (por ejemplo educativa o de estadísticos profesionales) y el *significado personal* adquirido por cada miembro de la institución. Los autores presentan el objeto matemático como un emergente de un sistema de prácticas ligadas a la evolución de *campos de problemas* específicos ligados al objeto. Las prácticas sociales dependerán de la institución (por tanto del momento histórico). Tanto los campos de problemas como las prácticas son *elementos del significado* (Godino, 2002) de un objeto matemático y contribuyen a caracterizarlo, además de otros elementos, como los procedimientos ligados a la resolución de problemas, el lenguaje, conceptos y propiedades y argumentos.

La mayor parte de aplicaciones actuales de la estadística en la ciencia y la ingeniería, se basan en obtención de conclusiones sobre una población, a partir de los datos disponibles de una muestra (inferencia estadística). La comprensión del TCL, es fundamental en los procedimientos inferenciales de estimación puntual de parámetros, estimación por intervalos de confianza y prueba de hipótesis, cuando no se conoce la distribución exacta del estadístico en el muestreo, ya que permite determinar las *distribuciones asintóticas* de la media y otros parámetros.

Mediante el TCL podemos también obtener distribuciones aproximadas de otras distribuciones clásicas, tales como la Binomial, Poisson, Chi-Cuadrado, mediante la consideración de la *suma de variables aleatorias*, ya que el teorema se cumple, independientemente de la distribución de partida de la población. Esto explica también que muchos métodos estadísticos requieren la condición de normalidad para su correcta aplicación y que se hayan desarrollado muchos métodos basados en dicha distribución.

Finalmente, desde el punto de vista didáctico el teorema reúne *elementos diversos* cuya comprensión debe adquirir el estudiante antes de abordarlo, tales como: variable aleatoria, distribución normal, muestra aleatoria, estadístico, distribuciones muestrales, lo que hace rico y complejo su estudio.

3. DESARROLLO HISTÓRICO

Desde el punto de vista histórico, el TCL es uno de los resultados más notables de la teoría estadística, debido al esfuerzo sostenido de varios matemáticos destacados, que fue establecido por primera vez en 1738 por Abraham De Moivre, bajo condiciones muy restringidas. A principios del siglo XIX, Laplace lo formuló de manera más general; pero hasta 1901 no se pudo establecer por Liapounov en condiciones generales ni proporcionar una demostración rigurosa, empleando herramientas matemáticas sofisticadas.

En lo que sigue analizamos este desarrollo. Las notaciones que usamos en adelante son las siguientes:

$X_1, X_2, ...$: variables aleatorias (v.a.);

v. a. i.i.d. : v. a. independientes e idénticamente distribuidas;

$S_n = \sum_{i=1}^{n} X_i$: suma de n variables aleatorias;

$\sigma_n^2 = \sum_{i=1}^{n} \sigma_i^2$: suma de *n* varianzas, donde X_i tiene varianza $\sigma_i^2 < \infty$;

$\varphi(t) = E\left(e^{itX}\right)$: función característica de *X*;

$X \approx B(n, p)$: la v.a. *X* tiene distribución binomial de parámetros *n* y *p*;

$N(\mu, \sigma)$: distribución normal de media μ y desviación típica σ;

$E(X_i)$: la esperanza o media de la v. a. X_i;

3.1. IDEAS PRECURSORAS

El origen del TCL surge de la necesidad de encontrar fórmulas de cálculo para la distribución Binomial y otras distribuciones discretas, en las que intervienen los términos factoriales. Un problema es que, al aumentar el valor de *n* estos términos crecen muy rápidamente, por lo que el cálculo de las probabilidades de los valores de la variable en las distribuciones exactas es demasiado laborioso, incluso hoy día. Mucho antes de la aparición de los ordenadores, este problema llevó a distintos matemáticos a tratar de encontrar valores aproximados de estas probabilidades para valores de *n* grandes y a estudiar las condiciones en que estas aproximaciones podrían utilizarse con un error acotado. Estos estudios inician el interés hacia los teoremas de límite y dan origen al primer campo de problemas relacionados con el mismo:

CP1: ¿En qué condiciones y mediante qué funciones o distribuciones de probabilidad pueden aproximarse otras distribuciones, cuando aumentamos progresivamente el valor de ciertos parámetros de las mismas, y, en particular, tamaño de una muestra?

Un problema particular dentro de este campo sería la búsqueda de una aproximación de la distribución binomial $B(n, p)$ para valores de *n* grandes. Según Xiuyu (2003), Abraham De Moivre (1667-1754) publicó comentarios sobre este tópico en 1730 en su "Miscellanea Analytica" y publica la solución en su "Doctrine of Chances" (1733), donde mostró en notación moderna, el siguiente resultado:

Sean $X \approx B(n, p)$, $P(x) = \binom{n}{x} p^x (1-p)^{n-x}$*, entonces, para valores grandes de n, X sigue una distribución aproximadamente normal* $N(\mu, \sigma)$ *de media* $\mu = n\,p$ *y desviación típica* $\sigma = \sqrt{npq}$ *; es decir la variable* $Z = \frac{x - np}{\sqrt{npq}}$ *sigue una distribución normal N (0,1).*

De Moivre investigó los límites de la distribución binomial cuando el número de ensayos aumenta, encontrando una relación con la función e^{-x^2}, pero no lo relaciona con la distribución normal. Intentó determinar, para valores grandes del número de ensayos *n*, la probabilidad de ocurrencia del valor central $\frac{n}{2}$ (n es el total de ensayos), aproximandola a la expresión $\frac{2}{(K\sqrt{n})}\cdot\frac{1}{2n}$. James Stirling descubrió que *K* es igual a $\sqrt{2\pi}$.

3.2. SUMA DE VARIABLES ALEATORIAS DISCRETAS IDENTICAMENTE DISTRIBUIDAS CON MEDIA Y VARIANZAS FINITAS

Puesto que la distribución binomial *B (n,p)* puede considerarse como la suma de variables aleatorias idénticamente distribuidas *B (1, p)* o variables de Bernoulli, la demostración que la distribución binomial se aproximaba mediante la distribución normal es un caso particular de lo que sería posteriormente el teorema central del limite. La suma de variables aleatorias independientes fue estudiada para diferentes distribuciones particulares de errores desde 1776, y daría lugar a un segundo campo de problemas:

CP2: Encontrar la distribución de la suma de variables aleatorias discretas independientes e idénticamente distribuidas (o de la media de variables aleatorias independientes e idénticamente distribuidas).

Un primer caso de este problema sería aquél en que las variables fueran uniformes discretas (CP2.1). El trabajo probabilístico de Laplace sobre la suma de variables aleatorias jugó un rol importante en lo que sería posteriormente el TCL. En un artículo publicado en 1776, estuvo trabajando en calcular distribuciones de probabilidad de la suma de ángulos de inclinación de meteoros. Afrontó el problema de la desviación entre la media aritmética de los datos (obtenido con errores observacionales) y los valores teóricos, suponiendo que todos estos ángulos distribuidos aleatoriamente siguen una distribución rectangular entre 0º y 90º. Puesto que la distribución de la media es equivalente a la de la suma (excepto un factor constante, que es el valor *n* por el que se divide), el problema se centró en hallar la distribución de la suma de los datos, todos los cuales tienen la misma distribución y son independientes. Debido a la considerable magnitud de dichas desviaciones, no fue capaz de representar un cálculo exacto y así necesitó de una aproximación (Fisher, 2000). Fue en el proceso de encontrarla, que Laplace eventualmente vino a formular la primera versión del TCL.

La distribución de la media aritmética había sido estudiada por muchas distribuciones de rango finito, por ejemplo, para una variable aleatoria discreta uniformemente distribuida entre los enteros *–a* y *a*, con resultados de fórmulas complicadas (véase Xiuyu, 2003, pág. 5). Laplace logró una aproximación más simple para $P(S_n = 0)$ en 1785, usando la función característica para calcularla (S_n es la suma de los *n* posibles errores) y llegando a la siguiente aproximación:

$$P(S_n = 0) \approx \frac{\sqrt{3}}{\sqrt{2\pi na(a+1)}}$$

Laplace generalizó en el año 1809, el trabajo de De Moivre referido a la distribución binomial a la suma de variables aleatorias discretas idénticamente distribuidas con media y varianza finita (CP2.2). Podemos ver en Feller (1975, pp. 191) una extensión de la aproximación a la binomial general con $p \neq 1/2$; en la cual el procedimiento de cálculo es más complicado. En 1810, publicó una memoria, donde estableció y probó más formalmente la siguiente versión del TCL:

Sea X una variable aleatoria valorada entera que toma valores enteros $\{-a, ..., 0, ...a\}$ *con probabilidad* $P(X = x) = \left(\frac{x}{2a}\right)$, *entonces para una muestra n grande, la suma* $S_n = \sum_{i=1}^{n} X_i$ *está distribuida aproximadamente en forma normal* N(μ, σ) de media nμ y desviación típica $\sqrt{n}\ \sigma$ siendo μ, σ *la media y desviación típica de la distribución de partida.*

Laplace dio una demostración "rigurosa" para el caso *discreto* del TCL usando la función característica $\varphi(t) = E\left(e^{itX}\right)$ que desarrolla en serie de potencias hasta el segundo término e invierte para hallar a qué distribución corresponde (técnica de la probabilidad inversa). La idea principal de su demostración puede encontrarse en Fisher (2000). Laplace intentó argumentar dos generalizaciones, no logrando demostraciones rigurosas:

- Debido a que la distribución límite es independiente de *a* (sólo depende de *a* a través de la media y la varianza), el resultado es válido también para una *distribución discreta con rango infinito*, en el caso de que existan los momentos finitos (Podemos considerar este otro caso particular del campo de problemas CP2.3).

- Los resultados también se cumplen para *variables aleatorias con distribución continua y acotada*. Consideraremos éste un nuevo campo de problemas (CP4) ya que se introducen nuevos conceptos como el de variable continua, función de densidad, etc. y se utilizan nuevas herramientas, tal como la integral.

Gauss derivó en 1809 la función de densidad normal $f(x)=\frac{1}{\sqrt{\pi}}e^{-x^2}$ en su trabajo "Teoría motus corporum celestium". Su descubrimiento estuvo abierto a muchas críticas, sin embargo tuvo un impacto grande en Laplace. Stigler (1986) sugiere que no hay indicios que Laplace haya pensado en que esta función pudiese representar una función de densidad para una distribución de probabilidad, sino que encontró el trabajo de Gauss de 1810 y se basó en el mismo.

3.3. SUMA DE VARIABLES NO IDÉNTICAMENTE DISTRIBUIDAS Y VARIABLES CONTINUAS

Contribución de Poisson

De todos los que aportaron al TCL en el siglo XIX, la contribución más decisiva fue la de Siméon Denis Poisson, quien publicó dos artículos (1824 y 1829) donde discutió el TCL. La idea de Poisson fue que todos los procedimientos en el mundo físico están gobernados por leyes matemáticas claras. En este espíritu, intentó suministrar un análisis matemático más exacto al teorema de Laplace. Las contribuciones de Poisson al TCL fue doble (Fisher, 2000):

- Mejoró y generalizó la demostración de Laplace. Comienza estableciendo una demostración del TCL para variables independientes distribuidas idénticamente; primero para una suma de éstas y luego para una combinación lineal de ellas. Luego generalizó su demostración a la *suma de variables aleatorias con distribuciones diferentes* (CP3). En esta generalización, la demostración considera las variables continuas, que da origen a otro campo de problemas (CP4).
- Discutió la validez del TCL, principalmente dando algunos contraejemplos, en el TCL no se cumple para la distribución de Cauchy. De este modo inicia un nuevo campo de problemas, esta vez estrictamente matemático que consistirá en *encontrar condiciones suficientes y necesarias para su validez* (CP6).

En 1824, Poisson da una demostración más rigurosa para una *variable continua* (de este modo sembró la semilla para el concepto de variables aleatorias). Los resultados generales de la última parte del artículo pueden ser resumidos de la siguiente forma (Mether, 2003):

Sean $X_1, ..., X_n$ *variables aleatorias independientes con funciones de densidad que desaparecen más allá del intervalo fijo [a , b]. Si la función característica está acotada para valores entre 0 y 1, entonces* $X_1 + ... + X_n$ *se distribuye aproximadamente como una normal* $N(\mu, \sigma)$*, siendo* $\mu = \sum E(X_i)$ *y* $\sigma^2 = \sum E(X_i)$*. Esta aproximación es mejor cuando n es grande y los valores del intervalo donde se calcula la aproximación se aproximan a la media (Fisher, 2000).*

Parece que el propósito principal del TCL para Poisson fue ser una herramienta en el cálculo de probabilidades clásicas, más que ser un teorema matemático en sí mismo. Por lo tanto, Poisson no formuló explícitamente alguna condición para el TCL. Parece claro desde su demostración y ejemplos, que él asumió que las varianzas de las componentes de la suma están acotadas para que la varianza de la suma sea de orden *n*. Aunque no indica explícitamente esta idea, sin embargo, discutió unos pocos contraejemplos donde el TCL no se cumple. Uno de estos ejemplos es la llamada variable distribuida- Cauchy donde la densidad de probabilidad toma la siguiente forma:

$f(x) = \dfrac{1}{\pi(1 + x^2)}$, que no puede aproximarse mediante una distribución normal.

Contribución de Dirichlet y Bessel

Dirichlet y Bessel siguieron principalmente los pasos de Laplace y Poisson, pero introdujeron el llamado "factor de discontinuidad" en la demostración, que les permitió establecer el resultado de Poisson para valores arbitrarios de *n*. Ellos siguieron caminos diferentes para lograr la demostración, pero la regla es la misma. Bessel además hizo hincapié en que Poisson usó el mismo factor de discontinuidad, aunque en una forma ligeramente diferente, cuando estableció la fórmula para *n=1.*

CP5: Encontrar estimaciones para el error cometido en las aproximaciones de las sumas de variables aleatorias mediante la distribución normal

Dirichlet fue el primero en intentar estimar el error de la aproximación, aunque no fue muy afortunado y su técnica fue diferente de la moderna. Dirichlet antes había estimado los errores para otras aproximaciones no probabilísticas, y mostró que estas mismas técnicas pueden ser aplicadas para la teoría de probabilidad tan bien como lo habían sido para la matemática pura.

Contribución de Cauchy

Cauchy fue uno de los primeros matemáticos que consideró seriamente la teoría de probabilidad como matemática pura. Su demostración del TCL sigue una línea diferente comparada con las demostraciones previas. Cauchy en su demostración usó la función característica para variables independientes e idénticamente distribuidas $X_1, X_2, ..., X_n$ con función de densidad simétrica y de rango finito. Siguió la línea de la demostración de Poisson pero consideró la distribución de la función $\sum_{i=1}^{n} w_i x_i$ aproximada asintóticamente a la distribución normal de media cero y varianza $\sigma^2 \sum_{i=1}^{n} w_i$, bajo la condición que la función lineal es el mejor estimador lineal insesgado de un parámetro, es decir, $\sum_{i=1}^{n} w_i = 1$. Primero estableció una cota superior para la diferencia entre el valor exacto y la aproximación y luego especificó condiciones para esta cota cuando tiende a cero (Hald, 1998).

3.4. DISTRIBUCIONES CON RANGO INFINITO

Con la demostración de Cauchy finalizó el llamado primer período (1810-1853) del TCL. La demostración presentada en este período fue insatisfactoria en tres aspectos (Hald, 1998):

- El teorema sólo fue probado para distribuciones de rango finito.
- No se explicitaron las condiciones, en términos de los momentos, bajo el cual el teorema se cumple.
- La razón de convergencia para el teorema no fue estudiada.

Estos problemas fueron eventualmente resueltos por matemáticos rusos, entre 1870 y 1910, usando dos soluciones: el “método de los momentos (Chebyshev y Markov) y la función característica (Liapounov). Nos centraremos en los probabilistas Chebyshev, Markov y Liapounov, reconocidos en el grupo de la escuela de San Petersburgo, aunque otros matemáticos rusos como Nekrosov y Sleshinsky se implicaron en el trabajo con el mismo problema (Seneta, 1984).

Contribución de Chebyshev y Markov

El artículo de Chebyshev en 1887, usualmente es considerado el primer intento (no conseguido) de demostración rigurosa del TCL. En él estableció lo siguiente (Seneta, 1984):

Sean $X_1, X_2, X_3, \ldots,$ *variables aleatorias independientes con densidad de probabilidad dada. Si*
i) $E(X_i) = 0 \quad \forall i$ *, la media de estas variables es cero y*
ii) $\left|E(X_i^k)\right| \leq C \quad \forall i, \quad k \geq 2$ *los momentos de orden superior a dos están acotados;*
Entonces cuando $n \to \infty$ *la distribución de la suma es aproximadamente normal* $N(\mu, \sigma)$, siendo $\mu = \sum_{i=1}^{n} X_i$ *y* $\sigma^2 = \sum_{i=1}^{n} Var(X_i)$.

Chebyshev usó el "método de los momentos", tempranamente desarrollado por él y simplificado y completado más tarde por Markov, quien también completó la demostración de Chebyshev del TCL en 1898. Markov estableció que: "se necesita añadir condiciones para que el teorema se cumpla". Primero propuso la condición de que la suma de varianzas está uniformemente acotada lejos de 0. Después reemplazó la condición con la de que las esperanzas $E\left(X_n^2\right)$ estén acotadas cuando *n* tienda a infinito.

Después de la demostración dada por Liapounov del TCL utilizando la función característica en 1901, Markov intentó lograr el mismo nivel de generalidad con el método de los momentos. Lo logró en 1913 cuando dio una demostración rigurosa del TCL bajo la condición de Liapounov usando el método de los momentos.

Teorema de Liapounov

La demostración de Liapounov, publicada en 1901, es considerada la primera demostración rigurosa del TCL. Demos una mirada al teorema establecido por Liapounov (Uspensky, 1937):

Sean $X_1, X_2, \ldots, X_n$ *variables aleatorias independientes con las siguientes propiedades:* $E(X_i) = 0 \quad \forall i$ y $E\left|X_i\right|^k \leq \infty \quad \forall i, \quad k \geq 2$ *, entonces usando el teorema de continuidad y unicidad de la función característica concluyó que la distribución de la suma es aproximadamente normal.*

La demostración de Liapounov sigue la idea de Laplace usando la función característica e introduce un ***lema fundamental*** que es la clave a la simplicidad y rigurosidad de su demostración:

Sea S_n *una variable que depende del entero n, con media o y varianza 1. Si la función característica de la variable* S_n *converge uniformemente a la función característica de la distribución normal en algún intervalo finito, entonces la función distribución de la variable* S_n *tiende uniformemente a la distribución normal.*

Liapounov no separó explícitamente este lema fundamental en su demostración, en la que está contenido implícitamente. Este lema fue usado por Linderberg y Lévy en su mejora del TCL (Uspensky,1937).

Podemos observar que en este período se trabajó el TCL bajo el supuesto sólo de variables independientes, a diferencia del primer período cuyo supuesto fue demostrar el teorema para variables aleatorias independientes e idénticamente distribuidas. En particular, una aplicación de la demostración del TCL dada por Liapounov al caso de variables independientes e idénticamente distribuidas es la siguiente:

Sean $X_1, X_2, X_3, \ldots,$ *variables aleatorias independientes e idénticamente distribuidas con media finita* μ *y varianza* $0 < \sigma^2 < \infty$*, entonces*

$\frac{S_n - n\mu}{\sqrt{n}\,\sigma} \to N(0,1)$ *cuando* $n \to \infty$.

La demostración de este teorema utiliza la función característica y la expansión de Taylor (Ver Xiuyu, 2003, pág. 14, Parzen, 1987, pág. 474. También ver Meyer, 1992, pág. 260 donde se demuestra mediante la función generadora de momentos).

3.5. ESTUDIO DE CONDICIONES NECESARIAS Y SUFICIENTES PARA EL TEOREMA (CP6)

El tercer y último capítulo en la historia del TCL de 1920 a 1937, comienza con la demostración de Linderberg y las demostraciones de Lévy y Féller que añadieron condiciones necesarias al teorema.

La Condición de Linderberg

El 1922, Linderberg publicó una demostración elemental que fue la base de los trabajos de Lévy y Feller sobre el mismo problema, del teorema siguiente (Cam, 1986):

*Teorema (**Condición de Linderberg**)*

Sean X_i *variables aleatorias independientes con esperanza* $E(X_i) = 0$ *y varianza* σ_i^2 *finitas. Si los momentos de tercer orden están acotados, entonces la distribución de la suma se aproxima a la distribución normal.*

La condición anterior, es conocida como de Linderberg y su formulación matemática se puede ver en textos de teoría de la probabilidad (Shiryayev (1984, pág.326, Loève 1976, pág. 274), usualmente usando la función característica. Linderberg usó una simple aproximación sin considerar la función característica. En su demostración, encontró una cota para el $\sup_X |F(x) - \Phi(x)|$, donde $\Phi(x)$ es la distribución de probabilidad normal, en términos del tercer momento. El método de demostración de Linderberg no se usó en toda la primera mitad del siglo XX, hasta que el TCL comenzó a ser estudiado para espacios de dimensión infinita, donde es fácilmente transferible por cuanto no usa la función característica. Actualmente es usada en muchos casos donde la convergencia a la distribución normal es considerada con variables distribuidas no idénticamente y su fuerza está principalmente en que puede ser aplicado en un contexto muy general y considera la razón de convergencia en el teorema límite.

Cabe señalar que el método de Linderberg no dio un orden óptimo de la razón de convergencia, en el caso de sumandos de valores reales independientes e idénticamente distribuida con tercer momento finito para el orden n (Paulauskas, 1999). Linderberg dio entonces una demostración rigurosa para la condición *suficiente* del TCL, pero no probó la condición *necesaria*. Esta carencia fue particularmente remediada por Lévy y Féller en 1935 y 1937. Por otro lado, Poisson muestra, alrededor del año 1824, que la aproximación a la distribución normal no siempre se cumple para variables independientes arbitrarias.

Teorema de Feller

Feller en un artículo publicado en 1935 da las condiciones necesaria y suficiente para el TCL, pero el resultado es algo limitado. Feller consideró, en la convergencia normal de la suma, una secuencia infinita X_i de variables aleatorias independientes y usó los segundos momentos finitos para cambiar orígenes y escala de valores de las sumas consecutivas de la forma $\frac{1}{a_n}(\sum X_i - c_n)$, a fin de evitar que se extiendan hacia valores infinitos. Luego dio las condiciones de restricción para cada variable x_k/a_n que tiende a cero en probabilidad. Puesto que Feller puso esta restricción, y porque sólo trató la suma de variables normales, su solución no puede ser considerada la solución final al TCL. Este teorema de Feller es frecuentemente llamado el Teorema de Linderberg-Feller, al usar la condición de Linderberg. La demostración puede verse en Shiryayev (1984, pág. 332) y está basada en la estimación de las colas de las varianzas en términos del comportamiento de la función característica en el origen.

Contribución de Lévy

Lévy probó la condición de Linderberg en 1925 usando la función característica, considerando esta demostración superior a la de Linderberg, (Cam, 1986). Lévy trató en primer lugar el nuevo caso de segundos momentos infinitos y de primeros momentos finitos o infinitos (CP2.3, que Laplace intento argumentar sin éxito). Publicó algunos artículos relacionados con el TCL entre los años 1925 y 1930, usando principalmente la función característica en su demostración. Después de 1930, evitó usar la función característica. En su artículo de 1935, presentado sólo unos pocos meses después de Feller, y a pesar que trató el mismo problema, ambos autores negaron tener algún contacto previo (Cam, 1986). Establece las siguientes ideas:

- Dio las condiciones necesarias y suficientes para la convergencia de sumas de variables aleatorias independientes e idénticamente distribuidas normales a la distribución normal.
- También dio las condiciones necesarias y suficientes para el caso general de sumas independientes.
- Por último, intentó dar las condiciones necesarias y suficientes para variables dependientes, de procesos estocásticos, caso Martingalas. Consideramos este caso el último campo de problemas (CP7) en la historia del TCL.

En este último, Lévy tuvo algunos problemas con la demostración para el caso martingalas, que no fue completamente satisfactoria y que no sustenta una prueba de rigurosidad. Su demostración de condición necesaria y suficiente para el caso general de variables independientes fue correcta, pero se basó en un lema hipotético, que no había sido probado todavía (como la demostración de Feller). El **lema** es el siguiente:

"Si la suma S = X + Y de dos variables aleatorias independientes (X e Y) tiene una distribución normal, entonces también lo son X e Y".

El año siguiente 1936, Crámer probó el lema (como un teorema) y éste fue establecido, mostrando la validez del uso de sumas normales y los teoremas de Lévy y Feller donde generalmente se aplica. En 1937, Feller y Lévy refinaron la demostración, después de usar el resultado de Crámer. El TCL fue así probado con las condiciones necesarias y suficientes (Cam, 1986) y se llega al final del desarrollo.

3.6. RESUMEN DEL ESTUDIO HISTÓRICO

En las secciones anteriores se ha mostrado el interés de matemáticos y estadísticos prestigiosos en el TCL, uno de los más importante teoremas en estadística matemática y teoría de la probabilidad. Su formulación final establece que si se cumplen algunas condiciones, entonces la *distribución de la media aritmética* de un número de variables aleatorias independientes se aproxima a la distribución normal según el número de variables crece indefinidamente. Es decir, independientemente de la distribución actual de las variables, a medida que crece el número de sumandos, la suma de éstas pueden ser tratadas como una distribución normal. En la Tabla1 se presenta un resumen cronológico de la evolución del Teorema Central del Límite, a partir de sus campos de problemas, las herramientas utilizadas y resultados obtenidos en cada paso.

Tabla 1. Desarrollo histórico del TCL

Sgnificados históricos	**Campo de problema**	**Métodos de solución (procedimientos y conceptos previos)**	**Resultados (Conceptos y propiedades)**
1713-1733 Aproximación de distribuciones	CP1: Aproximación de ciertas distribuciones (e.g. binomial) para valores grandes de sus parámetros	Aproximación de n! Paso al límite en sucesiones	Teorema de Bernoulli (De Moivre)
1785-1809 Suma de v.a. iid discretas	CP2: Suma de v.a. discretas independientes e idénticamente distribuidas	Distribución de la suma de v. a. Independencia y regla del producto Inversión de la función característica (Laplace) Desarrollo en serie	CP2.1. v.a. uniforme discreta (Laplace) CP2.2. v.a. discretas i.i.d., media y varianza acotadas (Laplace) Densidad normal (Inicio de estudio de teoría de los errores por Gauss y Laplace)
1810-1853 Suma de v.a. independientes	CP3: Suma v.a. discretas independientes no idénticamente distribuidas. CP4: Suma v.a. continuas	Cálculo de probabilidades mediante función de densidad Integración Introducción del factor de discontinuidad de Dirichlet (Bessel) Acotación de la diferencia y búsqueda de condiciones para que la cota tienda a cero (Cauchy) límite Teorema de la función inversa (Cauchy)	Generalización a v.a. independientes e idénticamente distribuida, para la suma y una combinación lineal (Poisson). V.a. independientes no i.d. (Poisson) V.a. continuas i.i.d. con densidad simétrica y rango finito (Cauchy) Búsqueda de casos donde no se cumple (Poisson)
1854-1919 Suma de v.a. cualesquiera, incluso dependientes	CP5: Estimaciones del error de aproximación CP6: Condiciones de validez del teorema CP7: Suma de variables	Método de los momentos (Chebychev y Markov) Lema de Liapounov (convergencia en f. característica) Continuidad de la f. característica (Liapounov) Función generadora	*Variables aleatorias con rango infinito* Comienzo de la demostración rigurosa del teorema (Chebychev) Condición necesaria para el teorema de Chebyshev (Markov) Liapounov demostró el TCL con la función característica Demostración rigurosa del teorema bajo la condición de Liapounov (Markov)

1920-1937 Suma de variables aleatorias en ciertas condiciones	aleatorias dependientes	Espacios Euclídeos y de Hilbert Razón de convergencia Sucesiones infinitas de v.a. Martingalas Convergencia en distribución	Condiciones necesarias y suficientes TCL aplicado a valores euclideanos, para vectores aleatorios y espacio de Hilbert (Linderberg) Prueba de la condición de Linderberg usando la función característica (Lévy) CP2.3. TCL para v.a. con rango infinito (Lévy) Condición necesaria y suficiente del TCL con resultado limitado (Feller) Prueba del lema (Cramer) Refinamiento de demostración (Feller-Levy)

Hemos seguido de cerca la historia del TCL, identificando diferentes campos de problemas, que podemos resumir en la forma siguiente:

CP1: Aproximación de la distribución binomial para valores grandes de n

CP2: Distribución de la suma de variables aleatorias discretas idénticamente distribuidas

- *CP2.1. Variables con distribución uniforme discreta*
- *CP2.2. Variables discretas acotadas*
- *CP2.3. Variables discretas no acotadas*

CP3. Distribución de la suma de variables aleatorias no idénticamente distribuidas

CP4. Distribución de la suma de variables aleatorias continuas

- *CP4.1. Variables acotadas*
- *CP4.2. Variables con momentos acotados*

CP5. Estimación del error de aproximación en el TCL

CP6. Búsqueda de condiciones necesarias y suficientes para el TCL

CP7. Distribución de la suma de variables aleatorias dependientes.

Comenzamos discutiendo los primeros campos de problemas, que parten de problemas prácticos, por ejemplo, la necesidad de aproximar la distribución de la suma de errores y la búsqueda de aproximaciones para la suma de variables aleatorias. Seguimos el razonamiento de Laplace, quien dio una primera formulación limitada del TCL, que varios autores trataron posteriormente de generalizar, originando campos de problemas cada vez más abstractos y ligados al método deductivo, propio de la matemática.

Poisson dio ejemplos de distribuciones que no pueden ser aproximadas por el TCL y mejora la demostración para el caso continuo. Dirichlet y Bessel introducen el factor de discontinuidad probaron el teorema de Poisson para un caso general, y Cauchy define la primera cota superior para la diferencia entre la distribución de la suma y la distribución normal. Después de esta primera etapa, los matemáticos rusos consiguieron la primera demostración rigurosa del TCL; primero Chebyshev y Markov con el método de los momentos; después, Liapounov usando la condición que lleva su nombre y la función característica. Linderberg finalizó la búsqueda de la condición suficiente, Feller y Lévy hallaron la condición necesaria para el TCL que fue rigurosamente demostrada por Cramer, llegando a la versión actual del TCL. La condición de Linderberg ha sido mejorada (Cam, 1986) y se han obtenido las condiciones necesaria y suficiente para varias variables dependientes, pero los principios básicos de Linderberg, Lévy y Feller aún permanecen.

Aunque este resumen es aún provisional, en el sentido que podemos completarlo con la descripción del lenguaje y argumentos utilizados en cada periodo, nos muestra ya un esbozo de la evolución del significado del TCL, que comienza simplemente como una herramienta de cálculo aproximado de distribuciones para ciertos valores grandes de sus parámetros y continúa como suma de variables aleatorias discretas independientes e idénticamente distribuidas. Progresivamente su significado se amplía, hasta el punto de aplicarse a variables aleatorias discretas y continuas, independientes o no, idénticamente distribuidas o no. Por un periodo parece tener validez general, hasta que se llegan a encontrar las condiciones necesarias y suficientes de su validez (restringiendo, por tanto su significado).

El estudio realizado destaca también las distintas herramientas matemáticas empleadas, algunas de ellas de gran complejidad (como función característica, convergencia en distribución o en función característica, martingalas o espacios de Hilbert). Asimismo nos muestra algunas dificultades en el desarrollo del TCL, que podrían ser compartidas por los estudiantes:

- Aceptar que una distribución discreta (por ejemplo la binomial) puede llegar a aproximarse mediante una distribución continua. Esto requiere pasar de considerar valores aislados de la variable a valores continuos y de la idea de función de probabilidad a la de función de densidad de probabilidad. El mismo De Moivre, aunque halla la aproximación, no relaciona la fórmula obtenida con la distribución normal, en el sentido de considerar la función obtenida (curva normal) con la función de densidad de una variable aleatoria (que sería el límite de la sucesión correspondiente de v.a. binomiales). Usualmente en las clases de probabilidad enseñamos a los alumnos a diferenciar entre variables discretas y continuas, como si la naturaleza de las variables fuese real, más que un modelo matemático que aplicamos para comprender una situación. Esta drástica división puede provocar *conflictos semióticos* en los alumnos, al entender que el límite de una sucesión (en este caso convergencia en distribución) de elementos de un conjunto (en este caso distribuciones discretas) ha de ser un elemento del mismo conjunto (lo que no se cumple en el TCL).

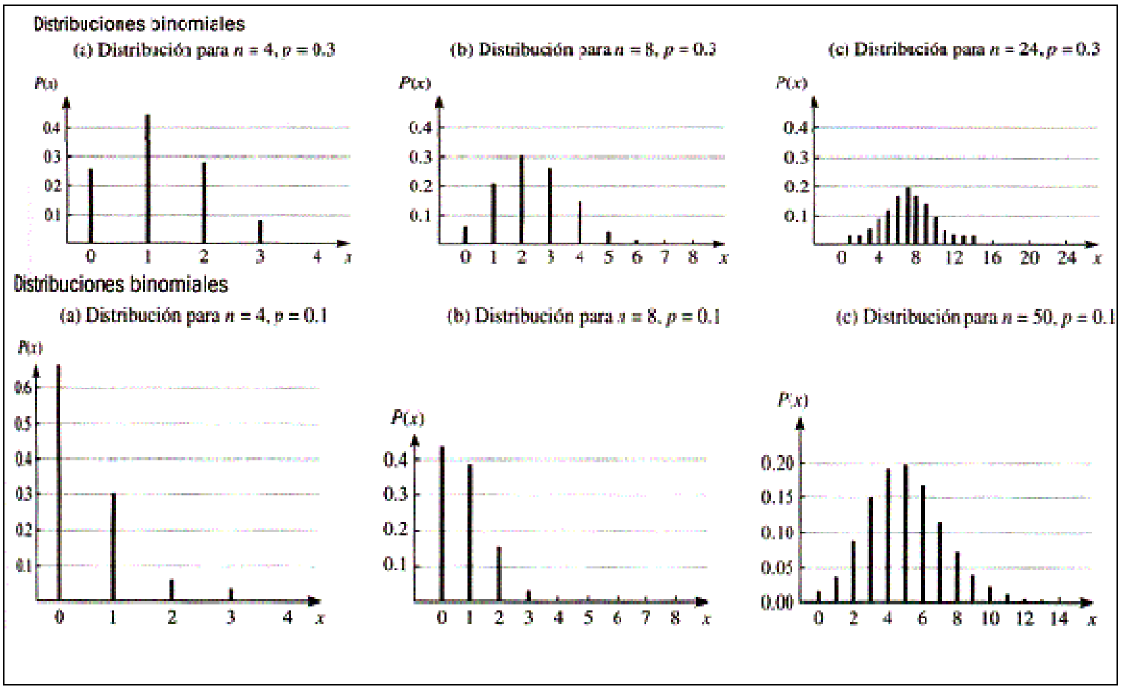
- Aceptar las sucesivas generalizaciones del TCL. Esto no ha sido sencillo y ha requerido numerosos pasos: de la distribución binomial a la uniforme discreta, de esta a variables independientes idénticamente distribuidas discretas, etc. Hoy día mostramos el TCL con total generalidad, pero la comprensión de esta generalidad de aplicación podría requerir un tiempo de maduración. Es decir el estudio ha procedido ejemplar a ejemplar, comprobando que el TCL se cumple en cada uno de ellos y sólo bastante más tarde se ha pasado a estudiar el tipo general (teoremas de límite generalizado).
- Comprender que el TCL también se aplica cuando las variables no son idénticamente distribuidas o no son independientes. La demostración posterior de estos casos sugiere que pueda ser más difícil o menos inmediato buscar ejemplos de aplicación. La suma de v.a. i.i.d. es natural en el contexto del cálculo de la media en el muestreo, al ampliar el tamaño de muestra. La suma de v.a. generales también se presenta en muchas circunstancias, cuando una magnitud puede ser explicada como la suma de un número grande de factores (por ejemplo, en muchas magnitudes biológicas), pero la determinación de estos factores, ni la consideración de la variable como suma de ellos no es tan evidente.
- Pensar que se puede aplicar el TCL en cualquier condición. La generalización progresiva lleva a pensar que el TCL tiene una validez general, y que sólo falta su demostración. Cauchy fue el primero que encontró algunos contraejemplos y el estudio de las condiciones necesarias y suficientes fue largo. Esto es un problema general de validez del razonamiento inductivo, donde la validez de una proposición para una serie de casos particulares no muestra la validez general. Los alumnos podrían presentar las mismas dificultades o confundir las condiciones necesarias con las suficientes.

4. IMPLICACIONES PARA LA ENSEÑANZA Y LA INVESTIGACIÓN

La historia del teorema como recurso didáctico nos da una perspectiva global que permite conocer la aparición de dificultades epistemológicas que presentan una gran similitud con las que atraviesan los estudiantes, y por tanto podría facilitar al profesor la comprensión del proceso de aprendizaje de los alumnos. La abstracción de los campos de problemas presentados y el aparato analítico utilizado en su solución plantea la pregunta de si es posible enseñar este teorema a los alumnos, debido a la gran complejidad de su aparato analítico.

Actualmente es posible mostrar intuitivamente el TCL en sus diversas formulaciones usando la simulación y graficación . Por ejemplo, para el caso de la aproximación de De Moivre los alumnos podrían utilizar algún software estadístico para representar gráficamente la distribución *B (n,p)* para distintos valores de sus parámetros *n* y *p.*

Figura 1: Aproximación de De Moivre



Por ejemplo, en la figura 1 se muestran valores para *p = 0.3 con n = 4, 8, 24 y para p = 0.1 con n = 4, 8 y 50.* Estas gráficas sugieren intuitivamente la convergencia que podría ser posteriormente justificada en forma analítica a los alumnos con suficiente base de cálculo. Similarmente es sencillo generar muestras aleatorias y estudiar la distribución empírica de las medias muestrales usando software estadístico, que puede descargarse de Internet. Un ejemplo es el desarrollado por delMas, Garfield. y Chance, (1999), quienes también proporcionan actividades orientadas al estudio de las distribuciones muestrales. Las evidencias empíricas del grado de comprensión de los estudiantes son, sin embargo muy limitadas, lo que sugiere el interés de continuar esta línea de investigación.

Este estudio, desde la perspectiva de las TFS, debe estar basado en el análisis del significado de referencia. El estudio histórico realizado nos permite comenzar a reconstruir el significado global del concepto, en el que, además de su dimensión matemática, tendríamos que incluir las componentes epistemológica, didáctica y cognitiva y será la base para el posterior diseño de la enseñanza del tema.

REFERENCIAS

Alvarado, H. y Batanero, C. (2004). Elementos del significado del teorema central del límite. *Actas del Octavo Simposio SEIEM, Sociedad Española de Investigación en Educación Matemática,* CD ROM Comunicaciones en los grupos de investigación. La Coruña: Sociedad Española de Investigación en Educación Matemática.

Alvarado, H. y Batanero, C. (2004). Elementos del Significado del teoerma central del límite. Granada: Sociedad Andaluza de Educación Matemática Thales.

delMas, R. C., Garfield, J. B. y Chance, B. L. (1999). A model of classroom research in action: developing simulation activities to improve students´ statistical reasoning. *Journal of Statistic Education*, 7, 3. On line: [http://www.amstat.org/publications/jse/secure/v7n3/delmas.cfm]

Feller, W. (1975). *Introducción a la teoría de la probabilidades y sus aplicaciones*. México: Limusa.

Fisher, H. (2000). *The central limit theorem from Laplace to Cauchy: Changes in stochastic objetives and in analytical methods*. On line: [http://mathsrv.ku-eichstaett.de/MGF/homes/didmath/seite/1850.pdf].

Godino, J. D. (2002). Un enfoque ontológico y semiótico de la cognición matemática. *Recherches en Didactique des Mathématiques,* 22 (2/3): 237-284.

Godino, J. D. (2003). *Teoría de las funciones semióticas. Un enfoque semiótico de la cognición e instrucción matemática*. Granada: El autor.

Godino, J. D. y Batanero, C. (1994). Significado institucional y personal de los objetos matemáticos. *Recherches en Didactique des Mathématiques*, 14 (3), 325-355.

Godino, J. D. y Batanero, C. (1998). Clarifying the meaning of mathematical objects as a priority area of research in Mathematics Education. En A. Sierpinska y J. Kilpatrick (Eds.), *Mathematics education as a research domain: A search for identity* (pp. 177-195). Dordrecht: Kluwer.

Godino, J. D. y Batanero, C. (1998). Funciones semióticas en la enseñanza y aprendizaje de las matemáticas. Seminario de Investigación en Education Matemática (SIEM) de la Sociedad Portuguesa de Investigación en Education Matemática.

Hald, A. (1998). *A history of mathematical statistics from 1750 a 1930*. New York: John Wiley.

Loève, M. (1976). *Teoría de la Probabilidad*. Madrid: Tecnos.

Mether, M. (2003). The history of the central limit theorem. *Sovelletun Matematiikan erikoistyöt*, Mat-2.108. On line: [http://www.sal.tkk.fi/Opinnot/Mat-2.108/pdf-files/emet03.pdf].

Meyer, P. (1992). *Probabilidad y aplicaciones estadística*. Segunda edición. México: Addison-Wesley.

Parzen, E. (1987). *Teoría moderna de probabilidades y sus aplicaciones*. Quinta edición. México: Limusa.

Paulauskas, V. (1999). J.W. Linderberg and the central limit theorem. *Statistics, registries and research- experiences from Finland*, (pp. 111-112). Helsinki: Statisticsl Findland.

Seneta, E. (1984). The central limit problem and linear least squares in prerevolutionary Russia: The background. *Mathematical Scientis*t 9, 37-77.

Shiryayev, A. N. (1984). *Probability*. New York: Springer.

Stigler, S. M. (1986). *The history of statistics the measurement of uncertainty before 1900*. Cambridge, MA: Belkap Press of Harvard University Press.

Tauber, L. (2001). *Significado y comprensión de la distribución normal a partir de actividades de análisis de datos.* Tesis Doctoral. Universidad de Sevilla.

Uspensky, J. V. (1937). *Introduction to mathematical probability*. New York: McGraw-Hill.

Wisniewski, P. y Velasco, G. (2001). *Problemario de probabilidad*. México: Thompson.

Xiuyu, J. (2003). Historical development of Central Limit Theorem. On line: [http://www.stat.rice.edu/~blairc/seminar/Files/julieTalk.pdf].