LA DISTRIBUCIÓN NORMAL UNA RÁPIDA REVISIÓN HISTÓRICA

Conde A. Gabriel*

**Universidad del Valle, Facultad de Ingeniería, Escueal de Estadística*

e-mail: *gabriel.conde@correounivalle.edu.co*

Abstract: This paper may be viewed more as a proposal, or a study guide, with the hope that the reader can approximate by him/herself on how the Normal distribution emerged historically. Two paths are shown: one from the development of probability concepts, which will be referred to as "the probability branch"(RPb), and the other from the search of the errors distribution on experimental measurements, which will be referred to as "the errors branch"(Ree). Each one of these two paths associates different concepts, epochs and people from which it is possible to explore and go deep into a lot of subjects such as the first versions of probability concepts, the Law of Large Numbers and the Central Limit Theorem, the J. Bernoulli's Binomial distribution, the original version of the Normal distribution due to A. De Moivre, and the ideas on the distributions of the errors from Galileo to Gauss, going through the ideas owing to T. Tompson, D. Bernoulli, L. Euler, P. S. Laplace, and A. L. Cauchy.

Keywords: Normal approximation of the Binomial distribution, Combinatory arguments, Recursive arguments, Errors curve, Errors distribution, The problem of points.

Resumen: Este articulo es más bien una propuesta, una guía de estudio, con la finalidad de que el lector haga una aproximación de cómo históricamente, surge la distribución normal. Se muestran dos vías: una desde el desarrollo de los conceptos de probabilidad, que llamaremos Rama de la Probabilidadz denotaremos RPb, y otra desde la búsqueda de la distribución de los errores en las mediciones experimentales, que llamaremos Rama de los erroresz denotaremos Ree. Cada una de estas vías asocia diferentes conceptos, épocas y personajes, a partir de los cuales es posible explorar y profundizar en muchos temas tales como los primeros conceptos de probabilidad, las primeras versiones de la ley de los grandes números y del teorema del límite central, la distribución binomial de J. Bernoulli, la primera versión de la distribución normal encontrada por A. De Moivre y las diferentes propuestas de distribución de los errores hechas desde Galileo hasta Gauss pasando por las ideas de T. Tompson, D. Bernoulli, L. Euler, P. S. Laplace y A. L. Cauchy

Palabras claves: Aproximación de la Binomial a la Normal, Argumentos combinatorios, Argumentos recursivos, Curva de los errores, Distribución de los errores, El problema de los puntos.

1. INTRODUCCIÓN

El propósito es reconstruir el surgimiento de la distribución normal. Por supuesto que esto no es tarea fácil, pero vale la pena este intento, porque al comprender la historia de las ideas, sus derroteros, sus dificultades y sobre todo la manera y los contextos en que surgieron, la asimilación (aprendizaje) y la "apropiación" del conocimiento desde todo punto de vista, teórico y práctico, se verá enriquecida con elementos que en los actuales currículos de las carreras profesionales no aparecen. La idea es proporcionar a la comunidad académica herramientas que, al ser valoradas, lleguen a sembrar inquietudes, de tal manera que sea posible dar sustento a expresiones como la que usa Parzen y que van en este mismo sentido:

"Comprenderemos mejor la naturaleza de los fenómenos aleatorios gobernados por una ley normal si examinamos la manera como surgieron"[Parzen, 1973].

Es necesario precisar que, desde lo histórico, puede haber muchas versiones o visiones de un mismo tema o de unos mismos hechos y por lo tanto debe ser claro que aquí se presenta una visión particular. Si el estudio histórico de la distribución normal conduce a este tipo de reconocimientos, creo que por lo menos, en lo tocante a esta nota, se habrá cumplido en buena parte al haber hecho el intento, así sea de una manera resumida, de seguir las ideas que llevaron a la conformación de un "descubrimiento" como este.

Este trabajo puede verse como una propuesta para estudiar algunos personajes que tuvieron mucho que ver con la aparición de la distribución normal, incluso antes de las publicaciones de Gauss en 1809, tales como Pascal, Fermat, Huygens, J. Bernoulli, De Moivre, Laplace y otros. Esto define por lo menos dos ramas o vertientes que están ligadas a las primeras ideas sobre probabilidad y que asocian, a su vez el trabajo de Bernoulli sobre la distribución binomial y más atrás, a Pascal y Fermat cuando abordan las soluciones al problema de los puntos (PP. Ver versión más adelante). A esta rama se le llamará "Rama de Probabilidad" (RPb) para diferenciarla de un segundo camino que se propone y que parte de la preocupación de la ciencia por determinar o encontrar una curva de distribución de los errores que se comenten en todo proceso de medición experimental y que se podría llamar "Rama de los Errores" (Ree). Por este lado será necesario referirse a Galileo Galilei, J. J. Tompson, D. Bernoulli, Laplace y Gauss. La segunda de estas ramas (Ree) está referenciada en el artículo de [Stahl, 2006].

La Figura 1 esquematiza un posible "encadenamiento" o "derrotero" visto en retrospectiva desde la época actual y pretende aclarar mediante un "mecanismo" de implicaciones como pudo emerger la distribución normal siguiendo estos dos caminos.

En lo que sigue se recorrerán estos caminos paso a paso, iniciando con la descripción detallada de la rama RPb seguida de la de la rama Ree y de algunas consideraciones finales.

2. EL CAMINO DE LA PROBABILIDAD

La aproximación normal a la distribución binomial tiene gran valor teórico y práctico, condujo al primer teorema límite. La primer versión del Teorema Central del Límite fue dada por De Moivre en su libro "The Doctrine of Chances" (1733), para el caso especial en el que el parámetro de la binomial es $p = \frac{1}{2}$ (cuadro Dp en la figura 1). Laplace generalizó al caso en que p es arbitrario (cuadro Ep en la figura 1).

Como se dijo, los caminos que se presentan en la Figura 1 son muy particulares y están asociados a la visión del autor. En un primer nivel (Nivel Ap) de la RPb se esquematizan los trabajos previos y que se pueden considerar la génesis de la teoría de las probabilidades, en esencia tocan con un problema muy sencillo sobre juegos: "el problema de los puntos" (PP). Se destacan los trabajos en el renacimiento, de Luca Pacioli, Tartaglia y G. Cardano (1494, 1556, 1565). Ver [Martín Casalderrey, 2000]. En esta misma referencia se puede encontrar una versión de problema de los puntos.

Hay registros sobre el PP desde 1380, es posible que su origen sea árabe. Luca Pacioli (1446 - 1517), en su obra *"Summa de arithmetica, geometria, proportioni et proportionalita"* publicada en 1494 ydonde encontramos temas de geometría, aritmética y contabilidad comercial, presenta la versión siguiente del PP: *"A y B participan en un juego de lanzamiento de pelota y acuerdan que el juego terminará hasta que alguno de ellos gane 6 rondas, pero el juego se suspende abruptamente cuando Aha ganado 5 rondas a su favor y B sólo tiene 2 rondas a su favor. ¿Cómo debería distribuirse la apuesta?"*.

El propio Pacioli da soluciones basadas en sencillas proporciones referidas al número de puntos obtenidos por cada jugador antes de suspender el juego.

Este problema es retomado en los trabajos que esquematizamos en el segundo nivel (Nivel Bp figura 1), donde referimos los trabajos de Pascal, Fermat y Huygens (1654, 1657). Los primeros desarrollos de la teoría de la probabilidad tuvieron un gran avance en Francia en la mitad del siglo XVII. Uno de los hechos importantes fue la correspondencia entre Blaise Pascal y Pierre de Fermat durante 1654. Los métodos usados por ellos eran novedosos en la época. Fermat utilizó lo que hoy llamamos métodos combinatorios.

* MSc en Ingeniería de Sistemas, Universidad del Valle. Matemático, Universidad del Valle. Estadístico, Universidad del Valle.

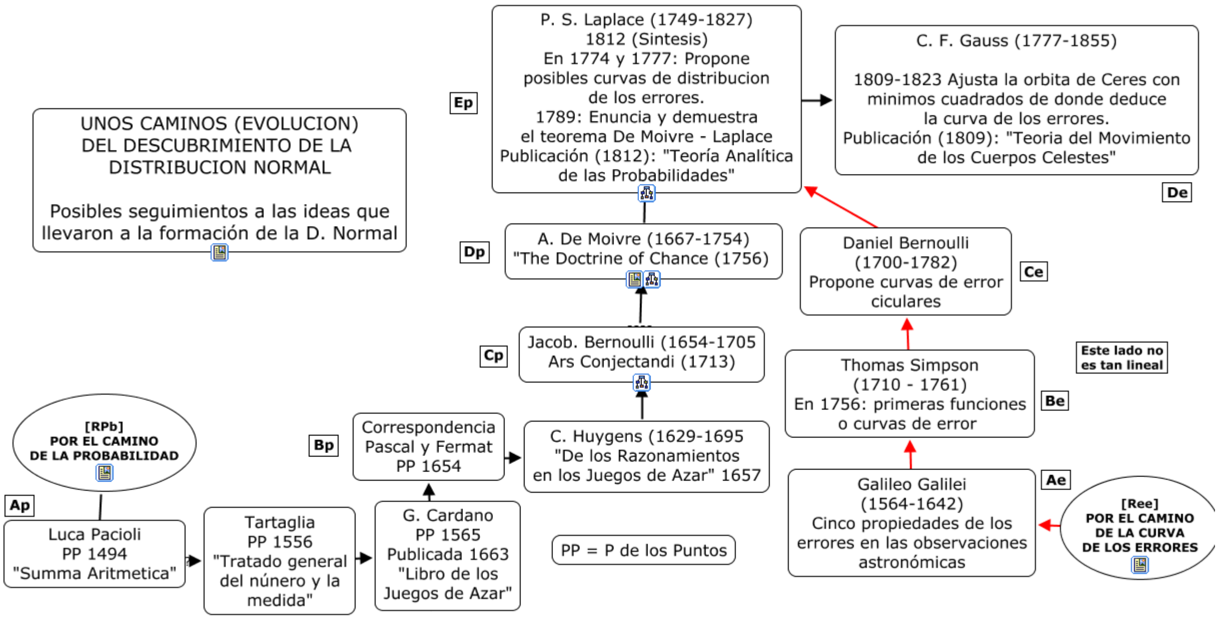

Figura 1. Dos posibles caminos del "descubrimiento" de la distribución normal. El camino RPb inicia en el óvalo de la parte inferior izquierda, mientras que el camino Ree inicia en el óvalo de la parte inferior derecha.

Los utilizó para calcular las posibilidades de los futuros lanzamientos de cada jugador en el PP (esto diferencia los métodos utilizados por Pastal y Fermat de los métodos usados por Paciolli y Tartaglia) con los que se trata de "estimar" las posibilidades que tiene cada jugador en referencia a la finalización del juego. Esta "posición futurista" para resolver el PP es la que relaciona los primeros conceptos de probabilidad.

En 1665, Pascal publica su "Tratado sobre el Triángulo Aritmético". El libro comienza con la construcción del llamado triángulo de Pascal, (aunque era conocido desde hacía más de 500 años en diversas partes del mundo). Pascal demostró algunas propiedades importantes del triángulo. Para hacer sus demostraciones usaba algo parecido al principio de inducción, pues demostraba un caso y a continuación, mostraba que eso implicaba el caso inmediatamente siguiente. Pascal aplicó estos resultados para dar una solución sistemática del problema de los puntos. Sus métodos se pueden denominar "métodos de argumentos recursivos". Fundamentalmente el triángulo de Pascal relaciona argumentos combinatorios con argumentos recursivos. Invitamos al lector a mirar y deducir o repasar algunas de estas relaciones con sólo mirar dicho triángulo o revisar el propio texto de Pascal "Tratado del Triángulo Aritmético" (Ver el compendio *"Pascal. Las Provinciales, Opúsculos, Cartas, Pensamientos, Obras Matemáticas, Obras Físicas"*. Editorial Gredos. Madrid (2012).

Por ejemplo, en el triángulo de Pascal de la figura 2 se puede apreciar la sucesión de Fibonacci con su "contenido recursivo": $F_n = F_{n-2} + F_{n-1}$. Esta relación entre lo combinatorio y lo recursivo es fundamental en la discusión epistolar entre Pascal y Fermat.

En 1655 Christian Huygens (holandés) conoció en París los problemas tratados en la correspondencia entre Pascal y Fermat, sobre la solución de este tipo de problemas y publicó en 1657 su obra *"De los razonamientos en los juegos de azar"* que es considera la primera obra escrita relacionada con el cálculo de probabilidades.

Huygens introdujo el concepto de esperanza matemática como una generalización de la media aritmética. En la época ya se utilizaba la media aritmética para calcular los precios o las ganancias medias. *"Para Huygens la esperanza significaba, por ejemplo, el precio medio de las oportunidades de ganar"* [Hald, 1990].

Estrechamente enlazado con el anterior, se define un tercer nivel histórico, representado en la figura 1 por Cp y hace referencia a Jacob Bernoulli (1654 – 1705) quien escribió el "Ars Conjectandi" publicado después de su muerte en 1713. En esta obra hace comentarios sobre los problemas que aborda Huygens. Sus métodos resaltan la importancia de la esperanza matemática. Analizó el problema de los puntos y concluyó que eran más importantes las posibilidades que tenía cada jugador para ganar y deberían no tener importancia las que ya se había realizado.

Bernoulli generaliza el problema de los puntos al caso en el que los dos jugadores no tuviesen la misma probabilidad de ganar. Llegando así a la distribución Binomial. Para hacer sus cálculos desarrolló "el arte de la combinatoria" que influenció en el desarrollo del concepto de probabilidad pues llevaron a problemas más interesantes relacionados con el azar.

La cuarta parte de su libro es considerada la parte más

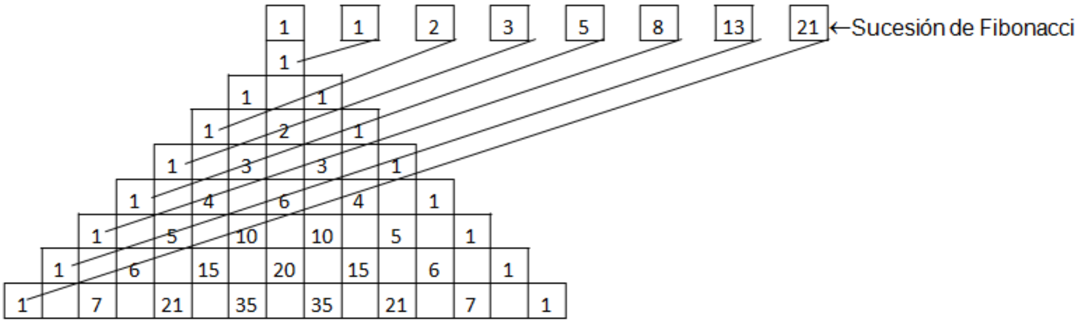

Figura 2. El Triángulo de Pascal y la sucesión de Fiboncci

fundamental de su trabajo, en ella toca con problemas de naturaleza distinta de la de los juegos de azar como problemas de tipo económico y social. Buena parte de las aplicaciones actuales hacen referencia al concepto estadístico de probabilidad (frecuencia) que aparece por primera vez en esta cuarta parte del "Ars Conjectandi" sustentada en la conocida ley de los grandes números (llamada así posteriormente). En sus métodos se reconocen ya la influencia de los "nuevos" conceptos que llegan con el desarrollo de las primeras ideas del cálculo a finales del siglo XVII. Más adelante en otros trabajos como los de De Moivre, se recogen estos métodos. El siguiente texto tomado de [Newman, 1968] muestra la idea central de Bernoulli con sus propias palabras

"... en estos casos tenemos otro camino y aquello que no es posible obtener a priori puede al menos obtenerse a posteriori, esto es, por múltiples observaciones de los resultados en ejemplos semejantes ... Aunque esto naturalmente es conocido por todos, la demostración construida sobre fundamentos científicos, en general, no es tan frecuente, por esto necesitamos exponerla. Precisamente se trata de investigar si ante tal aumento del número de observaciones, la probabilidad alcanzará realmente la relación entre el número de los casos bajo los cuales un cierto suceso puede ocurrir o no ocurrir ... o sea, si el problema tiene una asíntota...". [Newman, 1968].

El nivel Dp de la figura 1 ubica a Abraham De Moivre (1667 – 1754) con sus trabajos publicados que van desde 1712 hasta 1756. Para el estudio de los trabajos de De Moivre asociados con la distribución Normal proponemos una revisión de por lo menos tres textos, a saber *"The Doctrine of Chance" (1756), "Miscellanea Analytica" (1730) y los desarrollos que hace [Hald, 1990]* en su texto *"A History of Probability and Statistics and Their Applications befote 1730"*. En cuanto a *"The Doctrine of Chance"* salieron tres ediciones (tenemos una versión del 2005 de la 3^a edición de 1756 (*www.ibiblio.org/chance*). Ver comentarios en el libro de A. Hald (Cap 22, página 408) sobre estas tres ediciones. Para consultar el libro "Miscellanea Analytica" visitar: *http://www.math.dartmouth.edu/~euler/*, "The works of Leonhard Euler online". La figura 3 sirve de guía para el estudio de la obra de De Moivre.

De Moivre estudió detalladamente la binomial para n grande y asoció a sus expresiones las series infinitas convergentes (que recién se estudiaban). Llegó a la expresión fundamental que aparece en la ecuación 1 para distribuciones binomiales $b(x, n, p)$ simétricas con $p = \frac{1}{2}$, donde d es una distancia a la media np [esta notación es sugerida por [Hald, 1990]]

$$b(np + d, n, p) \cong \frac{1}{\sqrt{2\pi npq}} e^{\frac{-d^2}{2npq}} \tag{1}$$

Luego sin ser explícito deduce la integral de la ecuación 2. Aquí ya se esboza, por primera vez, la forma de la función de distribución (la normal) que aproxima la binomial de Bernoulli para valores de p que hacen simétrica tal distribución.

$$P_d = \sum_{|X-np|\leq d} b(x, n, p) \cong \frac{4}{\sqrt{2\pi}} \int_0^{\frac{d}{2}\sqrt{npq}} e^{-2y^2} dy \tag{2}$$

Tal demostración de De moivre, se puede ver como una "transición" importante entre dos grandes épocas. Una enmarcada por los cálculos combinatorios con personajes como B. Pascal, C. Huygens, James y Nicholas Bernoulli (segunda mitad siglo $XVII$ y principios del $XVIII$). (Aunque en el "Arte de las Conjeturas" aparecen algunos intentos de aplicar el naciente análisis matemático a los cálculos aproximados de los números combinatorios).

Realmente DeMoivre basa su deducción en cuatro teoremas ya conocidos en la época:

1. Producto infinito de $\frac{\pi}{2}$ (Wallis: 1655).
2. La serie infinita de Newton para el logaritmo: $ln\left[\frac{(1+x)}{(1-x)}\right]$ (decada 60's siglo $XVII$).
3. La fórmula de Bernoulli para la suma de potencias de enteros (1713).
4. La aproximación de Stirling (1725).

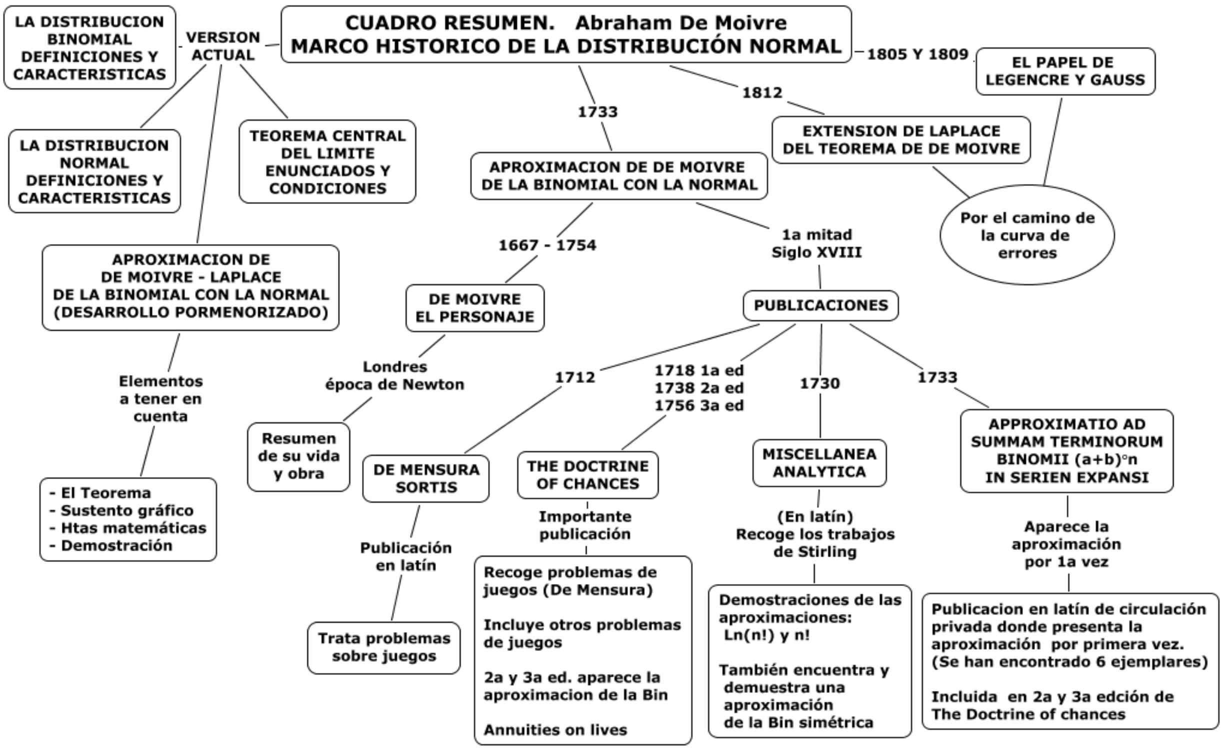

Figura 3. Guía de estudio de la obra de A. De Moivre

La otra época, principios del siglo XIX, llega con Laplace y Gauss (Nivel Ep, figura 1) quienes reafirman los métodos del análisis matemático, no solo para aproximar los cálculos de la binomial sino para hablar en propiedad de otra distribución (la normal) que desde los trabajos de Gauss tiene otro tipo de utilización (diferente a la aproximación binomial) como es el de modelar los errores experimentales en las mediciones astronómicas. Tema que venía siendo de interés por lo menos desde la época en que Galileo Galilei escribió los *"Diálogos Sobre los Dos Máximos Sistemas del Mundo"* y donde aparece por primera vez, la idea de desarrollar un estudio sistemático de los errores en las mediciones experimentales sobre todo las referidas a la ciencia de la astronomía.

Según [Hald, 1990], Galileo Galilei considera, que los errores en observaciones de fenómenos celestes, cumplen con los siguientes 5 puntos:

1. Hay solamente un número, el cual da la distancia de una estrella al centro de la Tierra ola distancia angular entre objetos. La verdadera distancia.

2. Todas las observaciones tienen errores debido al observador, los instrumentos y otras condiciones observacionales.

3. Las observaciones están distribuidas simétricamente alrededor del verdadero valor.

4. Errores pequeños ocurren más frecuentemente que errores grandes.

5. La distancia calculada es función de las observaciones angulares directas, tal que, pequeños ajustes en las observaciones pueden resultar en un gran ajuste de la distancia.

3. EL CAMINO DE LOS ERRORES

Se pasa de esta manera a la rama (ascendente) derecha del la Figura 1. Tal como se muestra en ella, se parte de las consideraciones hechas por Galileo Galilei. Si se sigue esta ruta (Niveles Ae, Be, Ce y De Figura 1) se identifican diferentes pensadores que de una u otra forma trabajaron y contribuyeron en la búsqueda de una curva que representara la distribución de los errores experimentales en las mediciones. Por ejemplo, sin presentar detalles, se muestran en la figura 4 las propuestas que elaboró Thomas Simpson's (1710-1761) hechas en 1756 a partir de series muy particulares.

Daniel Bernoulli (1700-1782) contribuyó en esta "búsqueda en la oscuridad" de un patrón de distribución de los errores experimentales proponiendo una curva circular para ésta distribución.

En esta misma línea se encuentra P. S. Laplace (1749-1827)(Nivel De, de la figura 1) quien paradójicamente, generalizó el teorema de De Movre para aproximar la binomial, pero no asoció estos resultados con la curva de los errores. En cambio, P. S. Laplace sí propuso algunas formas de distribución tal como $\phi(x) = (m/2)e^{-m/x}$ o como

Figura 4. Distribución de los errores. Dos distribuciones propuestas por Thomas Simpson's (1756)

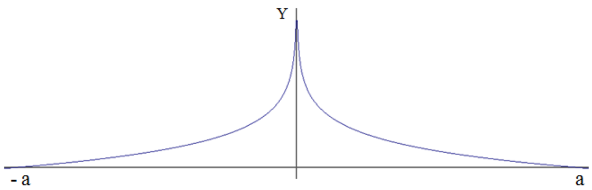

Figura 5. Segunda curva de error. Laplace

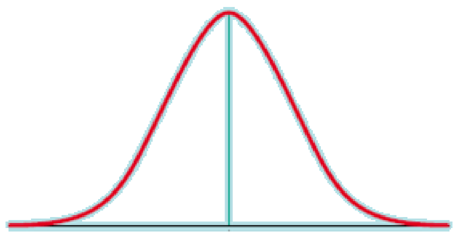

Figura 6. La Normal de Gauss de la Ecuación 3

$\phi(x) = (1/2a)ln(\frac{1}{|x|})$ que se representa en la Figura 5. De todas maneras Laplace contribuyó enormemente en la consolidación de la teoría de las probabilidades. En este sentido, organizó en forma coherente y sistemática las ideas dispersas sobre la probabilidad de la época. Se puede decir que Laplace conjuntamente con Gauss produjeron una síntesis de tales ideas.

El 1º de enero de 1801 fue descubierto el planetoide Ceres. Tiempo después era necesario calcular su órbita con base a las leyes de la mecánica celeste.

F. Gauss (1777 – 1855) hizo la mejor predicción de la posición de Ceres, en trabajos publicados posteriormente, uno en 1809, titulado *"Theoria motus corporum coelestium in sectionibus conicis solem ambientum"* y otro en 1821 titulado *"Theoria combinationis observationum erroribus minimis obnoxiae"*. Gauss usó el criterio de los mínimos cuadrados para calcular la órbita que mejor ajustara a las observacionesy se basó en las siguientes afirmaciones o principios:

1. Pequeños errores son más probables que grandes errores.
2. Las probabilidades de cometer errores de magnitud e y –e son iguales.
3. En presencia de varias medidas de la misma cantidad el valor más probable es la media.

De esta manera, partiendo de $\phi(x)$ como la función de densidad de probabilidad de los errores demostró que de todas las funciones $\phi(x)$ simétricas, diferenciables y con un solo máximo, existe una única distribución: la distribución normal.

A diferencia de Laplace y utilizando métodos análogos a los de la máxima verosimilitud, llega a la ecuación diferencial: $\frac{\phi'(x)}{\phi(x)} = kx$, donde k es una constante de proporcionalidad. Integrando y haciendo algunos cambios de variables, llega a la expresión dada en la Ecuación 3. Gauss interpretó h en la Ecuación 3 como una medida de la precisión de las observaciones.

$$\Phi(x) = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2} \tag{3}$$

4. ALGUNAS IDEAS ADICIONALES

Después de Gauss hasta bien entrado el siglo XX se encuentran muchos aportes incluso fuera del contexto de las ciencias naturales, que es necesario revisar o estudiar, referidos a la utilización de la normal como distribución de probabilidad.

Adolphe Quetelet (1796 – 1874) hizo aplicaciones a las ciencias sociales publicadas en: *"Sobre el hombre y el desarrollo de las facultades humanas: Ensayo sobre física social"* en 1835 y *"La antropometría, o medida de las diferentes facultades del hombre"* en 1871.

Poisson (en 1837) y Cauchy (en 1857) utilizaron posteriormente la función $f(x) = \frac{1}{\pi(1+x^2)}$ como curva de errores de las observaciones. Con esta función hoy se reconoce la función de densidad de una distribución Cauchy.

A mediados del siglo XIX se establece la Normal como curva generalizada de la distribución de los errores experimentales.

Franscis Galton (1822 - 1911), en el año 1873, construyó un ingenioso dispositivo hecho con clavos sobre un tablero, llamado *quincunx*. Su esquema aparece en la Figura 7 y se utiliza para comprobar cómo se

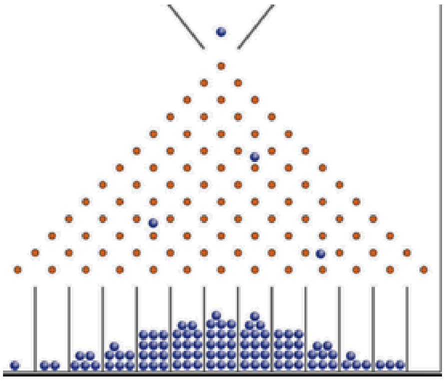

Figura 7. Un esquema del Quincunx, la "máquina de Galton". (Fuente:*http://crashoil.blogspot.com/2012/10/una-leccion-de-fisica-estadistica.html*)

distribuyen, en la parte inferior, las bolitas cuando se abre una llave que cierra el recipiente superior, donde están almacenadas inicialmente un número grande de bolitas. Estas chocan aleatoriamente contra los clavos (representados por puntos en la figura) y se depositan en la parte inferior formando una distribución muy especial. ¿Cuál cree el lector que es la forma de la distribución en que quedan las bolas en la parte inferior cuando todas han caído?

REFERENCIAS

[Hald, 1990] Hald, A. (1990). *A history of probability and statistics and their applications before 1750*. John Wiley & Sons.

[Martín Casalderrey, 2000] Martín Casalderrey, F. (2000). Cardano y tartaglia. las matemáticas en el renacimiento italiano. *La matemática en sus personajes*, 4.

[Newman, 1968] Newman, J. R. (1968). *Sigma: el mundo de las matemáticas*. Grijalbo, Barcelona, 10 edition.

[Parzen, 1973] Parzen, E. (1973). *Teoría moderna de probabilidades y sus aplicaciones*. Limusa.

[Stahl, 2006] Stahl, S. (2006). The evolution of normal distribution. *Mathematics Magazine*. University of Kansas.