Miscelánea Matemática **33** (2001) 69–92 SMM

La probabilidad en el siglo XX

Luis G. Gorostiza
Departamento de Matemáticas
Centro de Investigación y de Estudios Avanzados del I.P.N.
Apdo. Postal 14–740
07000 México, D.F.
gortega@servidor.unam.mx

Un coup de dés
jamais n'abolira le hasard
S. Mallarmé

Resumen

Se presenta un bosquejo del desarrollo de la teoría de la probabilidad durante el siglo XX.

Abstract

An outline of the development of the theory of probability during the 20th century is presented.

1. Introducción.

Durante el siglo XX la probabilidad tuvo un gran desarrollo. En los años treintas se transformó de "cálculo de probabilidades" en "teoría de la probabilidad". El cálculo de probabilidades ya tenía algunos de los ingredientes básicos de la teoría, pero consistía principalmente en una colección de problemas computacionales con ideas intuitivas y resultados poco precisos. La noción misma de probabilidad era confusa. Henri Poincaré afirmó en su libro "Calcul des probabilités" (1900) que no se podía dar una definición satisfactoria de probabilidad. Hoy en día la teoría de la probabilidad es una rama importante de las matemáticas, con sus propios conceptos, métodos y resultados. Tiene relaciones profundas con otros campos de las matemáticas e innumerables usos teóricos y aplicaciones prácticas en otras ciencias y en la tecnología, así como implicaciones que nos afectan más de lo que nos imaginamos.

En este artículo mencionaré algunas de las aportaciones de los siguientes protagonistas de la teoría de la probabilidad en el siglo XX: A. N. Kolmogorov, P. Lévy, N. Wiener y K. Itô. Otros destacados matemáticos han hecho contribuciones valiosas, pero las ideas de estos cuatro son las que han tenido mayor impacto en el desarrollo de la teoría.

También daré algunas ideas sobre la ubicación de la probabilidad dentro de las matemáticas y sus relaciones con otras disciplinas.

No iré más allá de un vistazo, pensando en lectores que no tienen conocimientos de probabilidad.

Sobra decir que este artículo refleja en parte mi visión de las cosas.

2. Un poco de historia.

¿Qué era la probabilidad antes y a principios del siglo XX? La historia es extensa pero seré muy breve.

En tiempos del primer emperador romano, Augusto (-63 a 14), ya eran comunes los juegos de azar y se hacían tablas de mortandad. Éste es el origen de la probabilidad y la estadística. Posteriormente estas dos disciplinas se fueron separando debido a sus distintos objetivos (como la hicieron también las matemáticas y la física), pero sin dejar de estar relacionadas.

En el Renacimiento Italiano (siglo XVI) había discusiones filosóficas sobre la probabilidad (Fra Luca Pacioli, Celio Calcagnini, Nicola Tartaglia) y Gerolamo Cardano fue uno de los primeros en hacer un tratamiento matemático del azar.

Entre los siglos XVII y XVIII, a la vez que surgía el cálculo infinitesimal y la idea de límite (Newton y Leibniz), se hicieron importantes avances en probabilidad (Blaise Pascal, Pierre de Fermat, Christianus Huygens, Jakob Bernoulli, Daniel Bernoulli). De este periodo sobresale la "ley de números grandes" (J. Bernoulli), que es un teorema de límite fundamental en muchas aplicaciones de la probabilidad. Una versión sencilla de dicha ley se encuentra en los lanzamientos sucesivos de una moneda: las proporciones de "águila" y "sol" tienden a estabilizarse a medida que aumenta el número de repeticiones.

Entre los siglos XVIII y XIX se desarrollaron los conceptos de independencia de eventos y probabilidad condicional, y surgió el "teorema de límite central" (Abraham De Moivre, Thomas Bayes, Pierre Simon de Laplace, Karl Friedrich Gauss, Simeon Denis Poisson; el nombre "teorema de límite central" se debe a G. Polya). La "ley Gaussiana", dada por la famosa "campana de Gauss", que aparece como límite en el teorema de límite central, no es la contribución matemática más notable de Gauss, pero no es exagerado decir que es la que más impacto ha tenido. Tan trascendente es, que aparece en el billete de 10 marcos alemanes junto a la efigie de Gauss. Esta ley fue obtenida y usada por Gauss en su "teoría de los errores", que desarrolló en relación con observaciones en astronomía y geodesia. (Cabe decir que la ley gaussiana se llama también "ley normal" y antes de Gauss ya era conocida por De Moivre y Laplace). Más adelante veremos una forma sencilla del teorema de límite central.

A fines del siglo XIX y principios del XX hubo contribuciones importantes de A.A. Markov, A.M. Liapunov, P.L. Chebyshev, H. Poincaré y otros.

A principios del siglo XX uno de los problemas científicos más importantes era comprender el "movimiento Browniano". Éste es un fenómeno que observó en el microscopio e intentó explicar (sin lograrlo) el botánico Robert Brown en 1828. Se trata del movimiento caótico de partículas de polen en el agua causado por choques con las moléculas vecinas. En 1905, sin conocer los trabajos de Brown, Albert Einstein elaboró una teoría que predecía este fenómeno. Su objetivo era demostrar la existencia de átomos de un tamaño definido. El trabajo de Einstein tuvo importantes repercusiones en la física. El lector interesado en este aspecto del movimiento Browniano puede consultar el libro de E. Nelson, "Dynamical Theories of Brownian Motion" (1967). Otro acontecimiento singular fue la tesis de Louis Bachelier (alumno de Poincaré) en 1900, que es la primera investigación matemática sobre el movimiento Browniano y fue hecha en relación con las fluctuaciones de los precios de las acciones (en la Bolsa de París). Este trabajo fue casi ignorado durante mucho tiempo, pero posteriormente dio lugar a la aplicación de la probabilidad en los mercados financieros, con consecuencias "globalizadoras" para todo el mundo.

3. Un poco de probabilidad y procesos estocásticos.

Con el fin de poder explicar algo de lo que hicieron Kolmogorov, Lévy, Wiener e Itô, es necesario presentar algunas ideas sobre la probabilidad y los procesos estocásticos. Supongo un mínimo de antecedentes por parte del lector, y por esta razón expondré pocos conceptos y no haré énfasis en la precisión matemática.

Los conceptos básicos de la teoría de la probabilidad son un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ y una función $X : \Omega \to S$, donde S es un espacio topológico.

El conjunto Ω es un espacio abstracto cuyos puntos ω se interpretan como "eventos elementales". $\mathcal{F}$ es una colección de subconjuntos de Ω que tiene cierta estructura (de σ-álgebra), cuyos elementos se llaman "eventos". P es una "medida de probabilidad" sobre $(\Omega, \mathcal{F})$, es decir, una función $P : \mathcal{F} \to [0, 1]$ con ciertas propiedades (de medida), en particular $P(\Omega) = 1$. La idea intuitiva de la probabilidad es que algún ser superior, por ejemplo la diosa del azar ("Fortuna" para los romanos, "Tyché" para los griegos), escoge un elemento $\omega \in \Omega$ de tal forma que para cualquier conjunto $F \in \mathcal{F}$, la probabilidad de que el ω escogido pertenezca a F es $P(F)$; es decir, $P(F)$ es la probabilidad de que "ocurra el evento $\omega \in F$".

La función $X : \Omega \to S$ es una función "medible" respecto a $\mathcal{F}$ y a una colección de subconjuntos $\mathcal{S}$ de S generada por la topología de S (la σ-álgebra generada por los conjuntos abiertos; los elementos de $\mathcal{S}$ se llaman "conjuntos de Borel"; se denota $\mathcal{S} = \mathcal{B}(S)$). Esto significa que para cualquier $A \in \mathcal{S}$, la imagen inversa $X^{-1}(A)$ pertenece a $\mathcal{F}$. Se dice que X es un "elemento aleatorio" de S. El término "aleatorio" proviene del latín y se refiere a los juegos de dados, o más generalmente a los juegos de azar.

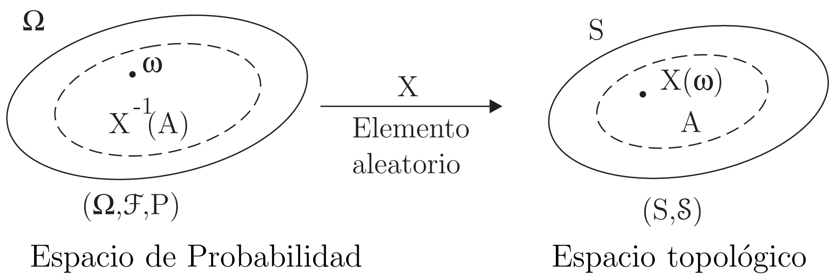
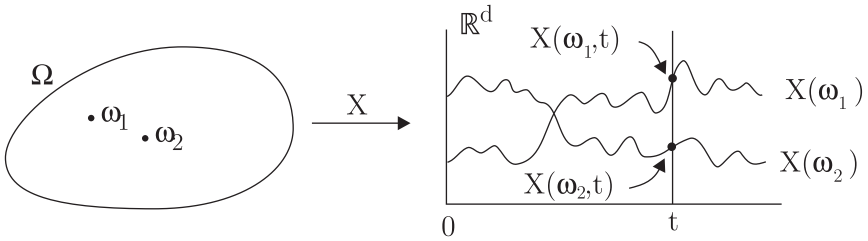


Figura 1.

En el caso de que S sea $\mathbb{R}$, los números reales, por tradición se dice que X es una "variable aleatoria", pero hay que tener presente que es una función. Si S es un espacio de funciones, por ejemplo $C([0,\infty),\mathbb{R}^d)$, el espacio de las funciones continuas de $[0,\infty)$ en $\mathbb{R}^d$, se dice que X es un "proceso estocástico". El término "estocástico" proviene del griego y se refiere a conjeturar. Un proceso estocástico X se puede ver como una función de dos variables, $\omega \in \Omega$ y $t \in [a,b]$ (algún intervalo de $\mathbb{R}$); frecuentemente t se refiere al tiempo. Así, por ejemplo, la función $t \mapsto X(\omega,t)$ es el elemento de $C([0,\infty),\mathbb{R}^d)$ que corresponde al ω escogido por la diosa del azar, y se llama la "trayectoria" del proceso X correspondiente a ω.



Proceso estocástico

Figura 2.

En general, para un elemento aleatorio X de un espacio topológico S, dado cualquier $A \in \mathcal{S}$, ya que $X^{-1}(A) \in \mathcal{F}$, está bien definida la probabilidad $P(X^{-1}(A))$, que se denota por $P[X \in A]$ y representa la probabilidad de que, habiendo escogido ω la diosa del azar, "ocurra el evento $X(\omega) \in A$". La medida $m_X(A) = P[X \in A]$, $A \in \mathcal{S}$, se llama la "distribución de X".

Como ejemplos de procesos estocásticos podemos mencionar los siguientes: la temperatura durante el día, la concentración de ozono en el aire durante el día, la presión arterial de una persona durante una semana, el precio de una acción en el mercado de valores durante una semana, los instantes de inicio y las duraciones de las llamadas telefónicas, las vibraciones de un edificio durante un sismo, el movimiento de un avión en aire turbulento, las exhalaciones del Popocatépetl durante un mes, y muchos más. A cada quien se le puede ocurrir un ejemplo de proceso estocástico con sólo reflexionar sobre los hechos que ocurren en la vida.

Dos constantes que dan una información importante sobre una variable aleatoria X (cuando existen) son el "valor medio" $E(X)$ y la "varianza" $\mathrm{Var}(X)$, que están definidas por las integrales

$$E(X) = \int_\Omega X dP = \int_{\mathbb{R}} x m_X(dx)$$

y

$$\mathrm{Var}(X) = \int_\Omega (X - E(X))^2 dP = \int_{\mathbb{R}} (x - E(X))^2 m_X(dx)\,.$$

En ocasiones se escribe $\mu = E(X)$ y $\sigma^2 = \mathrm{Var}(X)$. El valor medio es el "centro de masa" de la medida m_X en $\mathbb{R}$, y la varianza expresa la dispersión que tiene la masa de la medida m_X con respecto a su valor medio.

Con estas bases podemos mencionar el "teorema clásico de límite central": Sean $X_1, X_2, \ldots,$ variables aleatorias con la misma distribución e independientes. ¿Qué quiere decir esto? Que tengan la misma distribución significa que las medidas $m_{X_n}(A) = P[X_n \in A]$, $A \in \mathcal{B}(\mathbb{R})$, son iguales para toda n. Que sean independientes significa que $P[(X_1, X_2, \ldots, X_k) \in A_1 \times A_2 \times \cdots \times A_k] = \prod_{i=1}^{k} P[X_i \in A_i]$, para cada k y $A_1, \ldots, A_k \in \mathcal{B}(\mathbb{R})$. Por ejemplo, estas suposiciones se cumplen si las X_n son los resultados de lanzamientos sucesivos de una moneda. Denotemos por μ y σ^2 al valor medio y a la varianza comunes de las X_n.

El teorema de límite central afirma que bajo las hipótesis dadas se cumple

$$\lim_{n\to\infty} P\left[\frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} \in A\right] = \int_A \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx$$

para cualquier conjunto $A \in \mathcal{B}(\mathbb{R})$. Nótese que $n\mu$ y $\sqrt{n}\sigma$ son el valor medio y la varianza de la suma $\sum_{i=1}^n X_i$.

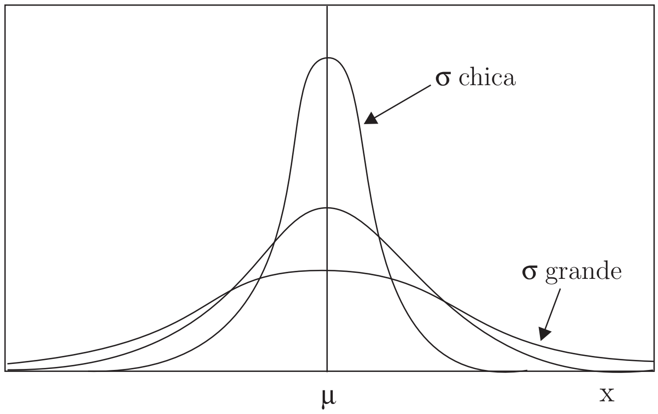
La medida

$$\Phi(A) = \int_A \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx, \qquad A \in \mathcal{B}(\mathbb{R})\,,$$

que aparece como límite, es una medida de probabilidad sobre $\mathbb{R}$ llamada la "distribución Gaussiana estándar". La función $e^{-x^2/2}/\sqrt{2\pi}$ bajo la integral es un caso especial de la función

$$\frac{e^{-(x-\mu)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}, \qquad x \in \mathbb{R}\,,$$

cuya gráfica es la "campana de Gauss". Esta función corresponde a la "distribución Gaussiana con valor medio μ y varianza σ^2".



"Campana" de Gauss

Figura 3.

El teorema de límite central y la medida Gaussiana tienen una importancia enorme. La forma que hemos visto del teorema es muy simple, pero en ella ya se aprecia la gran generalidad de este resultado: no depende de la distribución particular de las variables aleatorias X_n, sino solamente de las constantes μ y σ^2. El teorema de límite central y la medida Gaussiana ocurren con mayor generalidad: la distribución de las X_n no es necesariamente la misma, la hipótesis de independencia se puede relajar, y el espacio en el que toman valores las X_n puede ser más general. La idea intuitiva del teorema de límite central es que un fenómeno que resulta de la agregación de un gran número de causas aleatorias pequeñas semejantes exhibe un "comportamiento Gaussiano". Este tipo de comportamiento se observa en muchísimas y muy diversas situaciones. Por ejemplo, la distribución de las estaturas en una población, el sonido que emite la sección de violines de una orquesta (que es diferente al de un solo violín aumentado varias veces), y hay muchos ejemplos más. Se puede decir que el comportamiento Gaussiano es una de las leyes más universales.

Como veremos enseguida, el movimiento Browniano es una versión "funcional" de la medida Gaussiana. Más aún, también ocurre como una distribución límite (teorema de límite central funcional para caminatas aleatorias) y juega un papel tan universal como la ley Gaussiana.

El movimiento Browniano $B = \{B(t), t \geq 0\}$ en $\mathbb{R}^d$ es un proceso estocástico con trayectorias continuas, es decir, es un elemento aleatorio de $C([0, \infty), \mathbb{R}^d)$, tal que su distribución está determinada por la función de Gauss

$$p_t(x, y) = \frac{e^{-|y-x|^2/2t}}{(2\pi t)^{d/2}}, \quad x, y \in \mathbb{R}^d, t > 0\,,$$

de la manera siguiente: $B(0) = 0$ y, dados $0 < t_1 < t_2 < \ldots < t_n$ y $A_1, A_2, \ldots, A_n \in \mathcal{B}(\mathbb{R}^d)$, se tiene

$$P[B(t_1) \in A_1, B(t_2) \in A_2, \ldots, B(t_n) \in A_n]$$
$$= \int_{A_n} \cdots \int_{A_2} \int_{A_1} p_{t_1}(0, x_1) p_{t_2 - t_1}(x_1, x_2) \cdots p_{t_n - t_{n-1}}(x_{n-1}, x_n) dx_1 \cdots dx_n.$$

Ésta es la expresión de la "medida de Wiener" para un cilindro de $C([0,\infty), \mathbb{R}^d)$ con bases $A_1, \ldots, A_n$ en los puntos respectivos $t_1, \ldots, t_n$. Basta con que la medida esté definida en los cilindros para que esté determinada en todos los conjuntos de Borel de $C([0,\infty), \mathbb{R}^d)$.

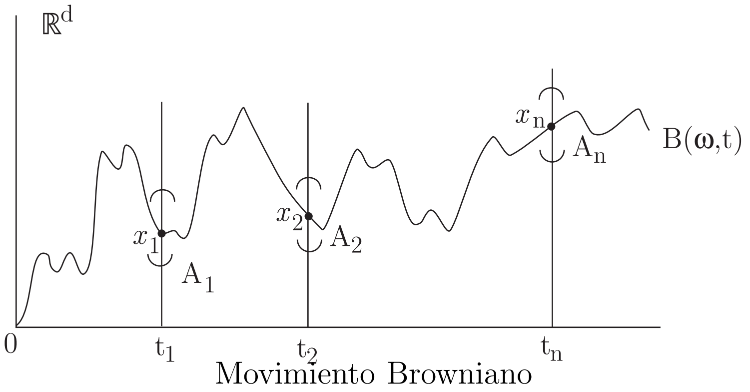


Figura 4.

Las trayectorias $t \mapsto B(\omega, t)$ del movimiento Browniano tienen muchas propiedades interesantes, una de las cuales es que no son diferenciables en ningún punto t (para P-casi todo ω).

La medida de Wiener parece muy extravagante, ya que está concentrada en las funciones no diferenciables (pero hay que recordar que las funciones diferenciables no son muchas). Sin embargo, es en cierta forma la medida de probabilidad más natural que se puede poner en $C([0,\infty), \mathbb{R}^d)$. Esto es una consecuencia de lo siguiente: Sea $X = (X_1, \ldots, X_n)$ un vector aleatorio con distribución uniforme P_n en la esfera $S^{n-1}(n^{1/2})$ de dimensión $n-1$ y radio $n^{1/2}$ ("uniforme" significa que es la medida de Lebesgue normalizada en la esfera). Sean $a_i < b_i, i = 1, \ldots k, k \leq n$, constantes cualesquiera. Se tiene entonces que

$$P_n[X_1 \in (a_1, b_1), X_2 \in (a_2, b_2), \ldots, X_k \in (a_k, b_k)] \to \prod_{i=1}^{k} \int_{a_i}^{b_i} \frac{e^{-x_i^2/2}}{\sqrt{2\pi}} dx_i$$

cuando $n \to \infty$.

Con este resultado se puede caracterizar a la medida de Wiener como la “medida de probabilidad uniforme en la esfera $S^\infty(\infty^{1/2})$ de dimensión ∞ y radio $\infty^{1/2}$”. Este interesante hecho tiene su antecedente en un trabajo de F.G. Mehler (1866) y ha sido investigado recientemente por medio del análisis no estándar. Véase N.J. Cutland y S.A. Ng, “The Wiener sphere and Wiener measure” (1993).

El movimiento Browniano es un proceso estocástico básico a partir del cual se pueden construir muchos otros procesos estocásticos, en particular los llamados “procesos de difusión”.

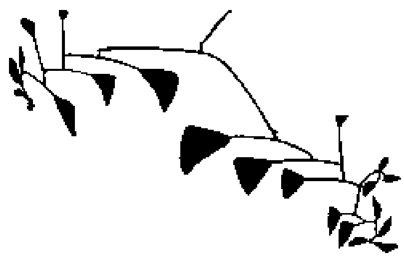
Un concepto muy importante en la teoría de la probabilidad es el de “martingala”.

Consideremos una sucesión de variables aleatorias $Z_1, Z_2, \ldots$ y la esperanza condicional $E[Z_{n+1}|Z_1, \ldots, Z_n]$ para cada n. Intuitivamente, esta esperanza condicional significa el valor medio de Z_{n+1} cuando se conocen los valores de $Z_1, \ldots, Z_n$. (Matemáticamente, la esperanza condicional es una derivada de Radon-Nikodým). Si se cumple que

$$E[Z_{n+1}|Z_1, \ldots, Z_n] = Z_n \quad \text{para toda } n\,,$$

se dice que la sucesión $Z_1, Z_2, \ldots$ es una martingala. Por ejemplo, si $Z_n = \sum_{i=1}^{n} X_i$, donde $X_1, X_2, \ldots$ son variables aleatorias independientes con valor medio 0, entonces $Z_1, Z_2, \ldots$ es una martingala. Un ejemplo de esta martingala es un “juego justo”, donde X_n es lo que se gana en la n-ésima jugada. La condición de martingala significa que después de cada nueva jugada el jugador tendrá en promedio lo que ya tenía al término de la jugada anterior.

Un modelo “físico” de una martingala son los “móviles” que se utilizan como adornos o para entretener a los bebés.



Una martingala: “móvil” (A. Calder)
Figura 5.

En el caso de tiempo continuo, un proceso estocástico $Z = \{Z(t), t \geq 0\}$ es una martingala si

$$E[Z(t)|Z(r), r \leq s] = Z(s) \quad \text{para todo} \quad s < t\,.$$

La interpretación es como en el caso de tiempo discreto visto arriba. Ejemplos de martingalas en tiempo continuo son el movimiento Browniano B y la integral estocástica con respecto a B, que veremos más adelante.

La calidad de martingala tiene propiedades muy útiles. En un problema de procesos estocásticos, si uno descubre alguna martingala escondida, ya ha dado un paso importante hacia su solución.

La teoría de las martingalas se inició con P. Lévy y comenzó a tener una gran difusión con el libro de J.L. Doob, "Stochastic Processes" (1953), que fue uno de los primeros textos importantes sobre procesos estocásticos.

Hay que mencionar también a W. Feller, cuyos libros "An Introduction to Probability Theory and Its Applications" I (1950), II (1966), dieron un gran impulso al desarrollo de la teoría de la probabilidad, en especial en los procesos de difusión.

Concluyendo, la teoría de la probabilidad es una disciplina matemática que permite decir cosas precisas acerca de los fenómenos aleatorios (azarosos, impredecibles, casuales, fortuitos, no deterministas), o por lo menos algunos de ellos. Esta descripción también vale para la teoría matemática de la estadística. Sin embargo, la probabilidad y la estadística tienen papeles distintos y en cierta forma complementarios. En la probabilidad el espacio de probabilidad ya está dado, y la teoría se refiere a la forma de hacer operaciones con elementos aleatorios definidos en ese espacio. La estadística tiene por objetivo principal establecer el espacio de probabilidad más adecuado como modelo para un fenómeno aleatorio dado. Es por ello que la estadística tiene una relación más directa con los problemas concretos y utiliza datos obtenidos por observaciones. La estadística emplea a la probabilidad como herramienta matemática, pero también tiene sus propios conceptos y métodos.

4. Las contribuciones de A. N. Kolmogorov, P. Lévy, N. Wiener y K. Itô.

Andrei Nikolaevich Kolmogorov (1903-1987).

Antes de Kolmogorov ya se había avanzado mucho en el cálculo de probabilidades, pero los conceptos y el significado de los resultados eran poco precisos. Como ya lo hemos visto, Poincaré consideraba que no se podía dar una definición satisfactoria de probabilidad. En su libro clásico, "Foundations of the Theory of Probability" (publicado por primera vez en alemán en 1933), Kolmogorov axiomatizó la teoría de la probabilidad por medio de la teoría de la medida, que estaba siendo desarrollada por varios de los matemáticos más destacados de la época (E. Borel, H. Lebesgue, M. Fréchet). Quedaron claros los conceptos de espacio de probabilidad, variable aleatoria como función medible, probabilidad condicional, independencia, la ley fuerte de grandes números y las reglas de operación del cálculo de probabilidades. Uno de los resultados más importantes de Kolmogorov es el "teorema de consistencia", que permite establecer la existencia de procesos estocásticos como elementos aleatorios de espacios de dimensión infinita. Este libro de Kolmogorov marca el inicio de la teoría de la probabilidad como una parte legítima y plenamente aceptada de las matemáticas. Otro de sus libros fundamentales es "Limit Distributions for Sums of Independent Random Variables", en colaboración con B.V. Gnedenko (1954).

Además de su trabajo sobre los fundamentos de la probabilidad, Kolmogorov hizo contribuciones importantes en varias partes de este campo, por ejemplo procesos de Markov, teoremas límites y procesos de ramificación. También realizó investigaciones en otras áreas de las matemáticas: series trigonométricas, teoría de medida e integración, teoría de aproximación, mecánica clásica, teoría ergódica, turbulencia, dinámica de poblaciones, teoría de información, estadística, teoría de autómatas, complejidad algorítmica. Además aportó en otros terrenos: metodología de la educación matemática, libros de texto, divulgación de la ciencia, análisis de historia y poesía rusas, artículos en periódicos. Sus publicaciones suman 518 y en el año 2000 apareció póstumamente un artículo suyo, en el que hace un análisis del ritmo en la poesía rusa.

Kolmogorov está entre los matemáticos más admirados del siglo XX, por la variedad de temas en los que hizo contribuciones y por lo extraordinario de éstas.

Paul Lévy (1886-1971).

En 1919 Lévy ya era un matemático muy renombrado, principalmente por su trabajo en análisis funcional, y en ese año incursionó en la probabilidad porque se le pidió que diera unas conferencias sobre las nociones del cálculo de probabilidades y el papel de la medida Gaussiana en la teoría de los errores, en la Escuela Politécnica de París.

Los trabajos de Lévy son de los más profundos y contribuyeron de manera notable a transformar el cálculo de probabilidades en teoría de la probabilidad. Una de sus principales motivaciones fue investigar la generalidad de la ley Gaussiana y del teorema de límite central, lo que condujo a formas más generales de este teorema.

Entre sus contribuciones más importantes están: Tipos de leyes límites (además de la Gaussiana), leyes estables, leyes infinitamente divisibles (representación de Lévy-Hinchin), martingalas (el concepto es creación de Lévy y el nombre fue puesto por J. Ville), procesos de Lévy (que incluyen al Browniano, al de Poisson y a los estables), teorema de continuidad de funciones características (transformadas de Fourier-Stieltjes), propiedades finas de trayectorias Brownianas.

Varias veces ha ocurrido que lo que se creían nuevos descrubimientos importantes en teoría de la probabilidad, ya estaban contenidos de alguna forma en los trabajos de Lévy.

Lévy escribió 10 libros y más de 270 artículos en probabilidad y otros campos. Sus libros más importantes en probabilidad son "Théorie de l'addition des variables aléatoires" (1954) y "Processus stochastiques et mouvement brownien" (1967). Su libro autobiográfico y filosófico "Quelques aspects de la pensée d'un mathématicien" (1970), es una lectura enriquecedora para cualquier persona interesada en las matemáticas.

Norbert Wiener (1894-1964).

A pesar de que ya en los años veintes se disponía de una metodología bastante desarrollada en el cálculo de probabilidades, ninguno de los grandes matemáticos de la época que lo intentaron (Borel, Lebesgue, Lévy, Banach, Fréchet, Kolmogorov) pudo hacer una construcción matemática rigurosa del modelo de Einstein del movimiento Browniano. Este problema fue resuelto por Wiener y condujo a la "medida de Wiener" en el espacio de trayectorias, en 1923. Como ya se ha dicho, la medida de Wiener en el espacio de trayectorias es tan universal como la ley Gaussiana lo es en dimensión finita.

Wiener también hizo aportaciones muy importantes en los siguientes temas de las matemáticas: teoría de distribuciones (que inspiró a L. Schwartz para desarrollar su teoría general), extensiones del análisis de Fourier y análisis armónico generalizado (para resolver problemas planteados por el procesamiento de señales eléctricas), teoremas Tauberianos, problemas aleatorios no lineales (teoría del caos homogéneo), teoría de predicción.

En 1933, Norbert Wiener y Arturo Rosenblueth (prestigiado fisiólogo mexicano que fue el fundador del Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional) iniciaron un seminario interdisciplinario en la Escuela de Medicina de Harvard, donde relacionaban sistemas mecánicos y sistemas fisiológicos. Estos trabajos llevaron a Wiener a la invención de la "cibernética", cuyo desarrollo tuvo lugar en parte en el Instituto Nacional de Cardiología de México. Su libro "Cybernetics, or Control and Communication in the Animal and the Machine" (1948) está dedicado a A. Rosenblueth[1].

En la Segunda Guerra Mundial (1939-1945), Wiener contribuyó a resolver el problema de mejorar el fuego antiaéreo (teoría de predicción no lineal y filtraje). Kolmogorov también desarrolló una teoría similar durante la guerra. Recuérdese que los aliados y la Unión Sovética luchaban contra la barbarie nazi.

Es importante destacar que las contribuciones de Wiener siempre tuvieron por motivo resolver problemas prácticos (su actitud no era la del matemático "puro" que no se quiere "ensuciar las manos"), pero trascendieron mucho más allá de la solución de dichos problemas, con ideas generales y abstractas de gran valor matemático y científico.

Wiener publicó más de 200 trabajos, entre los que se cuentan sus libros "Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications" (1958), "The Fourier Integral and Certain of its Applications" (1958). Sus libros de divulgación de la cibernética y sus libros autobiográficos son de gran interés general.

[1]La cibernética no llegó a constituir una teoría científica unificada porque involucra muchas disciplinas distintas que se han desarrollado de manera independiente. El término "cibernética" (del griego, ciencia de pilotear naves) ha caído en desuso excepto en algunos países del Este de Europa. A veces se escuchan términos como "ciberespacio" en películas de Hollywood o en boca de gente como comentaristas de televisión o políticos.

Kyosi Itô (1915–).

No se ha escrito mucho sobre las contribuciones de Itô porque aún vive, pero su legado en la teoría de la probabilidad es fundamental. Sus trabajos principales están en los siguientes temas: cálculo estocástico, también llamado "cálculo de Itô", ecuaciones diferenciales estocásticas, integrales múltiples de Wiener, procesos estocásticos generalizados (con valores en espacios de distribuciones). La "fórmula de Itô", que veremos más adelante, es posiblemente la fórmula más usada por los especialistas en análisis estocástico.

El cálculo estocástico es ahora una herramienta esencial en muchos campos de las matemáticas (ecuaciones diferenciales parciales, teoría de potencial, análisis armónico, geometría diferencial), así como en muchas aplicaciones (física teórica; biología, sobre todo genética y ecología; ingeniería, especialmente comunicaciones y control; matemáticas financieras).

Itô ha publicado 47 artículos y 6 libros (hasta 1987). Dos de sus libros principales son "Diffusion Processes and their Sample Paths", en colaboración con H.P. McKean, (1965), y "Foundations of Stochastic Differential Equations in Infinite Dimensional Spaces" (1984).

Es notable el hecho que Itô se inició en la teoría de la probabilidad cuando no había forma de estudiar esta disciplina en su país, el Japón, y empezó por buscar en diccionarios el significado de la palabra "estocástico". Con el tiempo formó en ese país una de las escuelas de probabilidad más importantes en la actualidad.

Debido a la importancia del cálculo estocástico, en la siguiente sección veremos de qué se trata.

5. Un poco de cálculo estocástico.

K. Itô desarrolló el cálculo estocástico por motivaciones puramente matemáticas. Quería entender y resolver rigurosamente ecuaciones diferenciales estocásticas de la forma

$$dX(t) = f(t, X(t))dt + g(t, X(t))\dot{B}(t)dt,$$

donde $\dot{B}(t)$ es un "ruido blanco", que se concibe como una derivada generalizada del movimiento Browniano $B(t)$. Esta clase de ecuaciones se presentan desde hace mucho en ingeniería electrónica y en física. Como sabemos que las trayectorias del movimiento Browniano no son diferenciables, no es claro qué sentido puede tener una ecuación de este tipo. Debido a que también aparece la variable ω en $B(\omega,t)$, que está regida por una medida de probabilidad P, no es cuestión solamente de interpretar a $B(\omega,t)$ como una derivada generalizada (en el sentido de distribuciones) de $B(\omega,t)$ con respecto a t para cada ω individual.

Escribiendo la ecuación en forma integral se tiene

$$X(t) = X(0) + \int_0^t f(s,X(s))ds + \int_0^t g(s,X(s))dB(s),$$

donde $\int_0^t g(s,X(s))dB(s)$ es una "integral estocástica". Itô definió de manera general esta clase de integrales estocásticas y estudió sus propiedades. Una de ellas es que la integral estocástica es una martingala.

La teoría desarrollada en torno a la integral estocástica llevó al "cálculo estocástico" como herramienta fundamental. Actualmente el ingrediente principal de dicho cálculo es el concepto de "semimartingala". Una semimartingala es la suma de una martingala "local" (algo un poco más general que una martingala) y un proceso estocástico con trayectorias de variación acotada y "adaptado" (una condición de medibilidad). Las semimartingalas $Z = \{Z(t), t \geq 0\}$ son los procesos estocásticos más generales con respecto a los cuales se pueden integrar procesos estocásticos, es decir, en los que está bien definida la integral estocástica $\int_0^t X(s)dZ(s)$ para una clase grande de procesos estocásticos X. Una buena referencia en español es el libro de T. Bojdecki, "Teoría General de Procesos e Integración Estocástica" (1995).

En el cálculo ordinario se tiene la regla de la cadena,

$$\frac{d}{dt}f(x(t)) = f'(x(t))x'(t),$$

que podemos expresar con la fórmula

$$f(x(t)) = f(x(0)) + \int_0^t f'(x(s))x'(s)ds.$$

En el cálculo estocástico la fórmula correspondiente a la regla de la cadena es la "fórmula de Itô" para una semimartingala $Z(\omega,t)$, $\omega \in \Omega$, $t \geq 0$:

$$\begin{aligned} f(Z(t)) = f(Z(0)) &+ \int_{(0,t]} f'(Z(s-))dZ(s) \\ &+ \frac{1}{2}\int_{(0,t]} f''(Z(s-))d[Z](s) \\ &+ \sum_{s\leq t}(\Delta f(Z(s)) - f'(Z(s-))\Delta Z(s) - \frac{1}{2}f''(Z(s-))(\Delta Z(s))^2). \end{aligned}$$

Unas aclaraciones sobre la notación: Las trayectorias de Z pueden tener saltos y se consideran continuas por la derecha; $Z(s-)$ significa el límite por la izquierda de Z en s y $\Delta Z(s) = Z(s) - Z(s-)$ es el salto de Z en s. Las integrales son sobre el intervalo semiabierto $(0, t]$. El proceso $\{[Z](t), t \geq 0\}$ se llama "proceso de variación cuadrática" de Z; se define de manera análoga a la variación de una función, pero tomando los cuadrados de los incrementos de los valores de la función en los intervalos. En la suma $\sum_{s \leq t}$ hay un conjunto de saltos que es a lo más numerable.

Si consideramos el caso especial del movimiento Browniano, $Z = B$, que es una martingala continua tal que $[B](s) = s$, la fórmula de Itô se reduce a

$$f(B(t)) = f(B(0)) + \int_0^t f'(B(s))dB(s) + \frac{1}{2}\int_0^t f''(B(s))ds.$$

Nótese que esta fórmula es como la del cálculo ordinario pero con el término adicional $\frac{1}{2}\int_0^t f''(B(s))ds$. Se tiene además el hecho de que el proceso $f(B(t)) - \frac{1}{2}\int_0^t f''(B(s))ds$ es una martingala, lo que resulta muy útil.

También cabe notar que si en la semimartingala Z la parte martingala es nula y la parte de variación acotada es determinista y diferenciable (por lo que su variación cuadrática es nula), entonces lo que queda de la fórmula de Itô es la regla de la cadena del cálculo ordinario. Así pues, el cálculo estocástico contiene al cálculo ordinario que todos conocemos.

6. La probabilidad dentro de las matemáticas.

La teoría de la probabilidad utiliza principalmente métodos del análisis matemático, sobre todo la teoría de la medida, y del análisis funcional, especialmente los espacios vectoriales topológicos. Sin embargo no es una parte de estas disciplinas porque tiene sus propios conceptos, métodos y resultados. Por ejemplo, las nociones relacionadas con las martingalas se pueden expresar de manera analítica, pero su significado cobra un sentido pleno solamente cuando se les ve a la luz de probabilidad.

La probabilidad se puede emplear como herramienta en varios campos de las matemáticas y nos da maneras nuevas de resolver problemas (deterministas, no aleatorios) por medio de modelos específicos. Entre estos campos están: ecuaciones diferenciales parciales (lineales y no lineales), teoría de potencial, análisis armónico, geometría de espacios de Banach, geometría de variedades Riemannianas, crecimiento de grupos, teoría de números (distribución de los primos), topología de variedades (teoremas de índice).

Como ejemplos clásicos de la probabilidad empleada en el análisis tenemos, para el caso de ecuaciones diferenciales parciales y teoría de potencial, el problema de Dirichlet y la fórmula de Feynman–Kac, que veremos a continuación.

El problema de Dirichlet.

Sea D un conjunto abierto y acotado en $\mathbb{R}^d$ con frontera ∂D "regular" (sin picos agudos). Sean dadas dos funciones, $g : D \to \mathbb{R}$ y $f : \partial D \to \mathbb{R}$, con ciertas condiciones de regularidad. Se trata de resolver la siguiente ecuación diferencial parcial con condición de frontera para una función $u(x)$,

$$\begin{aligned} \frac{1}{2}\Delta u &= -g(u) \quad \text{en} \quad D, \\ u &= f \quad \text{en} \quad \partial D \end{aligned}$$

(Δ es el Laplaciano: $\Delta u(x) = \sum_{i=1}^{d} \frac{\partial^2 u(x)}{\partial x_i^2}$). La solución (si existe) está dada por

$$u(x) = E_x\left(\int_0^{\tau_D} g(B(s))ds + E_x f(B(\tau_D))\right),$$

donde B es el movimiento Browniano en $\mathbb{R}^d$, τ_D es el tiempo de salida de $B(t)$ del conjunto D, es decir, $\tau_B = \inf\{t : B(t) \notin D\}$, y $E_x(\quad)$ significa el valor medio de la variable aleatoria que está dentro del paréntesis cuando el movimiento Browniano se inicia en el punto x, es decir, $B(0) = x$.

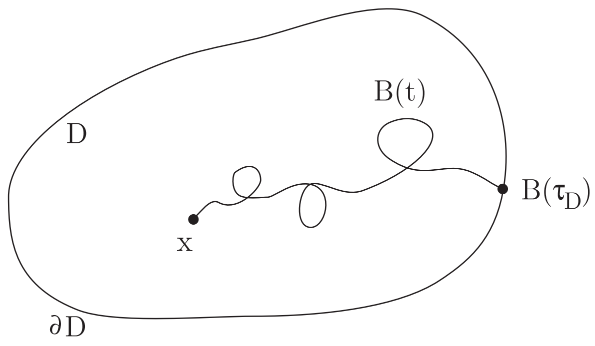


Figura 6.

Se dispone así de un modelo específico del problema de Dirichlet y de una expresión explícita para la solución, cosa que no se tiene en la teoría de potencial clásica.

La fórmula de Feynman–Kac.

Sea $V : \mathbb{R}^d \to \mathbb{R}$ una función continua acotada, y sea $f : \mathbb{R}^d \to \mathbb{R}$ una función con cierta regularidad. Consideremos la siguiente ecuación diferencial parcial con condición inicial para una función $u(x,t)$:

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{1}{2}\Delta u + Vu, \\ u(x,0) &= f(x). \end{aligned}$$

Es una "perturbación" de la ecuación del calor. La solución del problema está dada por

$$u(x,t) = E_x\bigg(f(B(t)) \exp\bigg\{ \int_0^t V(B(s))ds \bigg\} \bigg),$$

donde B es el movimiento Browniano y $E_x(\quad)$ es el valor medio condicionado a $B(0) = x$ (como en el ejemplo anterior).

Nuevamente, la solución probabilista nos da un modelo específico del problema y una solución explícita que resulta muy útil.

El meollo de la solución probabilista del problema de Dirichlet, de la formula de Feynman-Kac y de muchos otros problemas, está en el hecho de que la función

$$p_t(x) = \frac{e^{-|x|^2/2t}}{(2\pi t)^{d/2}},$$

que como recordamos determina a la distribución del movimiento Browniano, es la solución fundamental de la ecuación diferencial parcial (ecuación del calor)

$$\frac{\partial p}{\partial t} = \frac{1}{2}\Delta p\,.$$

Las versiones anteriores del problema de Dirichlet y la fórmula de Feynman–Kac tienen generalizaciones importantes y muchas aplicaciones. Estos temas se pueden consultar en el libro de O. Kallenberg, "Foundations of Modern Probability" (1997), que por cierto es uno de los mejores libros en la teoría moderna de la probabilidad.

Existen también modelos probabilistas para ecuaciones del calor no lineales y generalizaciones de ellas, así como para sistemas de tales ecuaciones. Estos modelos están basados en la idea de introducir un mecanismo de ramificación en las partículas que se mueven aleatoriamente.

Otro campo de interacción muy interesante es el de "caminatas aleatorias en grupos". Una caminata aleatoria es como un movimiento Browniano en tiempo discreto con saltos (recíprocamente, por medio del análisis no estándar se puede considerar al movimiento Browniano como una caminata aleatoria con saltos infinitesimales en intervalos de tiempo infinitesimales). En un grupo se pueden considerar las siguientes propiedades: el "crecimiento", que es un concepto geométrico, la existencia de "desigualdades de Sobolev", que es una propiedad analítica, y la "transitoriedad" o "recurrencia" de caminatas aleatorias, que son nociones probabilistas. Resulta que estas tres propiedades estan íntimamente ligadas, lo que es muy fructífero porque permite ver a cada una de ellas de distintas maneras y usar las técnicas de un campo para obtener resultados de otro. Véase por ejemplo el libro de N. Th. Varopoulos et al, "Analysis and Geometry on Groups" (1992).

7. Aplicaciones de la probabilidad.

No es sorprendente que la teoría de la probabilidad tenga muchos campos de aplicación, por sus orígenes prácticos y por la generalidad y riqueza de sus conceptos y métodos.

Algunos ejemplos de los campos en los que tiene aplicaciones la probabilidad son los siguientes:

Física (métodos matemáticos en física teórica, modelos en mecánica estadística, materiales).

Química (modelos de reacciones químicas, polímeros).

Biología (modelos en neurofisilogía, biología celular, bacteriología, genética, evolución, epidemiología, ecología, biodiversidad).

Ingeniería (comunicaciones, computación, redes y teoría de información, control y filtraje, optimización, tecnología de laser, telefonía celular).

Ciencias sociales (demografía, genética de poblaciones, sociología, antropología genética, impartición de justicia).

Finanzas y seguros (transacciones financieras, evaluación de riesgos).

Cabe destacar, por su importancia actual, que el cálculo estocástico se usa para obtener la fórmula de Black–Scholes, que se emplea en las transacciones financieras (R.C. Merton y M.S. Scholes recibieron el premio Nobel de Economía en 1997 por su uso del cáculo estocástico en finanzas). El lector interesado en este tema puede consultar el artículo de P. Protter, “A partial introduction to financial asset pricing theory” (2001), que está dirigido a un público matemático general.

En muchas de las aplicaciones actuales de la probabilidad se trata de estudiar el comportamiento global de sistemas formados por un gran número de componentes que interactúan. Esto lleva al análisis de procesos estocásticos con valores en espacios de dimensión infinita.

Una nueva rama de las matemáticas que está siendo desarrollada es la “probabilidad cuántica”, con diversas aplicaciones, entre ellas la “computación cuántica”.

En varios de los campos de aplicación, la probabilidad es indispensable porque hay en ellos fenómenos intrínsicamente aleatorios. Muchos problemas científicos y tecnológicos han sido resueltos con técnicas probabilistas y modelos estocásticos. Las comodidades de la vida moderna, con la variedad de aparatos y sistemas a los que consideramos como instrumentos usuales en la vida diaria (computadoras, automóviles, telecomunicaciones, equipo médico, etc.) no serían posibles sin las aplicaciones de la probabilidad y en general de las matemáticas. En contrapartida, también puede haber fallas por un uso inadecuado o poco cuidadoso de los modelos; piénsese por ejemplo en las crisis financieras que a todos nos afectan.

8. Conclusiones.

La teoría de la probabilidad se ha desarrollado desde la antigüedad gracias a grandes pensadores. Durante el siglo XX tuvo avances impresionantes debido a las mentes privilegiadas de A.N. Kolmogorov, P. Lévy, N. Wiener, K. Itô y otros destacados matemáticos, y tiene múltiples usos y aplicaciones.

Además de ser una teoría matemática, la probabilidad nos da una forma de pensar sobre las cosas que ocurren en el universo, desde las más abstractas hasta las más concretas, cuando tomamos conciencia de su naturaleza aleatoria. La ciencia ha demostrado que Einstein estaba equivocado en su convicción de que "Dios no juega a los dados". Los fenómenos aleatorios están en todas partes: los sucesos cósmicos, las mutaciones en nuestros genes, el vuelo de una mosca, las catástrofes de la naturaleza, la política, las crisis económicas, las relaciones amorosas, etc. Nadie escapa a la esencia azarosa de la existencia. Todos tomamos decisiones con base en estimaciones intuitivas de probabilidades que hacemos sin estar plenamente conscientes de ello. También podemos usar a la probabilidad para divertirnos con juegos o hacer composiciones musicales, como hizo Mozart.

En la probabilidad hay cosas atractivas para todos los gustos (matemáticas abstractas, aplicaciones concretas,...), y también hay mucho que pensar sobre qué es lo aleatorio y cuál es el significado de la probabilidad.

Espero que este artículo despierte el interés de los estudiantes de matemáticas por la teoría de la probabilidad. Aunque no quieran ser especialistas en probabilidad, no se arrepentirán porque tendrán una imaginación más rica y porque conocer distintas áreas de las matemáticas siembra la mente con nuevas ideas.

Agradecimiento.

Agradezco a Víctor Pérez-Abreu sus comentarios sobre una versión preliminar de este trabajo, que me ayudaron a mejorar su presentación.

Bibliografía.

Historia de la Probabilidad:

I. Todhunter, “A History of the Mathematical Theory of Probability”, Chelsea, New York (1962).

L.E. Maistrov, “Probability Theory, A Historical Sketch”, Academic Press, New York (1974).

A.N. Kolmogorov:

A.N. Kolmogorov, “Foundations of the Theory of Probability” Chelsea, New York (1956).

B.V. Gnedenko Y A.N. Kolmogorov, “Limit Distribution for Sums of Random Variables”, Addison-Wesley, Reading, Mass. (1968).

Annals of Probability, Vol. 17, No. 3 (1989). Artículos de Doob, Dynkin, Sinai, Shiryaev y otros.

“Kolmogorov in Perspective”, History of Mathematics, Vol. 20, Amer. Math. Soc. (2000). Artículos de Shiryaev, Arnol'd, Sinai y otros.

P. Lévy:

P. Lévy, “Théorie de l'addition des variables aléatoires”, Gauthier-Villars, Paris (1954).

P. Lévy, “Processus stochastiques et mouvement Brownian”, Gauthier-Villars, Paris (1967).

P. Lévy, “Quelques aspects de la pensée d'un mathématicien”, A. Blanchard, Paris (1970).

Annals of Probability, Vol. 1, No. 1 (1973), por M. Loève.

Annales de l'Institut Henri Poincaré, Vol. 23, número especial (1987), por J. Bretagnolle.

N. Wiener:

N. Wiener, “The Fourier Integral and Certain of its Applications”, Cambridge Univ. Press (1933).

N. Wiener, “Extrapolation, Interprolation, and Smoothing of Stationary Time Series with Engineering Applications”, M.I.T. Press, Cambridge, Mass. (1966).

N. Wiener,“Cybernetics, or Control and Communication in the Animal and the Machine”, Second Edition, The M.I.T. Press, Cambridge, Mass. (1961).

N. Wiener, “The Human Use of Human Beings. Cybernetics and Society”, Doubleday, Garden City, New York (1954).

N. Wiener, “I am a Mathematician. The Later Life of a Prodigy”, The M.I.T. Press, Cambridge, Mass. (1956).

N. Wiener, “God and Golem, Inc.”, The M.I.T. Press, Cambridge, Mass. (1964) (traducido al español por SIGLO XXI).

“Norbert Wiener 1894–1964”, Bulletin Amer. Math. Soc., Vol. 72 (1966), No. 1, part II.

“Norbert Wiener: Collected Works with Commentaries”, editado por P. Masani, (4 volúmenes), M.I.T. Press, Cambridge, Mass.

Notices of the A.M.S., Vol. 42, No. 4 (1995), por D. Jerison y D. Stroock.

Notices of the A.M.S., Vol. 42, No. 6 (1995), por V. Mandrekar.

“The Legacy of Norbert Wiener”, Proceedings Symposia in Pure Mathematics, Vol. 60. Amer. Math. Soc. (1997), editado por D. Jerison, I.M. Singer y D. Stroock.

K. Itô:

K. Itô y H. P. McKean, “Diffusion Processes and their Sample Paths”, Springer-Verlag, Berlin (1965).

K. Itô, “Foundations of Stochastic Differential Equations in Infinite Dimensional Spaces”, SIAM, Philadelphia (1984).

“K. Itô, Selected Papers”, Editado por D.W. Stroock y S.R.S. Varadhan, Springer–Verlag (1987).

Notices of A.M.S. Vol. 45, No. 8 (1988).

Otros textos citados:

T. Bojdecki, "Teoría General de Procesos e Integración Estocástica", Serie Textos, Vol. 6, Aportaciones Matemáticas, Soc. Mat. Mexicana (1995).

N.J. Cutland y S.A. Ng, "The Wiener sphere and Wiener measure", Ann. Probab., Vol. 21 (1993), 1-13.

J.L. Doob, "Stochastic Processes", Wiley, New York (1953).

W. Feller, "An Introduction to Probability Theory" I, II, Wiley, New York, (1950,1966).

O. Kallenberg, "Foundations of Modern Probability", Springer, New York (1997).

E. Nelson, "Dynamical Theories of Brownian Motion", Princeton University Press, Princeton, New Jersey (1967).

H. Poincaré, "Calcul des Probabilités", Carré, Paris (1900).

P. Protter, "A partial introduction to financial asset pricing", Stoch. Proc. Appl. Vol. 91 (2001), 169-203.

N. Th. Varopoulos, L. Saloff-Coste y T. Coulhon, "Analysis and Geometry on Groups", Cambridge University Press, Cambridge (1992).