Aportaciones Matemáticas
Serie Comunicaciones **35** (2005) 265–283
Artículo de Exposición

CONSTRUYENDO LA INTEGRAL ESTOCÁSTICA DE ITÔ

LUIS RINCÓN

Resumen. Se presenta de una forma sencilla y breve el concepto de integral estocástica de Itô respecto del movimiento Browniano. Con ello se ilustra además la noción de ecuación diferencial estocástica mediante algunos ejemplos.

Es cada vez más común hablar de modelos basados en ecuaciones diferenciales estocásticas en los últimos cursos de algunas carreras de ciencias. Nuestro objetivo en este trabajo es presentar de una forma sencilla y compacta el concepto de integral de Itô respecto del movimiento Browniano y con ello ilustrar el concepto de ecuación estocástica.

El plan de trabajo es el siguiente. Empezaremos recordando algunas ideas básicas de probabilidad y procesos estocásticos, particularmente mencionaremos algunas propiedades del movimiento Browniano. Definiremos después la integral de Itô respecto de este proceso en una serie de extensiones sucesivas, y finalmente mostraremos el concepto de ecuación estocástica.

1. Probabilidad y procesos estocásticos

El modelo matemático básico de la teoría de la probabilidad es el *espacio de probabilidad*, que consta de una terna ordenada $(\Omega, \mathcal{F}, P)$ en donde Ω es un conjunto arbitrario que convenientemente puede ser interpretado como el conjunto de todos los posibles resultados de un experimento aleatorio. Al conjunto Ω se le llama *espacio muestral*, y a un elemento típico de Ω se le denota por ω. El segundo elemento es una colección no vacía $\mathcal{F}$ de subconjuntos de Ω, llamada *σ-álgebra*, que es cerrada bajo las operaciones de tomar complementos y uniones numerables. A los elementos de $\mathcal{F}$, subconjuntos de Ω, se les llama *eventos* o *conjuntos medibles*. Finalmente el tercer elemento es una función $P : \mathcal{F} \to [0, 1]$, llamada *medida de probabilidad*, que cumple los siguientes axiomas (Kolmogorov, 1933): $P(\Omega) = 1$ y es σ-aditiva, es decir, si $A_1, A_2, \ldots$ es una sucesión de eventos, disjuntos dos a dos, entonces

$$P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n).$$

Figura 1. Andrey Nikolaevich Kolmogorov (Rusia, 1903-1987). Fuente: Archivo MacTutor, St. Andrews.

Las leyes del azar están representadas por las distintas medidas de probabilidad existentes. El número $P(A)$ es una medida de la frecuencia con la que se observa el evento A cuando se realiza el experimento aleatorio. Todo lo que se mencione en el resto de este trabajo tiene como estructura base un espacio de probabilidad $(\Omega, \mathcal{F}, P)$. A la pareja $(\Omega, \mathcal{F})$ se le llama *espacio medible.* En particular, si $\mathcal{B}(\mathbb{R})$ denota la σ-álgebra más pequeña que contiene a todos los intervalos abiertos de $\mathbb{R}$, entonces se tiene el espacio medible $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. A los elementos de la σ-álgebra $\mathcal{B}(\mathbb{R})$ se les llama *Borelianos* o *conjuntos Borel medibles.*

Ahora podemos recordar el concepto ubicuo de *variable aleatoria.* Una v.a. es una función $X : \Omega \to \mathbb{R}$ que transforma a los elementos de Ω en números reales y es tal que para cualquier $B \in \mathcal{B}(\mathbb{R})$, el conjunto

$$X^{-1}B = \{\omega \in \Omega : X(\omega) \in B\}$$

es un elemento de $\mathcal{F}$. En este caso también se dice que X es una *función medible* entre los espacios medibles $(\Omega, \mathcal{F})$ y $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, o simplemente que es $\mathcal{F}$-medible. Mediante una de estas funciones uno puede pensar que el azar no escoge elementos de Ω como resultados del experimento aleatorio, sino números reales. Las operaciones básicas de suma, diferencia, producto y cociente (cuando existe) de v.a.s producen v.a.s. Procesos límite de v.a.s (cuando existen) resultan también ser v.a.s.

El espacio medible $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ puede convertirse en un espacio de probabilidad con la ayuda de una variable aleatoria X de la siguiente manera. Para cada $B \in \mathcal{B}(\mathbb{R})$ se define $P_X(B) = P(X^{-1}B)$. La función P_X resulta ser una medida de probabilidad sobre $\mathcal{B}(\mathbb{R})$. Se le llama la *distribución* de X, y encierra en ella toda la información probabilística de X. Equivalentemente se estudia la función $F(x) : \mathbb{R} \to [0, 1]$ dada por $F(x) = P(X \leq x)$ y llamada *función de distribución.* Por ejemplo, la variable X tiene una distribución normal o gausiana con parámetros μ y $\sigma^2 > 0$ si su función de distribución es

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}}\ e^{-(u-\mu)^2/2\sigma^2}\, du.$$

En este caso se dice que X tiene distribución $\mathrm{N}(\mu, \sigma^2)$. De particular importancia es el concepto de esperanza de una variable aleatoria o más generalmente de una función de una variable aleatoria. Si g es una función real de variable real tal que la composición $g(X)$ es una variable aleatoria, entonces se define la esperanza de $g(X)$ como sigue

$$E(g(X)) = \int_{-\infty}^{\infty} g(x) dF(x),$$

en donde $F(x)$ es la función de distribución de X y la integral involucrada es una integral de Riemann-Stieltjes sobre la cual se asume su existencia. En particular, cuando $g(x) = x$ se obtiene la esperanza de X denotada por $E(X)$ y cuando $g(x) = (x - E(X))^2$ entonces se obtiene la varianza de X denotada por $\mathrm{Var}(X)$. Para la distribución normal mencionada antes puede demostrarse que $E(X) = \mu$ y $\mathrm{Var}(X) = \sigma^2$.

Más generalmente, la *esperanza condicional* de una variable aleatoria integrable X, dada una sub-σ-álgebra $\mathcal{G} \subseteq \mathcal{F}$, es una variable aleatoria, denotada usualmente por $E(X|\mathcal{G})$, que es integrable, $\mathcal{G}$-medible y satisface la igualdad

$$\int_G E(X|\mathcal{G}) dP = \int_G X dP,$$

para cualquier $G \in \mathcal{G}$. Estas tres propiedades caracterizan de manera única (en el sentido casi seguro) a la esperanza condicional. Particularmente haremos uso de las siguientes propiedades. Si X es $\mathcal{G}$-medible entonces X mismo cumple con la definición de esperanza condicional y por lo tanto $E(X|\mathcal{G}) = X$. Por otro lado también usaremos el hecho de que si X es independiente de $\mathcal{G}$ entonces la esperanza condicional $E(X|\mathcal{G})$ es la constante $E(X)$.

El siguiente concepto de nuestro interés es el de proceso estocástico. Un *proceso estocástico* es una colección de variables aleatorias $\{X_t : t \in \mathbb{T}\}$ parametrizada por un conjunto $\mathbb{T}$, usualmente interpretado como un conjunto de tiempos y llamado naturalmente espacio parametral. Se dice que el proceso es *a tiempo discreto* en caso de que el conjunto de índices $\mathbb{T}$ sea un conjunto discreto, por ejemplo $\mathbb{T} = \{0, 1, 2, \ldots\}$. En este caso el proceso consiste de una sucesión de variables aleatorias. En cambio se dice que el proceso es *a tiempo continuo* cuando $\mathbb{T}$ consiste de un subintervalo de $\mathbb{R}$, por ejemplo $\mathbb{T} = (a, b)$. En lo sucesivo consideraremos procesos en donde las v.a.s toman valores en $\mathbb{R}$ y el espacio parametral $\mathbb{T}$ es el intervalo $[0, \infty)$. Un proceso estocástico es entonces una función de dos variables

$$X : [0, \infty) \times \Omega \to \mathbb{R}$$

tal que para cada $t \geq 0$, la función $\omega \mapsto X_t(\omega)$ es una variable aleatoria, mientras que para cada ω en Ω, la función $t \mapsto X_t(\omega)$ es una *trayectoria* del proceso. En principio no hay ninguna condición sobre estas trayectorias, pueden ser continuas o no serlo, aunque una hipótesis común es suponer trayectorias *càdlàg*, es decir, continuas por la derecha con límite por la izquierda. Por simplicidad denotaremos un proceso por X_t anteponiendo tal adjetivo para evitar confusiones. Revisamos a continuación algunos conceptos técnicos generales relativos a procesos estocásticos.

Una familia $(\mathcal{F}_t)_{t\geq 0}$ de σ-álgebras es una *filtración* si para $0 \leq s \leq t$, se cumple $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$. Al espacio $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t\geq 0})$ se le llama *espacio de probabilidad filtrado.* Cuando X_t es $\mathcal{F}_t$-medible para cada $t \geq 0$ entonces se dice que el proceso es *adaptado* a la filtración. Todo proceso estocástico X_t determina una *filtración natural* dada por $\mathcal{F}_t = \sigma\{X_s : 0 \leq s \leq t\}$. Claramente todo proceso es adaptado a su filtración natural. En este caso a la σ-álgebra $\mathcal{F}_t$ se le interpreta como la "historia" del proceso al tiempo t, pues en ella se encuentran todos los posibles eventos o sucesos que el proceso haya tenido hasta ese momento. Adicionalmente se dice que una filtración es continua por la derecha cuando $\mathcal{F}_{t+} = \bigcap_{s>t} \mathcal{F}_s$ coincide con $\mathcal{F}_t$.

Es de utilidad también conocer alguna noción de igualdad entre procesos estocásticos. Dos procesos X_t y Y_t son *equivalentes*, o también se dice que uno es una *versión* (o *modificación*) del otro, si para cada $t \geq 0$ se cumple $P(X_t = Y_t) = 1$. Un tipo de igualdad más fuerte establece que los procesos son *indistinguibles* si

$$P(X_t = Y_t \text{ para cada } t \geq 0) = 1.$$

Esto significa que con probabilidad uno las trayectorias de los dos procesos son idénticas. Claramente la indistinguibilidad es más fuerte que la equivalencia. Sin embargo, cuando los procesos son continuos, es decir, cuando sus trayectorias son funciones continuas del parámetro, ambas nociones de igualdad coinciden.

Una característica importante que cumplen algunos procesos es la de tener *incrementos independientes*. Esta propiedad, que usaremos más adelante, puede escribirse de la siguiente forma. Para cualesquiera tiempos $t_0 < t_1 < \cdots < t_n$, las variables incremento $X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$ son independientes, es decir, la función de distribución conjunta de todas ellas coincide con el producto de las funciones de distribución individuales. Finalmente, para concluir esta breve sección mencionaremos a continuación dos tipos de procesos de conocida relevancia.

Un proceso estocástico X_t es *de Markov* (Markov, 1906) si para cada $0 \leq s \leq t$ y $A \in \mathcal{B}(\mathbb{R})$, con probabilidad uno se cumple

$$P(X_t \in A|\mathcal{F}_s) = P(X_t \in A|X_s).$$

Esta igualdad establece que el estado del proceso al tiempo futuro $t > s$ es independiente del pasado (tiempos antes de s) dado el estado del proceso al tiempo presente $s \geq 0$. Esta propiedad es equivalente a la que exhiben los sistemas dinámicos deterministas, cuya evolución queda perfectamente determinada una vez que se establece la ley de movimiento y un estado inicial del sistema, no influyendo lo sucedido antes del estado inicial. Un proceso de Markov determina por tanto una función de probabilidad de transición dada por

$$p(s, x, t, A) = P(X_t \in A | X_s = x),$$

en donde $0 \leq s \leq t$, $x \in \mathbb{R}$ y $A \in \mathcal{B}(\mathbb{R})$. Esta función de cuatro variables proporciona la probabilidad de que el proceso se encuentre en A al tiempo t dado que se encontró en el estado x en un tiempo anterior s. Inversamente, dada una función de transición de esta forma (junto con algunas hipótesis adicionales) y una distribución de probabilidad inicial, es posible construir un proceso de Markov cuya función de transición es la dada.

Otro ejemplo de proceso estocástico de interés es aquel conocido con el nombre de martingala. Un proceso X_t es una *martingala* (Lévy) si es adaptado, integrable y para $0 \leq s \leq t$, con probabilidad uno se cumple

$$E(X_t | \mathcal{F}_s) = X_s. \tag{1}$$

Las martingalas son procesos que están relacionados con los juegos justos. Si X_t representa la fortuna de un jugador que apuesta continuamente entonces la igualdad anterior se interpreta del siguiente modo. En promedio la fortuna del jugador al tiempo t dada toda la historia del juego hasta el tiempo $s \leq t$ es la fortuna del jugador al tiempo s, es decir, el juego es justo pues el jugador en promedio no pierde ni gana. Cuando en lugar de (1) se cumple $E(X_t | \mathcal{F}_s) \leq X_s$ se dice el proceso es una *supermartingala* (juego desfavorable al jugador pues en promedio su fortuna disminuye). En caso de la desigualdad contraria el proceso es una *submartingala* (juego favorable al jugador).

El clásico libro de Karlin y Taylor [8] constituye una excelente referencia general sobre el tema de los procesos estocásticos. En la siguiente sección estudiaremos muy brevemente el proceso estocástico de trayectorias continuas que posiblemente pueda considerarse de mayor importancia.

2. Movimiento Browniano

El fenómeno natural conocido ahora como movimiento Browniano tiene una larga e interesante historia. El primer registro, aunque no así la primera observación del fenómeno, data de 1828 (véase [1]) cuando el botánico Robert Brown reportó en una revista científica que granos de polen suspendidos en una cierta substancia y vistos a través de un microscopio, realizaban un movimiento irregular e inexplicable. Este extraño movimiento fue objeto de muchas discusiones, y muy diversas hipótesis fueron formuladas en ese entonces con la intención de dar una explicación al fenómeno observado. Hoy en día este movimiento es entendido y explicado a través de las múltiples colisiones aleatorias de las moléculas del líquido con los granos de polen. Llegar a tal aseveración tomó muchos años pues debió aceptarse la teoría cinético molecular de la materia, y el seminal trabajo de Einstein de 1905 sobre el movimiento Browniano [4] contribuyó decididamente a tal tarea. Las observaciones reales y directas del movimiento de los granos de polen u otras partículas sugieren que el fenómeno satisface las siguientes propiedades: (a) El movimiento es continuo. (b) Parece tener desplazamientos independientes en intervalos de tiempo disjuntos. (c) Debido al gran número de colisiones del grano de polen con las moléculas circundantes en longitudes de tiempo no pequeños, y teniendo en cuenta el teorema del límite central, los incrementos pueden modelarse como variables aleatorias gausianas.

La estructura matemática de un proceso estocástico, es decir una colección de variables aleatorias $\{B_t : t \geq 0\}$, ha resultado exitosa para modelar este tipo de fenómenos. La variable B_t puede entonces interpretarse como la posición de una partícula Browniana al tiempo t. La definición matemática, en el caso unidimensional, es la siguiente.

Definición 2.1. Un movimiento Browniano estándar unidimensional es un proceso estocástico $\{B_t : t \geq 0\}$ tal que

1. $B_0 = 0$ casi seguramente.
2. Las trayectorias $t \mapsto B_t$ son continuas.
3. El proceso tiene incrementos independientes.
4. La variable $B_t - B_s$ tiene distribución $N(0, t-s)$ para $0 \leq s < t$.

Las condiciones que aparecen en esta definición son consecuencia directa de las observaciones del fenómeno físico, pero ello no garantiza que tal objeto matemático exista. En 1923 el matemático norteamericano Norbert Wiener demostró la existencia de un proceso con tales condiciones. Es por esta razón que a menudo a este proceso también se le llama *proceso de Wiener* y se le denota también por $\{W_t : t \geq 0\}$. En sentido estricto el movimiento Browniano es el fenómeno físico mientras que su modelo matemático es el proceso de Wiener, aunque es común llamar a ambas cosas por el mismo nombre: movimiento Browniano.

FIGURA 2. Norbert Wiener (Estados Unidos, 1894-1964). Fuente: Archivo MacTutor, St. Andrews.

Se tiene entonces que cada variable aleatoria B_t tiene distribución $N(0,t)$ y por lo tanto $E(B_t) = 0$ y $\mathrm{Var}(B_t) = E(B_t^2) = t$. En particular para $0 \leq s < t$ se cumple $E|B_t - B_s|^2 = t - s$.

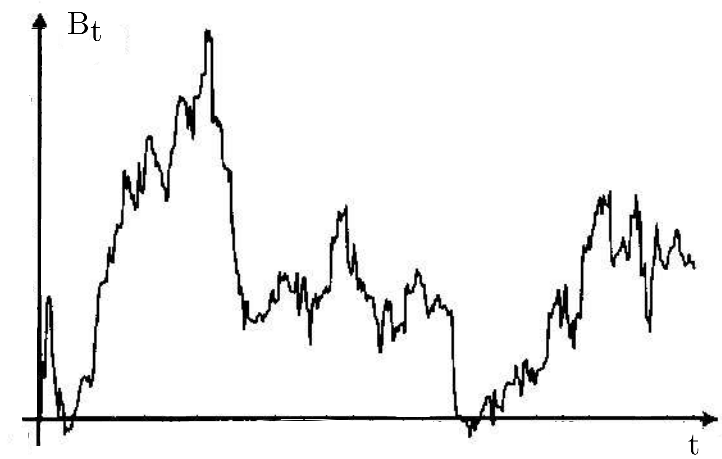

FIGURA 3. Una trayectoria Browniana.

El movimiento Browniano físico se presenta en tres dimensiones y es completamente errático. En la Figura 3 puede apreciarse una posible trayectoria Browniana cuando ésta se proyecta sobre una de sus coordenadas. Este movimiento tiene muchas propiedades y conexiones con otras ramas de las matemáticas, por ejemplo es un proceso de Markov y es también una martingala continua. Es interesante también mencionar que casi todas sus trayectorias son no diferenciables en ningún punto, las trayectorias Brownianas son entonces ejemplos de funciones, otrora consideradas extrañas, que son continuas pero no diferenciables en ningún punto. También puede demostrarse que sobre un intervalo de tiempo finito $[a,b]$, casi todas las trayectorias tienen variación no acotada. Esto es,

$$\sup_{\Delta} \sum_{i=1}^{n-1} |B_{t_{i+1}} - B_{t_i}| = \infty,$$

en donde Δ es una partición $a = t_0 < t_1 < \cdots < t_n = b$ del intervalo $[a,b]$. Esta propiedad es particularmente importante en el presente trabajo pues tiene como consecuencia el hecho de que no se pueden usar las trayectorias Brownianas como integradores en el sentido de Riemann-Stieltjes. Por otro lado puede demostrarse que la variación cuadrática sobre $[a,b]$ es

$$\sup_{\Delta} \sum_{i=1}^{n-1} |B_{t_{i+1}} - B_{t_i}|^2 = b - a.$$

Otro de los muchos resultados interesantes del movimiento Browniano es el *teorema de caracterización* de Paul Lévy que establece que un proceso cualquiera $\{X_t : t \geq 0\}$ es un movimiento Browniano si y sólo si tiene trayectorias continuas, empieza en cero, y tanto $\{X_t : t \geq 0\}$ como $\{X_t^2 - t : t \geq 0\}$ son martingalas. A través de este resultado o directamente de la definición, puede demostrarse que los siguientes procesos son versiones del movimiento Browniano: a) $X_t = \frac{1}{c} B_{c^2 t}$ con $c > 0$ constante, b) $X_t = t X_{1/t}$ para $t > 0$, con $X_0 = 0$, c) $X_t = B_{t+s} - B_s$ con $s \geq 0$ fijo.

El lector interesado puede encontrar una muy interesante y motivadora exposición histórica sobre el descubrimiento del movimiento Browniano en el excelente libro de Edward Nelson [12], ahora disponible en la red en formato electrónico en la página web del autor. Para una primera introducción a algunos aspectos matemáticos del movimiento Browniano puede consultarse por ejemplo [8].

3. Integración estocástica

Esta sección es la parte central de nuestro trabajo y la intención es definir la integral de Itô de un proceso estocástico $\{X_t : 0 \leq t \leq T\}$ respecto del movimiento Browniano, es decir, una integral de la forma

$$\int_0^T X_t dB_t. \tag{2}$$

Llevar a cabo tal tarea nos conducirá a enunciar sin demostración algunos resultados técnicos pero los esfuerzos tendrán su recompensa hacia el final del trabajo en donde haremos necesariamente uso de este tipo de integrales. Igualmente la justificación para desear definir integrales de la forma (2) se volverá evidente más adelante. Se define (2) en varios pasos. Primero para procesos simples y después, por aproximación, para procesos más generales. Consideraremos entonces como elementos iniciales un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ y un movimiento Browniano estándar unidimensional $\{B_t : t \geq 0\}$ junto con su filtración natural $(\mathcal{F}_t)_{t \geq 0}$. Asumiremos que el proceso $\{X_t : 0 \leq t \leq T\}$ visto como función $X : \Omega \times [0,T] \to \mathbb{R}$ es $\mathcal{F}_T \otimes \mathcal{B}[0,T]$-medible, y es además adaptado, es decir, para cada t en $[0,T]$, la función $X_t : \Omega \to \mathbb{R}$ es $\mathcal{F}_t$-medible.

Denotaremos por $L^2(P)$ al espacio vectorial de variables aleatorias X que son cuadrado integrables, es decir, que cumplen la condición

$$||X||_{L^2(P)} = (E|X|^2)^{1/2} < \infty.$$

La función $X \mapsto ||X||_{L^2(P)}$ define una norma[1] en $L^2(P)$ y este espacio es completo respecto de esta norma, es decir, es una espacio de Banach. Esto quiere decir que toda sucesión de Cauchy en este espacio tiene límite en él.

En lo que resta del trabajo consideraremos procesos con espacio parametral el intervalo $[0,T]$ con $T > 0$ fijo. También denotaremos por $L^2(P \times dt)$ al espacio de Banach de procesos $\{X_t : 0 \leq t \leq T\}$ que cumplen la condición

$$||X||_{L^2(P\times dt)} = (E\int_0^T |X_t|^2 dt)^{1/2} < \infty.$$

Paso 1: Integral para procesos simples. Sea $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = T$ una partición finita del intervalo $[0,T]$. Un proceso estocástico *simple* es un proceso de la forma

$$X_t = \sum_{k=0}^{n-1} X^k \cdot 1_{[t_k,t_{k+1})}(t),$$

en donde X^k es una variable aleatoria $\mathcal{F}_{t_k}$-medible y cuadrado integrable. La expresión $1_{[a,b)}(t)$ corresponde a la función indicadora del intervalo $[a,b)$. Un proceso simple es entonces un proceso "constante" por pedazos con trayectorias càdlàg, y las condiciones solicitadas garantizan que el proceso es adaptado y tiene trayectorias cuadrado integrables. Denotaremos por $\mathcal{H}_0^2$ al espacio vectorial de todos los procesos estocásticos simples. La integral estocástica de Itô de un proceso simple X respecto del movimiento Browniano, denotada por $I(X)$, se define entonces naturalmente como la variable aleatoria

$$I(X) = \sum_{k=0}^{n-1} X^k (B_{t_{k+1}} - B_{t_k}) = \int_0^T X_s dB_s.$$

Observe que esta integral es una variable aleatoria integrable pues es evidente que su esperanza es cero. Además es cuadrado integrable y cumple la siguiente igualdad fundamental llamada *isometría de Itô*,

$$||I(X)||_{L^2(P)} = ||X||_{L^2(P\times dt)}. \tag{3}$$

[1]Una norma es una función real denotada regularmente por $||\cdot||$ y definida sobre un espacio lineal que cumple: $||x|| \geq 0$, $||x|| = 0 \Leftrightarrow x = 0$, $||x+y|| \leq ||x|| + ||y||$ y $||\alpha x|| = |\alpha|||x||$.

En efecto, si ΔB_k denota la diferencia $B_{t_{k+1}} - B_{t_k}$, y $\Delta t_k = t_{k+1} - t_k$ entonces

$$\begin{aligned}
||I(X)||^2_{L^2(P)} &= E(|\sum_{k=0}^{n-1} X^k (B_{t_{k+1}} - B_{t_k})|^2) \\
&= E(\sum_{j,k=0}^{n-1} X^j X^k \Delta B_j \Delta B_k) \\
&= E(\sum_{k=0}^{n-1} (X^k)^2 (\Delta B_k)^2) \\
&= \sum_{k=0}^{n-1} E(X^k)^2 \Delta t_k \\
&= E(\int_0^T |X_t|^2 dt) \\
&= ||X||^2_{L^2(P\times dt)}.
\end{aligned}$$

Hemos usado el hecho de que X_k es $\mathcal{F}_{t_k}$-medible y teniendo el movimiento Browniano incrementos independientes, las variables X_k y ΔB_k resultan ser independientes. Esta igualdad juega un papel primordial en la definición de integral estocástica como se verá más adelante. La integral estocástica asigna entonces a cada elemento del espacio $\mathcal{H}^2_0$ una variable aleatoria dentro del espacio $L^2(P)$. De esta forma se tiene el mapeo lineal $I : \mathcal{H}^2_0 \to L^2(P)$, que resulta ser continuo por la isometría de Itô.

Paso 2: Extensión por aproximación. Ahora extendemos la integral estocástica a procesos un poco más generales. Sea $\mathcal{H}^2$ el espacio de todos los procesos $\{X_t : 0 \leq t \leq T\}$ medibles y adaptados tales que

$$E(\int_0^T |X_t|^2 dt) < \infty.$$

El espacio $\mathcal{H}^2$ resulta ser un subespacio lineal cerrado de $L^2(P \times dt)$. Observe que la única diferencia entre estos dos espacios es que a los elementos de $\mathcal{H}^2$ se les pide que sean medibles y adaptados. Claramente todo proceso simple es un elemento de $\mathcal{H}^2$. Tenemos entonces la contención de espacios $\mathcal{H}^2_0 \subset \mathcal{H}^2 \subset L^2(P \times dt)$, en donde puede probarse que $\mathcal{H}^2_0$ es denso en $\mathcal{H}^2$ respecto de la norma en $L^2(P \times dt)$. Esto significa que para cualquier proceso X en $\mathcal{H}^2$ existe un sucesión de procesos X^k en $\mathcal{H}^2_0$ tales que

$$\lim_{k\to\infty} ||X - X^k||_{L^2(P\times dt)} = 0. \tag{4}$$

Este procedimiento de aproximación puede llevarse a cabo de la siguiente forma. Mediante la técnica de truncación todo proceso en $\mathcal{H}^2$ puede ser aproximado por un proceso acotado. A su vez todo proceso en $\mathcal{H}^2$ que es acotado se puede aproximar por procesos acotados y continuos. Y éstos a su vez se aproximan por procesos simples de la forma

$$\sum_{j=0}^{n} X_{t_j} \cdot 1_{[t_j, t_{t_{j+1}})}(t),$$

en donde $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = T$ una partición finita de $[0,T]$. Los detalles completos de esta sucesión de aproximaciones pueden encontrarse en [13]. Por la isometría de Itô, la sucesión $I(X^k)$ es una sucesión de Cauchy en el espacio $L^2(P)$ pues

$$\begin{aligned} ||I(X^k) - I(X^l)||_{L^2(P)} &= ||I(X^k - X^l)||_{L^2(P)} \\ &= ||X^k - X^l||_{L^2(P\times dt)} \\ &\leq ||X - X^k||_{L^2(P\times dt)} + ||X - X^l||_{L^2(P\times dt)}. \end{aligned}$$

Debido a (4) la última expresión puede hacerse tan pequeña como se desee tomando índices k y l suficientemente grandes. Entonces de manera natural se define, para cada X en $\mathcal{H}^2$,

$$I(X) = \lim_{k\to\infty} I(X^k),$$

en donde el límite debe entenderse dentro del espacio $L^2(P)$. Esto significa que la variable aleatoria $I(X)$ es un elemento de $L^2(P)$ y es tal que $\lim_{n\to\infty} ||I(X) - I(X^k)||_{L^2(P)} = 0$. No es difícil verificar que tal definición es correcta en el sentido de que el límite no depende de la sucesión aproximante. La isometría de Itô y la propiedad de esperanza nula se cumplen también para procesos en $\mathcal{H}^2$. De esta forma tenemos ahora el mapeo lineal y continuo $I : \mathcal{H}^2 \to L^2(P)$.

Paso 3: La integral como un proceso. Hagamos ahora una pequeña extensión. Para cada t en $[0,T]$ y para X en $\mathcal{H}^2$ se define

$$I_t(X) = \int_0^T X_s \cdot 1_{[0,t]}(s) dB_s = \int_0^t X_s dB_s.$$

Esto permite ver a la integral estocástica no como una variable aleatoria sino como un proceso. Es claro que tal proceso no es necesariamente continuo sin embargo puede demostrarse que existe una versión continua de él, y que esa versión es una martingala respecto de la filtración natural del movimiento Browniano. Denotaremos por el mismo símbolo a tal martingala continua.

Paso 4: Extensión por localización. Mediante un procedimiento llamado de localización es posible extender la definición de integral de Itô a procesos medibles y adaptados que cumplen la condición más relajada

$$P(\int_0^T |X_t|^2 dt < \infty) = 1.$$

Denotaremos por $\mathcal{L}^2_{loc}$ el espacio de todos estos procesos. Este procedimiento de localización hace uso de tiempos de paro. Los tiempos de paro relativos a una filtración

$(\mathcal{F}_t)_{t\geq 0}$ son variables aleatorias $\tau : \Omega \to [0, \infty]$ tales que para cada $t \geq 0$, el evento $(\tau \leq t)$ pertenece a $\mathcal{F}_t$. Cuando la filtración es continua por la derecha, la condición anterior es equivalente a $(\tau \geq t) \in \mathcal{F}_t$. El nuevo espacio $\mathcal{L}^2_{loc}$ contiene a $\mathcal{H}^2$ y es tal que para cada proceso X en $\mathcal{L}^2_{loc}$ existe una sucesión creciente de tiempos de paro $\{\tau_n : n \geq 1\}$ tal que $X_t \cdot 1_{(\tau_n \geq t)} \in \mathcal{H}^2$ en donde $\tau_n \nearrow T$ cuando n crece a infinito. Se define entonces la integral estocástica

$$\int_0^t X_s dB_s = \lim_{n\to\infty} \int_0^T X_s \cdot 1_{[0,t]}(s) \cdot 1_{(\tau_n \geq t)}(\omega) dB_s.$$

Nuevamente es posible demostrar que tal límite existe, que existe una versión continua de tal límite, y que es independiente de la sucesión de tiempos de paro localizante. En este caso la integral ya no es una martingala sino una martingala local, es decir, el proceso $I_{t\wedge\tau_n}(X)$ es una martingala para cada natural n. En particular, para cualquier función continua f el proceso $\{f(B_t) : 0 \leq t \leq T\}$ tiene trayectorias continuas y acotadas, por lo tanto es un elemento de $\mathcal{L}^2_{loc}$ y tiene sentido la expresión

$$\int_0^t f(B_s) dB_s.$$

Finalmente mencionaremos que es posible demostrar la integral anterior puede ser calculada mediante el siguiente límite cuya convergencia es válida en probabilidad

$$\int_0^t f(B_s) dB_s = \lim_{n\to\infty} \sum_{i=0}^{n-1} f(B_{t_i})(B_{t_{i+1}} - B_{t_i}). \tag{5}$$

Con esto concluimos la serie de ideas generales bajo las cuales puede construirse la integral de Itô. Las demostraciones de algunos detalles técnicos que hemos simplemente mencionado se pueden encontrar en [14]. El esquema simplificado del procedimiento seguido para definir la integral estocástica se ilustra en la Figura 4.

$$\begin{tikzcd}
\mathcal{H}^2_0 \arrow[r, "I"] \arrow[d, dotted, "\text{Aproximación}"'] & L^2(P) \\
\mathcal{H}^2 \arrow[r, "I"] \arrow[d, dotted, "\text{Localización}"'] & L^2(P) \\
\mathcal{L}^2_{loc} \arrow[r, "I"] & L^2(P)
\end{tikzcd}$$

Figura 4. Diagrama para la definición de la integral estocástica.

Ejemplo. Con la ayuda de (5) calcularemos la integral estocástica

$$\int_0^t B_s dB_s.$$

Sea $0 = t_0 < t_1 < \cdots < t_n = t$ una partición uniforme de $[0,t]$, es decir $t_{i+1} - t_i = 1/n$. Usando la identidad

$$a(b-a) = \frac{1}{2}(b^2 - a^2) - \frac{1}{2}(a-b)^2$$

se obtiene

$$\begin{aligned}
\int_0^t B_s dB_s &= \lim_{n\to\infty} \sum_{j=0}^{n-1} B_{t_j}(B_{t_{j+1}} - B_{t_j}) \\
&= \lim_{n\to\infty} \sum_{j=0}^{n-1} [\frac{1}{2}(B_{t_{j+1}}^2 - B_{t_j}^2) - \frac{1}{2}(B_{t_{j+1}} - B_{t_j})^2] \\
&= \frac{1}{2}B_t^2 - \frac{1}{2}t.
\end{aligned}$$

Observe que el primer término del lado derecho corresponde a las reglas de integración usual. El término adicional se conoce como la *corrección de Itô*.

Ahora veamos el cambio en la solución cuando modificamos ligeramente la forma de calcular la integral, al hacer la evalución del integrando en el extremo derecho de cada subintervalo. Observe que en este caso el proceso a integrar ya no es adaptado y por lo tanto queda fuera de la teoría desarrollada antes. Usando la identidad

$$b(b-a) = \frac{1}{2}(b^2 - a^2) + \frac{1}{2}(a-b)^2$$

se obtiene

$$\begin{aligned}
\int_0^t B_s dB_s &= \lim_{n\to\infty} \sum_{j=0}^{n-1} B_{t_{j+1}}(B_{t_{j+1}} - B_{t_j}) \\
&= \lim_{n\to\infty} \sum_{j=0}^{n-1} [\frac{1}{2}(B_{t_{j+1}}^2 - B_{t_j}^2) + \frac{1}{2}(B_{t_{j+1}} - B_{t_j})^2] \\
&= \frac{1}{2}B_t^2 + \frac{1}{2}t.
\end{aligned}$$

El signo del segundo término cambió de negativo a positivo. Esto muestra que, a diferencia de la integral de Riemann, la integral estocástica es sensible al punto donde se evalúa el integrando. Al considerar en cambio el promedio de las evaluaciones en los extremos se obtiene la así llamada *integral de Stratonovich*, denotada de la forma siguiente

$$\int_0^t B_s \circ dB_s = \frac{1}{2}B_t^2.$$

Observe que el término adicional de la integral de Itô ha desaparecido. La integral de Stratonovich tiene algunas ventajas operacionales pues sigue algunas reglas usuales del cálculo integral, pero vista como proceso deja de ser una martingala.

FIGURA 5. Kiyosi Itô (Japón, 1915–). Fuente: Archivo MacTutor, St. Andrews.

Propiedades de la integral. La integral estocástica de Itô cumple varias propiedades aunque sólo mencionaremos algunas de ellas aquí a manera de resumen de las características mencionadas antes. Primeramente debemos mencionar que la integral $I_t : \mathcal{L}^2_{loc} \to L^2(P)$ es lineal, su esperanza es cero y se cumple la isometría de Itô

$$E \mid \int_0^t X_s dB_s|^2 = E \int_0^t |X_s|^2 ds.$$

En particular, para $X \in \mathcal{H}^2$ la integral $I_t(X)$ vista como un proceso es una martingala, es decir, es integrable, adaptado y para $0 \leq s \leq t$, se cumple $E(I_t(X)|\mathcal{F}_s) = I_s(X)$. Además existe una versión continua de tal proceso. En general para $X \in \mathcal{L}^2_{loc}$, la integral $I_t(X)$ ya no es una martingala sino una martingala local.

4. Ecuaciones diferenciales estocásticas

Una ecuación diferencial estocástica es una ecuación de la forma

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t \tag{6}$$

definida para t en $[0, T]$ y con condición inicial la variable aleatoria X_0 que se asume $\mathcal{F}_0$-medible e independiente del movimiento Browniano. La incógnita de esta ecuación es el proceso X_t y los coeficientes $b(t, x)$ y $\sigma(t, x)$ son funciones de $[0, T] \times \mathbb{R}$ en $\mathbb{R}$ y se conocen como los coeficientes de *tendencia* (*drift* en inglés o *deriva* en español) y de *difusión* respectivamente. La ecuación diferencial (6) se interpreta como la ecuación integral

$$X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s \tag{7}$$

en donde la primera es una integral de Riemann mientras que la segunda es una integral estocástica de Itô. Este proceso puede interpretarse como un sistema determinista gobernado por la parte no aleatoria de la ecuación pero perturbado por un ruido aditivo dado por la integral estocástica. A un proceso estocástico de la forma (7) se le llama *proceso de Itô* y para que esta ecuación tenga alguna solución se deben imponer condiciones en los coeficientes. De manera análoga al caso determinista, los teoremas básicos de existencia y unicidad para ecuaciones diferenciales estocásticas establecen que bajo ciertas condiciones de regularidad para los coeficientes b y σ, la ecuación (6) tiene una solución única, por ejemplo, si b y σ satisfacen la condición de Lipschitz en la variable x,

$$|b(t,x) - b(t,y)|^2 + |\sigma(t,x) - \sigma(t,y)|^2 \leq K|x-y|^2,$$

y la condición de crecimiento en x,

$$|b(t,x)|^2 + |\sigma(t,x)|^2 \leq K(1+|x|^2),$$

para alguna constante $K > 0$, entonces existe un proceso estocástico X_t solución de (6) que es adaptado, continuo, uniformemente acotado en $L^2(P)$, es decir, $\sup_{0\leq t\leq T} E(X_t^2) < \infty$, y es además único en el sentido de indistinguibilidad. En este caso a tal solución se le llama *solución fuerte* de la ecuación. Estos teoremas básicos no establecen la forma de encontrar la solución a una ecuación estocástica dada. La famosa *fórmula de Itô* es el resultado clave para realizar algunos de estos cálculos. Este importante resultado establece que si X_t un proceso de Itô dado por (6) y $f(t,x)$ es un función de clase C^1 en t y de clase C^2 en x, entonces el proceso $Y_t = f(t, X_t)$ es también un proceso de Itô y satisface la ecuación

$$dY_t = f_t(t,X_t)dt + f_x(t,X_t)dX_t + \frac{1}{2}f_{xx}(t,X_t)(dX_t)^2, \tag{8}$$

en donde los subíndices indican derivada y (6) se substituye en (8) usando la siguiente *tabla de multiplicación de McKean*,

$\times$	dt	dB_t
dt	0	0
dB_t	0	dt

Observe que como las derivadas involucradas son funciones continuas las integrales estocásticas resultantes están bien definidas. Por ejemplo, si se toma $f(t,x) = \frac{1}{2}x^2$ y el proceso $X_t = B_t$ entonces la fórmula de Itô establece que

$$\int_0^t B_s dB_s = \frac{1}{2}B_t^2 - \frac{1}{2}t.$$

Este mismo resultado había sido obtenido antes mediante un procedimiento límite. Veamos otro ejemplo sencillo. Para la función $f(t,x) = \frac{1}{3}x^3$ y $X_t = B_t$ se obtiene

$$\int_0^t B_s^2 dB_s = \frac{1}{3}B_t^3 - \int_0^t B_t dt.$$

Un ejemplo mas,

$$\int_0^t s dB_s = tB_t - \int_0^t B_t dt.$$

Para verificar esta formula puede tomarse el proceso $X_t = B_t$ y la función $f(t,x) = tx$. Entonces

$$\begin{aligned} d(f(t,B_t)) &= f_t(t,B_t)dt + f_x(t,B_t)dB_t + \frac{1}{2}f_{xx}(t,B_t)(dB_t)^2 \\ d(tB_t) &= B_t dt + t dB_t. \end{aligned}$$

Integrando se obtiene la fórmula enunciada.

5. Dos modelos simples

Estudiamos ahora dos ejemplos importantes de ecuaciones diferenciales estocásticas. El primero de ellos con aplicación en finanzas y el otro en física. Estos modelos son sencillos y podremos encontrarles solución explícita.

Movimiento Browniano geométrico. Suponga que el proceso X_t sigue una ley de movimiento dada por la ecuación estocástica

$$dX_t = \mu X_t\, dt + \sigma X_t\, dB_t \tag{9}$$

con condición inicial $X_0 = x_0 > 0$, en donde μ y $\sigma > 0$ son constantes. Este ecuación es de amplio uso en finanzas para modelar el precio de algunos bienes que fluctúan siguiendo los vaivenes de los mercados financieros.

La ecuación (9) puede interpretarse de la siguiente forma. En ausencia del término estocástico la ecuación se reduce a $dX_t = \mu X_t dt$ con solución $X_t = x_0 e^{\mu t}$. Esta solución representa el comportamiento de un capital inicial $x_0 > 0$ que crece de manera continua y determinista a una tasa efectiva del $100 \cdot \mu\,\%$ suponiendo $\mu > 0$ [2].

Por otro lado la parte estocástica corresponde a la volatilidad de una inversión con riesgo sujeta a las fluctuaciones de los mercados financieros. El modelo asume que dicha variabilidad es proporcional al valor de la inversión. A mayor valor de la inversión mayor variación en el precio. En la Figura 6 puede apreciarse una trayectoria de este proceso con una inversión inicial x_0 de una unidad monetaria y con parámetros $\mu = 1$ y $\sigma^2 = 1/2$. La curva creciente corresponde al crecimiento determinista de la inversión cuando no hay aleatoriedad, es decir cuando $\sigma^2 = 0$. Realmente se efectuaron

[2] Esto es equivalente a una tasa continua o nominal del $100 \cdot \ln(1+\mu)\,\%$. Por ejemplo si μ =0.1 entonces al final de una unidad de tiempo, el capital inicial x_0 crecerá a $x_0 \cdot e^{0.1} = x_0 \cdot$ 1.10517092= $x_0 \cdot$ (1+0.10517092) unidades monetarias.

varias simulaciones resultando trayectorias a veces por arriba y a veces por abajo de la curva determinista y eventualmente algún cruce, pero se decidió mostrar la presente pues en ella se observa que efectivamente la trayectoria estocástica sigue la curva determinista y oscila alrededor de ella. El lector interesado en la simulación de ecuaciones estocásticas puede consultar el clásico y muy completo libro de Kloeden y Platen [9].

Observe que los coeficientes de esta ecuación satisfacen las condiciones para la existencia y unicidad de la solución. Es posible resolver la ecuación (9) usando el método de *igualación de coeficientes*. Para ello se necesita encontrar una función $f(t,x)$ tal que al aplicar la fórmula de Itô al proceso $X_t = f(t, B_t)$ se obtenga la ecuación (9). Comparando entonces los coeficientes de

$$dX_t = f_t(t,X_t)dt + f_x(t,X_t)dB_t + \frac{1}{2}f_{xx}(t,X_t)dt$$

con los de (9) se obtienen las igualdades

$$\begin{aligned} \mu f(t,x) &= f_t(t,x) + \frac{1}{2}f_{xx}(t,x), \\ \sigma f(t,x) &= f_x(t,x). \end{aligned}$$

De la segunda ecuación se obtiene que $f(t,x) = \exp[\sigma x + g(t)]$ para alguna función $g(t)$. Substituyendo en la primera ecuación se obtiene $g'(t) = \mu - \frac{1}{2}\sigma^2$ cuya solución es $g(t) = (\mu - \frac{1}{2}\sigma^2)t$. Por lo tanto la solución de (9) es

$$X_t = x_0 \exp[(\mu - \frac{1}{2}\sigma^2)t + \sigma B_t].$$

A este proceso se le llama *movimiento Browniano geométrico* y también se le conoce como *movimiento Browniano exponencial*.

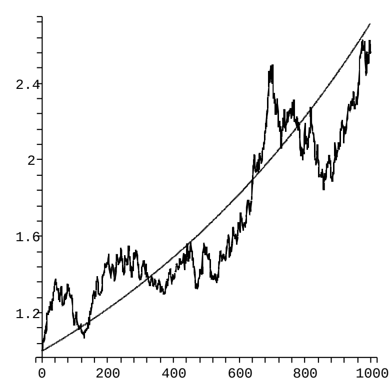

FIGURA 6. Simulación del movimiento Browniano geométrico.

Proceso de Ornstein-Uhlenbeck. Considere ahora la ecuación estocástica

$$dX_t = -\alpha X_t \, dt + \sigma \, dB_t \tag{10}$$

en donde α y σ son constantes positivas. Esta ecuación fue propuesta por Ornstein y Uhlenbeck para modelar la variación de la velocidad en el movimiento difuso de una partícula para tiempos pequeños. La variable X_t se interpreta entonces como la velocidad de la partícula al tiempo t. La parte determinista $-\alpha X_t$ corresponde a la fuerza de fricción y el sumando $\sigma \, dB_t$ es una perturbación aleatoria. Encontraremos la solución de (10) suponiendo la condición inicial $X_0 = x_0$. Considere una solución de la forma

$$X_t = a(t)[x_0 + \int_0^t b(s) dB_s], \tag{11}$$

en donde a y b son funciones diferenciables. Derivando (11) y usando la fórmula de Itô se obtiene

$$\begin{aligned} dX_t &= a'(t)[x_0 + \int_0^t b(s) dB_s] dt + a(t) b(t) dB_t \\ &= \frac{a'(t)}{a(t)} X_t dt + a(t) b(t) dB_t. \end{aligned}$$

Comparando con (10) las funciones a y b deben entonces cumplir

$$\frac{a'(t)}{a(t)} = -\alpha, \qquad a(t) b(t) = \sigma.$$

Suponiendo $a(0) = 1$ se obtiene $a(t) = \exp(-\alpha t)$ y $b(t) = \sigma \exp(\alpha t)$. Por lo tanto el proceso solución de (10), llamado *proceso de Ornstein-Uhlenbeck*, es

$$X_t = x_0 e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-s)} dB_s.$$

Otros modelos de ecuaciones estocásticas y una gran variedad de aplicaciones pueden encontrarse en el libro de Kloeden y Platen [9]. Para un estudio sistemático y detallado de estos temas pueden consultarse [3],[5],[7],[13], o [14].

Agradecimientos. Me permito agradecer sinceramente al revisor de este artículo por sus comentarios y sugerencias, así como a los editores del presente volumen por su asistencia en la preparación final del trabajo.

Referencias

[1] Brown R. (1828) *A brief account of Microscopical Observations made in the Months of June, July, and August, 1827, on the Particles contained in the Pollen of Plants; and on the general Existence of active Molecules in Organic and Inorganic Bodies*, Philosophical Magazine N. S. 4, 161-173.

[2] Brzeźniak Z. y Zastawniak T., Basic stochastic processes, Springer (1999).

[3] Chung K. L. y Williams R. J., Introduction to stochastic integration, Birkhäuser (1983).

[4] Einstein A., Investigations on the theory of the Brownian movement, Dover (1956).
[5] Gard T. C., Introduction to stochastic differential equations, Marcel Dekker, Inc. (1988).
[6] Hernández-Hernández D., *Movimiento Browniano y ecuaciones de Hamilton-Jacobi*, Carta Informativa, **42**, Soc. Mat. Mexicana (2004).
[7] Karatzas I. y Shreve S. E., Brownian motion and stochastic calculus, Springer (1991).
[8] Karlin S. y Taylor H. M., A first course in stochastic processes, Academic Press, Inc. (1975).
[9] Kloeden P. E. y Platen E., Numerical solution of stochastic differential equations, Springer–Verlag (1999).
[10] Korn R. y Korn E., Option pricing and portfolio optimization: modern methods of financial mathematics, Graduate Studies in Mathematics **31**. American Math. Soc. (2001).
[11] León J. A., *Una introducción a las ecuaciones diferenciales estocásticas*, Carta Informativa **31**, Soc. Mat. Mexicana. (2001).
[12] Nelson E., Dynamical theories of Brownian motion, Princeton University Press, (1967).
[13] Øksendal B., Stochastic differential equations: an introduction with applications, Springer–Verlag (1992).
[14] Steele J. M., Stochastic calculus and financial applications, Springer–Verlag (2001).

Departamento de Matemáticas, Facultad de Ciencias UNAM, Circuito Exterior, Ciudad Universitaria, 04510 México DF

E-mail address: `lars@fciencias.unam.mx`