

Procesos Estocásticos II

Diplomado en Riesgos

Gerónimo Uribe Bravo
Instituto de Matemáticas
Universidad Nacional Autónoma de México

CAPÍTULO 1

Martingalas

En este capítulo nos enfocaremos en el estudio de las martingalas. Esta es una clase de procesos fundamental para la teoría moderna de la probabilidad. Tanto así que la herramienta teórica sobre la cual se construye la teoría moderna de las finanzas matemáticas (el llamado cálculo estocástico) es una teoría basada en las martingalas.

1. Recordatorio sobre esperanza condicional

Si $(\Omega, \mathcal{F}, \mathbb{P})$ es un espacio de probabilidad y $B \in \mathcal{F}$ es tal que $\mathbb{P}(B) > 0$, podemos definir la probabilidad condicional de A dado B mediante la fórmula

$$\mathbb{P}(A | B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

que se puede entender a través de la interpretación frecuentista de la probabilidad. Así, para una variable aleatoria discreta¹ estamos acostumbrados a expresiones como $\mathbb{P}(A | X = j)$ y a la connotación que que se les ha dado. Desafortunadamente, una extensión del concepto de probabilidad condicional a eventos cualquiera no es tan inmediata², por lo que primero desarrollaremos algunas propiedades de la esperanza condicional que nos permitan entender la solución que se le ha dado a este problema de extensión, definiendo algunos conceptos y verificando algunas propiedades de las variables aleatorias que nos faciliten el camino.

1.1. Preliminares. A lo largo de la sección, $(\Omega, \mathcal{F}, \mathbb{P})$ designará a un espacio de probabilidad arbitrario y a las funciones medibles de Ω en \mathbb{R} las llamaremos variables aleatorias reales, aunque generalmente se omitirá la palabra reales. Si $X : \Omega \rightarrow \mathbb{R}$ es una variable aleatoria en $(\Omega, \mathcal{F}, \mathbb{P})$ se utilizará la notación

$$\{X \in B\} = X^{-1}(B).$$

También, m_n (m) representará a la medida de Lebesgue sobre los Boreelianos de \mathbb{R}^n (\mathbb{R}).

NOTA. Si $f : \mathbb{R} \rightarrow \mathbb{R}$ es Borel medible, entonces $f \circ X$ es borel medible, por lo que está definida su esperanza cuando la integral de la composición esté definida.

¹Esto es, una función $X : \Omega \rightarrow \mathbb{R}$ Borel medible tal que $X(\Omega)$ sea a lo más numerable.

²?Qué pasaría en el caso de eventos condicionantes de probabilidad cero?

DEFINICIÓN. Si $X : \Omega \rightarrow \mathbb{R}$ es una variable aleatoria, la medida de probabilidad inducida por X , es la función $\mathbb{P}_X : \mathcal{B}_{\mathbb{R}} \rightarrow [0, 1]$ dada por:

$$\mathbb{P}_X(B) = \mathbb{P}(X \in B).$$

NOTA. Si \mathbb{P}_X es absolutamente continua con respecto a la medida de Lebesgue sobre los Boreelianos de \mathbb{R} , diremos que X es absolutamente continua y en este caso existe una densidad $g_X : \mathbb{R} \rightarrow \mathbb{R}$ tal que

$$\mathbb{P}(X \in B) = \mathbb{P}_X(B) = \int_B g_X dm.$$

TEOREMA 1.1 (Teorema de Cambio de variable). *Si $X : \Omega \rightarrow \mathbb{R}$ es una variable aleatoria en $(\Omega, \mathcal{F}, \mathbb{P})$ y $f : \mathbb{R} \rightarrow \mathbb{R}$ es Borel medible tal que la integral de $f \circ X$ está definida, entonces:*

$$\mathbb{E}(f \circ X) = \int f \mathbb{P}_X(d).$$

EJERCICIO 1.1. Sea X una variable aleatoria normal centrada de varianza 1. Utilice el teorema de cambio de variable para calcular

$$\mathbb{E}(X^{2n})$$

para toda $n \in \mathbb{N}$.

PROPOSICIÓN 1.1. *Sea $Z : \Omega \rightarrow \mathbb{R}$ una variable aleatoria fija en $(\Omega, \mathcal{F}, \mathbb{P})$ y $\mathcal{G} = \sigma(Z)$. Si $X : \Omega \rightarrow \mathbb{R}$ es \mathcal{G} medible, entonces existe $f : \mathbb{R} \rightarrow \mathbb{R}$ Borel-medible, tal que $X = f \circ Z$.*

1.2. Esperanza Condicional. Si $Z : \Omega \rightarrow \mathbb{R}$ es una variable aleatoria simple en $(\Omega, \mathcal{F}, \mathbb{P})$ y $Y : \Omega \rightarrow \mathbb{R}$ es una variable aleatoria, una definición natural de la probabilidad condicional $\mathbb{P}(Y \in B | Z)$ es la siguiente:

$$\mathbb{P}(Y \in B | Z) = \sum_{i \in \mathcal{R}_Z} \mathbb{P}(Y \in B | Z = i) \mathbf{1}_{\{Z=i\}},$$

donde $\mathcal{R}_Z = Z(\Omega) \subset \mathbb{R}$ es un conjunto finito. Notemos que en este caso, la probabilidad condicional es una función de la variable aleatoria Z , por lo que resulta ser $\sigma(Z)$ -medible y que cumple la relación

$$\mathbb{P}(Y \in B, A) = \int_A \mathbb{P}(Y \in B | Z) d\mathbb{P}, \quad A \in \sigma(Z),$$

que es equivalente a

$$\int_A \mathbf{1}_{Y \in B} d\mathbb{P} = \int_A \mathbb{P}(Y \in B | Z) d\mathbb{P},$$

esto es, obtenemos información (la integral sobre un conjunto) de la variable $\mathbf{1}_{Y \in B}$, que no es necesariamente $\sigma(Z)$ -medible a través de la variable $\mathbb{P}(Y \in B | Z)$ que si lo es, aunque sea para una clase restringida de eventos ($\sigma(Z)$, que resulta ser una σ -álgebra). Además, en la propiedad anterior de probabilidad condicional, la

variable aleatoria Z solo juega un papel secundario, y la σ -álgebra $\sigma(Z)$ se torna imprescindible. Como un comentario adicional, recordemos que dos variables aleatorias Y y Z son iguales \mathbb{P} -p.s. si y solo si $\int_A Y d\mathbb{P} = \int_A Z d\mathbb{P}$ para todo $A \in \mathcal{F}$ (una propiedad parecida a la encontrada en la probabilidad condicional), por lo que la función que a cada elemento A de \mathcal{F} le asigna el número $\int_A Y d\mathbb{P}$ (que resulta ser una medida con signo si la integral de Y está definida) determina completamente a la variable aleatoria Y . El comentario anterior puede motivar la definición de esperanza condicional, de la cual la probabilidad condicional es un caso particular³, en la que se condiciona con respecto a una σ -álgebra:

DEFINICIÓN. Si X es una variable aleatoria en $(\Omega, \mathcal{F}, \mathbb{P})$, y $\mathcal{G} \subset \mathcal{F}$ es una σ -álgebra, la **esperanza condicional** de X dado \mathcal{G} , denotada por $\mathbb{E}(X | \mathcal{G})$, es una variable aleatoria \mathcal{G} -medible que cumple

$$\int_A X d\mathbb{P} = \int_A \mathbb{E}(X | \mathcal{G}) d\mathbb{P}$$

para todo $A \in \mathcal{G}$.

PROPOSICIÓN 1.2. Si $X : \Omega \rightarrow \mathbb{R}$ es una variable aleatoria en $(\Omega, \mathcal{F}, \mathbb{P})$ cuya integral está definida y $\mathcal{G} \subset \mathcal{F}$ es una σ -álgebra, entonces existe una variable aleatoria $Y : \Omega \rightarrow \mathbb{R}$ tal que $\int_A X d\mathbb{P} = \int_A Y d\mathbb{P}$, para $A \in \mathcal{G}$. Además, si Z cumple la misma propiedad, entonces $Y = Z$ casi seguramente respecto a $\mathbb{P}|_{\mathcal{G}}$.

EJERCICIO 1.2. Si (X, Y) son dos variables aleatorias con densidad conjunta $f(x, y)$, pruebe que:

$$\mathbb{E}(g(Y) | X) = \frac{\int f(X, y) g(y) dy}{\int f(X, y) dy}.$$

EJERCICIO 1.3. Sean X_1, X_2, \dots v.a.i.d.s. Sea K una variable aleatoria independiente de X_1, X_2, \dots y con valores en \mathbb{N} . Cacule

$$\mathbb{E}(X_1 + \dots + X_K | K).$$

Sugerencia: ¿Qué pasa cuando K toma sólo un valor?

1.3. Propiedades de la esperanza condicional. Las siguientes son algunas propiedades de la esperanza condicional, en las que consideramos $\mathcal{G} \subset \mathcal{F}$ una σ -álgebra. Si X y Y son variables aleatorias, la ecuación $Y = \mathbb{E}(X | \mathcal{G})$ significa que el lado izquierdo de la ecuación es \mathcal{G} -medible y que $\int_A Y d\mathbb{P} = \int_A X d\mathbb{P}$ para $A \in \mathcal{G}$. Consideraremos solo variables aleatorias cuya integral esté definida, por lo que la existencia de la esperanza condicional queda garantizada.

PROPIEDAD 1 (Linealidad de la esperanza condicional). Si X y Y son variables aleatorias integrables y $a, b \in \mathbb{R}$, entonces $\mathbb{E}(aX + bY | \mathcal{G})$ existe y es igual a $a\mathbb{E}(X | \mathcal{G}) + b\mathbb{E}(Y | \mathcal{G})$.

³Se utilizará la relación $\mathbb{P}(A) = \mathbb{E}(\mathbf{1}_A)$ para este efecto.

PROPIEDAD 2 (Monotonía de la esperanza condicional). *Si X es no negativa \mathbb{P} -p.s. entonces $\mathbb{E}(X | \mathcal{G})$ existe y es no negativa $\mathbb{P}|_{\mathcal{G}}$ -p.s..*

PROPIEDAD 3. *Si la integral de X está definida entonces $|\mathbb{E}(X | \mathcal{G})| \leq \mathbb{E}(|X| | \mathcal{G})$.*

PROPIEDAD 4. *Si X es \mathcal{G} -medible, entonces $\mathbb{E}(X | \mathcal{G}) = X$.*

PROPIEDAD 5. *Si X es independiente de \mathcal{G} entonces $\mathbb{E}(X | \mathcal{G}) = \mathbb{E}(X)$.*

PROPIEDAD 6. $\mathbb{E}(\mathbb{E}(X | \mathcal{G})) = \mathbb{E}(X)$.

PROPIEDAD 7 (Propiedad de torre). *Si $\mathcal{D} \subset \mathcal{G}$ y \mathcal{D} es σ -álgebra entonces*

$$\mathbb{E}(\mathbb{E}(X | \mathcal{G}) | \mathcal{D}) = \mathbb{E}(X | \mathcal{D}) = \mathbb{E}(\mathbb{E}(X | \mathcal{D}) | \mathcal{G}).$$

PROPIEDAD 8 (Teorema de Convergencia Monótona para la Esperanza Condicional). *Si $(X_n)_{n \in \mathbb{N}}$ son variables aleatorias tal que $0 \leq X_n \leq X_{n+1}$ y $X = \lim_{n \rightarrow \infty} X_n$, entonces $\mathbb{E}(X | \mathcal{G})$ existe y*

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_n | \mathcal{G}) = \mathbb{E}(X | \mathcal{G}).$$

PROPIEDAD 9 (Lema de Fatou para la Esperanza Condicional). *Si $X_n \geq 0$ para $n \in \mathbb{N}$ entonces existe*

$$\mathbb{E}\left(\liminf_{n \rightarrow \infty} X_n \mid \mathcal{G}\right)$$

y

$$\mathbb{E}\left(\liminf_{n \rightarrow \infty} X_n \mid \mathcal{G}\right) \leq \liminf_{n \rightarrow \infty} \mathbb{E}(X_n | \mathcal{G}).$$

PROPIEDAD 10 (Teorema de Convergencia Dominada para la Esperanza Condicional). *Si*

$$(X_n)_{n \in \mathbb{N}} \subset \mathcal{L}_1(\mathbb{P})$$

es puntualmente convergente y existe $Y \in \mathcal{L}_1(\mathbb{P})$ tal que $|X_n| \leq Y$ para $n \in \mathbb{N}$, entonces

$$\mathbb{E}\left(\lim_{n \rightarrow \infty} X_n \mid \mathcal{G}\right)$$

existe y es igual a $\lim_{n \rightarrow \infty} \mathbb{E}(X_n | \mathcal{G})$ (donde la existencia de este último límite solo se asegura $\mathbb{P}|_{\mathcal{G}}$ -p.s.).

PROPIEDAD 11 (\mathcal{G} -homogeneidad). *Si X_1 y X_2 son variables aleatorias integrables tales que $X_1 X_2$ es integrable y X_1 es \mathcal{G} -medible entonces*

$$\mathbb{E}(X_1 X_2 | \mathcal{G}) = X_1 \mathbb{E}(X_2 | \mathcal{G}).$$

(Note que la hipótesis de integrabilidad del producto quedaría garantizada si X_1 y X_2 pertenecen a $\mathcal{L}_2(\mathbb{P})$).

PROPIEDAD 12. *Si $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ es boreiana, la integral de $f(X, Y)$ existe, $Y \perp \mathcal{G}$ y X es \mathcal{G} -medible entonces $\mathbb{E}(f(X, Y) | \mathcal{G}) = \int_{\mathbb{R}} f(X, y) \mathbb{P}_Y(dy)$.*

PROPIEDAD 13. Si $\mathcal{H}, \mathcal{G} \subset \mathcal{F}$ son σ -álgebras y $\mathcal{H} \perp \sigma(\mathcal{G}, \sigma(X))$, entonces

$$\mathbb{E}(X | \sigma(\mathcal{H}, \mathcal{G})) = \mathbb{E}(X | \mathcal{G}).$$

PROPIEDAD 14 (Desigualdad de Jensen para la Esperanza Condicional). Si $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ es una función convexa, la integral de X existe y la integral de $\varphi \circ X$ está definida, entonces

$$\varphi(\mathbb{E}(X | \mathcal{G})) \leq \mathbb{E}(\varphi \circ X | \mathcal{G}) \quad \mathbb{P}|_{\mathcal{G}} - p.s.$$

Si $\psi : \mathbb{R} \rightarrow \mathbb{R}$ es una función cóncava, X es integrable y la integral de $\psi \circ X$ está definida, entonces

$$\psi(\mathbb{E}(X | \mathcal{G})) \geq \mathbb{E}(\psi \circ X | \mathcal{G}) \mathbb{P}|_{\mathcal{G}} - p.s..$$

2. Martingalas

Estudiaremos ahora una familia de procesos estocásticos que es importante dentro de la teoría de la probabilidad, principalmente por sus aplicaciones teóricas, tanto así, que su estudio resulta imprescindible para la teoría moderna de la probabilidad. Mediante su uso, verificaremos ciertos teoremas clásicos para caminatas aleatorias, como la ley 0 – 1 de Kolmogorov y la ley fuerte de los grandes números. En este capítulo solamente consideraremos procesos estocásticos indicados por un subconjunto de \mathbb{Z} .

Consideremos la siguiente situación: jugamos una serie de volados, obteniendo 1 si ganamos el n -ésimo y –1 si lo perdemos. El modelo matemático que consideraremos está conformado por una sucesión de variables aleatorias independientes e idénticamente distribuidas $(X_i)_{i=1}^{\infty}$, donde X_i representa el resultado del i -ésimo volado. Nuestra fortuna al tiempo n , S_n , está dada por

$$S_n = \sum_{i=1}^n X_i$$

para $n \geq 1$ y definiremos $S_0 = 0$. Para que el juego resulte justo para las dos personas que lo juegan, debemos pedir que $\mathbb{P}(X_i = 1) = 1/2$. Si este es el caso, podemos preguntarnos por la mejor aproximación a S_{n+1} que podemos dar al utilizar la información sobre el juego que conocemos hasta el tiempo n . La información al tiempo n la interpretaremos como

$$\mathcal{F}_n = \sigma(X_1, \dots, X_n),$$

puesto que esta σ -álgebra contiene a todos los conjuntos de la forma

$$\{X_1 = i_1, \dots, X_n = i_n\},$$

y así, lo que realmente buscamos es la esperanza condicional de S_{n+1} dada \mathcal{F}_n , que es sencilla de calcular, pues

$$\mathbb{E}(S_{n+1} | \mathcal{F}_n) = \mathbb{E}(S_n + X_{n+1} | \mathcal{F}_n) = S_n + \mathbb{E}(X_{n+1}) = S_n.$$

Como un ejercicio, el lector puede verificar que de hecho,

$$\mathbb{E}(S_{n+m} | \mathcal{F}_n) = S_n, \quad \forall m \geq 0,$$

por lo que al conocer la información hasta el tiempo n , solamente podemos afirmar que nos quedaremos con lo que tenemos, y como lo mismo sucede con el jugador contra el cual competimos, el juego resulta ser justo.

Informalmente, podemos definir una martingala como un proceso estocástico $(X_n)_{n \in \mathbb{N}}$ tal que X_n representa la ganancia al tiempo n de un jugador involucrado en un juego justo respecto a cierta información. Para precisar esta idea, necesitamos un ingrediente extra:

DEFINICIÓN. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y $(\mathcal{F}_n)_{n \in \mathbb{N}}$ una colección de σ -álgebras contenidas cada una en \mathcal{F} . Decimos que dicha familia es una **filtración** si $\mathcal{F}_n \subset \mathcal{F}_m$ cuando $n \leq m$.

Si $(\mathcal{F}_n)_{n \in \mathbb{N}}$ es una filtración, interpretaremos a \mathcal{F}_n como la información acumulada al tiempo n .

DEFINICIÓN. Sean $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y $(\mathcal{F}_n)_{n \in \mathbb{N}}$ una filtración en dicho espacio. Una colección de variables aleatorias reales $(X_n)_{n \in \mathbb{N}}$ es una **martingala** respecto a la filtración considerada si

- (1) X_n es \mathcal{F}_n -medible.
- (2) $X_n \in L_1$.
- (3) $\mathbb{E}(X_{n+1} | \mathcal{F}_n) = X_n$ para cualquier $n \in \mathbb{N}$.

La primera propiedad nos dice que conocemos la ganancia al tiempo n a partir de la información que se nos proporciona hasta ese instante, generalmente se dice que la sucesión de variables aleatorias es adaptada a la filtración. La segunda es una hipótesis técnica que nos permite utilizar a la esperanza condicional como un operador lineal y la tercera nos dice que el juego es justo respecto a la información proporcionada.

DEFINICIÓN. Supongamos ahora que $(X_n)_{n \in \mathbb{N}}$ satisface (1) y (2), pero en vez de tener una igualdad en (3), observamos una desigualdad:

$$\mathbb{E}(X_{n+1} | \mathcal{F}_n) \leq X_n.$$

Entonces le llamaremos a la sucesión una **supermartingala**. (Note que de acuerdo a nuestra interpretación de X_n como evolución de nuestra fortuna, una supermartingala no tiene nada de super...) Si se da la desigualdad contraria, esto es,

$$\mathbb{E}(X_{n+1} | \mathcal{F}_n) \geq X_n,$$

entonces a $(X_n)_{n \in \mathbb{N}}$ le llamamos **submartingala**.

Notemos que si $(X_n)_{n \in \mathbb{N}}$ es una martingala, entonces la sucesión $(\mathbb{E}(X_n))_{n \in \mathbb{N}}$ es constante. Para una supermartingala o una submartingala, la palabra constante

se debe substituir por decreciente o por creciente. Además, podemos inferir una propiedad más fuerte a partir de (3), a saber, que

$$\mathbb{E}(X_{n+m} \mid \mathcal{F}_n) = X_n$$

si $n, m \in \mathbb{N}$. Esto se sigue de las propiedades de la esperanza condicional, puesto que

$$\mathbb{E}(X_{n+m+1} \mid \mathcal{F}_n) = \mathbb{E}(\mathbb{E}(X_{n+m+1} \mid \mathcal{F}_{n+m}) \mid \mathcal{F}_n) = \mathbb{E}(X_{n+m} \mid \mathcal{F}_n).$$

Una afirmación similar es válida para supermartingalas o submartingalas al cambiar la igualdad por una desigualdad. Para concluir esta sección, veamos un método para construir submartingalas a partir de una martingala dada.

TEOREMA 1.2. *Si $(X_n)_{n \in \mathbb{N}}$ es una martingala respecto a la filtración $(\mathcal{F}_n)_{n \in \mathbb{N}}$ y $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ es una función convexa tal que $\varphi(X_n) \in L_1$, entonces $(\varphi(X_n))_{n \in \mathbb{N}}$ es una submartingala respecto a la misma filtración.*

DEMOSTRACIÓN. Como cualquier función convexa (sobre \mathbb{R}) es continua, entonces $\varphi(X_n)$ es \mathcal{F}_n -medible, que pertenece por hipótesis a L_1 . Finalmente, por la desigualdad de Jensen para la esperanza condicional,

$$\mathbb{E}(\varphi(X_{n+1}) \mid \mathcal{F}_n) \geq \varphi(\mathbb{E}(X_{n+1} \mid \mathcal{F}_n)) = \varphi(X_n). \quad \square$$

2.1. Ejemplos. En esta sección supondremos que $(\Omega, \mathcal{F}, \mathbb{P})$ es un espacio de probabilidad en el cual están definidas variables aleatorias con las características deseadas.

EJEMPLO 1.1. Supongamos que X es una variable aleatoria que pertenece a L_1 y $(\mathcal{F}_n)_{n \in \mathbb{N}}$ una filtración en $(\Omega, \mathcal{F}, \mathbb{P})$. Entonces la sucesión de variables aleatorias $(X_n)_{n \in \mathbb{N}}$ en la cual

$$X_n = \mathbb{E}(X \mid \mathcal{F}_n)$$

(sin importar la versión de la esperanza condicional) es una martingala respecto a $(\mathcal{F}_n)_{n \in \mathbb{N}}$.

Para verificar la veracidad de la anterior afirmación, notemos que por definición de esperanza condicional, X_n es \mathcal{F}_n -medible y que $X_n \in L_1$. Finalmente, como $\mathcal{F}_n \subset \mathcal{F}_{n+1}$, entonces

$$\mathbb{E}(X_{n+1} \mid \mathcal{F}_n) = \mathbb{E}(\mathbb{E}(X \mid \mathcal{F}_{n+1}) \mid \mathcal{F}_n) = \mathbb{E}(X \mid \mathcal{F}_n) = X_n.$$

Esta martingala es un ejemplo bastante general y muy importante. Posteriormente podremos determinar cuando una martingala es de este tipo. A este proceso se le conoce como la **martingala cerrada**.

EJEMPLO 1.2. Sean $(\xi_i)_{i=1}^{\infty}$ variables aleatorias independientes e idénticamente distribuidas, $\mathcal{F}_0 = \{\emptyset, \Omega\}$ y $\mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n)$ para $n \geq 1$. Entonces $(\mathcal{F}_n)_{n \in \mathbb{N}}$ es

una filtración.. Si

$$\begin{aligned} S_0 &= 0, \\ S_n &= \sum_{i=1}^n \xi_i \quad n \geq 1, \end{aligned}$$

y ξ_i tiene media finita μ , entonces $(X_n)_{n \in \mathbb{N}}$ es una martingala respecto a $(\mathcal{F}_n)_{n \in \mathbb{N}}$, donde

$$X_n = S_n - n\mu.$$

Si además, $\sigma^2 = \text{Var}(\xi_i) < \infty$, entonces $(Y_n)_{n \in \mathbb{N}}$ es martingala respecto a la misma filtración, donde

$$Y_n = (S_n - n\mu)^2 - n\sigma^2.$$

Por otro lado, si

$$\varphi(\theta) = \mathbb{E}(e^{\theta\xi_i}) < \infty$$

para $\theta \in \mathbb{R}$, definimos $Z_0 = 1$ y para $n \geq 1$:

$$Z_n = \frac{e^{\theta S_n}}{(\varphi(\theta))^n},$$

entonces $(Z_n)_{n \in \mathbb{N}}$ es una martingala respecto a la misma filtración.

Como X_n, Y_n y Z_n están dadas por $f(\xi_1, \dots, \xi_n)$, para una función continua $f : \mathbb{R}^n \rightarrow \mathbb{R}$ (una función distinta para cada variable aleatoria) y el vector aleatorio (ξ_1, \dots, ξ_n) es medible respecto a \mathcal{F}_n , se sigue que las tres variables consideradas son \mathcal{F}_n -medibles. Para ver que pertenecen a L_1 , notemos que X_n , es la diferencia de dos funciones en L_1 , por ser este último cerrado bajo la suma. Además, si ξ_i tiene momento de segundo orden finito, entonces a S_n le pasa lo mismo, por lo que $Y_n \in L_1$. Para Z_n , el argumento que utilizamos es el de independencia, puesto que esto implica que

$$\mathbb{E}(\exp(\theta S_n)) = (\varphi(\theta))^n < \infty.$$

Para verificar la última propiedad que define a las martingalas, notemos que

$$\mathbb{E}(X_{n+1} - X_n \mid \mathcal{F}_n) = \mathbb{E}(\xi_{n+1} - \mu \mid \mathcal{F}_n) = \mathbb{E}(\xi_{n+1} - \mu) = 0,$$

por lo que $(X_n)_{n \in \mathbb{N}}$ es efectivamente una martingala. Por otro lado,

$$\begin{aligned} \mathbb{E}(Y_{n+1} - Y_n \mid \mathcal{F}_n) &= \mathbb{E}\left((S_{n+1} - (n+1)\mu)^2 - (S_n - n\mu)^2 - \sigma^2 \mid \mathcal{F}_n\right) \\ &= \mathbb{E}(2(S_n - n\mu)(\xi_{n+1} - \mu) \mid \mathcal{F}_n) \\ &\quad + \mathbb{E}\left((\xi_{n+1} - \mu)^2 - \sigma^2 \mid \mathcal{F}_n\right) \\ &= 2(S_n - n\mu)\mathbb{E}(\xi_{n+1} - \mu) + \mathbb{E}((\xi_{n+1} - \mu)^2) - \sigma^2 \\ &= 0, \end{aligned}$$

por lo que $(Y_n)_{n \in \mathbb{N}}$ es una martingala. Finalmente se tiene que

$$\begin{aligned}\mathbb{E}(Z_{n+1} | \mathcal{F}_n) &= \mathbb{E}\left(Z_n \frac{e^{\theta \xi_{n+1}}}{\varphi(\theta)} \mid \mathcal{F}_n\right) \\ &= Z_n \mathbb{E}\left(\frac{e^{\theta \xi_{n+1}}}{\varphi(\theta)}\right) \\ &= Z_n,\end{aligned}$$

por lo que $(Z_n)_{n \in \mathbb{N}}$ es una martingala.

Más adelante, utilizaremos estas martingalas para hacer ciertos cálculos referentes a caminatas aleatorias.

EJEMPLO 1.3. Sea U una variable aleatoria uniforme en $(0, 1)$ y definamos a

$$X_n = 2^n \mathbf{1}_{U \leq 1/2^n}.$$

Entonces X_0, X_1, \dots es una martingala respecto de la filtración que genera.

EJERCICIO 1.4. Probar la afirmación anterior.

Notemos que en esta martingala, se tiene que $X_n \rightarrow 0$ \mathbb{P} -p.s., pero que sin embargo, $(X_n)_{n \in \mathbb{N}}$ no converge en L_1 a 0.

EJEMPLO 1.4. Consideremos el siguiente experimento aleatorio, se tiene una urna con r bolas rojas y v bolas verdes. Extraemos una bola, la reemplazamos junto con c bolas del mismo color, revolvemos la urna y volvemos a realizar el experimento. Sea X_0 la fracción inicial de bolas rojas en la urna y X_n la fracción de bolas rojas en la urna una vez realizado el experimento n veces. Entonces $(X_n)_{n \in \mathbb{N}}$ es una martingala con respecto a la filtración que genera esta sucesión. Antes de proceder a verificar la afirmación anterior, debemos considerar el modelo matemático preciso del experimento aleatorio en cuestión, para poder calcular las esperanzas condicionales. Notemos que al momento de la n -ésima extracción hay

$$b_n = r + v + nc$$

bolas en la urna. Sean (U_i) variables aleatorias independientes e idénticamente distribuidas, $r, v > 0$ y definamos $X_0 = r/(r + v)$ y para $n \geq 1$:

$$Y_{n+1} = \mathbf{1}_{U_{n+1} \leq X_n} \quad \text{y} \quad X_{n+1} = \frac{r + v + nc}{r + v + (n + 1)c} X_n + \frac{c}{r + v + (n + 1)c} Y_n.$$

Esta es la descripción matemática que utilizaremos del experimento considerado anteriormente y en él, la variable X_n es función de X_0, U_1, \dots, U_n para $n \geq 1$ (de hecho es función de X_{n-1} y U_n) y por lo tanto, U_{n+1} es independiente de \mathcal{F}_n , la σ -álgebra generada por X_0, \dots, X_n .

EJERCICIO 1.5. Verificar que la sucesión X es una martingala respecto de (\mathcal{F}_n) .

2.2. El teorema de muestreo opcional de Doob. Entre las razones por las cuales las martingalas son importantes, se encuentran los teoremas de convergencia de martingalas, que bajo ciertas condiciones de acotamiento nos permiten concluir la convergencia casi segura (o de otro tipo) de una martingala. Para abordar este resultado, es importante extender la igualdad $\mathbb{E}(X_n) = \mathbb{E}(X_0)$ para abarcar no sólo a tiempos deterministas como n , sino también a ciertos tiempos aleatorios, concepto que procedemos a discutir. Consideremos $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, $(\mathcal{F}_n)_{n \in \mathbb{N}}$ una filtración y $(X_n)_{n \in \mathbb{N}}$ una martingala respecto a la anterior filtración. Nuestro objetivo es observar a la martingala a un tiempo que a su vez es una variable aleatoria. Esto se logra como sigue: si $T : \Omega \rightarrow \mathbb{N}$ es una variable aleatoria y definimos a $X_T : \Omega \rightarrow \mathbb{R}$ por medio de

$$X_T(\omega) = X_{T(\omega)}(\omega),$$

entonces X_T resulta ser una variable aleatoria, puesto que si $B \in \mathcal{B}_{\mathbb{R}}$, entonces

$$X_T^{-1}(B) = \cup_{n \in \mathbb{N}} \{\omega \in \Omega : T(\omega) = n, X_n(\omega) \in B\}.$$

Mediante la anterior variable aleatoria, observamos a la martingala al tiempo aleatorio T . En realidad, trabajaremos con una clase más reducida de tiempos aleatorios, a saber, los tiempos de paro. Para explicarlos, pensemos que al instante n debemos decidir si parar a la martingala (definiendo a n como el valor de T) de acuerdo a la información que tenemos disponible (es decir \mathcal{F}_n). Esto motiva la siguiente

DEFINICIÓN. Sea $T : \Omega \rightarrow \mathbb{N} \cup \{\infty\}$ una variable aleatoria. Decimos que T es un tiempo de paro respecto a la filtración $(\mathcal{F}_n)_{n \in \mathbb{N}}$ si

$$\{\omega \in \Omega : T(\omega) = n\} \in \mathcal{F}_n \quad \forall n \in \mathbb{N}.$$

El lector puede verificar que T es un tiempo de paro respecto a la filtración $(\mathcal{F}_n)_{n \in \mathbb{N}}$ si y sólo si $\{T \leq n\} \in \mathcal{F}_n$.

TEOREMA 1.3. *Sea X una submartingala. Si T es tiempo de paro respecto a $(\mathcal{F}_n)_{n \in \mathbb{N}}$ y T está acotado por N entonces*

$$\mathbb{E}(X_T) \leq \mathbb{E}(X_N).$$

Si X es una martingala entonces

$$\mathbb{E}(X_T) = \mathbb{E}(X_N).$$

DEMOSTRACIÓN. Por hipótesis, existe un natural $N > 0$ tal que $T \leq N$. Así,

$$\mathbb{E}(X_T) = \sum_{n=0}^N \mathbb{E}(X_n \mathbf{1}_{(T=n)}) ,$$

pero como el conjunto $\{T = n\}$ pertenece a \mathcal{F}_n

$$\mathbb{E}(X_n \mathbf{1}_{(T=n)}) \leq \mathbb{E}(\mathbb{E}(X_N \mid \mathcal{F}_n) \mathbf{1}_{(T=n)}) = \mathbb{E}(X_N \mathbf{1}_{(T=n)}) ,$$

por lo que

$$\mathbb{E}(X_T) \leq \sum_{n=0}^N \mathbb{E}(X_N \mathbf{1}_{(T=n)}) = \mathbb{E}(X_N).$$

Si X es una martingala, la desigualdad que utilizamos es una igualdad. \square

El teorema anterior vale para tiempos de paro acotados y posteriormente, al hacer un análisis más a fondo de las martingalas y de los tiempos de paro podremos extender el teorema anterior a una familia más amplia de tiempos de paro.

EJERCICIO 1.6. Sea X una supermartingala. Pruebe que si T es un tiempo de paro acotado por N entonces

$$\mathbb{E}(X_T) \geq \mathbb{E}(X_N).$$

2.3. El teorema de muestreo opcional de Doob y el problema de la ruina. En esta sección aplicaremos el teorema de muestreo opcional para tiempos de paro acotados para resolver algunas preguntas concernientes a un problema clásico dentro de la probabilidad, el problema de la ruina. Utilizaremos las martingalas del ejemplo (??). Supongamos que dos personas, A y B, juegan a los volados, donde A gana el n -ésimo volado con probabilidad $p \in (0, 1)$. Si A cuenta con una fortuna inicial de a pesos, B una de b pesos y apuestan en una serie de volados, un peso cada volado, hasta que uno de los dos quede sin dinero, ‘¿cuál es la probabilidad de que A se quede con la fortuna de B?’ y ‘¿cuál es la duración esperada del juego?’

Para responder a dichas preguntas, primero las formularemos en términos de la caminata aleatoria simple de la siguiente manera: Sean $(X_i)_{i=1}^\infty$ variables aleatorias independientes que toman el valor 1 con probabilidad p y -1 con probabilidad $1-p$. Así, X_i toma el valor 1 si A le gana un peso a B y el valor -1 si pierde un peso en el i -ésimo volado. Sean

$$S_0 = 1 \quad \text{y} \quad S_n = X_1 + \cdots + X_n \quad \text{para } n \geq 1.$$

Así, $a + S_n$ representa a la fortuna de A después de n volados, y por lo tanto, si

$$T_x = \inf \{n \geq 1 : S_n = x\},$$

A le gana su fortuna a B si $T_b < T_{-a}$. Para responder la primera pregunta, debemos calcular $\mathbb{P}(T_b < T_{-a})$.

La cantidad de volados que juegan hasta que alguien se quede sin dinero es $T_b \wedge T_{-a}$, por lo que para responder a la segunda pregunta, debemos

$$\text{calcular } \mathbb{E}(T_b \wedge T_{-a}).$$

El análisis es distinto si se trata de un juego justo ($p = 1/2$) o no. Haremos el caso $p = 1/2$ y el caso $p \neq 1/2$ se dejará indicado como ejercicio.

Necesitamos un resultado preliminar.

PROPOSICIÓN 1.3. *Para cualquier $a, b > 0$, $\mathbb{P}(T_b \wedge T_{-a} < \infty) = 1$.*

DEMOSTRACIÓN. Sea K un entero mayor a $a+b$. Notemos que $\mathbb{P}(|S_K| \geq a \vee b) > 0$ y que, como los eventos

$$|S_K| \geq a \vee b, |S_{2K} - S_K| \geq a \vee b, \dots$$

son independientes y tienen la misma probabilidad, el lema de Borel-Cantelli nos dice que $|S_{nK} - S_{(n-1)k}| \geq a \vee b$ para una infinidad de índices n casi seguramente. Por otra parte, vemos que si $|S_{nK} - S_{(n-1)k}| \geq a \vee b$ entonces $T \leq nK$. \square

Caso $p = 1/2$: Como $(S_n)_{n \in \mathbb{N}}$ es martingala respecto a la filtración que genera, tiene media cero y $T_{-a} \wedge T_b \wedge n$ es un tiempo de paro acotado, se tiene que

$$0 = \mathbb{E}(S_{T_{-a} \wedge T_b \wedge n}).$$

Además, $(S_{T_{-a} \wedge T_b \wedge n})_{n \geq 1}$ converge a $S_{T_{-a} \wedge T_b}$ y los elementos de dicha sucesión están acotados por $a \vee b$. Por el teorema de convergencia acotada,

$$0 = \mathbb{E}(S_{T_{-a} \wedge T_b}) = -a\mathbb{P}(T_{-a} < T_b) + b\mathbb{P}(T_b < T_{-a}),$$

de donde

$$\mathbb{P}(T_b < T_{-a}) = \frac{a}{a+b}.$$

Para responder a la segunda pregunta en este caso, notemos que $(S_n^2 - n)_{n \in \mathbb{N}}$ es martingala respecto a la filtración que genera, ya que $\mathbb{E}(X_i) = 0$ y $\text{Var}(X_i) = 1$. Al utilizar el tiempo de paro acotado $T_{-a} \wedge T_b \wedge n$, vemos que

$$\mathbb{E}(S_{T_{-a} \wedge T_b \wedge n}^2) = \mathbb{E}(T_{-a} \wedge T_b \wedge n).$$

Como $(S_{T_{-a} \wedge T_b \wedge n})_{n \geq 1}$ es una sucesión acotada por $a^2 \vee b^2$ y converge a $S_{T_{-a} \wedge T_b}$, por lo que podemos aplicar el teorema de convergencia acotada para concluir que

$$\mathbb{E}(S_{T_{-a} \wedge T_b}) = \lim_{n \rightarrow \infty} \mathbb{E}(S_{T_{-a} \wedge T_b \wedge n}) = \lim_{n \rightarrow \infty} \mathbb{E}(T_{-a} \wedge T_b \wedge n).$$

Como $(T_{-a} \wedge T_b \wedge n)_{n \in \mathbb{N}}$ es una sucesión creciente de variables aleatorias no negativas, que converge a $T_{-a} \wedge T_b$, por el teorema de convergencia monótona, se tiene que

$$\mathbb{E}(T_{-a} \wedge T_b) = \lim_{n \rightarrow \infty} \mathbb{E}(T_{-a} \wedge T_b \wedge n) = \lim_{n \rightarrow \infty} \mathbb{E}(S_{T_{-a} \wedge T_b \wedge n}^2) = \mathbb{E}(S_{T_{-a} \wedge T_b}^2).$$

Finalmente, al utilizar el valor de $\mathbb{P}(T_b < T_{-a})$, vemos que

$$\mathbb{E}(T_{-a} \wedge T_b) = a^2 \frac{b}{a+b} + b^2 \frac{a}{a+b} = ab,$$

por lo que la cantidad esperada de volados hasta la ruina de alguno de los dos jugadores es ab .

EJERCICIO 1.7. Suponga que $p > 1 - p$.

- (1) Sea $\phi(x) = (p/q)^x$ y pruebe que $(\phi(S_n))_{n \in \mathbb{N}}$ es martingala respecto a la filtración que genera.
- (2) Note que al aplicar el teorema de muestreo opcional de Doob al tiempo de paro acotado $T_{-a} \wedge T_b \wedge n$ se obtiene

$$1 = \mathbb{E}(\phi(S_{T_{-a} \wedge T_b \wedge n})) .$$

Utilice alguna propiedad de la esperanza para pasar al límite conforme $n \rightarrow \infty$ y concluir que

$$1 = \mathbb{E}(\phi(S_{T_{-a} \wedge T_b})) = \phi(-a)\mathbb{P}(T_{-a} < T_b) + \phi(b)\mathbb{P}(T_b < T_{-a}) .$$

Concluya con el cálculo explícito de $\mathbb{P}(T_b < T_{-a})$.

- (3) Pruebe que $(S_n - n(2p - 1))_{n \in \mathbb{N}}$ es una martingala.
- (4) Note que al aplicar muestreo opcional al tiempo de paro $T_{-a} \wedge T_b \wedge n$ se obtiene

$$\mathbb{E}(S_{T_{-a} \wedge T_b \wedge n}) = (2p - 1)\mathbb{E}(T_{-a} \wedge T_b \wedge n) .$$

Aplique propiedades de la esperanza al lado derecho y de la probabilidad al lado derecho que permitan pasar al límite conforme $n \rightarrow \infty$ en la expresión anterior y obtener:

$$\begin{aligned}\mathbb{E}(T_{-a} \wedge T_b) &= \frac{1}{2p-1}\mathbb{E}(S_{T_{-a} \wedge T_b}) \\ &= \frac{1}{2p-1}(-a\mathbb{P}(T_{-a} < T_b) + b\mathbb{P}(T_b < T_{-a}))\end{aligned}$$

y calcule explícitamente $\mathbb{E}(T_{-a} \wedge T_b)$.

2.4. El teorema de convergencia casi segura. Para proceder a estudiar la convergencia casi segura de las martingalas, necesitamos una caracterización de la convergencia de sucesiones. Para esto, consideremos una sucesión real $\{x_n\}_{n=0}^{\infty}$. Para verificar si esta sucesión es convergente en \mathbb{R} , es necesario y suficiente probar que $\liminf_{n \rightarrow \infty} x_n = \limsup_{n \rightarrow \infty} x_n$ y que esta cantidad pertenece a \mathbb{R} . A continuación veremos una manera de concluir que el límite superior y el límite inferior de la sucesión coinciden: si $a < b$ son dos racionales, veamos cuantas veces cruzan hacia arriba los puntos x_0, x_1, \dots a $[a, b]$, cantidad que denotamos por $U_{[a,b]}(x_0, x_1, \dots)$ y cuyo cálculo procedemos a explicar: sean

$$\begin{aligned}A_1 &= \{k \in \mathbb{N} : x_k \leq a\} , \\ T_1(x_0, x_1, \dots) &= \begin{cases} \min A_1 & \text{si } A_1 \neq \emptyset \\ \infty & \text{si } A_1 = \emptyset \end{cases}\end{aligned}$$

y de manera recursiva, para $j \geq 1$

$$\begin{aligned} A_{2j} &= \{k \in \mathbb{N} : T_{2j-1} \leq k, x_k \geq b\}, \\ T_{2j}(x_0, x_1, \dots) &= \begin{cases} \min A_{2j} & \text{si } A_{2j} \neq \emptyset \\ \infty & \text{si } A_{2j} = \emptyset \end{cases} \\ A_{2j+1} &= \{k \in \mathbb{N} : T_{2j} \leq k, x_k \leq a\}, \\ T_{2j+1}(x_0, x_1, \dots) &= \begin{cases} \min A_{2j+1} & \text{si } A_{2j+1} \neq \emptyset \\ \infty & \text{si } A_{2j+1} = \emptyset \end{cases}. \end{aligned}$$

A partir de las anteriores cantidades, definimos

$$\begin{aligned} U_{[a,b]}(x_0, \dots, x_n) &= \sup \{k \in \mathbb{N} : k \geq 1, A_{2k} \neq \emptyset\} \\ &= \sup \{k \in \mathbb{N} : k \geq 1, T_{2k} < \infty\}, \end{aligned}$$

que es la cantidad de cruces hacia arriba de la sucesión en $[a, b]$ pues si que $T_{2k} < \infty$ entonces la sucesión ha cruzado $[a, b]$ hacia arriba de menos k veces, por definición de T_{2k} .

LEMA 1. *El límite inferior de la sucesión $(x_n)_{n \in \mathbb{N}}$ coincide con el límite superior de la misma si y sólo si para cualquier pareja de racionales $a < b$, la cantidad $U_{[a,b]}(x_0, x_1, \dots)$ es finita.*

DEMOSTRACIÓN. Si el límite inferior de la sucesión, l , no coincide con el límite superior de la misma, L , entonces existen dos racionales $a < b$ tales que $l < a < b < L$. Por definición de límite superior, para cada $n \in \mathbb{N}$ existe $m_n \in \mathbb{N}$ mayor que n tal que $x_{m_n} > b$ y similarmente, existe $m'_n \in \mathbb{N}$ mayor que n tal que $x_{m'_n} < a$. De lo anterior podemos concluir que $T_k < \infty$ para cada $k \in \mathbb{N}$, puesto que $T_1 \leq m'_1$, de lo cual $T_2 \leq m_{T_1}$ y si $T_{2k} < \infty$, entonces $T_{2k+1} < m'_{T_{2k}}$ y $T_{2k+2} < m_{T_{2k+1}}$. Así, como la sucesión $(T_k)_{k \geq 1}$ es estrictamente creciente, pues $a < b$, se sigue que el conjunto cuyo supremo es $\bar{U}_{[a,b]}(x_0, x_1, \dots)$ es no acotado y por lo tanto esta última cantidad es igual a ∞ .

Por otro lado, si $U_{[a,b]}(x_0, x_1, \dots) = \infty$ para alguna pareja de racionales $a < b$, entonces los conjuntos

$$\{n \in \mathbb{N} : x_n \leq a\} \quad \text{y} \quad \{n \in \mathbb{N} : x_n \geq b\}$$

son infinitos, por lo que el límite superior de la sucesión es mayor o igual a b y el inferior, menor o igual a a , y por lo tanto el límite superior y el inferior difieren. \square

TEOREMA 1.4. *Si $\sup_{n \in \mathbb{N}} \mathbb{E}(|X_n|) < \infty$, entonces $(X_n)_{n \in \mathbb{N}}$ converge casi seguramente a una variable aleatoria X que pertenece a L_1 .*

DEMOSTRACIÓN. Por el lema anterior, notemos que

$$\left\{ \liminf_{n \rightarrow \infty} X_n = \limsup_{n \rightarrow \infty} X_n \right\} = \bigcap_{\substack{a,b \in \mathbb{Q} \\ a < b}} \{U_{[a,b]}(X_0, X_1, \dots) < \infty\},$$

por lo que para demostrar la afirmación del teorema, veremos primero que $U_{[a,b]}$ ⁴ es una variable aleatoria finita casi seguramente. De esto se desprenderá que el conjunto del lado derecho de la anterior igualdad pertenece a \mathcal{F} y tiene probabilidad 1, por lo que el límite superior y el inferior de la martingala coinciden casi seguramente y por el lema de Fatou, obtendremos que el valor absoluto del límite inferior (y por lo tanto el del límite de la sucesión) tiene esperanza finita, las conclusiones del teorema.

LEMA 2. *para cada $k \in \mathbb{N}$ mayor o igual a 1, T_k es un tiempo de paro respecto a la filtración $(\mathcal{F}_n)_{n \in \mathbb{N}}$.*

DEMOSTRACIÓN. La prueba se hará por inducción: Para $k = 1$, la afirmación se deduce de la igualdad,

$$\{T_1 \leq n\} = \bigcup_{i=0}^n \{X_i \leq a\},$$

válida para cualquier $n \in \mathbb{N}$. Si $i \leq n$, el conjunto $\{X_i \leq a\}$ pertenece a \mathcal{F}_i y por lo tanto a \mathcal{F}_n , por lo que $\{T_1 \leq n\}$ pertenece a \mathcal{F}_n y por lo tanto T_1 es tiempo de paro respecto a la filtración.

Por otro lado, si T_k es tiempo de paro y $k = 2l$ es par ($l \geq 1$), entonces para $n \geq 2$

$$\begin{aligned} \{T_{k+1} \leq n\} &= \{T_k < n\} \cap \{T_{k+1} \leq n\} \\ &= \bigcup_{i=1}^{n-1} \left(\{T_k = i\} \cap \bigcup_{j=i+1}^n \{X_j \leq a\} \right) \in \mathcal{F}_n. \end{aligned}$$

Como $T_{k+1} \geq 2$, entonces T_{k+1} es tiempo de paro. Por otro lado, si $k = 2l + 1$ es tiempo de paro ($l \geq 0$), entonces para cada $n \geq 2$,

$$\begin{aligned} \{T_{k+1} \leq n\} &= \{T_k < n\} \cap \{T_{k+1} \leq n\} \\ &= \bigcup_{i=1}^{n-1} \left(\{T_k = i\} \cap \bigcup_{j=i+1}^n \{X_j \geq b\} \right) \in \mathcal{F}_n. \end{aligned}$$

De nueva cuenta $T_{k+1} \geq 2$, por lo que T_{k+1} es tiempo de paro. □

En particular, el lema anterior nos permite afirmar que T_k es una variable aleatoria (posiblemente extendida, pues puede tomar el valor ∞ .) Para continuar con la prueba del teorema, verifiquemos ahora que $U_{[a,b]}$ es una variable aleatoria, lo cual se desprende de manera inmediata del lema anterior pues

$$U_{[a,b]} = \sum_{k=1}^{\infty} \mathbf{1}_{(T_{2k} < \infty)}.$$

⁴Para la prueba de este teorema, las cantidades $U_{[a,b]}$ y T_k las evaluaremos en X_0, X_1, \dots sin indicarlo.

Ahora veamos que $U_{[a,b]}$ finita casi seguramente: para esto, introducimos a las variables aleatorias

$$U_{[a,b]}^n \sum_{k=1}^{\lfloor n/2 \rfloor} \mathbf{1}_{(T_{2k} < \infty)},$$

la cantidad de cruces hacia arriba de x_0, \dots, x_n en $[a, b]$. Como $(U_{[a,b]}^n)_{n \geq 1}$ es una sucesión creciente de variables aleatorias no-negativas que converge a $U_{[a,b]}$, entonces se sigue que

$$(1) \quad \mathbb{E}(U_{[a,b]}^n) \rightarrow \mathbb{E}(U_{[a,b]}).$$

Si $S_n = T_n \wedge n$, entonces S_n es un tiempo de paro acotado puesto que

$$\{S_n \leq k\} = \begin{cases} \{T_n \leq k\} & \text{si } k \leq n \\ \Omega & \text{si } k > n \end{cases},$$

por lo que la esperanza de la variable aleatoria

$$V_n = \sum_{k=1}^n X_{S_{2k}} - X_{S_{2k-1}}$$

es igual a cero. En la definición de la variable aleatoria V_n , puede haber muchos sumandos iguales a cero, puesto que para $k > \lfloor n/2 \rfloor$, $S_{2k} = S_{2k-1}$. Sabemos que (para cada $\omega \in \Omega$) existe $k \geq 1$ tal que $T_k \leq n < T_{k+1}$. Si k es par, entonces $V_n \geq (b-a) U_{[a,b]}^n$, por lo que en este caso,

$$(b-a) U_{[a,b]}^n - V_n \leq 0,$$

mientras que si $k = 2l + 1$ es impar, entonces

$$V_n \geq (b-a) U_{[a,b]}^n + X_n - X_{T_{2l+1}} \geq (b-a) U_{[a,b]}^n + X_n - a,$$

por lo que en este caso

$$(b-a) U_{[a,b]}^n - V_n \leq a - X_n,$$

de donde obtenemos una cota para cualquier $\omega \in \Omega$:

$$(b-a) U_{[a,b]}^n - V_n \leq (a - X_n)^+$$

Como V_n tiene esperanza 0,

$$(b-a) \mathbb{E}(U_{[a,b]}^n) = \mathbb{E}((b-a) U_{[a,b]}^n - V_n) \leq \mathbb{E}((a - X_n)^+),$$

por lo que

$$\mathbb{E}(U_{[a,b]}^n) \leq \frac{1}{b-a} \mathbb{E}((a - X_n)^+)$$

Esta es la clásica desigualdad de Doob, que nos permitirá terminar con la prueba del teorema, puesto de acuerdo a (??),

$$\begin{aligned}\mathbb{E}(U_{[a,b]}) &= \lim_{n \rightarrow \infty} \mathbb{E}(U_{[a,b]}^n) \\ &\leq \sup_{n \geq 1} \frac{1}{b-a} \mathbb{E}((a-X_n)^+) \\ &\leq \frac{1}{b-a} \left(\sup_{n \geq 1} \mathbb{E}(|X_n|) + |a| \right) < \infty.\end{aligned}$$

Así, la variable $U_{[a,b]}$ es finita \mathbb{P} -p.s., pues pertenece a L_1 . \square

Por ejemplo, como la esperanza es constante para una martingala, entonces una martingala no-negativa satisface las condiciones del teorema anterior. En los ejemplos que hemos analizado, vemos que la martingala $(Z_n)_{n \in \mathbb{N}}$ del ejemplo (??) converge casi seguramente, así como las martingalas de los ejemplos (??) y (??). Veamos que la martingala del ejemplo (??) también converge \mathbb{P} -p.s.: Por la desigualdad de Jensen para la esperanza condicional,

$$\mathbb{E}(|X_n|) = \mathbb{E}(|\mathbb{E}(X | \mathcal{F}_n)|) \leq \mathbb{E}(\mathbb{E}(|X| | \mathcal{F}_n)) = \mathbb{E}(|X|),$$

de donde

$$\sup_{n \in \mathbb{N}} \mathbb{E}(|X_n|) \leq \mathbb{E}(|X|) < \infty.$$

Consideremos a la martingala X del ejemplo (??) es un caso particular de la del ejemplo (??). En efecto, (X_n) converge casi seguramente (digamos a X_∞) y es una sucesión acotada. Por lo tanto, podemos aplicar el teorema de convergencia acotada para probar que si $A \in \mathcal{F}$ entonces

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_n \mathbf{1}_A) = \mathbb{E}(X_\infty \mathbf{1}_A).$$

Por otra parte, si $A \in \mathcal{F}_m$ y $m \leq n$ entonces

$$\mathbb{E}(X_m \mathbf{1}_A) = \mathbb{E}(X_n \mathbf{1}_A)$$

puesto que X es una martingala. Así, también vemos que

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_n \mathbf{1}_A) = \mathbb{E}(X_m \mathbf{1}_A).$$

Se concluye que para todo $A \in \mathcal{F}_m$

$$\mathbb{E}(X_m \mathbf{1}_A) = \mathbb{E}(X_\infty \mathbf{1}_A)$$

y que por lo tanto

$$X_m = \mathbb{E}(X_\infty | \mathcal{F}_m).$$

A continuación se presentarán las desigualdades de Doob, que permiten que las martingalas acotadas en L_p para alguna $p > 1$ de hecho son martingalas cerradas por su límite casi seguro.

2.5. Desigualdades maximales de Doob. Ahora veremos un criterio sencillo para verificar que una martingala converge no sólo casi seguramente sino también en L_p para algún $p > 1$. Para esto, estudiaremos al máximo valor que toma una martingala. Con una cota adecuada para el máximo, se puede entonces simplemente aplicar convergencia dominada para verificar la convergencia en L_p . Sea $M = (M_n, n \geq 0)$ una (\mathcal{F}_n) -submartingala. Definamos a

$$\overline{M}_n^+ = \max_{1 \leq i \leq n} M_i^+.$$

PROPOSICIÓN 1.4 (Desigualdad maximal de Doob). *Para toda $\lambda > 0$,*

$$\lambda \mathbb{P}(\overline{M}_n^+ > \lambda) \leq \mathbb{E}(M_n^+).$$

La cota obvia, obtenida al aplicar la desigualdad de Markov, es

$$\lambda \mathbb{P}(\overline{M}_n^+ > \lambda) \leq \mathbb{E}(\overline{M}_n^+);$$

el contenido no trivial de la desigualdad maximal de Doob es que de hecho podemos acotar la cola de la distribución del supremo de la martingala al utilizar la martingala misma.

DEMOSTRACIÓN. Recordemos que M_n^+ es una sub-martingala. Definamos a

$$A = \{\overline{M}_n^+ > \lambda\} \quad \text{y} \quad T = \min \{k \geq 0 : M_k^+ > \lambda\} \wedge n.$$

Notemos que

$$A \cap \{T = k\} = \{M_i^+ \leq \lambda \text{ para } i < k, M_k^+ > \lambda\} \in \mathcal{F}_k.$$

Por lo tanto

$$\lambda \mathbf{1}_{A \cap \{T=k\}} \leq M_k^+ \mathbf{1}_{A \cap \{T=k\}}.$$

Entonces

$$\begin{aligned} \lambda \mathbb{P}(A) &= \sum_{k=0}^n \lambda \mathbb{P}(A \cap \{T = k\}) \\ &\leq \sum_{k=0}^n \mathbb{E}(M_k^+ \mathbf{1}_{A \cap \{T=k\}}) \\ &\leq \sum_{k=0}^n \mathbb{E}(M_n^+ \mathbf{1}_{A \cap \{T=k\}}) \\ &= \mathbb{E}(M_n^+ \mathbf{1}_A) \\ &\leq \mathbb{E}(M_n^+). \end{aligned}$$
□

A partir de la desigualdad anterior, veremos que las normas p de \overline{M}_n^+ y de M_n^+ son comparables. Notemos que obviamente

$$\mathbb{E}(M_n^+) \leq \mathbb{E}(\overline{M}_n^+).$$

El contenido del siguiente resultado es establecer una especie de desigualdad recíproca.

PROPOSICIÓN 1.5 (Desigualdad L_p de Doob). *Para cualquier $p \in (1, \infty)$:*

$$\|\bar{M}_n^+\|_p \leq \frac{p}{p-1} \|M_n^+\|_p.$$

DEMOSTRACIÓN. Consideraremos una constante $K > 0$ y escribamos:

$$\begin{aligned} \mathbb{E}\left(\left(\bar{M}_n^+ \wedge K\right)^p\right) &= \int_0^K p\lambda^{p-1} \mathbb{P}\left(\bar{M}_n^+ > \lambda\right) d\lambda \\ &\leq \int_0^K p\lambda^{p-2} \int M_n^+ \mathbf{1}_{\bar{M}_n^+ > \lambda} d\mathbb{P} d\lambda \\ &= \int M_n^+ \int_0^{\bar{M}_n^+ \wedge K} p\lambda^{p-2} d\lambda d\mathbb{P} \\ &= \frac{p}{p-1} \int M_n^+ \left(\bar{M}_n^+ \wedge K\right)^{p-1} d\mathbb{P}. \end{aligned}$$

Al utilizar la desigualdad de Hölder, utilizando el exponente conjugado $q = p/(p-1)$ se obtiene:

$$\mathbb{E}\left(\left(\bar{M}_n^+ \wedge K\right)^p\right) \leq \frac{p}{p-1} \|X_n^+\|_p \left\| \left(\bar{M}_n^+ \wedge K\right)^q \right\|_q,$$

por lo que despejando se obtiene

$$\|\bar{M}_n^+ \wedge K\|_p \leq \frac{p}{p-1} \|M_n^+\|_p.$$

La demostración termina al tomar el límite conforme $K \rightarrow \infty$. \square

Finalmente, podemos obtener un criterio de convergencia en L_p para martingalas.

TEOREMA 1.5. *Si M_n es una martingala con $\sup_n \mathbb{E}(|M_n|^p) < \infty$ para alguna $p > 1$, X_n converge casi seguramente y en L_p a una variable M_∞ y se tiene que*

$$M_n = \mathbb{E}(M_\infty | \mathcal{F}_n).$$

DEMOSTRACIÓN. La hipótesis implica que $\sup_n \mathbb{E}(|M_n|) < \infty$, por lo que el teorema de convergencia casi segura de martingalas nos permite afirmar que M_n converge casi seguramente a M_∞ . Por otra parte, vemos que

$$\mathbb{E}\left(\sup_n |M_n|^p\right) = \lim_{n \rightarrow \infty} \mathbb{E}\left(\sup_{m \leq n} |M_m|^p\right) \leq \left(\frac{p}{p-1}\right)^p \sup_n \mathbb{E}(|M_n|^p) < \infty.$$

Puesto que

$$|M_n - M_\infty|^p \leq \left(2 \sup_n |M_n|\right)^p \in L_1,$$

podemos aplicar el teorema de convergencia dominada para ver que $M_n \rightarrow M_\infty$ en L_p conforme $n \rightarrow \infty$.

Finalmente, puesto que M_n converge a M_∞ en L_p también converge en L_1 y por lo tanto si $A \in \mathcal{F}$ entonces

$$\lim_{n \rightarrow \infty} \mathbb{E}(M_n \mathbf{1}_A) = \mathbb{E}(M_\infty \mathbf{1}_A).$$

Por otra parte, si $A \in \mathcal{F}_m$ y $m \leq n$ entonces

$$\mathbb{E}(M_m \mathbf{1}_A) = \mathbb{E}(M_n \mathbf{1}_A)$$

puesto que X es una martingala. Así, también vemos que

$$\lim_{n \rightarrow \infty} \mathbb{E}(M_n \mathbf{1}_A) = \mathbb{E}(M_m \mathbf{1}_A).$$

Se concluye que para todo $A \in \mathcal{F}_m$

$$\mathbb{E}(M_m \mathbf{1}_A) = \mathbb{E}(M_\infty \mathbf{1}_A)$$

y que por lo tanto

$$M_m = \mathbb{E}(M_\infty | \mathcal{F}_m).$$

□

2.6. La transformada martingala. Sea $M = (M_n)$ una martingala. Recorremos que nuestra interpretación es que $M_n - M_{n-1}$ es la ganancia que obtenemos de apostar en un juego justo nuestra fortuna al tiempo $n-1$. Si ahora decidimos apostar la fracción C_n de nuestra fortuna, entonces nuestra ganancia será $C_n(M_n - M_{n-1})$. Así, a las cantidades C_0, C_1, \dots la podemos pensar como la estrategia de apuesta y C_n obviamente dependerá de la información que tengamos al tiempo $n-1$, que la habíamos interpretado como \mathcal{F}_{n-1} . En otras palabras, se requiere que C_n sea \mathcal{F}_{n-1} -medible. Esta condición define a lo que se conoce como un **proceso predecible**. Nuestra ganancia al tiempo n al seguir la estrategia de apuesta $C = C_1, C_2, \dots$, que denotaremos por $(C \cdot M)_n$, está dada por

$$(C \cdot M)_0 = 0 \quad \text{y} \quad (C \cdot M)_n = \sum_{m=1}^n C_m (M_m - M_{m-1}).$$

TEOREMA 1.6. *Sea M una (sub)martingala y C un proceso predecible y acotado entonces $C \cdot M$ es una (sub)martingala.*

EJERCICIO 1.8. Pruebe el teorema anterior.

El teorema anterior es otra manifestación del hecho de que no es posible generar ganancias en un juego justo.

Por ejemplo, consideremos la siguiente estrategia: sean $a < b$ dos reales fijos y apostaremos ya sea todo lo que tengamos o nada con las siguientes reglas. Nos fijamos en M y esperamos hasta que M se encuentre por debajo de a , ahí comenzamos a apostar, deteniéndonos cuando M se encuentre por arriba de b . Repetimos al infinito. Obviamente esta estrategia trata de utilizar a los cruces hacia arriba de la martingala en el intervalo $[a, b]$ para producir ganancias. La definición formal de la estrategia es como sigue:

$$C_1 = \mathbf{1}_{M_0 \leq a} \quad \text{y} \quad C_n = \mathbf{1}_{C_{n-1}=0} \mathbf{1}_{M_{n-1} \leq a} + \mathbf{1}_{C_{n-1}=1} \mathbf{1}_{M_{n-1} \leq b}.$$

Sea $Y = C \cdot M$.

EJERCICIO 1.9. Sea U_n la cantidad de cruces hacia arriba que hace el proceso M en el intervalo $[a, b]$ antes de n . Argumente que

$$Y_n \geq (b - a) U_n + (M_n - a)^-.$$

Al tomar esperanzas verifique que se satisface la desigualdad de cruces de Doob

$$\mathbb{E}(U_n) \leq \frac{1}{b-a} \mathbb{E}\left((a - M_n)^+\right).$$

3. Martingalas e integrabilidad uniforme

El objetivo de esta sección es analizar el concepto de integrabilidad uniforme de una familia de variables aleatorias integrables. El interés de esta noción, para el estudio de las martingalas, es que permite caracterizar a las martingalas que son de la forma $X_n = \mathbb{E}(X | \mathcal{F}_n)$, y con esto, permite dar una versión muy general del teorema de paro opcional de Doob.

DEFINICIÓN. Sea $\{X_t\}_{t \in T}$ es una familia de variables aleatorias reales. Decimos que es uniformemente integrable si

$$\lim_{c \rightarrow \infty} \sup_{t \in T} \mathbb{E}(|X_t| \mathbf{1}_{|X_t| \geq c}) = 0.$$

EJEMPLO 1.5. La familia que consta de un sólo elemento $X \in L_1$ es uniformemente integrable. Esto se sigue de aplicar el teorema de convergencia dominada para concluir que

$$\lim_{c \rightarrow \infty} \mathbb{E}(|X| \mathbf{1}_{|X| > c}) = 0$$

al ser X casi seguramente finita.

EJEMPLO 1.6. Si $\{X_t\}_{t \in T}$ es tal que $\sup_{t \in T} \mathbb{E}(|X_t|^p) < \infty$ para alguna $p > 1$ entonces dicha familia es uniformemente integrable. En efecto, basta notar que

$$c^{p-1} \mathbb{E}(|X_t| \mathbf{1}_{|X_t| > c}) \leq \mathbb{E}(|X_t|^p).$$

EJEMPLO 1.7. Para cada $X \in L_1$, a la familia

$$E = \{\mathbb{E}(X | \mathcal{G}) : \mathcal{G} \text{ es sub}\sigma\text{-álgebra de } \mathcal{F}\}.$$

Se afirma que E es uniformemente integrable. En efecto, la desigualdad de Jensen implica que

$$\mathbb{E}(|\mathbb{E}(X | \mathcal{G})| \mathbf{1}_{|\mathbb{E}(X | \mathcal{G})| > c}) \leq \mathbb{E}(\mathbb{E}(|X| | \mathcal{G}) \mathbf{1}_{|\mathbb{E}(X | \mathcal{G})| > c}) = \mathbb{E}(X \mathbf{1}_{|\mathbb{E}(X | \mathcal{G})| > c}).$$

Por la desigualdad de Markov, vemos que

$$\mathbb{P}(|\mathbb{E}(X | \mathcal{G})| > c) \leq \frac{1}{c} \mathbb{E}(|\mathbb{E}(X | \mathcal{G})|) \leq \frac{1}{c} \mathbb{E}(|X|),$$

por lo cual

$$\lim_{c \rightarrow \infty} \sup_{\mathcal{G} \subset \mathcal{F}} \mathbb{P}(|\mathbb{E}(X | \mathcal{G})| > c) = 0.$$

Finalmente, se afirma que para cada $\varepsilon > 0$ existe $\delta > 0$ tal que si $\mathbb{P}(E) < \delta$, entonces $\mathbb{E}(|X| \mathbf{1}_E) < \varepsilon$. Esto se prueba a partir de la desigualdad

$$\mathbb{E}(|X| \mathbf{1}_E) \leq c\mathbb{P}(E) + \mathbb{E}(|X| \mathbf{1}_{|X|>c}).$$

Por convergencia dominada, el segundo término del lado derecho tiende a cero conforme $c \rightarrow \infty$. Así, dada $\varepsilon > 0$, escogemos $c > 0$ tal que el segundo sumando del lado derecho sea menor a $\varepsilon/2$. Basta entonces tomar $\delta < \varepsilon/2c$. Esto termina la prueba de que E es uniformemente integrable, puesto que dada $\varepsilon > 0$, escogemos δ tal que si $\mathbb{P}(A) < \delta$ entonces $\mathbb{E}(|X| \mathbf{1}_A) < \varepsilon$ y finalmente C tal que para toda $c \geq C$ y toda \mathcal{G} sub- σ -álgebra de \mathcal{F} tengamos

$$\mathbb{P}(|\mathbb{E}(X | \mathcal{G})| > c) < \delta.$$

Entonces

$$\mathbb{E}(|\mathbb{E}(X | \mathcal{G})| \mathbf{1}_{|\mathbb{E}(X | \mathcal{G})|>c}) < \varepsilon$$

para toda $c > C$.

En vista del ejemplo anterior, si $X_n = \mathbb{E}(X | \mathcal{F}_n)$ y $(\mathcal{F}_n, n \in \mathbb{N})$ es una filtración entonces la martingala (X_n) es uniformemente integrable. Un ejemplo de una martingala que no es uniformemente integrable es el siguiente: si U es una variable uniforme en $(0, 1)$ y $X_n = 2^n \mathbf{1}_{U \leq 2^{-n}}$, entonces X_n es una martingala respecto a la filtración que genera. Puesto que $U > 0$ casi seguramente, se sigue que $X_n \rightarrow 0$ casi seguramente. Sin embargo, $\lim_{c \rightarrow \infty} \sup_n \mathbb{E}(X_n \mathbf{1}_{X_n \geq c}) = 1$, por lo que no es uniformemente integrable.

Si $\{X_t\}_{t \in T}$ es uniformemente integrable, sea $c > 0$ tal que

$$\sup_{t \in T} \mathbb{E}(|X_t| \mathbf{1}_{X_t \geq c}) \leq 1.$$

Vemos que entonces

$$\sup_{t \in T} \mathbb{E}(|X_t|) \leq c + 1 < \infty,$$

por lo que la familia $\{X_t\}_{t \in T}$ es acotada en L_1 .

La importancia de la integrabilidad uniforme es que nos permite relacionar dos modos de convergencia, la casi segura y la convergencia en L_1 :

TEOREMA 1.7. *Si $\{X_n\}_{n \in \mathbb{N}}$ y X son variables aleatorias integrables tales que $X_n \rightarrow X$ casi seguramente, entonces las siguientes condiciones son equivalentes:*

- a) $\{X_n\}_{n \in \mathbb{N}}$ es uniformemente integrable.
- b) $X \in L_1$ y $X_n \rightarrow X$ en L_1 .

Como hemos visto anteriormente, una condición necesaria y suficiente para que una sucesión convergente casi seguramente también sea convergente en L_1 es que la sucesión sea uniformemente integrable, por lo que ahora estudiaremos martingalas uniformemente integrables para abarcar otro modo de convergencia en el estudio de las martingalas.

Si $(X_n)_{n \in \mathbb{N}}$ es una martingala uniformemente integrable (respecto a la filtración $(\mathcal{F}_n)_{n \in \mathbb{N}}$) entonces el conjunto $\{\mathbb{E}(|X_n|) : n \in \mathbb{N}\}$ es acotado, por lo que se satisfacen

las condiciones del teorema de convergencia de martingalas y por lo tanto existe una variable aleatoria integrable X a la que la sucesión converge casi seguramente conforme $n \rightarrow \infty$. Por ser la martingala uniformemente integrable, la convergencia también se dá en L_1 . Si A es un elemento de \mathcal{F}_n , la tercera condición que define a las martingalas nos permite afirmar que

$$\mathbb{E}(X_n \mathbf{1}_A) = \mathbb{E}(X_m \mathbf{1}_A) \quad \forall m \geq n,$$

y como

$$\mathbb{E}(X_n \mathbf{1}_A) \rightarrow \mathbb{E}(X \mathbf{1}_A)$$

por la convergencia de $(X_n)_{n \in \mathbb{N}}$ a X en L_1 , entonces

$$\mathbb{E}(X_n \mathbf{1}_A) = \mathbb{E}(X \mathbf{1}_A) \quad \forall A \in \mathcal{F}_n,$$

de donde se concluye que $X_n = \mathbb{E}(X | \mathcal{F}_n)$ (pues X_n es \mathcal{F}_n -medible) y por lo tanto, la martingala original era una martingala cerrada. De hecho:

TEOREMA 1.8. *Sea $(X_n)_{n \in \mathbb{N}}$ una martingala respecto a la filtración $(\mathcal{F}_n)_{n \in \mathbb{N}}$. Entonces existe una variable aleatoria integrable X tal que $X_n = \mathbb{E}(X | \mathcal{F}_n)$ si y sólo si $\{X_n : n \in \mathbb{N}\}$ es uniformemente integrable. Además, si se cumple alguna de las condiciones anteriores, $(X_n)_{n \in \mathbb{N}}$ converge casi seguramente y en L_1 a $\mathbb{E}(X | \mathcal{F}_\infty)$, donde*

$$\mathcal{F}_\infty = \sigma\left(\bigcup_{n \in \mathbb{N}} \mathcal{F}_n\right).$$

NOTA. Para una martingala cerrada, $(\mathbb{E}(X | \mathcal{F}_n))_{n \in \mathbb{N}}$ el límite casi seguro no tiene porque ser igual a X ; sin embargo, si $\mathcal{F} = \mathcal{F}_\infty$, entonces el límite casi seguro sí es igual a X .

DEMOSTRACIÓN. En el párrafo anterior, hemos visto como para cualquier martingala uniformemente integrable existe una variable aleatoria integrable X que la convierte en una martingala cerrada. Así, sólo hace falta verificar que una martingala cerrada es uniformemente integrable. Pero esto es inmediato, pues hemos verificado que si Σ es una familia de σ -álgebras en Ω contenidas en \mathcal{F} , entonces la familia de variables aleatorias $\{\mathbb{E}(X | G) : G \in \Sigma\}$ es uniformemente integrable, por lo que cualquier martingala cerrada lo es.

Si se satisfacen alguna de las dos condiciones, sea Y el límite casi seguro y en L_1 para la martingala $(Y_n = \mathbb{E}(X | \mathcal{F}_n))_{n \in \mathbb{N}}$. Como Y_n es \mathcal{F}_∞ -medible para cada $n \in \mathbb{N}$, se sigue que Y también lo es. Además, por la convergencia en L_1 , se sigue que para todo $A \in \mathcal{F}_n$,

$$\mathbb{E}(X \mathbf{1}_A) = \mathbb{E}(Y_n \mathbf{1}_A) = \lim_{m \rightarrow \infty} \mathbb{E}(Y_{n+m} \mathbf{1}_A) = \mathbb{E}(Y \mathbf{1}_A).$$

Sea

$$\mathcal{C} = \{A \in \mathcal{F}_\infty : \mathbb{E}(X \mathbf{1}_A) = \mathbb{E}(Y \mathbf{1}_A)\}.$$

Hemos visto que

$$\bigcup_{n \in \mathbb{N}} \mathcal{F}_n \subset \mathcal{C} \subset \mathcal{F}_\infty$$

y la anterior unión de σ -álgebras es un álgebra. Además, \mathcal{C} es una clase monótona, puesto que si $(A_n)_{n \in \mathbb{N}}$ es una sucesión creciente o decreciente de elementos de \mathcal{C} y A es el límite de la anterior sucesión de conjuntos (igual al límite superior o al inferior, que coinciden) entonces el teorema de convergencia dominada nos permite afirmar que

$$\mathbb{E}(X \mathbf{1}_A) = \lim_{n \rightarrow \infty} \mathbb{E}(X \mathbf{1}_{A_n}) = \lim_{n \rightarrow \infty} \mathbb{E}(Y \mathbf{1}_{A_n}) = \mathbb{E}(Y \mathbf{1}_A),$$

por lo que A pertenece a \mathcal{C} . Así, por el lema de clases monótonas, $\mathcal{C} = \mathcal{F}_\infty$, lo cual nos dice que

$$Y = \mathbb{E}(X \mid \mathcal{F}_\infty).$$

□

Bajo la hipótesis de integrabilidad uniforme, también podemos dar una primera extensión del teorema (??):

TEOREMA 1.9. *Si $(X_n)_{n \in \mathbb{N}}$ es una martingala uniformemente integrable respecto a la filtración $(\mathcal{F}_n)_{n \in \mathbb{N}}$ y T es un tiempo de paro respecto a la misma filtración entonces X_T es una variable aleatoria integrable y $\mathbb{E}(X_T) = \mathbb{E}(X_0)$.*

DEMOSTRACIÓN. Puesto que X_n converge conforme $n \rightarrow \infty$, digamos a X_∞ , podemos definir a X_T aún cuando T no sea finito. Para ver que X_T es integrable, notemos que para cada $A \in \mathcal{F}_n$:

$$\mathbb{E}(|X_n| \mathbf{1}_A) \leq \mathbb{E}(\mathbb{E}(|X_\infty| \mid \mathcal{F}_n) \mathbf{1}_A) = \mathbb{E}(|X_\infty| \mathbf{1}_A).$$

Dado que T es tiempo de paro, el evento $\{T = n\}$ pertenece a \mathcal{F}_n , por lo que de acuerdo a la desigualdad anterior

$$\begin{aligned} \mathbb{E}(|X_T|) &= \mathbb{E}(|X_\infty| \mathbf{1}_{T=\infty}) + \sum_{n \in \mathbb{N}} \mathbb{E}(|X_n| \mathbf{1}_{(T=n)}) \\ &\leq \sum_{n \in \mathbb{N}} \mathbb{E}(|X| \mathbf{1}_{(T=n)}) \\ &= \mathbb{E}(|X|) < \infty. \end{aligned}$$

Finalmente, sea

$$Y_n = X_\infty \mathbf{1}_{T=\infty} + \sum_{i=0}^n X_i \mathbf{1}_{(T=i)},$$

por lo que $(Y_n)_{n \in \mathbb{N}}$ converge casi seguramente a X_T . Por las desigualdades

$$|Y_n| \leq |X_\infty| \mathbf{1}_{T=\infty} + \sum_{i=0}^n |X_i| \mathbf{1}_{(T=i)} \leq |X_T| \in L_1,$$

podemos aplicar el teorema de convergencia dominada para concluir que

$$\mathbb{E}(Y_n) \rightarrow \mathbb{E}(X_T)$$

y como

$$\begin{aligned}\mathbb{E}(Y_n) &= \mathbb{E}(X_\infty \mathbf{1}_{T=\infty}) + \sum_{i=0}^n \mathbb{E}(X_i \mathbf{1}_{(T=i)}) \\ &= \mathbb{E}(X \mathbf{1}_{T=\infty}) + \sum_{i=0}^n \mathbb{E}(X \mathbf{1}_{(T=i)}) \\ &= \mathbb{E}(X \mathbf{1}_{T \leq n} \text{ ó } T=\infty),\end{aligned}$$

se concluye, mediante el uso del teorema de convergencia dominada que

$$\mathbb{E}(Y_n) \rightarrow \mathbb{E}(X) = \mathbb{E}(X_0)$$

y por lo tanto

$$\mathbb{E}(X_T) = \mathbb{E}(X_0).$$

□

La integrabilidad uniforme nos da un criterio importante para ver si podemos aplicar el teorema de muestreo opcional de Doob. En efecto, si $X = (X_n, n \in \mathbb{N})$ es cualquier martingala y T es un tiempo de paro finito, la integrabilidad uniforme de la martingala detenida $X^T = (X_{n \wedge T}, n \in \mathbb{N})$ implica la igualdad $\mathbb{E}(X_T) = \mathbb{E}(X_0)$.

4. La ley 0 – 1 de Kolmogorov

En esta sección veremos una primera aplicación de los teoremas de convergencia de martingalas a sucesiones de variables aleatorias independientes. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad en el cual están definidas una sucesión de variables aleatorias independientes $(X_i)_{i \in \mathbb{N}}$. Definiremos a

$$\mathcal{F}_n = \sigma(X_i : i \leq n) \quad \text{y a} \quad \mathcal{F}_\infty = \sigma(X_i : i \in \mathbb{N}).$$

El resultado que probaremos, la ley 0 – 1 de Kolmogorov, nos permite concluir bajo ciertas hipótesis, que un elemento de \mathcal{F}_∞ tiene probabilidad 0 ó 1. Para esto, sean

$$\mathcal{G}_n = \sigma(X_i : i > n) \quad \text{y} \quad \mathcal{T} = \bigcap_{n \in \mathbb{N}} \mathcal{G}_n.$$

Recordemos que \mathcal{T} es una σ -álgebra, pues es intersección de σ -álgebras, a la cual llamaremos la σ -álgebra cola. Por la hipótesis acerca de la independencia de las variables aleatorias, es natural esperar que \mathcal{F}_n y \mathcal{G}_n sean independientes y a continuación, veremos que efectivamente esto sucede, pero necesitamos un resultado antes para concluirlo.

LEMA 3. *Sea $\mathcal{G}_{n,k} = \sigma(X_{n+1}, \dots, X_{n+k})$. Entonces*

$$\mathcal{G}_n = \sigma\left(\bigcup_{k \geq 1} \mathcal{G}_{n,k}\right)$$

y las σ -álgebras \mathcal{F}_n y $\mathcal{G}_{n,k}$ son independientes.

El lema anterior es parte de uno más general que ya se verificó en cursos anteriores.

PROPOSICIÓN 1.6. *Las σ -álgebras \mathcal{F}_n y \mathcal{G}_n son independientes. Además, la σ -álgebra τ es independiente de \mathcal{F}_n para toda $n \in \mathbb{N}$.*

DEMOSTRACIÓN. Sea

$$\mathcal{A} = \sigma \left(\bigcup_{k \geq 1} \mathcal{G}_{k,n} \right).$$

Entonces \mathcal{A} es un álgebra que genera a la σ -álgebra \mathcal{G}_n . Si

$$\mathcal{M} = \{A \in \mathcal{G}_n : \mathbb{P}(A, B) = \mathbb{P}(A) \mathbb{P}(B) \ \forall B \in \mathcal{F}_n\},$$

entonces \mathcal{M} es una clase monótona que por el lema anterior contiene a \mathcal{A} . Por el lema de clases monótonas, también contiene a $\sigma(\mathcal{A})$, lo cual implica que \mathcal{F}_n es independiente de \mathcal{G}_n .

Para la segunda afirmación, notemos que $\tau \subset \mathcal{G}_n$ para toda $n \in \mathbb{N}$, por lo que τ es independiente de \mathcal{F}_n para toda $n \in \mathbb{N}$. \square

Finalmente, podemos enunciar y demostrar la Ley 0-1 de Kolmogorov:

TEOREMA 1.10. *Si $A \in \tau$, entonces $\mathbb{P}(A) \in \{0, 1\}$.*

DEMOSTRACIÓN. Como $A \in \tau$, entonces A es independiente de \mathcal{F}_n para cualquier $n \in \mathbb{N}$, de donde

$$\mathbb{E}(\mathbf{1}_A | \mathcal{F}_n) = \mathbb{P}(A) \quad \forall n \in \mathbb{N}.$$

Por otro lado, sabemos que

$$\mathbb{E}(\mathbf{1}_A | \mathcal{F}_n) \rightarrow \mathbb{E}(\mathbf{1}_A | \mathcal{F}_\infty) \quad \mathbb{P} - p.s.$$

y como $A \in \mathcal{F}_\infty$, entonces

$$\mathbb{P}(A) = \mathbf{1}_A \quad \mathbb{P} - p.s.,$$

de donde $\mathbb{P}(A) \in \{0, 1\}$. \square

5. Martingalas reversas y la ley fuerte de los grandes números

En esta sección, exploraremos una situación análoga a la contenida en el teorema de convergencia de martingalas naturales. Lo que se consideró en ese teorema fué la existencia del límite casi seguro y en L_1 de la sucesión

$$(\mathbb{E}(X | \mathcal{F}_n))_{n \in \mathbb{N}},$$

donde $(\mathcal{F}_n)_{n \in \mathbb{N}}$ es una subsucesión creciente de σ -álgebras de \mathcal{F} . Si ahora consideramos una colección $(\mathcal{F}_n)_{n \in \mathbb{Z}}$ creciente, ¿podemos concluir la existencia casi segura y en L_1 de $\mathbb{E}(X | \mathcal{F}_n)$ conforme $n \rightarrow -\infty$? Si $\mathcal{G}_n = \mathcal{F}_{-n}$ para $n \in \mathbb{N}$, por lo que $\mathcal{G}_{n+1} \subset \mathcal{G}_n$, lo que quisieramos afirmar es la existencia del límite casi seguro y en L_1 de $\mathbb{E}(X | \mathcal{G}_n)$ conforme $n \rightarrow \infty$.

Para indicar la relevancia de tal afirmación, consideremos una sucesión $(X_i)_{i=1}^{\infty}$ de variables aleatorias independientes e idénticamente distribuidas con media cero, $(S_n)_{n \in \mathbb{N}}$ la sucesión de sumas parciales asociadas y para $n \in \mathbb{N}$, $\mathcal{G}_n = \sigma(S_n, S_{n+1}, \dots)$. Se deja como ejercicio al lector comprobar que

$$\mathbb{E}(X_i | \mathcal{G}_n) = \frac{S_n}{n}, \quad i = 1, \dots, n, n \geq 1$$

por lo que

$$\mathbb{E}(S_n | \mathcal{G}_{n+1}) = \frac{n}{n+1} S_{n+1},$$

de donde

$$\mathbb{E}\left(\frac{S_n}{n} \mid \mathcal{G}_{n+1}\right) = \frac{S_{n+1}}{n+1}.$$

De aquí se sigue que

$$\frac{S_n}{n} = \mathbb{E}(X_1 | \mathcal{G}_n)$$

puesto que $\mathcal{G}_{n+1} \subset \mathcal{G}_n$, por lo cual la pregunta que formulamos anteriormente es acerca del límite casi seguro y en L_1 de S_n/n , resultado conocido como la ley fuerte de los grandes números.

Utilizaremos a continuación los resultados ya verificados sobre martingalas para atacar la pregunta que nos concierne, para lo cual necesitamos precisar cuales son los procesos con los cuales vamos a trabajar.

DEFINICIÓN. Sea $(X_n)_{n \in \mathbb{N}}$ una sucesión de variables aleatorias y $(\mathcal{G}_n)_{n \in \mathbb{N}}$ una sucesión decreciente de σ -álgebras contenidas en \mathcal{F} . Decimos que $(X_n)_{n \in \mathbb{N}}$ es una martingala reversa respecto a $(\mathcal{G}_n)_{n \in \mathbb{N}}$ si

- (1) $X_n \in L_1$ para toda $n \in \mathbb{N}$,
- (2) X_n es \mathcal{G}_n -medible para toda $n \in \mathbb{N}$ y
- (3) $\mathbb{E}(X_n | \mathcal{G}_{n+1}) = X_{n+1}$.

Notemos que de las propiedades de la esperanza condicional se sigue la igualdad

$$X_{n+2} = \mathbb{E}(X_{n+1} | \mathcal{G}_{n+2}) = \mathbb{E}(\mathbb{E}(X_n | \mathcal{G}_{n+1}) | \mathcal{G}_{n+2}) = \mathbb{E}(X_n | \mathcal{G}_{n+2}),$$

por lo que de manera inductiva se verifica

$$X_n = \mathbb{E}(X_0 | \mathcal{G}_n).$$

Así, la pregunta formulada anteriormente es simplemente verificar si una martingala reversa tiene un límite casi seguro y en L_1 .

TEOREMA 1.11. *Sea $(X_n)_{n \in \mathbb{N}}$ una martingala reversa respecto a $(\mathcal{G}_n)_{n \in \mathbb{N}}$. Entonces $(X_n)_{n \in \mathbb{N}}$ converge casi seguramente conforme $n \rightarrow \infty$ a una variable aleatoria X que pertenece a L_1 .*

NOTA. Nos basaremos en el teorema de convergencia casi segura para martingalas. En dicho teorema, se vió que para demostrar la existencia del límite casi seguro, era suficiente verificar que la cantidad de cruces hacia arriba

de X_0, X_1, \dots en $[a, b]$, denotada por $U_{[a,b]}(X_0, X_1, \dots)$, era finita casi seguramente para cualquier pareja de racionales a, b tal que $a < b$. En la prueba del teorema, vimos que la cantidad de cruces hacia arriba de X_0, X_1, \dots en $[a, b]$ así como la cantidad de cruces hacia arriba de X_0, \dots, X_n en $[a, b]$, $U_{[a,b]}^n(X_0, \dots, X_n)$ eran variables aleatorias y se demostró la desigualdad clásica de Doob.

DEMOSTRACIÓN. Sea $m \in \mathbb{N}$, $Y_n = -X_{(m-n)\vee 0}$ y $\mathcal{H}_n = \mathcal{G}_{(m-n)\vee 0}$ para $n \in \mathbb{N}$. Verifiquemos que $(Y_n)_{n \in \mathbb{N}}$ es una martingala respecto a $(\mathcal{H}_n)_{n \in \mathbb{N}}$: si $n \in \mathbb{N}$, entonces

$$\begin{aligned}\mathbb{E}(Y_{n+1} | \mathcal{H}_n) &= \begin{cases} \mathbb{E}(-X_{m-n-1} | \mathcal{G}_{n-m}) & \text{si } n < m \\ \mathbb{E}(-X_0 | \mathcal{G}_n) & \text{si } n \geq m \end{cases} \\ &= \begin{cases} -X_{n-m} & \text{si } n < m \\ -X_0 & \text{si } n \geq m \end{cases} \\ &= Y_n.\end{aligned}$$

Así, por la desigualdad clásica de Doob, y al utilizar la igualdad

$$U_{[-b,-a]}^n(Y_0, \dots, Y_n) = U_{[a,b]}^n(X_0, \dots, X_n)$$

se tiene que

$$\mathbb{E}\left(U_{[a,b]}^n(X_0, \dots, X_n)\right) \leq \frac{1}{b-a} (|b| + \mathbb{E}(|Y_n|)) = \frac{1}{b-a} (|b| + \mathbb{E}(|X_0|)).$$

De esta manera, vemos que

$$\mathbb{E}\left(U_{[a,b]}(X_0, X_1, \dots)\right) \leq \frac{1}{b-a} (|b| + \mathbb{E}(|X_0|)),$$

por lo que $U_{[a,b]}(X_0, X_1, \dots) < \infty$ \mathbb{P} -p.s.. Esto nos dice que existe el límite casi seguro de X_n conforme $n \rightarrow \infty$ y para ver que pertenece a L_1 , aplicamos el lema de Fatou:

$$\mathbb{E}\left(\left|\lim_{n \rightarrow \infty} X_n\right|\right) \leq \liminf_{n \rightarrow \infty} \mathbb{E}(|X_n|) = \liminf_{n \rightarrow \infty} \mathbb{E}(|\mathbb{E}(X_0 | \mathcal{G}_n)|) \leq \liminf_{n \rightarrow \infty} \mathbb{E}(|X_0|) < \infty.$$

□

Ahora, veamos que toda martingala reversa es uniformemente integrable, por lo que la convergencia casi segura nos permitirá concluir la convergencia en L_1 .

TEOREMA 1.12. *Si $(X_n)_{n \in \mathbb{N}}$ es una martingala reversa respecto a $(\mathcal{G}_n)_{n \in \mathbb{N}}$, entonces es uniformemente integrable.*

DEMOSTRACIÓN. Como $X_n = \mathbb{E}(X_0 | \mathcal{G}_n)$ y hemos visto que

$$\{\mathbb{E}(X | \mathcal{G}) : \mathcal{G} \in G\},$$

con G una familia de σ -álgebras contenidas en \mathcal{F} y X un elemento de L_1 es uniformemente integrable, se sigue que $\{X_n\}_{n \in \mathbb{N}}$ es uniformemente integrable. □

Pasaremos a la identificación del límite:

TEOREMA 1.13. *Sea $(X_n)_{n \in \mathbb{N}}$ una martingala reversa respecto a $(\mathcal{G}_n)_{n \in \mathbb{N}}$. Entonces*

$$\lim_{n \rightarrow \infty} X_n = \mathbb{E}\left(X_n \mid \bigcap_{n \in \mathbb{N}} \mathcal{G}_n\right).$$

DEMOSTRACIÓN. Sea X el límite casi seguro y en L_1 de $(X_n)_{n \in \mathbb{N}}$. Como X_m es \mathcal{G}_n -medible para cualquier $m \geq n$, se sigue que

$$X = \lim_{m \rightarrow \infty} X_{n+m}$$

es \mathcal{G}_n -medible para toda $n \in \mathbb{N}$, por lo que es medible respecto a la intersección de dichas σ -álgebras. Por otro lado, si $A \in \mathcal{G}_n$ para toda $n \in \mathbb{N}$, entonces

$$\mathbb{E}(X_n \mathbf{1}_A) = \mathbb{E}(X_0 \mathbf{1}_A)$$

y por la convergencia de X_n a X en L_1 ,

$$\int_A X \mathbb{P}(d=) \lim_{n \rightarrow \infty} \int_A X_n \mathbb{P}(d=) \int_A X_0 \mathbb{P}(d.)$$

□

Para finalizar esta sección, daremos una prueba de la ley fuerte de los grandes números que utiliza las ideas que se han desarrollado.

TEOREMA 1.14. *Sea $(X_i)_{i \in \mathbb{N}}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas tales que $X_i \in L_1$. Si $(S_n)_{n=0}^{\infty}$ denota a la sucesión de sumas parciales asociada, entonces*

$$\lim_{n \rightarrow \infty} \frac{S_n}{n} = \mathbb{E}(X_1) \quad \mathbb{P}\text{-p.s..}$$

DEMOSTRACIÓN. Sea $\mathcal{G}_n = \sigma(S_k : k \geq n)$ y $\mu = \mathbb{E}(X_1)$, por lo que $(S_n/n - \mu)_{n \geq 1}$ es una martingala reversa respecto a $(\mathcal{G}_n)_{n \in \mathbb{N}}$. Esto nos dice que $S_n/n - \mu$ tiene un límite conforme $n \rightarrow \infty$ y como

$$\lim_{n \rightarrow \infty} \frac{S_n}{n} - \mu = \lim_{n \rightarrow \infty} \frac{S_n - S_m}{n} - \mu = \lim_{n \rightarrow \infty} \frac{X_{m+1} + \cdots + X_n}{n} - \mu,$$

se sigue que dicho límite es medible respecto a $\sigma(X_k : k \geq m)$ para toda $m \in \mathbb{N}$, de donde es medible respecto a la σ -álgebra cola asociada a $(X_i)_{i \in \mathbb{N}}$ y por lo tanto es constante. Para determinar la constante⁵, recordemos que S_n/n también converge en L_1 y que una variable aleatoria constante es igual a su esperanza, por lo que

$$\lim_{n \rightarrow \infty} \frac{S_n}{n} - \mu = \mathbb{E}\left(\lim_{n \rightarrow \infty} \frac{S_n}{n} - \mu\right) = \lim_{n \rightarrow \infty} \mathbb{E}\left(\frac{S_n}{n} - \mu\right) = \mathbb{E}(S_1 - \mu) = 0,$$

de donde S_n/n converge a μ .

□

⁵Para no utilizar la integrabilidad uniforme en lo que sigue, se puede usar la ley débil de los grandes números.

6. Urnas de Pólya y el teorema de de Finetti

Comencemos por analizar con mayor profundidad el ejemplo de las urnas de Pólya. Sean (U_i) variables aleatorias independientes e idénticamente distribuidas, $r, v > 0$ y definamos $X_0 = r/(r + v)$ y para $n \geq 0$:

$$Y_{n+1} = \mathbf{1}_{U_{n+1} \leq X_n} \quad \text{y} \quad X_{n+1} = \frac{r + v + nc}{r + v + (n + 1)c} X_n + \frac{c}{r + v + (n + 1)c} Y_n.$$

Hemos interpretado a la sucesión $(X_n)_{n \in \mathbb{N}}$ como la fracciones sucesivas de bolas rojas en una urna que inicialmente contiene r bolas rojas y v bolas verdes (aunque con esta construcción podemos considerar a r y a v como reales positivos) y tal que, en cada unidad de tiempo, se revuelve, se extrae una bola y se regresa con c bolas del mismo color. Además, hemos visto que si $\mathcal{F}_n = \sigma(U_1, \dots, U_n)$, entonces (X_n) es una (\mathcal{F}_n) -martingala acotada. Esto implica que converge casi seguramente y en L_∞ a una variable aleatoria X_∞ que se puede interpretar como la proporción límite de la urna. Ahora determinaremos la distribución de X_∞ mediante una técnica importante que básicamente generaliza nuestra prueba de la ley fuerte de los grandes números.

Analicemos ahora a la sucesión (Y_n) : la variable Y_n se interpreta como la indicadora de que en la enésima extracción se obtuvo una bola roja. Ahora calcularemos la distribución conjunta de (Y_1, \dots, Y_n) , para lo cual necesitamos la notación del factorial ascendente

$$a^{(n)} = a(a + 1) \cdots (a + n - 1).$$

PROPOSICIÓN 1.7. *Las variables aleatorias (Y_n) son intercambiables y si $i_1, \dots, i_n \in \{0, 1\}$ y $s_n = i_1 + \dots + i_n$ entonces*

$$\mathbb{P}(Y_1 = i_1, \dots, Y_n = i_n) = \frac{(r/c)^{(s_n)}(v/c)^{(n-s_n)}}{((r+v)/c)^{(n)}}.$$

EJERCICIO 1.10. Pruebe la proposición anterior. Sugerencia, utilice el principio de inducción.

Definiremos a $S_n = Y_1 + \dots + Y_n$, por lo que

$$X_n = \frac{r + v + cS_n}{r + v + nc}$$

y por lo tanto $S_n/n \rightarrow X_\infty$. Más adelante, justificaremos el hecho de que

$$\mathbb{E}(Y_j | S_n, Y_{n+1}, Y_{n+2}, \dots) = \mathbb{E}(Y_1 | S_n, Y_{n+1}, Y_{n+2}, \dots) \text{ si } 1 \leq j \leq n,$$

de lo cual obtendremos

$$\mathbb{E}(Y_1 | S_n, Y_{n+1}, \dots) = S_n/n.$$

Al tomar el límite conforme $n \rightarrow \infty$, vemos que

$$\frac{r}{r+v} = \mathbb{E}(Y_1) = \mathbb{E}(X_\infty).$$

Procederemos análogamente para el cálculo de los momentos de X_∞ . Sea \mathcal{G}_n la σ -álgebra generada por las variables $f(Y_1, \dots, Y_n)$, donde f es una función (medible y) simétrica. Sea $\mathcal{H}_n = \sigma(\mathcal{G}_n, Y_{n+1}, Y_{n+2}, \dots)$.

PROPOSICIÓN 1.8. *Si π es una permutación de los índices 1 al n entonces*

$$\mathbb{E}(f(Y_1, \dots, Y_n) | \mathcal{H}_n) = \mathbb{E}(f(Y_{\pi_1}, \dots, Y_{\pi_n}) | \mathcal{H}_n).$$

La proposición anterior nos permite hacer utilizar la intercambiabilidad para hacer cálculos. Por ejemplo, al utilizar $f(y_1, \dots, y_n) = y_1$, vemos que $\mathbb{E}(Y_1 | \mathcal{H}_n) = S_n/n$. Otro ejemplo interesante es

$$\mathbb{E}(Y_1 Y_2 | \mathcal{H}_n) = \frac{1}{n(n-1)} \sum_{\substack{1 \leq i, j \leq n \\ i \neq j}} Y_j Y_j = \frac{1}{n(n-1)} \left(S_n^2 - \sum_{i=1}^n Y_i^2 \right) \rightarrow X_\infty^2.$$

Por lo tanto, vemos que

$$\mathbb{E}(X_\infty^2) = \mathbb{E}(Y_1 Y_2) = \mathbb{P}(Y_1 = 1, Y_2 = 1) = \frac{r(r+c)}{(r+v)(r+v+c)} = \frac{\text{B}(r/c+1, v/c)}{\text{B}(r/c, v/c)}.$$

El mismo argumento, muestra que

$$\mathbb{E}(X_\infty^n) = \frac{\text{B}(r/c+n, v/c)}{\text{B}(r/c, v/c)},$$

de lo cual se deduce que X_∞ tiene los mismos momentos que una variable B de parámetros r/c y v/c . Al ser la variable B acotada, cualquier variable aleatoria que tenga los mismos momentos tendrá dicha distribución. Se concluye que X_∞ tiene distribución B de parámetros r/c y v/c . Sin embargo, también obtenemos una consecuencia sorprendente: aunque las variables Y_1, Y_2, \dots disten mucho de ser iid, vemos que si $\mathcal{H}_\infty = \bigcap_n \mathcal{H}_n$, $i_1, \dots, i_n \in \{0, 1\}$ y $s_n = i_1 + \dots + i_n$ entonces:

$$\mathbb{P}(Y_1 = i_1, \dots, Y_n = i_n | \mathcal{H}_\infty) = X_\infty^{s_n} (1 - X_\infty)^{n-s_n} = \prod_{j=1}^n \mathbb{P}(Y_1 = i_j | \mathcal{H}_\infty),$$

por lo cual la sucesión Y_1, Y_2, \dots es iid, pero condicionalmente a \mathcal{H}_∞ (y por lo tanto, también condicionalmente a X_∞).

Este es un caso particular del teorema de de Finetti que afirma que toda sucesión de variables intercambiables es condicionalmente iid.

7. Regularización de martingalas

Ahora daremos una extensión adicional del teorema de convergencia de martingalas y probaremos el teorema de regularización de martingalas. Este último es útil a la construcción de una gran familia de procesos estocásticos entre los cuales se encuentran los procesos de Feller y en particular los procesos de Lévy. Nos centraremos en procesos a tiempo continuo.

DEFINICIÓN. Una **filtración** a tiempo continuo es una colección $(\mathcal{F}_t)_{t \geq 0}$ de sub σ -álgebras de \mathcal{F} tales que si $s \leq t$ entonces $\mathcal{F}_s \leq \mathcal{F}_t$. Decimos que la filtración es **continua por la derecha** si, al definir

$$\mathcal{F}_{t+} = \bigcap_{u > t} \mathcal{F}_u,$$

se tiene que $\mathcal{F}_t = \mathcal{F}_{t+}$. Decimos que la filtración es **completa** si \mathcal{F}_0 (y por lo tanto también cada \mathcal{F}_t con $t > 0$) contienen a los conjuntos \mathbb{P} nulos de \mathcal{F}_∞ . Decimos que la filtración satisface las **hipótesis habituales** si es continua por la derecha y completa.

Una colección de variables aleatorias $(X_t)_{t \geq 0}$ es una **martingala** respecto de $(\mathcal{F}_t)_{t \geq 0}$ si

- (1) X_t es \mathcal{F}_t -medible.
- (2) X_t es integrable.
- (3) Si $s \leq t$ entonces $\mathbb{E}(X_t | \mathcal{F}_s) = X_s$.

Análogamente se definen las nociones de supermartingala y submartingala al reemplazar la igualdad por \leq y \geq respectivamente.

Considere dos colecciones de variables aleatorias $(X_t, t \geq 0)$ y $(Y_t, t \geq 0)$, decimos que Y es una **modificación** de X si $\mathbb{P}(X_t = Y_t) = 1$ para toda $t \geq 0$.

Extenderemos ahora la noción de cantidad de cruces de una función $f : [0, \infty) \rightarrow \mathbb{R}$: recordemos que si $F \subset [0, \infty)$ es finito, ya tenemos definida la noción de la cantidad de cruces hacia arriba de $(f(t))_{t \in F}$ en el intervalo $[a, b]$, llamémosle $U_F(f, a, b)$. Si $T \subset \mathbb{R}$ es arbitrario, podemos definir

$$U_T(f, a, b) = \sup_{F \subset T, F \text{ finito}} U_F(f, a, b).$$

Es claro que si T es numerable y X es un proceso estocástico entonces $U_T(X, a, b)$ es una variable aleatoria. Por otra parte, si $T = [u, v] \cap \mathbb{Q}$, entonces para todo $t \in [u, v]$ existen los límites

$$f(t+) = \lim_{s \downarrow t, s \in T} \quad t \in [u, v)$$

y

$$f(t-) = \lim_{s \downarrow t, s \in T} \quad t \in (u, v]$$

si y sólo si $U_T(f, a, b) < \infty$ para cualquier pareja de racionales $a < b$. En este caso, si f es acotada entonces los límites por la derecha y por la izquierda son finitos.

TEOREMA 1.15 (Desigualdad de cruces de Doob). *Si $(X_t)_{t \geq 0}$ es una (\mathcal{F}_t) -supermartingala y $T \subset [0, \infty)$ es numerable entonces*

$$\mathbb{E}(U_T(X)) \leq \sup_{t \in T} \mathbb{E}((a - X_t)^-).$$

El teorema anterior se sigue de la desigualdad de cruces de Doob que ya demostramos al tomar supremos. Nuestro objetivo ahora será demostrar la siguiente proposición.

TEOREMA 1.16. *Sea $(X_t, t \geq 0)$ una martingala respecto a una filtración $(\mathcal{F}_t, t \geq 0)$ continua por la derecha y completa. Entonces existe una modificación Y de X que también es una martingala respecto de $(\mathcal{F}_t, t \geq 0)$ y tal que Y tiene trayectorias càdlàg casi seguramente.*

DEMOSTRACIÓN. Veamos primero que $\sup_{t \in [0, n] \cap \mathbb{Q}} |X_t| < \infty$ casi seguramente. En efecto, al pasar al límite (sobre conjuntos finitos que vayan creciendo a $[0, n] \cap \mathbb{Q}$, la desigualdad maximal de Doob nos dice que

$$\mathbb{P}\left(\sup_{t \in [0, n] \cap \mathbb{Q}} |X_s| = \infty\right) = \lim_{\lambda \rightarrow \infty} \mathbb{P}\left(\sup_{t \in [0, n] \cap \mathbb{Q}} |X_s| > \lambda\right) \leq \lim_{\lambda \rightarrow \infty} \frac{\mathbb{E}(|X_n|)}{\lambda} = 0.$$

Para cualquier $n \in \mathbb{N}$ y $a < b$, la desigualdad de cruces de Doob nos dice que

$$\mathbb{E}(U_{[0, n] \cap \mathbb{Q}}(X, a, b)) \leq |a| + \mathbb{E}(|X_n|) < \infty,$$

por lo cual

$$\mathbb{P}(U_{[0, n] \cap \mathbb{Q}}(X, a, b) < \infty) = 1.$$

Por σ -subaditividad, vemos que

$$\mathbb{P}(U_{[0, n] \cap \mathbb{Q}}(X, a, b) < \infty \text{ si } a, b \in \mathbb{Q}, a < b \text{ y } n \in \mathbb{N}) = 1.$$

En dicho conjunto, que denotaremos por N^c , X admite límites por la izquierda y por la derecha en t para todo $t \geq 0$, mismos que son finitos, y por lo tanto podemos definir a

$$\tilde{X}_t(\omega) = \begin{cases} X_{t+}(\omega) & \text{si } \omega \in N^c \\ 0 & \text{si } \omega \in N \end{cases}.$$

Como X_{t+} es \mathcal{F}_{t+} -medible y $\mathcal{F}_{t+} = \mathcal{F}_t$ entonces X_{t+} es \mathcal{F}_t -medible y puesto que N pertenece a \mathcal{F}_∞ y tiene probabilidad cero, entonces $N \in \mathcal{F}_t$ y por lo tanto \tilde{X}_t es \mathcal{F}_t -medible. Además, \tilde{X} es continuo por la derecha en N^c por el argumento siguiente: si $\varepsilon > 0$ entonces existe $\delta > 0$ tal que si $r \in [t, t + \delta] \cap \mathbb{Q}$ y $\omega \in N^c$ entonces $|\tilde{X}_t(\omega) - X_r(\omega)| < \varepsilon$; al tomar límite conforme $r \rightarrow s \in [t, t + \delta]$, vemos que $|\tilde{X}_t(\omega) - \tilde{X}_s(\omega)| \leq \varepsilon$. Una argumento análogo muestra que \tilde{X} admite límites por la izquierda en N^c .

Si $t_1 < t_2$ y s_n es una sucesión de racionales que decrecen a t_1 , sabemos que

$$\mathbb{E}(X_{t_2} | \mathcal{F}_{s_n}) = X_{s_n}$$

y por el teorema de convergencia de Lévy hacia abajo, vemos que casi seguramente y en L_1 :

$$\tilde{X}_{t_1} = \lim_{n \rightarrow \infty} X_{s_n} = \lim_{n \rightarrow \infty} \mathbb{E}(X_{t_2} | \mathcal{F}_{s_n}) = \mathbb{E}(X_{t_2} | \mathcal{F}_{t_1+}) = \mathbb{E}(X_{t_2} | \mathcal{F}_{t_1}) = X_{t_1},$$

por lo que \tilde{X} es una modificación de X .

Consideremos ahora $t_1 < t_2$ y s_n una sucesión de racionales que decrezcan a t_2 . Puesto que X_{s_n} converge casi seguramente y en L_1 a \tilde{X}_{t_2} , como vimos en el párrafo anterior, el teorema de convergencia dominada para la esperanza condicional nos dice que

$$\tilde{X}_{s_1} = \mathbb{E}(X_{s_n} \mid \mathcal{F}_{s_1}) \rightarrow \mathbb{E}\left(\tilde{X}_{t_2} \mid \mathcal{F}_{s_1}\right),$$

por lo que \tilde{X} es una \mathcal{F}_t -martingala. \square

CAPÍTULO 2

Movimiento Browniano

Consideremos una caminata aleatoria simple y simétrica $S = (S_n, n \in \mathbb{N})$. El teorema límite central afirma que S_n/\sqrt{n} converge débilmente a una variable normal estándar. Una manera de interpretar al movimiento browniano es como una extensión multidimensional (inclusive infinito-dimensional o funcional) del teorema límite central. En efecto, si S se extiende por interpolación lineal en cada intervalo $[n, n+1]$ y consideramos al proceso estocástico S^n dado por $S_t^n = S_{nt}/\sqrt{n}$, vemos que S_t^n converge débilmente a una normal de media 0 y varianza t . Por otra parte, como S tiene incrementos independientes y estacionarios (cuando nos restringimos a instantes de tiempo naturales) entonces si $0 = t_0 < t_1 < \dots < t_m$ entonces para n suficientemente grande los incrementos $S_{t_i}^n - S_{t_{i-1}}^n$, con $1 \leq i \leq m$ son independientes. Por lo tanto, vemos que dichos incrementos convergen débilmente a un vector aleatorio con entradas gaussianas independientes de varianzas respectivas $t_i - t_{i-1}$ para $1 \leq i \leq m$. El movimiento browniano es justamente un proceso estocástico que recoge este comportamiento límite de las caminatas aleatorias.

DEFINICIÓN. Un **movimiento browniano en ley** es un proceso estocástico $B = (B_t, t \geq 0)$ tal que:

- (1) $B_0 = 0$
- (2) B tiene incrementos independientes: si $0 = t_0 < t_1 < \dots < t_m$ entonces $B_{t_i} - B_{t_{i-1}}$, $1 \leq i \leq m$ son independientes
- (3) B tiene incrementos estacionarios: $B_{t+s} - B_t$ tiene la misma distribución que B_s y
- (4) la distribución de B_t es normal de media 0 y varianza t .

Un **movimiento browniano** es un movimiento browniano en ley que tiene trayectorias continuas.

1. Martingalas y procesos asociados

Continuaremos con algunos procesos asociados al Browniano que resultan ser útiles para su análisis. Comenzaremos con algunas martingalas.

PROPOSICIÓN 2.1. *Sea B un movimiento browniano. Entonces los siguientes procesos son martingalas.*

- (1) $B_t, t \geq 0$,

- (2) $B_t^2 - t, t \geq 0,$
 (3) $e^{\lambda B_t - \lambda^2 t/2} y$
 (4) $\cosh(\lambda B_t) e^{-\lambda^2 t/2}.$

DEMOSTRACIÓN. Se tiene que $B_t - B_s$ es independiente de \mathcal{F}_s para $s \leq t$; se deduce lo anterior pues por una parte $B_t - B_s$ es independiente de $(B_{s_i} - B_{s_{i-1}})_{i=0}^n$ para cualquier $n \geq 0$ y cualquier colección de reales

$$0 = s_0 \leq s_1 \leq \cdots \leq s_n \leq s.$$

y por otra, dichas variables aleatorias generan \mathcal{F}_s . (Luego, se aplica el lema de clases de Dynkin.)

- (1) Vemos que

$$0 = \mathbb{E}(B_t - B_s | \mathcal{F}_s) = \mathbb{E}(B_t | \mathcal{F}_s) - B_s,$$

pues B_s es \mathcal{F}_s medible. Se concluye que B es una $(\mathcal{F}_t)_{t \geq 0}$ -martingala.

- (2) Al ser B una martingala y $B_t - B_s$ independiente de \mathcal{F}_s , se tiene que

$$\begin{aligned} t - s &= \mathbb{E}(B_t - B_s) \\ &= \mathbb{E}\left((B_t - B_s)^2 \mid \mathcal{F}_t\right) \\ &= \mathbb{E}(B_t^2 \mid \mathcal{F}_t) - 2\mathbb{E}(B_t B_s \mid \mathcal{F}_t) + B_s^2 \\ &= B_s^2 - \mathbb{E}(B_t^2 \mid \mathcal{F}_s), \end{aligned}$$

de acuerdo a las propiedades de la esperanza condicional.

- (3) Basta recordar que el cálculo de la transformada de Laplace de una variable normal estandar y utilizar el que $B_t - B_s$ es independiente de \mathcal{F}_s para $s \leq t$ y se distribuye $N(0, t - s)$ pues entonces:

$$e^{\lambda^2(t-s)/2} = \mathbb{E}\left(e^{\lambda(B_t - B_s)}\right) = \mathbb{E}\left(e^{\lambda(B_t - B_s)} \mid \mathcal{F}_t\right) = \mathbb{E}\left(e^{\lambda B_t} \mid \mathcal{F}_t\right) e^{-\lambda B_s}. \quad \square$$

Contruyamos ahora una martingala a dos parámetros con el movimiento browniano: consideremos

$$M_{t,s} = B_t - B_s$$

para $0 \leq s < t$ y $\mathcal{F}_{s,t} = \sigma(B_u - B_s : u \in [s, t]).$ Entonces, como $\mathcal{F}_{s,t}$ es independiente de \mathcal{F}_s (por la propiedad de incrementos independientes de B) y está contenida en \mathcal{F}_t , si $0 \leq u \leq s < t \leq v$, se tiene que

$$\begin{aligned} \mathbb{E}(M_{u,v} \mid \mathcal{F}_{s,t}) &= \mathbb{E}(B_v - B_u \mid \mathcal{F}_{s,t}) \\ &= \mathbb{E}(B_v - B_t \mid \mathcal{F}_{s,t}) + \mathbb{E}(B_t - B_u \mid \mathcal{F}_{s,t}) + B_t - B_s \\ &= B_t - B_s = M_{t,s}. \end{aligned}$$

Ahora analizaremos cuatro procesos importantes que ilustran propiedades de invariancia de la distribución del movimiento browniano.

PROPOSICIÓN 2.2. *El movimiento browniano B tiene las siguientes propiedades de invariancia.*

Simetría: $-B$ es un movimiento browniano

Homogeneidad temporal: Para toda $t \geq 0$ el proceso B^t dado por $B_s^t = B_{t+s} - B_t$ es un movimiento browniano independiente de $\sigma(B_s : s \leq t)$.

Autosimilitud: Para toda $c > 0$ el proceso $B_{ct}/\sqrt{c}, t \geq 0$ es un movimiento browniano.

Inversión temporal: El proceso

$$X_t = \begin{cases} 0 & t = 0 \\ tB_{1/t} & t > 0 \end{cases},$$

para $t \geq 0$, es un movimiento browniano.

DEMOSTRACIÓN.

- (1) Los incrementos de $-B$ son iguales a menos los incrementos de B . Por lo tanto, los primeros serán independientes y estacionarios. Las trayectorias de $-B$ son continuas y comienzan en cero. Finalmente, puesto que la distribución normal centrada es invariante ante la transformación $x \mapsto -x$, vemos que $-B_t$ y B_t tienen la misma distribución y por lo tanto $-B$ es un movimiento browniano.
- (2) Notemos que las trayectorias de B^t son continuas y comienzan en cero. Si $0 = s_0 < s_1 < \dots < s_n$, entonces

$$(B_{s_1}^t - B_{s_0}^t, \dots, B_{s_n}^t - B_{s_{n-1}}^t) = (B_{t+s_1} - B_t, \dots, B_{t+s_n} - B_{t+s_{n-1}});$$

puesto que los incrementos de B son independientes y estacionarios, vemos que los de B^t también lo son. Además, ya que $B_s^t = B_{t+s} - B_t$, vemos B_s^t tiene distribución normal $(0, s)$. Finalmente, para verificar que B^t es independiente de \mathcal{F}_t , notemos que por la propiedad de incrementos independientes de B , $B_{t_1}^t, \dots, B_{t_n}^t$ es independiente de $(B_{s_1}, \dots, B_{s_n})$ si $s_1, \dots, s_n \leq t$. Por clases monótonas, se verifica entonces que B^t es independiente de \mathcal{F}_t .

- (3) Se omite. Buen ejercicio
- (4) Puesto el proceso de interés es gaussiano, se verifica mediante un cálculo de varianzas-covarianzas, que $(X_{t_1}, \dots, X_{t_n})$ y $(B_{t_1}, \dots, B_{t_n})$ si $t_1, \dots, t_n \geq 0$. Por lo tanto, X es un movimiento browniano en ley. Sin embargo, no es nada trivial es que el proceso de interés tiene trayectorias continuas, en particular en cero. Ofrecemos dos pruebas: la primera es notar que B satisface la ley fuerte de los grandes números: $B_t/t \rightarrow 0$ conforme $t \rightarrow \infty$ casi seguramente. Esto se verifica al notar que $B_n, n \in \mathbb{N}$ es una caminata aleatoria con distribución de salto integrable y media cero por lo cual $B_n/n \rightarrow 0$ casi seguramente. Por otro lado, las variables $\sup_{s \in [0,1]} |B_{s+n} - B_n|$ independientes (como se prueba a partir de la propiedad de incrementos independientes de B) e idénticamente distribuidas (al utilizar la homogeneidad temporal del movimiento browniano). Por otro lado, la desigualdad L_2 de

Doob aplicada a la submartingala $|B|$ implica que

$$\mathbb{P}\left(\sup_{s \leq 1} |B_s| > n\lambda\right) \leq \frac{4}{n^2\lambda^2} \mathbb{E}(|B_1|^2).$$

Por lo tanto

$$\sum_n \mathbb{P}\left(\sup_{s \leq 1} |B_{n+s} - B_n| > n\lambda\right) < \infty$$

y por el lema de Borel-Cantelli observamos que casi seguramente existe un índice N tal que

$$\frac{\sup_{s \leq 1} |B_{s+n} - B_s|}{n} \leq \lambda$$

para $n \geq N$. Se sigue que

$$\limsup_{n \rightarrow \infty} \frac{\sup_{s \leq 1} |B_{s+n} - B_s|}{n} \leq \lambda$$

casi seguramente y como esto sucede para toda $\lambda > 0$ vemos que

$$\limsup_{n \rightarrow \infty} \frac{\sup_{s \leq 1} |B_{s+n} - B_s|}{n} = 0.$$

Finalmente, concluimos que

$$0 \leq \limsup_{s \rightarrow \infty} \frac{|B_s|}{s} \leq \limsup_{s \rightarrow \infty} \frac{|B_{\lfloor s \rfloor} + \sup_{t \leq 1} |B_{t+\lfloor s \rfloor} - B_{\lfloor s \rfloor}||}{\lfloor s \rfloor} = 0.$$

Así, se ha probado que $\lim_{s \rightarrow \infty} B_s/s = 0$. La segunda prueba comienza con notar que B y X tienen las mismas distribuciones finito-dimensionales y trayectorias continuas en $(0, \infty)$. Luego, si s_1^k, s_2^k, \dots es una enumeración de los racionales en $[0, 1/k]$ para $k \geq 1$, se escribe

$$\left\{ \lim_{t \rightarrow 0} B_t = 0 \right\} = \bigcap_{n \geq 1} \bigcup_{k \geq 1} \bigcap_i \left\{ |B_{s_i^k}| < 1/n \right\},$$

y se tiene una expresión similar para $\{\lim_{t \rightarrow 0} X_t = 0\}$. Por continuidad de \mathbb{P} , se sigue entonces que

$$\begin{aligned} 1 &= \mathbb{P}\left(\lim_{t \rightarrow 0} B_t = 0\right) \\ &= \lim_{n \rightarrow \infty} \lim_{k \rightarrow \infty} \lim_{i \rightarrow \infty} \mathbb{P}\left(|B_{s_i^k}| < 1/n, \dots, |B_{s_i^k}| < 1/n\right) \\ &= \lim_{n \rightarrow \infty} \lim_{k \rightarrow \infty} \lim_{i \rightarrow \infty} \mathbb{P}\left(|X_{s_i^k}| < 1/n, \dots, |X_{s_i^k}| < 1/n\right) \\ &= \mathbb{P}\left(\lim_{t \rightarrow 0} X_t = 0\right). \end{aligned} \quad \square$$

Finalmente, estudiaremos algunos otros procesos importantes que se definen a partir del movimiento browniano.

Proceso de calor: Es el proceso estocástico $(t, B_t), t \geq 0$.

Movimiento browniano multidimensional: Si $d \geq 1$, sean B^1, \dots, B^d d movimientos brownianos independientes. Entonces $B = (B^1, \dots, B^d)$ es el llamado movimiento browniano en dimensión d .

Procesos de Bessel de dimensión entera: Si B es un movimiento browniano d dimensional, el proceso R dado por $R_t = \|B_t\|$ es el llamado proceso de Bessel d -dimensional.

Máximo acumulativo: Si B es un movimiento browniano unidimensional, su máximo acumulativo es el proceso \bar{B} dado por $\bar{B}_t = \max_{s \leq t} B_s$. Es un proceso adaptado respecto a la filtración canónica de B y tiene trayectorias continuas.

Proceso de tiempos de arribo: Se trata del inverso generalizado del máximo acumulativo; formalmente se trata del proceso T dado por

$$T_a = \inf \{t \geq 0 : B_t > a\}$$

para $a \geq 0$. Es un proceso con trayectorias no decrecientes y continuas por la derecha. De hecho, veremos que es un subordinador estable. Al utilizar la martingala exponencial del movimiento browniano es fácil calcular su transformada de Laplace.

Proceso de tiempo de positividad: Sea $A_t = \int_0^t \mathbf{1}_{B_s > 0} ds$. Entonces A_t es una variable aleatoria, lo cual se prueba al notar que la función $(t, \omega) \mapsto B_t(\omega)$ es medible en el espacio producto, lo que es consecuencia de que las trayectorias de B sean continuas. Entonces se puede aplicar Tonelli para concluir que A_t es variable aleatoria. Con esto, se deduce que A es un proceso con trayectorias continuas. Podrémos calcular explícitamente la distribución de A_t .

Proceso de edad de las excursiones: Sea $g_t = \sup \{s \leq t : B_s = 0\}$. Puesto que para cada $t \geq 0$, $B_t \neq 0$ casi seguramente, entonces casi seguramente $g_t < t$.

2. Vectores gaussianos y la distribución normal multivariada

A continuación, trabajaremos con variables aleatorias con valores en \mathbb{R}^n ; a los elementos de \mathbb{R}^n los tomaremos como vectores columna. Si $x \in \mathbb{R}^n$ e $i \in \{1, \dots, n\}$, denotaremos por x_i a su i -ésima coordenada y si x' es el vector transpuesto de x , escribiremos $x' = (x_1, \dots, x_n)$ ó $x = (x_1, \dots, x_n)'$.

DEFINICIÓN. Un **vector gaussiano** es una variable aleatoria X definida en un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ y con valores en \mathbb{R}^n y tal que para cualquier $\lambda \in \mathbb{R}^n$, la combinación lineal

$$\lambda \cdot X = \sum_i \lambda_i X_i$$

tiene distribución normal. Asociado a un vector gaussiano está el vector de medias $\mu = (\mathbb{E}(X_1), \dots, \mathbb{E}(X_n))'$ y la matriz de varianzas-covarianzas Σ (de tamaño $n \times n$) tal que $\Sigma_{i,j} = \Sigma_{j,i} = Cov(X_i, X_j)$.

Primero recordaremos, o más bien formalizaremos, algunos cálculos para el caso unidimensional. Sea X una variable aleatoria normal de media μ y varianza σ^2 . Entonces X tiene la misma distribución que $\sigma N + \mu$ donde N es una variable normal estándar.

- (1) Calculemos la función generadora de momentos de X en términos de la de N :

$$\mathbb{E}(e^{uX}) = \mathbb{E}(e^{u\sigma N + u\mu}) = e^{u\mu} \mathbb{E}(e^{u\sigma N}).$$

- (2) Calculemos ahora la función generadora de momentos de N :

$$\mathbb{E}(e^{uN}) = \int_{-\infty}^{\infty} e^{ux} e^{-x^2/2} \frac{1}{\sqrt{2\pi}} dx = \int_{-\infty}^{\infty} e^{u^2/2} e^{-(x-u)^2/2} \frac{1}{\sqrt{2\pi}} dx = e^{u^2/2}.$$

- (3) Concluimos que

$$\mathbb{E}(e^{uX}) = e^{u\mu} e^{-u^2\sigma^2/2}.$$

- (4) Probemos la desigualdad

$$\mathbb{P}(N > x) \leq e^{-x^2/2}$$

si $x > 0$. Ésta desigualdad se sigue del siguiente razonamiento: para $x, \lambda > 0$:

$$e^{\lambda x} \mathbb{P}(\xi > x) \leq \mathbb{E}(e^{\lambda \xi} \mathbf{1}_{\xi > x}) \leq \mathbb{E}(e^{\lambda \xi}) = e^{-\lambda^2/2},$$

por lo cual para cualquier $\lambda > 0$,

$$\mathbb{P}(\xi > x) \leq e^{-\lambda x + \lambda^2/2}.$$

Al minimizar el lado derecho de la expresión anterior sobre $\lambda > 0$ (el mínimo ocurre cuando $\lambda = x$), se obtiene la desigualdad deseada.

- (5) Calculemos ahora los momentos de N ; como su distribución es simétrica, los momentos de orden impar son cero. Los de orden par los calculamos como sigue: todos los momentos son finitos pues los momentos exponenciales son finitos; esto es, por el teorema de convergencia monótona

$$\sum_n \mathbb{E}(N^{2n}) \frac{1}{2n!} = \mathbb{E}(e^{uN}) < \infty$$

por lo cual todos los momentos de orden par son finitos. Otra forma de verlo es puesto que $x^n e^{-x^2/2} = x^n e^{-x^2/4} e^{-x^2/4}$, y $x \mapsto x^n e^{-x^2/4}$ es acotada y $x \mapsto e^{-x^2/4}$ es integrable, se sigue que todos los momentos son finitos. Esto implica que la función generadora de momentos es infinitamente diferenciable y que su enésima derivada en cero es el momento de orden n de N . Esto se verifica al utilizar los criterios para intercambiar derivadas e integrales del libro de Bartle. Sea ϕ_N la generadora de

momentos de la gaussiana. Hemos visto que $\phi_N(u) = e^{u^2/2}$, por lo que $\phi_N^{(2n)}(0) = 2n!/n!2^n$. Así:

$$\mathbb{E}(N^{2n}) = \frac{2n!}{n!2^n}.$$

Ahora, como la serie de momentos de N es absolutamente convergence, el teorema de convergencia dominada nos permite afirmar que

$$\mathbb{E}(e^{iuN}) = \sum_{n=0}^{\infty} \mathbb{E}(N^n) \frac{u^n}{n!} (i)^n = e^{-u^2/2}.$$

- (6) Un caso particular muy útil es que $\mathbb{E}(N^4) = 3$.
- (7) Ahora calculemos los momentos de $|N|$, ya tenemos a los momentos de orden par; los de orden impar se calculan de manera distinta:

$$\begin{aligned} \mathbb{E}(|N|^{2n+1}) &= \int_0^{\infty} \frac{e^{-x^2/2}}{\sqrt{2\pi}} x^{2n} 2x \, dx \\ &= \int_0^{\infty} \frac{e^{-y/2}}{\sqrt{2\pi}} y^n \, dy \\ &= \frac{2^{n+1}}{\sqrt{2\pi}} n! \\ &= 2^{n+1/2} n! / \sqrt{\pi}. \end{aligned}$$

- (8) Ahora calcularemos la función característica de X , al utilizar los cálculos sobre momentos pares y el siguiente razonamiento: al maximizar, se tiene que

$$e^{ux} e^{-x^2/2} = e^{ux-x^2/4} e^{-x^2/4} \leq e^{u^2} e^{-x^2/4},$$

por lo cual $\mathbb{E}(e^{u|N|}) < \infty$. Al utilizar el teorema de convergencia dominada, se sigue que

$$\mathbb{E}(e^{iuN}) = \sum_n \mathbb{E}(N^n) \frac{(iu)^n}{n!}.$$

Puesto que los momentos de orden impar de n son cero, se sigue que

$$\mathbb{E}(e^{iuN}) = \sum_n \frac{(2n)!}{n!2^n} \frac{u^{2n} (-1)^n}{(2n)!} = \sum_n \left(\frac{-u^2}{2} \right)^n \frac{1}{n!} = e^{-u^2/2}.$$

Ahora veremos que la distribución de un vector gaussiano está determinada por μ y A , tal como la distribución gaussiana está determinada por la media y la varianza. Para esto, sea $\lambda \in \mathbb{R}^n$ y calculemos la media y la varianza de $\lambda \cdot X$: la media es

$$\mathbb{E}(\lambda \cdot X) = \sum_{i=1}^n \lambda_i \mathbb{E}(X_i) = \sum_{i=1}^n \lambda_i \mu_i = \lambda \cdot \mu$$

y la varianza es

$$\begin{aligned}\text{Var}(\lambda \cdot X) &= \text{Var}\left(\sum_{i=1}^n \lambda_i X_i\right) \\ &= \sum_{i,j=1}^n \mathbb{E}(\lambda_i (X_i - \mu_i) \lambda_j (X_j - \mu_j)) \\ &= \sum_{i,j=1}^n \lambda_i \lambda_j A_{i,j} \\ &= \lambda' A \lambda.\end{aligned}$$

Recordemos que las variables aleatorias en \mathbb{R}^n están determinadas por su función característica. Como $\lambda \cdot X$ es una variable aleatoria gaussiana, se sigue que

$$\mathbb{E}(e^{i\lambda \cdot X}) = e^{i\lambda \cdot \mu} e^{-\lambda' A \lambda / 2}.$$

Se sigue por lo tanto que X tiene una distribución normal multivariada con media μ y matriz de varianzas-covarianzas A . Se deduce el siguiente corolario importante.

COROLARIO 1. *Las entradas de un vector gaussiano son independientes si y sólo si son no-correlacionadas.*

La prueba se basa en notar que si las entradas de un vector gaussiano son no-correlacionadas entonces la matriz de varianzas-covarianzas es diagonal lo cual implica que la función característica se factoriza y que por lo tanto las entradas son independientes.

Necesitaremos ver que las sucesiones débilmente convergentes de variables aleatorias gaussianas.

PROPOSICIÓN 2.3. *Si X_n es una sucesión de variables gaussianas que converge débilmente a una variable aleatoria X entonces X es gaussiana, $|X_n|^p$ es uniformemente integrable para toda $p > 0$ y $\mathbb{E}(X_n^p) \rightarrow \mathbb{E}(X^p)$.*

DEMOSTRACIÓN. Sean $\mu_n = \mathbb{E}(X_n)$ y $\sigma_n^2 = \text{Var}(X_n)$. Por hipótesis

$$e^{iu\mu_n - \sigma_n^2 u^2 / 2} = \mathbb{E}(e^{iuX_n}) \rightarrow \mathbb{E}(e^{iuX})$$

para toda $u \in \mathbb{R}$. Vemos entonces que

$$e^{-\sigma_n^2 u^2 / 2} = |\mathbb{E}(e^{iuX_n})| \rightarrow |\mathbb{E}(e^{iuX})|,$$

por lo que $e^{-\sigma_n^2}$ es convergente y como la función característica es distinta de cero en una vecindad de cero, vemos que σ_n^2 es una sucesión acotada y convergente, digamos a σ^2 . Esto nos muestra que μ_n es también una sucesión acotada. Pero entonces, como $e^{iu\mu_n}$ es convergente para cualquier $u \in \mathbb{R}$, esto no muestra que cualesquiera dos límites subsucesionales μ^1 y μ^2 de μ_n satisfacen $\mu^1 - \mu^2 = 2k\pi/u$ para todo

$u \in \mathbb{R}$ donde $k \in \mathbb{Z}$, lo cual forza la igualdad $k = 0$ y por lo tanto $\mu^1 = \mu^2$. Esto implica que μ_n converge, digamos a μ , y por lo tanto

$$\mathbb{E}(e^{iuX}) = e^{iu\mu - \sigma^2 u^2/2},$$

por lo cual X es normal de media μ y varianza σ^2 . Vemos que además,

$$\mathbb{E}(e^{uX_n}) \rightarrow \mathbb{E}(e^{uX}) < \infty,$$

por lo que para toda $p > 1$ se tiene que $\sup_n \mathbb{E}(|X_n|^p) < \infty$ y por lo tanto $(|X_n|^p)$ es uniformemente integrable para toda $p \geq 1$, lo cual a su vez implica que $\mathbb{E}(|X_n|^p) \rightarrow \mathbb{E}(|X|^p)$ para todo $p \geq 1$. \square

Ahora haremos algunos cálculos con la distribución gaussiana. Primero, calculemos la distribución de N^2 :

$$\mathbb{P}(N^2 \leq x) = 2\mathbb{P}(0 \leq N \leq \sqrt{x}),$$

por lo que

$$f_{N^2}(x) = \frac{e^{-x/2}}{\sqrt{2\pi}} \frac{1}{\sqrt{x}}.$$

Se concluye que N^2 tiene distribución Γ de parámetros $(1/2, 1/2)$, donde el primer parámetro es el de posición y el segundo el de escala. (Si $\gamma_{a,b}$ tiene distribución Γ de parámetros (a, b) entonces

$$\mathbb{P}(\gamma_{a,b} \in dx) = \frac{1}{\Gamma(a)} (bx)^{a-1} be^{-bx} dx,$$

por lo que $c\gamma_{a,b} \sim \Gamma(a, b/c)$ y esto último explica el nombre de parámetro de escala.)

Pasemos al cálculo de la distribución de N_2/N_1 , donde N_1 y N_2 son gaussianas independientes: primero calculamos la densidad de $(N_1, N_2/N_1)$ al utilizar la transformación $(x, z) \mapsto (x, zx)$, cuyo jacobiano es x , vemos que

$$f_{N_1, N_2/N_1}(x, z) = f_{N_1, N_2}(x, zx) x = e^{-x^2(1+z^2)/2} \frac{x}{2\pi}.$$

Al integrar z en la expresión anterior, utilizando el cambio de variable $y = x^2/2$, obtenemos:

$$f_{N_2/N_1}(z) = \int e^{-x^2(1+z^2)/2} \frac{x}{2\pi} dx = 2 \int e^{-y(1+z^2)} \frac{1}{2\pi} dy = \frac{1}{\pi(1+z^2)}.$$

Se sigue que N_2/N_1 tiene ley Cauchy; la distribución asociada se puede explicitar en términos de la función arcoseno.

Sea C una variable aleatoria Cauchy; ahora caracterizaremos a la distribución de $A = 1/(1 + C^2)$. Como

$$\mathbb{P}(A \geq x) = \mathbb{P}(1/x \geq 1 + C^2) = \mathbb{P}(1/x - 1 \geq C^2) = 2\mathbb{P}(0 \leq C \leq \sqrt{1/x - 1}),$$

entonces

$$f_A(x) = -2f_C\left(\sqrt{(1-x)/x}\right) \frac{1}{2\sqrt{(1-x)/x}} \frac{-1}{x^2} = \frac{1}{\pi\sqrt{x(1-x)}}.$$

Por lo tanto, A tiene distribución Beta de parámetros $1/2, 1/2$, que es la llamada distribución arcoseno. Así, vemos también que $N_1^2 / (N_1^2 + N_2^2)$ tiene distribución arcoseno.

Cuando X_1, \dots, X_δ son independientes y normales estándar, se puede calcular la distribución de $X_1 / \|\mathbf{X}\|$ mediante el siguiente razonamiento: como para $x > 0$

$$2\mathbb{P}(0 \leq X_1 / \|\mathbf{X}\| < x) = \mathbb{P}(X_1^2 < (X_2^2 + \dots + X_\delta^2) x^2 / (1 - x^2)),$$

entonces

$$\begin{aligned} & 2f_{X_1 / \|\mathbf{X}\|}(x) \\ &= \frac{\partial}{\partial x} \int_0^\infty dy \frac{1}{2^{\nu+1/2} \Gamma(\nu + 1/2)} y^{\nu-1/2} e^{-y/2} \int_0^{yx^2/(1-x^2)} dz \frac{1}{\sqrt{2} \Gamma(1/2)} z^{-1/2} e^{-z/2} \\ &= \int_0^\infty dy \frac{1}{2^{\nu+1/2} \Gamma(\nu + 1/2)} y^{\nu-1/2} e^{-y/2} y \frac{2x}{(1-x^2)^2} \frac{1}{\sqrt{2} \Gamma(1/2)} \left(y \frac{x^2}{1-x^2} \right)^{-1/2} e^{-y \frac{x^2}{2(1-x^2)}} \\ &= 2 \frac{1}{\Gamma(\nu + 1/2) \Gamma(1/2)} (1-x^2)^{-3/2} \int_0^\infty dy \frac{1}{2^{\nu+1}} y^{\nu-1} e^{y \frac{1}{2x^2}} \\ &= 2 \frac{\Gamma(\nu + 1)}{\Gamma(\nu + 1/2) \Gamma(1/2)} (1-x^2)^{\nu-1/2}. \end{aligned}$$

Notemos $X_1 / \|\mathbf{X}\|$ y $\|\mathbf{X}\|$ son independientes pues la distribución de \mathbf{X} es invariantante ante transformaciones ortogonales. La interpretación ahora es clara: como X_1^2 tiene distribución Γ de parámetros $1/2$ y $1/2$ se sigue que $\|\mathbf{X}\|$ tiene distribución Γ de parámetros $\delta/2$ y $1/2$, es independiente de $X_1 / \|\mathbf{X}\|$ cuya distribución es la de S_δ , por lo que se ha verificado la factorización de la distribución normal cuando δ es un entero positivo. La interpretación de la ley arcoseno es ahora clara: se trata de la distribución de $|X_1| / \|\mathbf{X}\|$ cuando $\delta = 2$.

3. Existencia del movimiento browniano

3.1. El método de Lévy. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad en el que están definidas una colección de variables aleatorias independientes e idénticamente distribuidas

$$(\xi_{i,n})_{0 \leq i \leq 2^n, n \geq 0}$$

de distribución $N(0, 1)$.

Definamos $X_0(0) = 0$, $X_0(1) = \xi_{0,0}$ y extendamos linealmente la definición de X_0 al intervalo $[0, 1]$. Definiremos una sucesión de procesos continuos con trayectorias continuas $(X_n)_{n \geq 0}$ postulando que X_n sea lineal sobre los intervalos de la forma $[k/2^n, (k+1)/2^n]$ y que

$$X_n\left(\frac{2j}{2^n}\right) = X_{n-1}\left(\frac{2j}{2^n}\right) \quad \text{y} \quad X_n\left(\frac{2j+1}{2^n}\right) = X_{n-1}\left(\frac{2j+1}{2^n}\right) + \frac{\xi_{2j+1,n}}{2^{(n+1)/2}}.$$

Una visualización de los procesos X_n se puede ver en la Figura ??

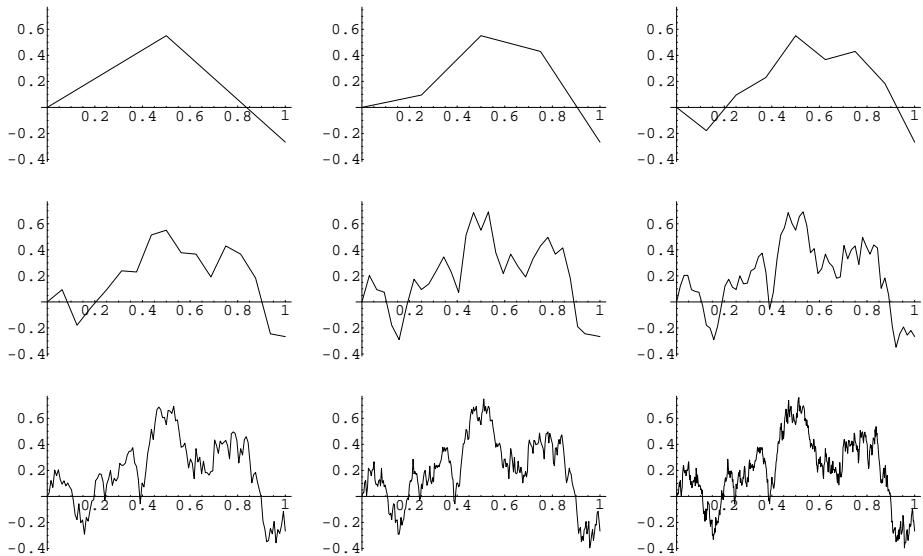


FIGURA 1. El procedimiento recursivo de Lévy para definir al movimiento browniano

Notemos que para toda $n \geq 0$, $Y_n = (X_n(k/2^n))_{0 \leq k \leq 2^n}$ es un vector aleatorio gaussiano, por lo que el proceso X_n es un proceso gaussiano ya que $X_n(t)$ es una combinación lineal de las entradas de Y^n . Para determinar a dicho proceso es suficiente explicitar su función de covarianza, que a su vez se obtiene por interpolación lineal en cada intervalo $[j/2^n, (j+1)/2^n]$ a partir de las cantidades:

$$\mathbb{E}\left(X_n\left(\frac{k}{2^n}\right) X_n\left(\frac{l}{2^n}\right)\right);$$

notemos que la función de media es cero.

LEMA 4. *Se tiene la igualdad*

$$\mathbb{E}\left(X_n\left(\frac{k}{2^n}\right) X_n\left(\frac{l}{2^n}\right)\right) = \frac{k \wedge l}{2^n}$$

DEMOSTRACIÓN. La prueba se hará por inducción sobre n , siendo la base inductiva ($n = 0$) inmediata. Si el lema es cierto para $n - 1$ y k y l son pares, entonces también es válido para n . Por otra parte, si $k = 2j + 1$ y l es par entonces, al utilizar la independencia entre $\xi_{2j+1,n}$ y X_{n-1} , se obtiene

$$X_n\left(\frac{2j+1}{2^n}\right) = \frac{1}{2}X_{n-1}\left(\frac{j}{2^{n-1}}\right) + \frac{1}{2}X_{n-1}\left(\frac{j+1}{2^{n-1}}\right) + \frac{\xi_{2j+1,n}}{2^{(n+1)/2}},$$

por lo que

$$\begin{aligned} & \mathbb{E}\left(X_n\left(\frac{k}{2^n}\right) X_n\left(\frac{l}{2^n}\right)\right) \\ &= \mathbb{E}\left(\frac{1}{2} X_{n-1}\left(\frac{j}{2^{n-1}}\right) X_{n-1}\left(\frac{l/2}{2^{n-1}}\right)\right) + \mathbb{E}\left(\frac{1}{2} X_{n-1}\left(\frac{j+1}{2^{n-1}}\right) X_{n-1}\left(\frac{l/2}{2^{n-1}}\right)\right) + 0 \\ &= \frac{1}{2} \frac{j \wedge (l/2)}{2^{n-1}} + \frac{1}{2} \frac{(j+1) \wedge (l/2)}{2^{n-1}}. \end{aligned}$$

Al analizar los distintos casos que pueden darse, nos damos cuenta de que

$$\mathbb{E}\left(X_n\left(\frac{k}{2^n}\right) X_n\left(\frac{l}{2^n}\right)\right) = \frac{(2j+1) \wedge l}{2^n}.$$

por otra parte, si tanto l como k son impares pero distintos, el análisis es análogo. Finalmente, si $k = l = 2j+1$, al escribir a $X_n(k/2^n)$ en términos de X_{n-1} y utilizar la hipótesis de inducción y la independencia entre $\xi_{k,n}$ y X_{n-1} , se observa que

$$\begin{aligned} \mathbb{E}\left(X_n\left(\frac{k}{2^n}\right)\right) &= \frac{1}{4} \frac{j}{2^{n-1}} + \frac{1}{4} \frac{j+1}{2^{n-1}} + 2 \frac{1}{4} \frac{j}{2^{n-1}} + \mathbb{E}\left(\left(\frac{\xi_{k,n}}{2^{(n+1)/2}}\right)^2\right) \\ &= \frac{4j+1}{2^{n+1}} + \frac{1}{2^{n+1}} = \frac{2j+1}{2^n}. \end{aligned} \quad \square$$

Verifiquemos ahora que la sucesión de procesos $(X_n)_{n \in \mathbb{N}}$ converge uniformemente. Para ésto, consideremos el evento

$$\begin{aligned} A_n &= \left\{ \sup_{t \in [0,1]} |X_n(t) - X_{n-1}(t)| > 2^{-n/4} \right\} \\ &= \left\{ \max_{0 \leq j \leq 2^{n-1}-1} \left| X_n\left(\frac{2j+1}{2^n}\right) - X_{n-1}\left(\frac{2j+1}{2^n}\right) \right| > 2^{-n/4} \right\} \\ &= \left\{ \max_{0 \leq j \leq 2^{n-1}-1} |\xi_{2j+1,n}| > 2^{(n+2)/4} \right\}. \end{aligned}$$

Entonces, por la subaditividad de \mathbb{P} , el hecho de que las variables $\xi_{i,j}$ tengan distribución $N(0, 1)$ y la cota para la cola de la distribución normal estandar:

$$\begin{aligned} \mathbb{P}(A_n) &\leq \sum_{j=0}^{2^{n-1}-1} \mathbb{P}\left(|\xi_{1,1}| > 2^{(n+2)/4}\right) \\ &\leq 2^{n-1} \times 2 \times e^{-2^{(n+2)/2}/2}. \end{aligned}$$

De la cota anterior, se concluye la convergencia de la serie $\sum_i \mathbb{P}(A_i)$, por lo cual, el lema de Borel-Cantelli nos permite afirmar que existe $E \in \mathcal{F}$ tal que $\mathbb{P}(E) = 1$ tal que si $\omega \in E$, existe $n_0 = n_0(\omega)$ tal que para $n \geq n_0$ se tiene que

$$|X_n(t) - X_{n-1}(t)| \leq 2^{-n/4},$$

de lo cual se deduce la convergencia uniforme de la sucesión $(X_n)_{n \in \mathbb{N}}$ hacia un límite $X = (X_t)_{t \in [0,1]}$ que es entonces continuo. La prueba estará (basicamente) terminada cuando verifiquemos que X es un movimiento browniano en $[0, 1]$; sin embargo, ésto se sigue del hecho de que una sucesión de variables aleatorias gaussianas que converge en probabilidad (lo cual está implicado por la convergencia casi segura) también converge en L^p para toda $p \geq 1$, se puede tomar el límite cuando $n \rightarrow \infty$ con $k = \lfloor 2^n s \rfloor / 2^n$ y $l = \lfloor 2^n t \rfloor / 2^n$ en el lema anterior para concluir que la función de media de X es cero y que

$$\mathbb{E}(X_t X_s) = t \wedge s.$$

Si $s_1 \leq s_2 \leq t_1 \leq t_2$, al igualdad anterior implica que

$$\mathbb{E}((X_{s_2} - X_{s_1})(X_{t_2} - X_{t_1})) = s_2 - s_1 - s_2 + s_1 = 0,$$

por lo que X tiene incrementos independientes (recordemos que se trata de un proceso gaussiano) y como tiene trayectorias continuas, empieza en cero y X_t tiene distribución $N(0, 1)$, se sigue que X es un movimiento browniano.

Para concluir, falta construir un movimiento browniano en $[0, \infty)$ en vez de en $[0, 1]$, pero ésto se puede lograr considerando una sucesión de movimientos brownianos independientes en $[0, 1]$ y concatenando sus trayectorias.

3.2. El método de Kolmogorov. Esbozaremos uno de los métodos más generales para construir procesos con trayectorias continuas, basados en el criterio de continuidad de Kolmogorov. Si $0 \leq t_1 < \dots < t_n$, definamos μ_{t_1, \dots, t_n} como la distribución normal multivariada de media cero y matriz de varianza-covarianza Σ dada por $\Sigma_{i,j} = t_j \wedge t_j$. Al quitar la coordenada i a un vector aleatorio con distribución μ_{t_1, \dots, t_n} obtenemos un vector aleatorio cuya distribución es $\mu_{t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n}$, por lo cual se puede aplicar el teorema de consistencia de Kolmogorov para concluir que existe un espacio de probabilidad en el que están definido un proceso estocástico $(B_t, t \geq 0)$ tal que la distribución de $(B_{t_1}, \dots, B_{t_n})$ es μ_{t_1, \dots, t_n} . Es claro que entonces B es un movimiento browniano en ley.

TEOREMA 2.1 (Criterio de continuidad de Kolmogorov). *Sea X un proceso estocástico real tal que existen constantes $\alpha, \beta, K \geq 0$ tales que*

$$\mathbb{E}(|X_t - X_s|^\alpha) \leq K(t - s)^{1+\beta}.$$

Entonces existe un proceso estocástico \tilde{X} tal que X es modificación de \tilde{X} , es decir que $\mathbb{P}(X_t = \tilde{X}_t) = 1$ para toda $t \geq 0$, y cuyas trayectorias son continuas.

Puesto que

$$\mathbb{E}(|B_t - B_s|^4) = 3(t - s)^2$$

si $0 \leq s \leq t$, vemos que el criterio de Kolmogorov aplica para construir una modificación de B con trayectorias continuas. Dicha modificación es un movimiento browniano.

4. La propiedad de Markov

La propiedad de homogeneidad temporal del movimiento browniano (que también comparte con el proceso de Poisson y otros procesos de Lévy) se puede interpretar también como una propiedad de Markov.

PROPOSICIÓN 2.4 (Propiedad de Markov para el movimiento browniano). *Sea B un movimiento browniano y $\mathcal{F}_t, t \geq 0$ su filtración canónica. Entonces para toda $t > 0$, el proceso B^t dado por $B_s^t = B_{t+s} - B_t$ es un movimiento browniano independiente de \mathcal{F}_t .*

De igual manera, la homogeneidad se puede extender a tiempos de paro e interpretar como una propiedad de Markov fuerte. Recordemos que un tiempo de paro es una función $T : \Omega \rightarrow [0, \infty]$ tal que $\{T \leq t\} \in \mathcal{F}_t$. Dado un tiempo de paro, podemos definir a la σ -álgebra \mathcal{F}_T mediante:

$$\mathcal{F}_T = \{A \in \mathcal{F} : A \cap T \leq t \in \mathcal{F}_t\}.$$

EJERCICIO 2.1. Probar que \mathcal{F}_T es una σ -álgebra.

PROPOSICIÓN 2.5 (Propiedad de Markov fuerte para el movimiento browniano). *Sea B un movimiento browniano y $\mathcal{F}_t, t \geq 0$ su filtración canónica. Si T es un tiempo de paro finito, el proceso B^T dado por $B_s^T = B_{T+s} - B_T$ es un movimiento browniano independiente de \mathcal{F}_T .*

DEMOSTRACIÓN. Consideremos a $T^n = \lceil 2^n T \rceil / 2^n$. Entonces T^n es un tiempo de paro puesto que

$$\left\{ T^n = \frac{k+1}{2^n} \right\} = \left\{ \frac{k}{2^n} < T \leq \frac{k+1}{2^n} \right\}.$$

Además, T^n decrece a T conforme $n \rightarrow \infty$. Notemos además que si $A \in \mathcal{F}_T$ entonces

$$A \cap \left\{ T^n \leq \frac{k+1}{2^n} \right\} = A \cap \left\{ T \leq \frac{k+1}{2^n} \right\} \in \mathcal{F}_{\frac{k+1}{2^n}},$$

por lo que en particular $A \in \mathcal{F}_{T^n}$.

Ahora descomponemos sobre el valor de T^n para calcular

$$\mathbb{E}\left(\mathbf{1}_A e^{iu_1 B_{t_1}^{T^n} + \dots + iu_m B_{t_m}^{T^n}}\right) = \sum_{k \in \mathbb{N}} \mathbb{E}\left(\mathbf{1}_{A \cap \left\{ T^n = \frac{k+1}{2^n} \right\}} e^{iu_1 B_{t_1}^{\frac{k+1}{2^n}} + \dots + iu_m B_{t_m}^{\frac{k+1}{2^n}}}\right)$$

que por la propiedad de Markov usual, se reexpresa como:

$$= \sum_{k \in \mathbb{N}} \mathbb{P}\left(\mathbf{1}_{A \cap \left\{ T^n = \frac{k+1}{2^n} \right\}}\right) \mathbb{E}(e^{iu_1 B_{t_1} + \dots + iu_m B_{t_m}}) = \mathbb{P}(A) \mathbb{E}(e^{iu_1 B_{t_1} + \dots + iu_m B_{t_m}})$$

Al pasar al límite conforme $n \rightarrow \infty$ se obtiene

$$\mathbb{E}\left(\mathbf{1}_A e^{iu_1 B_{t_1}^T + \dots + iu_m B_{t_m}^T}\right) = \mathbb{P}(A) \mathbb{E}(e^{iu_1 B_{t_1} + \dots + iu_m B_{t_m}}).$$

Puesto que la función característica determina a la distribución de un vector aleatorio, vemos que

$$\mathbb{P}(B_{t_1}^T \leq x_1, \dots, B_{t_m}^T \leq x_m \mid A) = \mathbb{P}(B_{t_1} \leq x_1, \dots, B_{t_m} \leq x_m),$$

por lo que B^T es un movimiento browniano independiente de A y así concluimos que B^T es un movimiento browniano independiente de \mathcal{F}_T . \square

5. Algunos cálculos distribucionales

En esta sección utilizaremos la propiedad de Markov y de Markov fuerte del movimiento browniano para calcular ciertos aspectos distribucionales de este proceso.

Sea B un movimiento browniano, \bar{B} su proceso de máximo acumulativo y T su proceso de tiempos de arriba.

PROPOSICIÓN 2.6. *El proceso T es un subordinador autosimilar. Además,*

$$\mathbb{E}(e^{-\lambda T_a}) = e^{-a\sqrt{2\lambda}} = e^{-a \int_0^\infty (1-e^{-\lambda x}) \nu(dx)}$$

donde

$$\nu(dx) = \frac{1}{\sqrt{2\pi x^3}}.$$

Finalmente, para cada $a > 0$, T_a tiene la misma distribución que a/\bar{B}_1^2 .

DEMOSTRACIÓN. Comenzamos con el cálculo de la transformada de Laplace. Al aplicar muestreo opcional a la martingala $M_t = e^{\lambda B_t - \lambda^2 t}$ al tiempo de paro T_a (hasta el cual M permanece acotada) se ve que

$$\mathbb{E}(e^{\lambda T_a}) = e^{-a\sqrt{2\lambda}}.$$

Por una parte

$$\int_0^\infty (1 - e^{-\lambda x}) \nu(dx) = \lambda \int_0^\infty e^{-\lambda y} \bar{\nu}(y) dy,$$

donde $\bar{\nu}(y) = \int_y^\infty \nu(dx)$, y, utilizando la definición de la función Γ , vemos que

$$\sqrt{\frac{2}{\lambda}} = \int_0^\infty \frac{\sqrt{2}e^{-\lambda x}}{\sqrt{\pi x}} dx,$$

de lo cual se deduce que si

$$\int_0^\infty (1 - e^{-\lambda x}) \nu(dx) = \sqrt{2\lambda}$$

entonces

$$\bar{\nu}(y) = \sqrt{\frac{2}{\pi x}}$$

y por lo tanto

$$\nu(dx) = \frac{1}{\sqrt{2\pi x^3}} \mathbf{1}_{x>0} dx.$$

Por construcción, T es el inverso continuo por la derecha de S , por lo que es no-decreciente y continuo por la derecha y por lo tanto càdlàg. Para ver que $T_0 = 0$ casi seguramente, notemos que

$$\mathbb{E}(e^{-\lambda T_0}) = 1,$$

y como $e^{-\lambda T_0} \leq 1$ entonces $e^{-\lambda T_0} = 1$ casi seguramente.

Veamos ahora que T tiene incrementos independientes y estacionarios. Puesto que $T_{a+b} - T_a$ es el tiempo que transcurre para que $B_{+T_a} - a$ sobrepase b , por lo que la propiedad de Markov fuerte nos dice que $T_{a+b} - T_a$ tiene la misma distribución que T_b y además es independiente de $\mathcal{F}_{T_a}^B$. Por otra parte, $\mathcal{F}_a^T \subset \mathcal{F}_{T_a}^B$, por lo que T tiene incrementos independientes. Se concluye que T es un subordinador.

Finalmente, debemos ver que es un subordinador autosimilar. Veremos específicamente que T_{ca} y $c^2 T_a$ tienen la misma distribución. Esto se deduce de que ambas variables tienen transformada de Laplace $\lambda \mapsto e^{-ac\sqrt{2\lambda}}$. Una prueba basada en la autosimilitud del movimiento browniano además nos dice que los procesos $T_{ca}, a \geq 0$ y $c^2 T_a, a \geq 0$ tienen la misma distribución. En efecto, recordemos que puesto que $B_{ct}/\sqrt{c}, t \geq 0$ y $\sqrt{c}B_t, t \geq 0$ tienen la misma distribución. T es el proceso de tiempos de arribo del segundo mientras que $cT_{\sqrt{c}a}, a \geq 0$ es el proceso de tiempos de arribo del primero.

Finalmente, notamos que por autosimilitud y la relación entre \overline{B} y T se sigue que

$$\mathbb{P}(T_a \leq t) = \mathbb{P}(a \leq \overline{B}_t) = \mathbb{P}\left(a/\sqrt{t} \leq \overline{B}_1\right) = \mathbb{P}\left(a^2/\overline{B}_1^2 \leq t\right). \quad \square$$

Veamos ahora que la distribución de \overline{B}_1 se conoce explícitamente.

PROPOSICIÓN 2.7 (Principio de reflexión). *El proceso estocástico B^b dado por*

$$B_t^b = \begin{cases} B_t & t < T_b \\ 2b - B_t & t \geq T_b \end{cases}$$

es un movimiento browniano. En consecuencia, la variable \overline{B}_1 tiene la misma distribución que $|B_1|$,

$$f_{B_1, \overline{B}_1}(a, b) = \mathbf{1}_{a < b} \frac{2b - a}{\sqrt{2\pi t^3}} e^{-(2b-a)^2/2t}$$

y

$$f_{T_a}(t) = \mathbf{1}_{t > 0} \frac{1}{\sqrt{2\pi t^3}} e^{-a^2/2t}.$$

DEMOSTRACIÓN. Si $\tilde{B}_t = B_{t+T_b} - b$, entonces \tilde{B} es un movimiento browniano independiente de $D_t = B_{t \wedge T_b}, t \geq 0$. Notemos que B se puede reconstruir a partir de D y \tilde{B} a partir de la igualdad

$$B_t = \begin{cases} D_t & t < T_b \\ b + \tilde{B}_{t-T_b} & t > T_b \end{cases}.$$

Puesto que $-\tilde{B}$ también es un movimiento browniano, vemos que entonces el proceso

$$B_t^b = \begin{cases} D_t & t < T_b \\ b - \tilde{B}_{t-T_b} & t > T_b \end{cases} = \begin{cases} D_t & t < T_b \\ 2b - B_t & t \geq T_b \end{cases}$$

es un movimiento browniano.

A través de la igualdades de conjuntos

$$\{T_b \leq t, B_t \leq b\} = \{T_b \leq t, B_t^b \geq b\} = \{B_t^b \geq b\},$$

vemos que

$$\begin{aligned} \mathbb{P}(T_b \leq t) &= \mathbb{P}(B_t \geq b) + \mathbb{P}(B_t \leq b, T_b \leq t) \\ &= \mathbb{P}(B_t \geq b) + \mathbb{P}(B_t^b \geq b) \\ &= 2\mathbb{P}(B_t \geq b) \\ &= \mathbb{P}(|B_t| \geq b). \end{aligned}$$

Si $a \leq b$, apliquemos un argumento similar con los conjuntos

$$\{\bar{B}_t \geq b, B_t \leq a\} = \{T_b \leq t, B_t \leq a\} = \{B_t^b \geq 2b - a\}$$

para obtener

$$f_{B_t, \bar{B}_t}(a, b) = \mathbf{1}_{a < b} \frac{\partial}{\partial a} \frac{2e^{-(2b-a)^2/2t}}{\sqrt{2\pi t}} = \mathbf{1}_{a < b} \frac{2(2b-a)e^{-(2b-a)^2/2t}}{\sqrt{2\pi t^3}}. \quad \square$$

Recordemos que una variable Cauchy estándar tiene densidad

$$\frac{1}{\pi(1+x^2)}.$$

En este caso la función de distribución es

$$\frac{1}{2} + \frac{\arctan(x)}{\pi}.$$

También se sabe que la distribución Cauchy es la misma que la del cociente de dos gaussianas estándar.

PROPOSICIÓN 2.8. *Sean B^1 y B^2 dos movimientos brownianos independientes y T^1 el proceso de tiempos de arribo de B^1 . Entonces $B_{T_a}^2$ tiene la misma distribución que aC.*

Veamos ahora un resultado clásico conocido como primera ley arcoseno de Paul Lévy. Sean

$$d_t = \inf \{s \geq t : B_s = 0\} \quad \text{y} \quad g_t = \sup \{s \leq t : B_s < 0\}.$$

PROPOSICIÓN 2.9. *Sea C una variable Cauchy; entonces d_t tiene la misma distribución que $t(1+C^2)$ y g_1 tiene la misma distribución que $1/(1+C^2)$.*

DEMOSTRACIÓN. Notemos que $d_t = t + T_0 \circ \theta_t$. Por lo tanto, si N es una gaussiana estándar independiente de B , se tiene que

$$\begin{aligned}\mathbb{P}(d_t > r) &= \mathbb{P}(t + T_0 \circ \theta_t > r) \\ &= \mathbb{P}(t + B_t^2/N^2 > r) \\ &= \mathbb{P}(t(1 + B_1^2/N^2) > r) \\ &= \mathbb{P}(t(1 + C^2) > r).\end{aligned}$$

Por otra parte, vemos que

$$\mathbb{P}(g_1 < t) = \mathbb{P}(d_t > 1) = \mathbb{P}(t(1 + C^2) > 1) = \mathbb{P}\left(\frac{1}{1+C^2} < t\right).$$

Al derivar, obtenemos la densidad arcoseno para g_1 :

$$f_{g_1}(t) = 2 \frac{1}{\pi(1/t)} \frac{1}{2} \sqrt{\frac{t}{t-1}} \frac{1}{t^2} = \frac{1}{\pi \sqrt{t} \sqrt{1-t}}.$$

□

CAPÍTULO 3

Integración estocástica

En este capítulo se dará una introducción al cálculo estocástico. Específicamente, veremos por qué la integral de Lebesgue-Stieltjes es insuficiente para integrar respecto del movimiento browniano y cómo se puede sortear este problema. Se introducirá la célebre fórmula de Itô y veremos cómo resolver ecuaciones diferenciales estocásticas conducidas por el movimiento browniano. Como veremos, la teoría de las martingalas es fundamental para desarrollar esta teoría.

El proceso estocástico más importante en estas notas es el movimiento browniano. Recordemos su definición.

DEFINICIÓN. Un **movimiento browniano** es una colección de variables aleatorias $B = (B_t, t \geq 0)$ tales que:

- (1) B comienza en cero
- (2) B tiene trayectorias continuas
- (3) B tiene incrementos independientes y estacionarios
- (4) La distribución de B_t es normal centrada de varianza t .

El teorema límite central nos permite entender a la distribución normal como una forma de aproximar a la suma de muchas contribuciones independientes (e idénticas desde el punto de vista probabilístico). Así, cada incremento del movimiento browniano se podría interpretar como una perturbación aleatoria obtenida de sumar muchas contribuciones pequeñas. Así, una forma de agregarle una fuente de error a una ecuación diferencial del tipo

$$dc_t = \varphi(c_t) dt$$

es considerar a

$$dc_t = \varphi(c_t) dt + dB_t \quad \text{ó equivalentemente} \quad c_t = x + \int_0^t \varphi(c_s) ds + B_t.$$

Un ejemplo muy concreto es el del proceso de Ornstein-Uhlenbeck, en el que $b(X_s) = \lambda X_s$. La interpretación en este caso es, si B denota nuestra ganancia o pérdida al jugar (continuamente) con un capital inicial x , X_t será nuestra ganancia si se paga interés a tasa λ por un préstamo ó un impuesto de la misma tasa.

Las ecuaciones diferenciales estocásticas surgen de la idea de hacer que la magnitud de la perturbación aleatoria dependan de la posición del sistema para llegar

a ecuaciones del tipo

$$dX_t = b(X_t) dt + \sigma(X_t) dB_t \quad \text{ó} \quad X_t = \int_0^t b(X_s) ds + \int_0^t \sigma(X_s) dB_s.$$

Sin embargo, surge inmediatamente el problema de interpretar a la integral respecto del movimiento browniano. Puesto que, como veremos, las trayectorias de B tienen variación infinita en cualquier intervalo, entonces no se puede interpretar como una integral de Lebesgue-Stieltjes. Aunque no sea posible definir a la integral estocástica como límite de sumas de Riemann trayectoria por trayectoria, sí se le puede definir como límite en probabilidad para una clase adecuada de integrandos (que no ven el futuro). La contribución de Itô fue darse cuenta de esto y entonces profundizar en el estudio de las similitudes y diferencias entre la integral usual y la integral estocástica. Al reinterpretar el artículo de Itô, algunos otros probabilistas como Doob se dieron cuenta de la similitud existente entre la integral estocástica y la transformada de martingala a tiempo discreto, lo cual ha marcó el desarrollo posterior del cálculo estocástico como una teoría apoyada fundamentalmente en las martingalas.

La integral estocástica respecto del movimiento browniano fue introducida en [?]. En este artículo, Itô introduce la integral estocástica respecto del movimiento browniano con la idea de utilizarla para darle sentido a ecuaciones diferenciales estocásticas (para las cuales da un teorema de existencia y unicidad) y utilizar a estas para dar construcciones de procesos de Markov con trayectorias continuas. Seguiremos de cerca este artículo.

1. Introducción a la integral de Lebesgue-Stieltjes

Comencemos por recordar la definición de integral de Lebesgue-Stieltjes cuando el integrando es una función de variación acotada.

DEFINICIÓN. Decimos que una función $f : [0, \infty) \rightarrow \mathbb{R}$ es de **variación acotada en el intervalo** $[0, t]$ si

$$\sup_{\pi=\{0=t_0 < \dots < t_n=t\}} \sum_i |f(t_i) - f(t_{i-1})| < \infty,$$

donde el supremo es sobre todas las particiones de $[0, t]$. Denotaremos por $V_t(f)$ al supremo anterior.

Diremos que f tiene **variación localmente acotada** si $V_t(f) < \infty$ para toda $t \geq 0$ y que f tiene variación acotada si $\lim_{t \rightarrow \infty} V_t(f) < \infty$.

Si f es una función de variación localmente acotada, la célebre descomposición de Jordan nos afirma que la podemos escribir como diferencia de dos funciones no decrecientes. En efecto, es fácil verificar que

$$t \mapsto \frac{V_t(f) + f(t)}{2}, \frac{V_t(f) - f(t)}{2}$$

son no decrecientes y su suma es $f(t)$. Una forma más intuitiva puede ser el definir la variación positiva o negativa de una función de variación acotada como

$$\sup_{\pi=\{0=t_0 < \dots < t_n=t\}} \sum_i |f(t_i) - f(t_{i-1})|^\pm \leq V_t(f) < \infty$$

y notar que ambas funciones son no decrecientes. Por otra parte, se tiene la descomposición

$$f(t) - f(0) = \sum |f(t_i) - f(t_{i-1})|^+ - |f(t_i) - f(t_{i-1})|^-,$$

lo cual nos hace sospechar que $f(t) - f(0) = V_t^+(f) - V_t^-(f)$. (Esto se puede justificar al escribir a las variaciones como límites al considerar una sucesión de particiones que se refinen y que alcancen los supremos.) Aún más, este argumento nos indica que si f es también continua por la derecha entonces f se puede escribir como la diferencia de dos funciones no-decrecientes continuas por la derecha. Puesto que cada una de estas funciones no decrecientes está asociada a una medida, vemos que existe una medida con signo μ tal que $f(t) = \mu([0, t])$. Si $g : [0, \infty) \rightarrow \mathbb{R}$ es localmente acotada y medible, podemos entonces definir a la **integral de Lebesgue-Stieltjes** de g respecto de f en $[0, t]$, denotada $\int_0^t g \, df$ mediante

$$\int_0^t g \, df = \int_0^t g \, d\mu.$$

El teorema siguiente nos afirma que las funciones de variación acotada aparecen naturalmente al tratar de considerar integrales elementales y extenderlas por "continuidad". Si nuestra función f es función continua por la derecha y π_n es una sucesión de particiones de $[0, t]$ que se refinan y cuyo paso tiende a cero, definamos

$$S^n(h) = \sum_{\pi_n} h(t_i) (f(t_{i+1}) - f(t_i)).$$

Si h es continua y f es de variación acotada en $[0, t]$ entonces $S^n(h) \rightarrow \int_0^t h \, df$.

TEOREMA 3.1. *La sucesión $(S^n(h), n \in \mathbb{N})$ converge a un límite para toda función continua h si y sólo si f es de variación acotada en $[0, t]$.*

DEMOSTRACIÓN. Ya hemos mostrado una de las implicaciones. Para la recíproca, supongamos que $(S^n(h), n \in \mathbb{N})$ converge a un límite para toda función continua h en $[0, t]$. Es bien sabido que el espacio de funciones continuas en $[0, t]$, al dotarlo de la norma uniforme, es un espacio de Banach. Además, T^n es una funcional lineal continua definido en él. Es fácil construir una función h de norma uniforme 1 tal que

$$T_n(h) = \sum_{\pi_n} |f(t_i) - f(t_{i-1})|$$

(por ejemplo al imponer que $h(t_i) = \operatorname{sgn}(f(t_i) - f(t_{i-1}))$ e interpolar linealmente). Por lo tanto,

$$V_t(f) \leq \sup_n \|T^n\|.$$

Por otra parte, para toda h continua se tiene que

$$\sup_n |T^n(h)| < \infty,$$

por lo que el principio de acotamiento uniforme (conocido también como teorema de Banach-Steinhaus) implica que

$$\sup_n \|T^n\| < \infty.$$

Así, vemos que f tiene variación acotada. \square

Análogamente a como se definió la noción de variación de una función, podemos definir el de p -variación para $p > 0$.

DEFINICIÓN. Sean $f : [0, \infty) \rightarrow \mathbb{R}$ y $p > 0$. Se define la p -variación de f en $[0, t]$ como

$$\sup_{\pi=\{0=t_0 < \dots < t_n=t\}} \sum_i |f(t_i) - f(t_{i-1})|^p.$$

La denotaremos por $V_t^p(f)$.

PROPOSICIÓN 3.1. Si f es una función continua de p -variación finita en $[0, t]$ entonces su p' -variación es cero si $p' > p$ e infinito si $p' < p$.

DEMOSTRACIÓN. Sean π_n una sucesión de particiones cuyo paso tiende a cero y tales que

$$\sum_{\pi_n} |f(t_i) - f(t_{i-1})|^p \rightarrow V_t^p(f) < \infty,$$

Si $p < p'$, notemos que

$$\sum_{\pi_n} |f(t_i) - f(t_{i-1})|^{p'} \leq \max_{\pi_n} |f(t_i) - f(t_{i-1})|^{p'-p} \sum_{\pi_n} |f(t_i) - f(t_{i-1})|^p.$$

Como f es continua, el máximo del lado derecho tiende a cero, mientras que la suma del lado derecho tiene un límite finito. Por lo tanto, el lado izquierdo tiende a cero. Un razonamiento análogo concluye la prueba de la afirmación restante. \square

2. La variación cuadrática del movimiento browniano

En esta sección veremos por qué la integral de Lebesgue-Stieltjes no se puede aplicar para definir una integral respecto del movimiento browniano. La razón principal está en el concepto de variación cuadrática del movimiento browniano.

PROPOSICIÓN 3.2. *Sea B un movimiento browniano en ley. Para toda $t \geq 0$, si Δ_n es una sucesión de particiones de $[0, t]$ cuya norma tiende a cero, entonces la sucesión de variables aleatorias*

$$\sum_{t_i \in \Delta_n} (B_{t_i} - B_{t_{i-1}})^2$$

converge en L_2 a t .

PRUEBA DE LA PROPOSICIÓN ??. Calculemos simplemente la norma $\mathcal{L}_2(\mathbb{P})$ al cuadrado de la diferencia entre las variables de interés:

$$\begin{aligned} \mathbb{E} \left(\left(\sum_i (B_{t_i} - B_{t_{i-1}})^2 - t \right)^2 \right) &= \mathbb{E} \left(\left(\sum_i \left((B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1}) \right) \right)^2 \right) \\ &= \mathbb{E} \left(\sum_i \left[(B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1}) \right]^2 \right), \end{aligned}$$

donde la última igualdad se justifica pues las variables

$$\left((B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1}) \right)_i$$

son independientes y tienen media cero. Si ahora utilizamos el hecho de que $\mathbb{E}(X^4) = 3\sigma^4$ si X tiene distribución $N(0, \sigma^2)$, entonces

$$\begin{aligned} \sum_i \mathbb{E} \left(\left((B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1}) \right)^2 \right) &= 2 \sum_i (t_i - t_{i-1})^2 \\ &\leq 2 |\Delta_n| t \rightarrow 0, \end{aligned}$$

lo cual demuestra la afirmación de la proposición. \square

Como veremos, la proposición anterior es un caso particular de un resultado análogo para martingalas continuas que conforma uno de los pilares sobre los que se sostiene la teoría del cálculo estocástico. El teorema al que se hace referencia es el Teorema 1.3 del capítulo 4 de [?] y la demostración que se encuentra ahí utiliza solamente resultados básicos sobre martingalas como la desigualdad maximal de Doob. En [?] se encuentra el mismo teorema (teorema 5.1 del capítulo 1) para martingalas continuas de cuadrado integrable. En particular, se sigue que $\langle B \rangle : t \mapsto t$ es el único proceso creciente y adaptado a la filtración (aumentada) de B tal que $B^2 - \langle B \rangle$ es una martingala.

PROPOSICIÓN 3.3. *Casi seguramente, las trayectorias de B son de variación infinita en $[u, v]$ para todo $u \leq v$.*

DEMOSTRACIÓN. Basta probar que para cada $u \leq v$ fijos, las trayectorias de B tienen variación infinita en $[u, v]$. Sin embargo, por la Proposición ?? aplicada al

movimiento browniano $B_{\cdot+u} - B_u$, vemos que si Δ_n es una sucesión de particiones de $[u, v]$ cuya norma tiende a cero entonces

$$\sum_{\Delta_n} |B_{t_i} - B_{t_{i-1}}| \rightarrow \infty$$

en probabilidad. Así, al pasar a una subsucesión para que la convergencia sea casi segura, vemos que la variación de B en $[u, v]$ es casi seguramente infinita. \square

Podríamos sin embargo, considerar a las sumas tipo Riemann

$$\sum_{\Delta_n} H_{t_i} (B_{t_i} - B_{t_{i-1}})$$

y considerar el límite conforme $n \rightarrow \infty$. Sin embargo, puesto que las trayectorias de B son casi seguramente de variación infinita, el Teorema ?? implica que habrá funciones H para las cuáles no hay convergencia.

3. Propiedades de la integral de Lebesgue-Stieltjes

En esta sección se nos concentraremos en algunas propiedades de la integral de Lebesgue-Stieltjes que posteriormente contrastaremos con propiedades de la integral estocástica. Comenzamos con la regla de asociatividad, la fórmula de integración por partes y la regla de la cadena.

La regla de asociatividad nos dice que, como función del intervalo de integración, una integral de Lebesgue-Stieltjes es de variación acotada y permite reexpresar la integral respecto de ella. Formalmente:

PROPOSICIÓN 3.4. *Sea f una función de variación localmente acotada, g una función medible y acotada y definamos*

$$h(t) = \int_0^t g \, df.$$

Entonces h es de variación localmente finita y si \tilde{g} es medible y acotada entonces

$$\int_0^t \tilde{g} \, dh = \int_0^t \tilde{g} g \, df.$$

Con la notación $h = g \cdot f$, podemos escribir la fórmula de manera compacta:

$$\tilde{g} \cdot (g \cdot f) = (\tilde{g}g) \cdot f.$$

Pasemos ahora a la fórmula de integración por partes.

PROPOSICIÓN 3.5. *Sean f y g dos funciones de variación localmente acotada. Para toda $t > 0$ se tiene que*

$$f(t) g(t) = f(0) g(0) + \int_0^t g_- \, df + \int_0^t f \, dg.$$

DEMOSTRACIÓN. Sean μ_f y μ_g las medidas con signo asociadas a f y a g y $\mu = \mu_f \otimes \mu_g$. Entonces

$$\mu([0, t]^2) = f(t) g(t).$$

Por otra parte, puesto que

$$[0, t]^2 = \{(0, 0)\} \cup \{(r, s) : 0 \leq r \leq s \leq t\} \cup \{(r, s) : 0 \leq s < r \leq t\},$$

podemos utilizar el teorema de Tonelli-Fubini para concluir que

$$\mu([0, t]^2) = f(0) g(0) + \int_0^t f dg + \int_0^t g_- df. \quad \square$$

La fórmula de integración por partes admite la siguiente forma más simétrica:

$$f(t) g(t) = f(0) g(0) + \int_0^t f_- dg + \int_0^t g_- df + \sum_{s \leq t} \Delta f(s) \Delta g(s).$$

Un caso, que resultará interesante contrastar con la integral estocástica, es

$$f(t)^2 = f(0)^2 + \int_0^t 2f_- df.$$

Aunque el siguiente resultado admite una versión para funciones discontinuas, nos concentraremos en el caso continuo.

PROPOSICIÓN 3.6 (Regla de la cadena). *Si f es una función continua de variación localmente acotada y F es de clase C_1 entonces $F \circ f$ es continua, de variación localmente acotada y*

$$F \circ f(t) = F \circ f(0) + \int_0^t F' \circ f df.$$

DEMOSTRACIÓN. La idea de la prueba es notar que si la fórmula es válida para F y para G y $\alpha, \beta \in \mathbb{R}$ entonces la fórmula también es válida para la función $\alpha \text{Id } F + \beta G$. Puesto que la fórmula es válida para la función identidad, también lo será para polinomios y, al aproximar a F' (y por ende a F) uniformemente en $[0, t]$ por una sucesión de polinomios, podremos concluir la fórmula para toda F continuamente diferenciable.

Basta entonces probar que si

$$F \circ f(t) = F \circ f(0) + \int_0^t F' \circ f df.$$

entonces

$$tF \circ f(t) = \int_0^t \text{Id}(F' \circ f) + F \circ f df.$$

Sin embargo, de acuerdo a la fórmula de integración por partes, vemos que si λ es la medida de Lebesgue

$$\begin{aligned} tF \circ f(t) &= \int_0^t F \circ f \, d\lambda + \int_0^t \text{Id } dF \circ f \\ &= \int_0^t F \circ f \, d\lambda + \int_0^t \text{Id } F' \circ f \, df \\ &= \int_0^t F \circ f + \text{Id } F' \circ f \, df \\ &= \int_0^t (\text{Id } F)' \circ f \, df. \end{aligned}$$

□

Finalmente, analizaremos la llamada fórmula de cambio de variable, al introducir un concepto fundamental llamado inverso continuo por la derecha y estudiar la técnica de cambios de tiempo.

Sea $f : [0, \infty) \rightarrow [0, \infty]$ no decreciente y continua por la derecha (por lo que también admite límites por la izquierda). Definimos a f^{-1} mediante

$$f^{-1}(t) = \inf \{s \geq 0 : f(s) > t\}.$$

La idea de la construcción es que la gráfica de f^{-1} se obtiene al intercambiar los ejes en la gráfica de f : los intervalos de constancia de f se convierten en saltos de f^{-1} y los saltos de f se convierten en intervalos de constancia de f^{-1} .

PROPOSICIÓN 3.7. *La función f^{-1} es no decreciente, continua por la derecha, $f \circ f^{-1} \geq \text{Id}$ y además*

$$f(t) = \inf \{s > 0 : f^{-1}(s) > t\}.$$

Si f es continua entonces $f \circ f^{-1} = \text{Id}$.

Informalmente podemos escribir la fórmula para f en términos de f^{-1} como $(f^{-1})^{-1} = f$.

DEMOSTRACIÓN. Si $t \leq \tilde{t}$ entonces

$$\{s \geq 0 : f(s) > t\} \supset \{s \geq 0 : f(s) > \tilde{t}\}.$$

Al tomar ínfimos obtenemos la monotonía de f^{-1} .

Si $t_n \downarrow t$ entonces

$$\{s \geq 0 : f(s) > t\} = \bigcup_n \{s \geq 0 : f(s) > t_n\}.$$

Al tomar ínfimos, concluimos que $f^{-1}(t) = \lim_n f^{-1}(t_n)$. En efecto, es claro que $t^{-1}(f) \leq \lim_n t_n^{-1}(f)$ y por definición de $f^{-1}(t)$, vemos que para toda $\varepsilon > 0$

$$t < f(f^{-1}(t) + \varepsilon),$$

por lo cual

$$t_n < f(f^{-1}(t) + \varepsilon)$$

para n suficientemente grande y por lo tanto

$$\lim_n f^{-1}(t_n) \leq f^{-1}(t) + \varepsilon.$$

Por definición de $f^{-1}(t)$, si $\varepsilon > 0$ se satisface

$$f(f^{-1}(t) + \varepsilon) > t$$

y por continuidad por la derecha de f obtenemos

$$f \circ f^{-1}(t) \geq t.$$

Si $s^{-1}(f) > t$ entonces $f(t) \leq s$. Por otra parte, si $\varepsilon > 0$, al ser f continua por la derecha, vemos que existe una vecindad derecha de t en la cual $f \leq f(t) + \varepsilon$. Por lo tanto $f^{-1}(f(t) + \varepsilon) > t$ y así vemos que

$$f(t) = \inf \{s \geq 0 : f^{-1}(s) > t\}.$$

Finalmente, si f es continua y no decreciente, entonces es sobre. Si (x_n) decrece estrictamente a x , existe (t_n) estrictamente decreciente (digamos con límite t) tal que $f(t_n) = x_n$. Por continuidad, se sigue que $f(t) = x$ y que a la derecha de t $f > x$. Vemos que entonces $f^{-1}(x) = t$ y que $f \circ f^{-1}(x) = f(t) = x$. \square

Ahora podemos enunciar la fórmula de cambio de variable.

PROPOSICIÓN 3.8. *Si $g : [0, \infty) \rightarrow [0, \infty)$ es Borel-medible y f es no-decreciente entonces*

$$\int_{[0, \infty)} g \, df = \int_{[0, \infty)} g \circ f^{-1} \mathbf{1}_{f^{-1} < \infty} \, d\lambda.$$

DEMOSTRACIÓN. Se sigue del procedimiento de aproximación estándar de funciones medibles al notar que la proposición es válida para $g = \mathbf{1}_{[0, t]}$; en efecto, basta notar que $\{x \geq 0 : f^{-1}(x) \leq t\}$ es un intervalo de tamaño $f(t)$. \square

La razón de llamar a una tal fórmula cambio de variable es el siguiente ejercicio.

EJERCICIO 3.1. Pruebe que si f es no decreciente y continuamente diferenciable entonces

$$\int_0^t g \, df = \int_0^t g f' \, d\lambda.$$

4. La integral estocástica respecto del movimiento browniano

El objetivo de esta sección será construir una integral respecto de las trayectorias brownianas. Esto es, definir integrales del tipo

$$\int_0^t H_s dB_s.$$

Aquí, H sería una función continua, que podría ser aleatoria o no. El problema para definirla como una integral de Lebesgue-Stieltjes, aún cuando H sea determinista, es el siguiente.

PROPOSICIÓN 3.9. *Casi seguramente, las trayectorias de B son de variación infinita en $[u, v]$ para todo $u \leq v$.*

DEMOSTRACIÓN. Basta probar que para cada $u \leq v$ fijos, las trayectorias de B tienen variación infinita en $[u, v]$. Sin embargo, por la Proposición ?? aplicada al movimiento browniano $B_{\cdot+u} - B_u$, vemos que si Δ_n es una sucesión de particiones de $[u, v]$ cuya norma tiende a cero entonces

$$\sum_{\Delta_n} |B_{t_i} - B_{t_{i-1}}| \rightarrow \infty$$

en probabilidad. Así, al pasar a una subsucesión para que la convergencia sea casi segura, vemos que la variación de B en $[u, v]$ es casi seguramente infinita. \square

Podríamos sin embargo, considerar a las sumas tipo Riemann

$$\sum_{\Delta_n} H_{t_i} (B_{t_i} - B_{t_{i-1}})$$

y considerar el límite conforme $n \rightarrow \infty$. Sin embargo, puesto que las trayectorias de B son casi seguramente de variación infinita, el Teorema ?? implica que habrá funciones H para las cuáles no hay convergencia. La idea de Itô fue considerar un modo de convergencia distinto para las sumas tipo Riemann (la convergencia en probabilidad) además de restringir a los integrandos posibles. Por ejemplo, podemos inmediatamente integrar a B respecto de B (donde $H_t = B_t$): puesto que

$$\sum_{\Delta_n} B_{t_{i-1}} (B_{t_i} - B_{t_{i-1}}) = \frac{1}{2} \sum_{\Delta_n} (B_{t_i}^2 - B_{t_{i-1}}^2) - (B_{t_i} - B_{t_{i-1}})^2,$$

la Proposición ?? implica que

$$\sum_{\Delta_n} B_{t_{i-1}} (B_{t_i} - B_{t_{i-1}}) \rightarrow \frac{B_t^2 - t}{2}$$

en probabilidad. Aquí ya nos podemos dar cuenta de que necesariamente habrá una diferencia entre la integral de Lebesgue-Stieltjes y la integral respecto de las

trayectorias brownianas, puesto que si f es continua y de variación localmente finita y $f(0) = 0$, entonces

$$\int_0^t f \, df = \frac{f(t)^2}{2}$$

de acuerdo a la fórmula de integración por partes. La diferencia es por supuesto la existencia de la variación cuadrática para las trayectorias brownianas.

Sea Δ una $2n+1$ -upla $t_0, \dots, t_n, \tau_1, \dots, \tau_n$ de reales en $[0, t]$, donde $0 = t_0 < t_1 < \dots < t_n = t$, $0 \leq \tau_1, \dots, \tau_{n-1} \leq t$ y $\tau_i \leq t_{i-1}$. (La idea es que t_1, \dots, t_n forman la partición de $[0, t]$ sobre la cual se consideran los incrementos del movimiento browniano y a las τ_i las utilizamos para evaluar al integrando.) A una tal $2n+1$ -upla le llamaremos partición con instantes de evaluación. Definimos

$$d(\Delta) = \max_{1 \leq i \leq n-1} (t_i - \tau_i).$$

Notemos que entonces

$$\tau_i \leq t_{i-1} \leq t_i \leq \tau_i + d(\Delta).$$

Sea H un proceso estocástico adaptado a $\mathcal{F}_t^B, t \geq 0$ con trayectorias continuas. Definamos entonces

$$Y_t^\Delta = \sum_{\Delta} H_{\tau_i} (B_{t \wedge t_i} - B_{t \wedge t_{i-1}}).$$

EJERCICIO 3.2. Pruebe que Y^Δ es una martingala respecto de la filtración canónica generada por el movimiento browniano. Pruebe además que si H está acotado por la constante C entonces

$$\mathbb{E}([Y_t^\Delta]^2) \leq C^2 t.$$

TEOREMA 3.2. Para toda $\varepsilon, \eta > 0$, existe $\delta > 0$ tal que si $d(\Delta), d(\Delta') < \delta$ entonces

$$\mathbb{P}(|Y_s^\Delta - Y_s^{\Delta'}| > \varepsilon) < \eta.$$

La forma en que Itô escribe el teorema anterior es que Y^Δ converge en probabilidad conforme $d(\Delta) \rightarrow 0$.

DEMOSTRACIÓN. Sea Δ una sucesión con instantes de evaluación consistente de la partición s_0, \dots, s_m y de los instantes de evaluación $\sigma_1, \dots, \sigma_m$.

Si t_0, \dots, t_n es un refinamiento de s_0, \dots, s_m definamos

$$\tau_j = \sigma_i \quad \text{si } (t_{j-1}, t_j) \subset (s_{i-1}, s_i).$$

Definamos entonces Δ' como la partición con evaluadores conformada por t_0, \dots, t_n y τ_1, \dots, τ_n , por lo que

$$Y^\Delta = Y^{\Delta'}.$$

Notemos que

$$d(\Delta') = \max t_j - \tau_{j+1} \leq \max s_i - \sigma_{i+1} = d(\Delta),$$

por lo que Δ' es una especie de refinamiento de la partición con evaluadores.

Si Δ y $\tilde{\Delta}$ son dos particiones con evaluadores distintas, podemos entonces expresar a Y^Δ y a $Y^{\tilde{\Delta}}$ en términos de la misma partición, pero con instantes de evaluación distintos y escribir por tanto

$$Y^\Delta - Y^{\tilde{\Delta}} = \sum_i (H_{\tau_i} - H_{\tilde{\tau}_i}) (B_{t_i} - B_{t_{i+1}}).$$

Como H tiene trayectorias continuas, entonces para toda $\varepsilon, \eta > 0$ existe $\delta > 0$ tal que

$$\mathbb{P}(|H_t - H_s| < \varepsilon \text{ si } |t - s| \leq \delta, s, t \leq 1) > 1 - \eta.$$

(Razón: se puede discretizar al evento anterior al considerar s, t racionales y utilizar la continuidad uniforme en el intervalo $[0, 1]$.)

Sean $\varepsilon, \eta > 0$. Si $d(\Delta), d(\tilde{\Delta}) < \delta$, donde $\delta = \delta(\varepsilon, \eta)$. Definamos

$$C_i = (H_{\tau_i} - H_{\tilde{\tau}_i}) \mathbf{1}_{|H_{\tau_i} - H_{\tilde{\tau}_i}| < \varepsilon},$$

por lo que

$$\mathbb{P}\left(Y^\Delta - Y^{\tilde{\Delta}} \neq \sum C_i (B_{t_i} - B_{t_{i-1}})\right) < \eta.$$

Por otra parte

$$\mathbb{E}\left(\left[\sum C_i (B_{t_i} - B_{t_{i-1}})\right]^2\right) = \sum \mathbb{E}(C_i^2) (t_i - t_{i-1}) \leq \varepsilon^2 T.$$

Por lo tanto

$$\mathbb{P}\left(\left|\sum C_i (B_{t_i} - B_{t_{i-1}})\right| > \sqrt{\varepsilon}\right) \leq \varepsilon T$$

y por lo tanto

$$\mathbb{P}\left(\left|Y^\Delta - Y^{\tilde{\Delta}}\right| > \sqrt{\varepsilon}\right) \leq \varepsilon + \eta.$$

Es entonces sencillo ver (al considerar una sucesión de particiones evaluadoras Δ_n con $d(\Delta_n) \rightarrow 0$ y tales que Y^{Δ_n} converja en probabilidad) que existe una variable aleatoria Y tal que para toda $\varepsilon, \eta > 0$ existe $\delta > 0$ tal que si $d(\Delta) < \delta$ entonces

$$\mathbb{P}(|Y^\Delta - Y| > \varepsilon) < \eta.$$

□

DEFINICIÓN. La **integral estocástica** de H respecto de B en el intervalo $[0, 1]$ es el límite en probabilidad de las sumas de Riemann

$$\sum_{\Delta} H_{\tau_i} (B_{t_i} - B_{t_{i-1}})$$

(donde Δ es una partición evaluadora de $[0, 1]$) conforme $d(\Delta) \rightarrow 0$. Se denotará por $\int_0^1 H_s dB_s$.

Claramente, los argumentos funcionan de igual manera en el intervalo $[u, v]$ en vez de $[0, 1]$. Además, de acuerdo a la definición, tenemos que

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

Itô introdujo su célebre fórmula para calcular otras integrales estocásticas.

El argumento anterior de hecho puede mejorarse para poder considerar una integral indefinida y probar propiedades de continuidad respecto del intervalo de integración. Antes de eso, enunciemos algunas propiedades de la integral estocástica.

PROPOSICIÓN 3.10. *Sean F y G dos procesos continuos y adaptados respecto de la filtración browniana. Entonces:*

- (1) $\int_s^t dB_r = B_t - B_s$,
- (2) si $\lambda, \mu \in \mathbb{R}$ entonces $\int_s^t [\lambda F_s + \mu G_s] dB_s = \lambda \int_s^t F_s dB_s + \mu \int_s^t G_s dB_s$ y
- (3) si $r < s < t$ entonces

$$\int_r^s F_u dB_u + \int_s^t F_u dB_u = \int_r^t F_u dB_u.$$

EJERCICIO 3.3. Pruebe la proposición anterior.

EJERCICIO 3.4. Mediante aproximación por sumas tipo Riemann, pruebe las siguientes dos igualdades:

- (1) $\int_0^t s dB_s = t B_t - \int_0^t B_s ds$ y
- (2) $\int_0^t B_s^2 dB_s = B_t^3/3 - \int_0^t B_s ds$.

PROPOSICIÓN 3.11. *Sea H un proceso adaptado con trayectorias continuas. Si existe una función continua tal que $\mathbb{E}(H_t^2) \leq M_t$ entonces*

$$\mathbb{E}\left(\left[\int_s^t H_r dB_r\right]^2\right) \leq \int_s^t M_r dr.$$

Sean H^1, H^2, \dots procesos estocásticos adaptados con trayectorias continuas y suponga que H^n converge a H uniformemente en probabilidad, es decir,

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\sup_{s \leq t} |H_s^n - H_s| > \varepsilon\right) = 0$$

para toda $\varepsilon > 0$. Entonces

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\left|\int_0^t H_s^n dB_s - \int_0^t H_s dB_s\right| > \varepsilon\right) = 0.$$

DEMOSTRACIÓN. Se probará el resultado cuando $s = 0$. Sea (Δ_n) una sucesión cuyo paso tiende a cero y tal que la sucesión de integrales estocásticas elementales

$$I_n = \sum_{\Delta_n} H_{t_{i-1}} (B_{t_i} - B_{t_{i-1}})$$

converge casi seguramente a la integral estocástica $\int_0^t H_r dB_r$. Entonces, por el lema de Fatou:

$$\mathbb{E} \left(\left[\int_0^t H_r dB_r \right]^2 \right) \leq \liminf_{n \rightarrow \infty} \mathbb{E} ([I_n]^2) \leq \liminf_{n \rightarrow \infty} \sum_n M_{t_{i-1}} (t_i - t_{i-1}) = \int_0^t M_r dr$$

puesto que M es Riemann integrable al ser continua.

Sea

$$C_s^n = \begin{cases} -\varepsilon & H_s - H_s^n \leq -\varepsilon \\ H_s - H_s^n & -\varepsilon < H_s - H_s^n < \varepsilon \\ \varepsilon & H_s - H_s^n \geq \varepsilon \end{cases}.$$

Al utilizar particiones aproximantes, es fácil ver que

$$\mathbb{P} \left(\int_0^t C_s^n dB_s \neq \int_0^t [H_s - H_s^n] dB_s \right) \leq \mathbb{P} \left(\sup_{s \leq t} |H_s - H_s^n| > \varepsilon \right).$$

Puesto que $C_s^n \leq \varepsilon$, se sigue que

$$\mathbb{P} \left(\left| \int_0^t C_s^n dB_s \right| \geq \sqrt{\varepsilon} \right) \leq \frac{1}{\varepsilon} \mathbb{E} \left(\left[\int_0^t C_s^n dB_s \right]^2 \right) \leq \varepsilon t.$$

Por lo tanto

$$\mathbb{P} \left(\left| \int_0^t H_s B_s - \int_0^t H_s^n dB_s \right| > \varepsilon + \sqrt{\varepsilon} \right) \leq \mathbb{P} \left(\sup_{s \leq t} |H_s - H_s^n| > \varepsilon \right) + \varepsilon t$$

□

Para cada partición evaluadora definamos

$$Y_t^\Delta = \sum H_{\tau_i} (B_{t \wedge t_i} - B_{t \wedge t_{i-1}}).$$

Esta sería una versión discreta de la integral indefinida. Su principal característica es ser un proceso con trayectorias continuas y una martingala.

PROPOSICIÓN 3.12. *Para toda $\varepsilon, \eta > 0$, existe $\delta > 0$ tal que si $d(\Delta), d(\Delta') < \delta$ entonces*

$$\mathbb{P} \left(\sup_{s \leq t} |Y_s^\Delta - Y_s^{\Delta'}| > \varepsilon \right) < \eta.$$

DEMOSTRACIÓN. Sean s_1, s_2, \dots densos en $[0, 1]$ y sean t_1^m, \dots, t_n^m los puntos obtenidos al refinar a Δ y a $\tilde{\Delta}$ con s_1, \dots, s_m . Entonces, podremos escribir

$$Y_t^\Delta - Y_t^{\tilde{\Delta}} = \sum (H_{\tau_i} - H_{\tilde{\tau}_i}) (B_{t_i^m} - B_{t_{i-1}^m}).$$

Podemos seguir la prueba del Teorema ?? para deducir de la desigualdad de Doob (aplicada a la transformada de martingala de C por B) que

$$\mathbb{P} \left(\max_{1 \leq i \leq n} \left| \sum_{i=1}^k C_i (B_{t_i^m} - B_{t_{i-1}^m}) \right| > \sqrt{\varepsilon} \right) \leq \varepsilon,$$

lo cual implicará que

$$\mathbb{P}\left(\sup_i \left|Y_{t_i^m}^\Delta - Y_{t_i^m}^{\tilde{\Delta}}\right| > \sqrt{\varepsilon}\right) \leq \varepsilon + \eta.$$

Al tomar el límite conforme $m \rightarrow \infty$ y utilizar la densidad de $\{s_1, s_2, \dots\}$, obtenemos

$$\mathbb{P}\left(\sup_{s \leq 1} \left|Y_s^\Delta - Y_s^{\tilde{\Delta}}\right| > \sqrt{\varepsilon}\right) \leq \varepsilon + \eta,$$

lo cual nos permite concluir de manera análoga que en el Teorema ??.

□

COROLARIO 2. *Existe un proceso estocástico con trayectorias continuas $H \cdot B$ tal que para todo $t \geq 0$*

$$\mathbb{P}\left((H \cdot B)_t = \int_0^t H_s dB_s\right) = 1.$$

Una propiedad importante de la integral estocástica respecto de su parámetro, que ya se ha utilizado implícitamente es su carácter de martingala.

PROPOSICIÓN 3.13. *Si H es adaptado, continuo y existe una función continua M tal que $\mathbb{E}(H_t^2) \leq M_t$ entonces $H \cdot B$ es una martingala cuadrado integrable con trayectorias continuas.*

DEMOSTRACIÓN. Con nuestra hipótesis:

$$\mathbb{E}\left([H \cdot B_t]^2\right) \leq \int_0^t M_s ds < \infty.$$

Para verificar que $H \cdot B$ es martingala, basta probar que

$$\mathbb{E}\left(\int_s^t H_r dB_r \mid \mathcal{F}_s\right) = 0.$$

Si Δ_n es una sucesión de particiones de $[s, t]$ tal que la sucesión

$$I_n = \sum_{\Delta_n} H_{t_{i-1}} (B_{t_i} - B_{t_{i-1}})$$

converge casi seguramente a $\int_s^t H_r dB_r$, notemos que I_n es medible respecto de $\sigma(B_{r+s} - B_s : r \geq 0)$, que es independiente de \mathcal{F}_s . Puesto que las variables I_n son acotadas en L_2 , son uniformemente integrables y por lo tanto

$$0 = \mathbb{E}(I_n \mid \mathcal{F}_s) \rightarrow \mathbb{E}\left(\int_s^t H_r dB_r \mid \mathcal{F}_s\right).$$

□

Pasemos ahora a la fórmula de Itô, que nos da una clase muy grande y útil de ejemplos de integrales estocásticas. Se trata de la versión estocástica del teorema de cambio de variable.

TEOREMA 3.3. *Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ una función de clase C_2 . Entonces*

$$f(B_t) = f(B_0) + \int_0^t f'(B_s) dB_s + \frac{1}{2} \int_0^t f''(B_s) ds.$$

La heurística de la prueba es sencilla: para cualquier partición $0 = t_0 < \dots < t_n = t$ suficientemente fina, al realizar una expansión de Taylor de orden 2 se obtiene

$$f(B_{t_i}) - f(B_{t_{i-1}}) \approx f'(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) + \frac{1}{2} f''(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}})^2.$$

Al sumar sobre i se obtiene

$$f(B_t) - f(B_{t_0}) = \sum_i f'(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) + \sum_i \frac{1}{2} f''(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}})^2.$$

El primer sumando del lado derecho converge a la integral estocástica

$$\int_0^t f'(B_s) dB_s$$

mientras que el segundo debería converger hacia

$$\int_0^t \frac{1}{2} f''(B_s) ds.$$

DEMOSTRACIÓN. Trabajaremos en $[0, 1]$.

Puesto que B tiene trayectorias continuas, el máximo de su valor absoluto en $[0, 1]$ es finito casi seguramente y por eso, dada $\eta > 0$ existe $M > 0$ tal que

$$\mathbb{P}\left(\sup_{s \leq 1} |B_s| \geq M\right) < \eta.$$

Puesto que f'' es continua en $[-M, M]$, dada $\varepsilon > 0$ existe $\delta > 0$ tal que $|f''(y) - f''(x)| < \varepsilon$ si $x, y \in [-M, M]$ y $|y - x| < \delta$.

Puesto que las trayectorias de B son uniformemente continuas, existe $\gamma > 0$ tal que

$$\mathbb{P}\left(\sup_{\substack{|t-s| < \gamma \\ 0 \leq s, t \leq 1}} |B_t - B_s| \geq \delta\right) < \eta.$$

Por lo tanto, si

$$\Omega' = \left\{ \sup_{s \leq 1} |B_s| < M, \sup_{\substack{|t-s| < \gamma \\ 0 \leq s, t \leq 1}} |B_t - B_s| < \delta \right\}$$

entonces $\mathbb{P}(\Omega') > 1 - 2\eta$.

Al realizar una expansión de Taylor de orden 2 se obtiene

$$f(B_t) - f(B_s) = f'(B_s)(B_t - B_s) + \frac{1}{2}f''(B_s)(B_t - B_s)^2 + \frac{1}{2}R_{s,t}(B_t - B_s)^2$$

para $0 \leq s \leq t \leq 1$, donde

$$R_{s,t} = f''(B_s + \theta(B_t - B_s)) - f''(B_s)$$

para alguna $\theta \in [0, 1]$ (que es aleatoria).

Definamos ahora a los truncamientos

$$C_{s,t} = \frac{1}{2}R_{s,t}(B_t - B_s)^2 \mathbf{1}_{|R_{s,t}| \leq \varepsilon}$$

y

$$E_t = f''(B_t) \mathbf{1}_{|f''(B_t)| \leq R}$$

donde $R = \max_{|x| \leq M} |f(x)|$. Notemos que en Ω' , si $s < t < s + \gamma$ entonces

$$\begin{aligned} f(B_t) - f(B_s) &= f'(B_s)(B_t - B_s) + \frac{1}{2}f''(B_s)(t - s) \\ &\quad + \frac{1}{2}R_{s,t}(B_t - B_s)^2 + \frac{1}{2}E_s[(B_t - B_s)^2 - (t - s)]. \end{aligned}$$

Si $0 = t_0 < \dots < t_n = 1$ es cualquier partición de paso menor a γ , podemos por lo tanto escribir en Ω'

$$\begin{aligned} f(B_1) - f(0) &= \sum_i f'(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) + \frac{1}{2}f''(B_{t_{i-1}})(t_i - t_{i-1}) \\ &\quad + \sum_i \frac{1}{2}C_{t_i, t_{i-1}}(B_{t_i} - B_{t_{i-1}})^2 \\ &\quad + \frac{1}{2}E_{t_{i-1}}[(B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})]. \end{aligned}$$

Escojamos ahora la norma de la partición de tal manera que

$$\mathbb{P}\left(\left|\sum_i f'(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) - \int_0^1 f'(B_s) dB_s\right| > \varepsilon\right) < \eta$$

y

$$\mathbb{P}\left(\left|\sum_i \frac{1}{2}f''(B_{t_{i-1}})(t_i - t_{i-1}) - \int_0^1 f''(B_s) ds\right| > \varepsilon\right) < \eta.$$

Por otra parte, por definición de C_t vemos que

$$\mathbb{E}\left(\left|\sum_i \frac{1}{2}C_{t_{i-1}, t_i}(B_{t_i} - B_{t_{i-1}})^2\right|\right) \leq \varepsilon.$$

Si además imponemos que $|t_i - t_{i-1}| < \varepsilon/R^2$ entonces

$$\begin{aligned} \mathbb{E}\left(\left|\sum_i \frac{1}{2}E_{t_{i-1}}\left[\left(B_{t_i} - B_{t_{i-1}}\right)^2 - (t_i - t_{i-1})\right]\right|^2\right) \\ \leq \frac{R^2}{4} \sum_i 2(t_i - t_{i-1})^2 \leq \frac{\varepsilon}{2} \leq \varepsilon. \end{aligned}$$

Así, podemos concluir que

$$\mathbb{P}\left(\left|\sum_i \frac{1}{2}C_{t_{i-1}, t_i} (B_{t_i} - B_{t_{i-1}})^2\right| > \varepsilon^{1/2}\right) < \sqrt{\varepsilon}$$

y

$$\mathbb{P}\left(\left|\sum_i \frac{1}{2}E_{t_{i-1}}\left[\left(B_{t_i} - B_{t_{i-1}}\right)^2 - (t_i - t_{i-1})\right]\right| > \sqrt{\varepsilon}\right) \leq \varepsilon.$$

Por lo tanto

$$\mathbb{P}\left(\left|f(B_t) - f(B_0) - \int_0^t f'(B_s) dB_s - \frac{1}{2} \int_0^t f''(B_s) ds\right| > 3\varepsilon + \sqrt{\varepsilon}\right) \leq 4\varepsilon. \quad \square$$

EJERCICIO 3.5 (Tomado del libro de Oksendal). Utilice la fórmula de Itô para escribir a los siguientes procesos Y en la forma estándar

$$dY_t = u(t, \omega) dt + v(t, \omega) dB_t.$$

- (1) $Y_t = B_t^2$
- (2) $Y_t = 2 + t + e^{B_t}$.

Utilice la fórmula de Itô para verificar que los siguientes procesos son martingalas

- (1) $X_t = e^{t/2} \cos(B_t)$
- (2) $X_t = e^{t/2} \sin(B_t)$
- (3) $X_t = (B_t + t) e^{-B_t - t/2}$.

5. Ecuaciones diferenciales estocásticas conducidas por el movimiento browniano

Equipados con una noción de integral estocástica y conscientes de su diferencia fundamental con la integral de Lebesgue-Stieltjes, podemos analizar el concepto de ecuación diferencial estocástica. Cabe mencionar que ésta era la motivación original del [?] para introducir la integral estocástica pues estaba interesado en verificar su intuición de que las soluciones a ecuaciones diferenciales estocásticas deberían ser procesos de Markov.

En este capítulo, nos interesaremos principalmente en ecuaciones del tipo

$$dX_t = \sigma(t, X_t) dB_t + b(t, X_t) dt.$$

La interpretación es que buscamos un proceso X con trayectorias continuas y adaptado a la filtración de B tal que para toda $t \geq 0$ se tenga que

$$(2) \quad X_t = x + \int_0^t \sigma(s, X_s) dB_s + \int_0^t b(s, X_s) ds.$$

Cuando σ es idénticamente igual a cero, nos reducimos al caso de las ecuaciones diferenciales ordinarias. Así puesto, la primera tarea que tendremos será la de estudiar existencia y unicidad para tales ecuaciones.

Intuitivamente, al discretizar una ecuación diferencial estocástica al utilizar una partición $0 = t_0 < t_1 < \dots$ obtenemos una relación de recurrencia del tipo:

$$X_{t_{i+1}} = X_{t_i} + \sigma(t_i, X_{t_i}) [B_{t_{i+1}} - B_{t_i}] + b(t_i) [t_{i+1} - t_i].$$

(Formalmente, estaríamos aplicando el método de Euler a la ecuación (??).) Un tal proceso se puede pensar como la adición del ruido $\sigma(t_i, X_{t_i}) [B_{t_{i+1}} - B_{t_i}]$ a la evolución determinista

$$X_{t_{i+1}} = X_{t_i} + b(t_i) [t_{i+1} - t_i].$$

Por esta razón, un tal proceso tiene la propiedad de Markov a tiempo discreto. Por supuesto, hay dos preguntas naturales. La primera es si la propiedad de Markov también se vale para las soluciones a (??) y la segunda es si, cuando la norma de la partición tiende a cero, la solución de la ecuación de recurrencia converge en algún sentido a la solución de (??). Esto es, la pregunta sería sobre la convergencia del método de Euler (en algún sentido).

Pensemos en la siguiente situación como motivación: recordemos que hemos interpretado a una martingala como el proceso de ganancias que tenemos al someter cierto capital en un juego de apuestas justo. Por ejemplo, pensemos que $B_{t+s} - B_s$ es la ganancia (o pérdida en caso de que su signo sea negativo) en el intervalo $[s, s+t]$, aunque pensamos que el juego se desarrolla a tiempo continuo. Si el gobierno nos cobra (continuamente) impuestos sobre nuestras ganancias a tasa λ_+ y, en caso de tener capital negativo pedimos dinero prestado que nos genera un interés infinitesimal de tasa λ_+ , estaríamos tentados a escribir nuestro capital al tiempo t como la solución a la ecuación diferencial estocástica

$$X_t = x + B_t - \int_0^t \lambda_- \mathbf{1}_{X_s < 0} X_s ds - \int_0^t \lambda_+ X_s ds.$$

Cuando $\lambda_+ = \lambda_- = -\lambda$ obtenemos al célebre proceso de Ornstein-Uhlenbeck, que es la solución de

$$X_t = x + B_t + \lambda \int_0^t X_s ds.$$

Otro ejemplo concreto de ecuación diferencial estocástica es el de la exponencial estocástica del movimiento browniano. Se trata de la solución a la ecuación

diferencial estocástica (EDE)

$$(3) \quad X_t = x + \int_0^t X_s dB_s.$$

EJERCICIO 3.6. Probar que si f es una función continua de variación acotada entonces g satisface

$$g(t) = x + \int_0^t g(s) f(ds)$$

si y sólo si

$$g(t) = xe^{f(t)}.$$

Sugerencia: utilice integración por partes para ver que ge^{-f} es constante.

El caso del browniano es un tanto distinto.

EJERCICIO 3.7. Pruebe que $X_t = xe^{B_t - t/2}$ satisface la ecuación (??). *Sugerencia:* Aplique la fórmula de Itô para procesos del tipo $f(t, B_t)$. Note que para obtener una conclusión de unicidad como la del caso de variación finita, hace falta una fórmula de integración por partes para integrales estocásticas.

Una manera de resolver ecuaciones diferenciales ordinarias es mediante el método de Picard. Esta misma idea funciona para ecuaciones diferenciales estocásticas, como hizo notar Itô. Para simplificar el enunciado del teorema de existencia y unicidad, haremos un supuesto de Lipschitz global en los coeficientes σ y b de la ecuación (??).

TEOREMA 3.4. *Suponga que σ y b son continuas y que existen constantes K y $t \geq 0$ tal que para toda $x, y \in \mathbb{R}$ se tiene que*

$$|\sigma(t, y) - \sigma(t, x)| \leq K |y - x| \quad y \quad |b(t, y) - b(t, x)| \leq K |y - x|.$$

Entonces

- (1) *Dada $x \in \mathbb{R}$ existe un proceso X continuo y adaptado a la filtración de B tal que*

$$X_t = x + \int_0^t \sigma(s, X_s) dB_s + \int_0^t b(s, X_s) ds.$$

- (2) *Si \tilde{X} es continuo, adaptado y*

$$\tilde{X}_t = x + \int_0^t \sigma(s, \tilde{X}_s) dB_s + \int_0^t b(s, \tilde{X}_s) ds.$$

entonces X y \tilde{X} son indistinguibles.

La prueba utilizará fuertemente el siguiente resultado conocido como la desigualdad (o lema) de Gronwall.

LEMÁ 5. Sea $T > 0$ y $g : [0, T] \rightarrow [0, \infty)$ medible y acotada. Si existen constantes $A, B \geq 0$ tales que

$$g(t) \leq A + B \int_0^t g(s) \, ds \text{ para toda } t \in [0, T]$$

entonces

$$g(t) \leq Ae^{Bt} \text{ para toda } t \in [0, T].$$

Note que en particular, si $A = 0$ entonces $g = 0$.

DEMOSTRACIÓN. Al iterar la desigualdad, vemos que

$$g(t) \leq A + B \int_0^t \int_0^{t_1} A + Bg(t_2) \, dt_2 \, dt_1 = A + ABt + B \int_0^t g(t_2)(t - t_2) \, dt_2$$

y al continuar se obtiene recursivamente

$$g(t) \leq A + ABt + A \frac{B^2 t^2}{2} + \cdots + A \frac{B^n t^n}{n!} + B^{n+1} \int_0^t g(t_{n+1}) \frac{(t - t_{n+1})^n}{n!} \, dt_{n+1}.$$

Puesto que g es acotada, vemos que la integral del lado derecho de la desigualdad anterior converge a cero conforme $n \rightarrow \infty$ y por lo tanto

$$g(t) \leq \sum_{n=0}^{\infty} A \frac{B^n t^n}{n!} = Ae^{Bt}. \quad \square$$

Es ilustrativo hacer el argumento de existencia y unicidad en el caso determinístico.

PRUEBA DEL TEOREMA ?? SI $\sigma = 0$. Como veremos en la prueba del caso general, podemos suponer que existe una constante K' tal que $|b(t, x)| \leq K' + K|x|$ (aunque sea en conjuntos acotados del parámetro temporal).

Para la existencia, sea $X_t^0 = x$ y para $n \geq 0$ definimos recursivamente

$$X_t^{n+1} = x + \int_0^t b(s, X_s^n) \, ds.$$

Se sigue entonces que (X^n) es una sucesión de funciones continuas. La prueba terminará si mostramos que convergen uniformemente en compactos a una función (necesariamente) continua X . En efecto, podemos entonces tomar el límite en la definición de X^{n+1} y concluir por el teorema de convergencia acotada que

$$X_t = x + \int_0^t b(s, X_s^n) \, ds.$$

Para mostrar que la sucesión (X^n) converge uniformemente en compactos, utilizamos la hipótesis de Lipschitz para concluir que

$$\sup_{s \leq t} |X_s^{n+1} - X_s^n| \leq \int_0^t K |X_s^n - X_s^{n-1}| \, ds$$

y puesto que

$$|X_t^1 - X_t^0| \leq \int_0^t K' + K |x| \, ds = K'' t,$$

se sigue que

$$\sup_{s \leq t} |X_s^{n+1} - X_s^n| \leq K'' \frac{t^n}{n!}.$$

Podemos concluir entonces que

$$\sum_n \sup_{s \leq t} |X_s^{n+1} - X_s^n| < \infty$$

lo cual implica que X^n converge uniformemente en $[0, t]$.

Para la unicidad, suponemos que X y \tilde{X} son dos soluciones a la ecuación diferencial ?? con $\sigma = 0$. Al utilizar la hipótesis de Lipschitz se obtiene

$$|X_t - \tilde{X}_t| \leq \int_0^t K |X_s - \tilde{X}_s| \, ds,$$

lo cual implica, por el lema de Gronwall, que $|X_t - \tilde{X}_t| = 0$ para toda t . □

PRUEBA DEL TEOREMA ??. Trabajaremos en el intervalo fijo $[0, 1]$. Notemos que existe una constante K' tal que $|\sigma(s, y)| \leq K |y| + K'$ para $s \leq 1$ y $y \in \mathbb{R}$ (y análogamente para b). En efecto, puesto que σ es continua entonces $K' = \sup_{s \leq 1} |\sigma(s, 0)| < \infty$. La hipótesis Lipschitz global implica que

$$|\sigma(s, y)| \leq |\sigma(s, 0)| + K |y|.$$

Probemos primero la unicidad. Se asumirá primero que σ y b son acotadas, digamos por M . Sean X y \tilde{X} dos procesos continuos y adaptados que satisfacen la ecuación (??). Sea

$$\tilde{g}(t) = \mathbb{E} \left(\sup_{s \leq t} |X_s - \tilde{X}_s| \right).$$

Entonces, al utilizar la cota para σ y B , así como la desigualdad de Doob y la desigualdad $(a + b)^2 \leq 2(a^2 + b^2)$ obtenemos

$$\begin{aligned} \tilde{g}(t) &\leq 2\mathbb{E} \left(\sup_{s \leq t} \left[\int_0^s \sigma(r, X_r) - \sigma(r, \tilde{X}_r) \, dB_r \right]^2 + \sup_{s \leq t} \left[\int_0^t b(r, X_r) - b(r, \tilde{X}_r) \right]^2 \right) \\ &\leq 32M^2 t + 8M^2 t^2 \leq 40M^2 t < \infty. \end{aligned}$$

Definamos ahora a

$$g(t) = \mathbb{E} \left([X_t - \tilde{X}_t]^2 \right).$$

Puesto que $X - \tilde{X}$ tiene trayectorias continuas y su supremo en $[0, t]$ tiene momento de orden dos finito, vemos que g es una función acotada y continua.

Por otra parte, al utilizar la hipótesis de Lipschitz global y la cota para el segundo momento de una integral estocástica implica que

$$g(t) \leq 2K^2(1+t) \int_0^t g(s) ds;$$

si en el intervalo $[0, 1]$ acotamos a $t \mapsto (1+t)$ por 2, vemos que se puede aplicar el lema de Gronwall y concluir que $g(t) = 0$. Así, hemos probado que para toda t , $X_t = \tilde{X}_t$ casi seguramente (esto es, que X es modificación de \tilde{X}). Sin embargo, como ambos procesos tienen trayectorias continuas, entonces son indistinguibles.

Cuando σ y b son sólo continuas y no acotadas y suponemos que hay dos procesos X y \tilde{X} continuos y adaptados que satisfacen (??) entonces definimos

$$\Omega_K = \left\{ \sup_{t \leq T} |X_t| \leq K, \sup_{t \leq T} |\tilde{X}_t| \leq K \right\}.$$

Se tiene que

$$\lim_{K \rightarrow \infty} \mathbb{P}(\Omega_K) \rightarrow 1.$$

Si M es una cota para b y σ en $[0, T] \times [-M, M]$, definamos

$$b_M(t, y) = \begin{cases} -M & b(t, y) \leq -M \\ b(t, y) & -M \leq b(t, y) \leq M \\ M & b(t, y) \geq M \end{cases}$$

y analogamente se define a σ_M . Puesto que en Ω_K , X y \tilde{X} satisfacen la ecuación diferencial estocástica con coeficientes acotados B_M y σ_M , entonces $X_t = \tilde{X}_t$ casi seguramente en Ω_K . Así:

$$\mathbb{P}(X_t \neq \tilde{X}_t) \leq 1 - \mathbb{P}(\Omega_K) \rightarrow_{K \rightarrow \infty} 0.$$

Para la existencia, definamos $X_t^0 = x$ para toda $t \geq 0$ y recursivamente al proceso adaptado y continuo

$$X_t^{n+1} = x + \int_0^t \sigma(s, X_s^n) dB_s + \int_0^t b(s, X_s^n) ds.$$

Primero probaremos que la sucesión $X_t^k, k \geq 0$ converge en L_2 uniformemente en t en el intervalo $[0, 1]$. Sea M una cota para σ y b en $[0, 1] \times \{x\}$. Entonces la desigualdad de Jensen implica

$$\mathbb{E} \left(\sup_{s \leq t} \left[\int_0^s b(r, X_r^0) dr \right]^2 \right) \leq M^2 t$$

y por desigualdad L_2 de Doob aplicada a la integral estocástica (que es una martingala) y la cota para el segundo momento de la integral estocástica

$$\mathbb{E}\left(\sup_{s \leq t} \left[\int_0^s \sigma(r, X_r^0) dB_r\right]^2\right) \leq 4M^2t.$$

La desigualdad $(a+b)^2 \leq 2(a^2 + b^2)$ implica entonces que

$$\mathbb{E}\left(\sup_{s \leq t} [X_s^1 - X_s^0]^2\right) \leq 10M^2t.$$

Un argumento análogo, que además utiliza la hipótesis de Lipschitz global, muestra que

$$\mathbb{E}\left(\sup_{s \leq t} [X_s^{n+1} - X_s^n]^2\right) \leq 10K^2 \int_0^t \mathbb{E}\left(\sup_{r \leq s} [X_r^n - X_r^{n-1}]^2\right) ds$$

por lo que inductivamente se verifica la desigualdad

$$\mathbb{E}\left(\sup_{s \leq t} [X_s^{n+1} - X_s^n]^2\right) \leq (10K^2)^n 10M^2 \frac{t^{n+1}}{(n+1)!}.$$

Puesto que

$$\sum_n \mathbb{E}\left(\sup_{s \leq 1} [X_s^{n+1} - X_s^n]^2\right)^{1/2} < \infty,$$

podemos concluir que X_t^n converge en L_2 a

$$X_t = x + \sum_{n=1}^{\infty} (X_t^n - X_t^{n-1})$$

y que además

$$\mathbb{P}\left(\sup_{t \leq 1} |X_t^n - X_t| > \varepsilon\right) \leq \frac{1}{\varepsilon^2} \left[\sum_{k=n+1}^{\infty} \mathbb{E}\left(\left[\sup_{t \leq 1} |X_t^n - X_t| \right]^2\right) \right] \xrightarrow{n \rightarrow \infty} 0.$$

Por lo tanto, existe una subsucesión n_k tal que casi seguramente

$$\sup_{s \leq 1} |X_t^{n_k} - X_t| \xrightarrow{n \rightarrow \infty} 0$$

por lo que X tiene trayectorias continuas y es adaptado. Por la convergencia uniforme, vemos que

$$\lim_{k \rightarrow \infty} \sup_{t \leq 1} \left| \int_0^t b(s, X_s) ds - \int_0^t b(s, X_s^{n_k}) ds \right| = 0$$

casi seguramente. Finalmente, puesto que $\sigma(t, X_t^{n_k})$ converge uniformemente en probabilidad hacia $\sigma(t, X_t)$, se puede aplicar el teorema de convergencia de integrales estocásticas para concluir que

$$\lim_{n \rightarrow \infty} \int_0^t \sigma(s, X_s^{n_k}) dB_s = \int_0^t \sigma(s, X_s) dB_s$$

en probabilidad (y al pasar a una nueva subsucesión podemos suponer que la convergencia es casi segura) y que por lo tanto X satisface la ecuación (??). \square

En particular, el caso en que $\sigma = 0$ nos reduce a la ecuación diferencial ordinaria

$$X'_t = b(t, X_t).$$

El teorema de existencia y unicidad aplica cuando bajo una condición de Lipschitz global sobre b . Hay un célebre teorema, debido a Peano que afirma que hay existencia local para la ecuación anterior con la sola hipótesis de que b sea continua. Respecto a la unicidad, y para darse una idea de lo que permite probar el esquema de iteración de Picard y lo que es verdad, veamos que si b no depende de la variable temporal y es continua, entonces hay unicidad con la sola hipótesis $b > 0$. En efecto, notemos que si f satisface $f' = b \circ f$ con b positiva, entonces f es (continua y) estrictamente creciente. Si $i = f^{-1}$, se puede calcular la derivada de i y concluir que

$$i'(x) = \frac{1}{f' \circ i(x)} = \frac{1}{b \circ f \circ i(x)} = \frac{1}{b(x)}.$$

Concluimos que si f y \tilde{f} ambas satisfacen la ecuación que nos interesa, sus inversas tienen la misma derivada, por lo que son iguales y por lo tanto $f = \tilde{f}$. De hecho, probemos unicidad en un contexto más general cuando $\sigma = 0$: si b es continua, estrictamente positiva y $t \mapsto b(t, x)$ es no-decreciente. Sean x^1 y x^2 dos funciones diferenciables con derivadas y^1 y y^2 que satisfagan $y^i = b(t, x_t^i)$. Sea $x_t^3 = x_{\alpha t}^2$ con $\alpha > 1$ y notemos que su derivada, denotada y^3 está dada por $y_t^3 = ab(\alpha t, x_t^3)$. Sea

$$\tau = \inf \{t \geq 0 : x_t^1 > x_t^3\}.$$

Si τ fuera finito entonces, puesto que $x_\tau^3 = x_\tau^1$ por continuidad y definición de τ , vemos que

$$y_\tau^3 = ab(\alpha \tau, x_\tau^3) > b(\tau, x_\tau^3) = b(\tau, x_\tau^1) = y_\tau^1.$$

Lo anterior implica que $x^3 > x^1$ en una vecindad derecha de τ contradiciendo la definición de τ . Vemos por lo tanto que $x^1 \leq x^3$ y, al considerar $\alpha \rightarrow 1$, vemos que $x^1 \leq x^2$. Al intercambiar los roles de x^1 y x^2 nos damos cuenta de que $x^1 = x^2$.

Continuaremos con la razón fundamental por la cual Itô introdujo a la integral estocástica y a las ecuaciones diferenciales estocásticas asociadas: la construcción de procesos de Markov. La idea es que cuando los coeficientes σ y b que conducen a una ecuación diferencial estocástica no dependen de la variable temporal entonces la solución es un proceso de Markov.

Comencemos con la definición de un proceso de Markov y, de hecho, mejor motivémosla en el caso del movimiento browniano. La idea es ver al movimiento browniano como un proceso de Markov y para esto, quisieramos definir al browniano que comienza en cualquier $x \in \mathbb{R}$ y definir un análogo de la matriz de transición. El browniano que comienza en x se define como el proceso $x + B$ y el sentido de esta definición es que como B_s y $B^s = (B_{t+s} - B_s, t \geq 0)$ son independientes y B^s es un browniano, si queremos la distribución condicional de $B_{t+s}, t \geq 0$ dado que

$B_s = x$, estará dada por la distribución de $x + B^s$ que es la de $x + B$. Por otra parte, recordemos que si X es cadena de Markov con matriz de transición P entonces $\mathbb{P}(X_{n+m} = j \mid X_m = i) = P_{i,j}^n$ y que esto caracteriza a la cadena. El problema es que para el movimiento browniano, aunque definamos a $\mathbb{P}(B_{t+s} = y \mid B_s = x)$ como $\mathbb{P}(B_t + x = y)$, esta probabilidad será cero. Por esta razón se define al núcleo de transición P_t de tal manera que para todo $x \in \mathbb{R}$, $P_t(x, \cdot)$ es la medida dada por

$$P_t(x, A) = \mathbb{P}(x + B_t \in A).$$

En general, se define a un **núcleo de medidas de probabilidad en \mathbb{R}** como una función $N : \mathbb{R} \times \mathcal{B}_{\mathbb{R}} \rightarrow [0, \infty)$ tal que

- para toda $x \in \mathbb{R}$ $N(x, \cdot)$ es una medida de probabilidad en $\mathcal{B}_{\mathbb{R}}$ y
- para toda $A \in \mathcal{B}_{\mathbb{R}}$, $N(\cdot, A)$ es una función medible.

En el caso browniano no hemos probado la segunda condición. Sin embargo, notemos que la medibilidad es cierta cuando $A = (-\infty, y]$ para toda $y \in \mathbb{R}$. El lema de clases de Dynkin nos permite entonces obtener la medibilidad deseada. Como en el caso de matrices de transición, a las que podemos pensar como núcleos de medidas de probabilidad en algún conjunto finito, podemos definir el producto de núcleos de medidas de probabilidad, que no será en general commutativo. Si M y N son núcleos de probabilidad en \mathbb{R} , definimos al núcleo NM por medio de la fórmula

$$NM(x, A) = \int N(y, A) M(x, dy).$$

En el caso del movimiento browniano, vemos que

$$P_t P_s(x, (-\infty, z]) = \int \mathbb{P}(y + B_t \leq z) \mathbb{P}(x + B_s \in dy)$$

y entonces vemos que $P_t P_s(x, \cdot)$ es la convolución de una distribución normal con media x y varianza s con una normal centrada de varianza t . Se obtiene por lo tanto una normal con media x y varianza $s+t$, que por supuesto es la medida de probabilidad $P_{t+s}(x, \cdot)$. Por lo tanto, se obtiene la igualdad

$$P_{t+s} = P_t P_s,$$

que podemos interpretar como una versión de las ecuaciones de Chapman-Kolmogorov. A $(P_t, t \geq 0)$ se le conoce como semigrupo de transición del movimiento browniano. Para definir a un proceso de Markov (homogéneo) con valores en \mathbb{R} se hace algo similar.

DEFINICIÓN. Un semigrupo de transición en \mathbb{R} es una colección de núcleos de transición $N = (N_t, t \geq 0)$ tal que $N_t N_s = N_{t+s}$.

Un proceso estocástico X definido en el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ es un **proceso de Markov** con semigrupo de transición N si

$$\mathbb{P}(X_{t+s} \in A \mid \mathcal{F}_s^X) = N_t(X_s, A).$$

Equivalentemente, si para toda función $f : \mathbb{R} \rightarrow \mathbb{R}$ medible y acotada

$$\mathbb{E}(f(X_{t+s}) \mid \mathcal{F}_s^X) = \int f(y) N_t(X_s, dy).$$

La anterior definición se puede escribir de manera más compacta al definir el semigrupo de operadores de transición asociado a N . Primero, dado N_t y una función medible y acotada, podemos definir a la función acotada $N_t f$ de la siguiente manera:

$$N_t f(x) = \int f(y) N_t(x, dy).$$

Esta función será medible; para probarlo, sólo notamos que cuando $f = \mathbf{1}_A$, entonces $N_t f$ es medible por definición de núcleo de medidas de probabilidad. Luego, se extiende el resultado a funciones simples. Finalmente se aproxima a cualquier función medible y acotada por una sucesión de funciones medibles y uniformemente acotadas y se aplica el teorema de convergencia dominada para concluir que $N_t f$ es el límite de una sucesión de funciones medibles y por lo tanto medible. La definición de proceso de Markov se puede entonces escribir de la manera siguiente:

$$\mathbb{E}(f(X_{t+s}) \mid \mathcal{F}_s^X) = N_t f(X_s).$$

Finalmente, podemos enunciar el teorema de Itô.

TEOREMA 3.5. *Bajo las hipótesis del Teorema ??, sea X_t^x la (única) solución a la ecuación diferencial estocástica*

$$X_t^x = x + \int_0^t \sigma(X_s^x) dB_s + \int_0^t b(X_s^x) ds.$$

Entonces, X^x es un proceso de Markov homogéneo.

Vale la pena contrastar con el caso determinista en el que $\sigma = 0$. En este caso, notemos que

$$X_{t+s}^x = X_s^x + \int_0^t b(X_{r+s}^x) dr,$$

por lo que obtenemos la igualdad

$$X_{t+s}^x = X_t^{X_t^x}.$$

Esto se puede escribir como una propiedad de flujo de la función $F : (t, x) \mapsto X_t^x$:

$$F(t+s, x) = F(t, F(s, x)).$$

Algo similar ocurre en el caso estocástico; sin embargo, como las funciones F_t también dependen del azar, debemos también pensar en cuestiones de medibilidad.

PRUEBA DEL TEOREMA ??. Dada una función medible y acotada $f : \mathbb{R} \rightarrow \mathbb{R}$, definimos $P_t f(x) = \mathbb{E}(f(X_t^x))$. A continuación, probaremos que si $\mathcal{F}_t = \sigma(B_s : s \leq t)$ entonces

$$(4) \quad \mathbb{E}(f(X_{t+s}^x) \mid \mathcal{F}_s) = P_t f(X_s^x).$$

Puesto que X^x es adaptado a (\mathcal{F}_s) , la ecuación (??) implica que X^x es un proceso de Markov con semigrupo (P_t) .

Para probar (??), se requieren algunos preliminares. Sea $B_t^s = B_{t+s} - B_s$. Entonces B^s es un movimiento browniano independiente de \mathcal{F}_s . Sea por otra parte \tilde{X}^x el único proceso continuo y adaptado a la filtración canónica de B^s que resuelve la ecuación diferencial estocástica

$$\tilde{X}_t^x = x + \int_0^t \sigma(r, X_r^x) dB_r^s + \int_0^t b(r, \tilde{X}_r^x) dr.$$

Entonces $\tilde{X}_t^x = F_t(x, \omega)$ donde F_t es $\mathcal{B}_{\mathbb{R}} \otimes \mathcal{F}_t^{B^s}$ -medible; esto se puede ver a partir del procedimiento de aproximación dado en la prueba del Teorema ???. Consideraremos al proceso \overline{X} dado por

$$\overline{X}_t^x(\omega) = \begin{cases} X_t^x(\omega) & t < s \\ F_{t-s}(X_s^x(\omega, \omega)) & t > s \end{cases}.$$

Entonces \overline{X}^x satisface la misma ecuación diferencial que X^x , pues es fácil verificar la igualdad casi segura

$$\int_0^t \sigma(r, \tilde{X}_r^x) dB_r^s = \int_s^{t+s} \sigma(r, \tilde{X}_{r+s}^x) dB_r.$$

Por el Teorema ?? vemos que $\overline{X}_{t+s}^x = X_{t+s}^x$ casi seguramente. Finalmente, puesto que \mathcal{F}_s es independiente de $\mathcal{F}_t^{B^s}$ y X_s^x es \mathcal{F}_s -medible, entonces

$$\mathbb{E}(f(X_{t+s}^x) | \mathcal{F}_s) = \mathbb{E}(f(F_t(X_s^x, \cdot)) | \mathcal{F}_s) = P_t f(X_s^x).$$

□

Pasamos a un fenómeno con el que se debe tener cuidado al tratar de aproximar ya sea integrales estocásticas o ecuaciones diferenciales ordinarias al aproximar al browniano por un proceso de variación finita. Un ejemplo de esto sería el substituir al browniano por la sucesión de procesos gaussianos lineales por pedazos como lo hace Paul Lévy. El fenómeno que ilustraremos a continuación se conoce con el nombre de Wong-Zakai quienes lo introdujeron en [?]. Supongamos que B^n es una sucesión de procesos estocásticos cuyas trayectorias tienen casi seguramente variación finita y $B_0^n = 0$. Entonces, por la regla de la cadena, se sigue que

$$(B_t^n)^2 = 2 \int_0^t B_s^n dB_s^n.$$

Esto implica que

$$\lim_{n \rightarrow \infty} 2 \int_0^t B_s^n dB_s^n = B_t^2 \neq \int_0^t B_s dB_s.$$

El siguiente teorema no es más que una elaboración de esta idea.

TEOREMA 3.6. *Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ con derivada continua y B_n una sucesión de procesos con trayectorias de variación finita que comienzan en cero y convergen casi*

seguramente al movimiento browniano además de ser casi seguramente uniformemente acotados en compactos. Entonces

$$\lim_{n \rightarrow \infty} \int_0^t f(B_n(s)) dB_n(s) = \int_0^t f(B_s) dB_s - \frac{1}{2} \int_0^t f'(B_s) ds.$$

De igual manera, tenemos la versión para ecuaciones diferenciales estocásticas.

TEOREMA 3.7. *Suponga que σ y b no dependen de la variable temporal, que σ es derivable y que tanto σ , b como σ' son globalmente Lipschitz. Suponga además que existe $\varepsilon > 0$ tal que $\sigma \geq \varepsilon$. Sea B_n una sucesión de procesos con trayectorias de variación acotada que comienzan en cero y convergen casi seguramente al movimiento browniano uniformemente en compactos. Sea X_n la única solución a la ecuación diferencial ordinaria*

$$X_n(t) = x + \int_0^t \sigma(X_n(s)) dB_n(s) + \int_0^t b(X_n(s)) ds.$$

Entonces X_n converge casi seguramente y uniformemente en compactos al único proceso X que satisface la ecuación diferencial estocástica

$$dX_t = \sigma(X_t) dB_t + b(X_t) + \sigma'(X_t) dt.$$

CAPÍTULO 4

La integral estocástica respecto de semimartingalas continuas

1. Martingalas continuas y su variación cuadrática

Bajo ciertas condiciones, hemos visto que una integral estocástica es una martingala y que tiene trayectorias continuas. En esta sección abordaremos el estudio general de las martingalas con trayectorias continuas definidas en un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ con filtración $(\mathcal{F}_t, t \geq 0)$. Por el momento no asumiremos las condiciones habituales, aunque el lector debe guardar en mente que serán fundamentales para construir la integral estocástica.

Sorprendentemente, salvo las martingalas continuas triviales, todas tienen variación no acotada y esto imposibilita la definición de una integral estocástica tipo Lebesgue-Stieltjes.

PROPOSICIÓN 4.1. *Sea $(M_t, t \geq 0)$ una martingala continua con trayectorias de variación acotada en intervalos compactos casi seguramente. Entonces M tiene trayectorias constantes casi seguramente.*

DEMOSTRACIÓN. Al restar el valor inicial, podemos suponer que $M_0 = 0$.

Supongamos primero que la variación V_t de M en $[0, t]$ es casi seguramente menor o igual que la constante $K > 0$. Esto implica la cota

$$\sup_{\substack{|s_2-s_1| \leq \delta \\ s_1, s_2 \leq t}} |M_{s_2} - M_{s_1}| \leq 2K.$$

Consideremos particiones Δ de $[0, t]$ de norma menor o igual a δ . Entonces, puesto que los incrementos de una martingala no tienen correlación:

$$\begin{aligned} \mathbb{E}(M_t^2) &= \mathbb{E}\left(\left[\sum_{\Delta} M_{t_i} - M_{t_{i-1}}\right]^2\right) \\ &= \mathbb{E}\left(\sum_{\Delta} [M_{t_i} - M_{t_{i-1}}]^2\right) \\ &\leq \mathbb{E}\left(V_t \sup_{|s_1-s_2| \leq \delta} |M_{s_2} - M_{s_1}|\right). \end{aligned}$$

Por el teorema de convergencia acotada y la continuidad de las trayectorias de M , vemos que el lado derecho tiende a cero conforme la norma de la partición tiende a cero. Por lo tanto $M_t = 0$ casi seguramente. Así, vemos que M es una modificación del proceso $t \mapsto 0$ y al tener ambos trayectorias continuas, entonces M y 0 son indistinguibles.

Cuando la variación de M es sólo finita casi seguramente y no acotada por una constante, entonces consideramos a los tiempos de paro

$$S_K = \inf \{t \geq 0 : V_t > K\}.$$

Al utilizar la conclusión del párrafo anterior, notamos que la martingala M^{S_K} tiene trayectorias constantes y puesto que $S_K \wedge t \rightarrow t$ conforme $K \rightarrow \infty$ (pues la variación de M en $[0, t]$ es finita casi seguramente) vemos que M tiene trayectorias constantes casi seguramente. \square

Sin embargo, justo como en el caso browniano, las martingalas continuas tienen variación cuadrática finita. La idea de la prueba es considerar primero el caso de martingalas cuadrado integrables y luego, descomponer a la martingala sobre una partición Δ como sigue:

$$M_t^2 - M_0^2 = 2 \sum_{\Delta} M_{t_{i-1} \wedge t} [M_{t_i \wedge t} - M_{t_{i-1} \wedge t}] + \sum_{\Delta} [M_{t_i} - M_{t_{i-1}}]^2.$$

Si I_t^Δ denota al primer sumando del lado derecho y T_t^Δ al segundo, notemos que I^Δ es automáticamente una martingala. Por cálculos directos, se muestra que para cada t fija, I_t^Δ converge en L_2 conforme $|\Delta| \rightarrow 0$ y por la desigualdad de Doob, se obtiene que la convergencia es uniforme sobre compactos y que por lo tanto el límite es un proceso continuo y creciente. A dicho proceso lo denominaremos por $\langle M \rangle$ y lo caracterizaremos como el único proceso tal que $M^2 - \langle M \rangle$ es una martingala.

TEOREMA 4.1. *Si M es una martingala continua y acotada, entonces existe un único proceso creciente, continuo, adaptado y nulo en cero, denominado por $\langle M \rangle$, tal que $M^2 - \langle M \rangle$ es una martingala. Además, para cualquier sucesión de particiones Δ_n cuyo paso tienda a cero, se tiene la convergencia en probabilidad*

$$\mathbb{P} - \lim_{n \rightarrow \infty} \sum_{\Delta_n} [M_{t_i} - M_{t_{i-1}}]^2 = \langle M \rangle.$$

El Teorema ?? podría parecer limitado pues, al imponer que la martingala sea acotada, deja fuera incluso al movimiento browniano. Una sencilla técnica, llamada de localización, nos permite extender el teorema anterior a las llamadas martingalas locales y a las semimartingalas. Recordemos que si T es un tiempo aleatorio y X es un proceso estocástico entonces X^T denota al proceso X detenido en T , dado por $X_t^T = X_{t \wedge T}$.

PROPOSICIÓN 4.2. *Si M es una martingala continua y acotada y T es un (\mathcal{F}_t) -tiempo de paro entonces $\langle M^T \rangle = \langle M \rangle^T$.*

DEMOSTRACIÓN. Si M es acotada, entonces M^T es también una martingala acotada (respecto a la misma filtración), por lo que tiene sentido cuestionarse sobre su variación cuadrática. Puesto que por definición $M^2 - \langle M \rangle$ es una martingala entonces $(M^2 - \langle M \rangle)^T = (M^T)^2 - \langle M \rangle^T$ es una martingala y, por unicidad de la variación cuadrática, $\langle M \rangle^T = \langle M^T \rangle$. \square

Con la propiedad anterior podremos dar una primera extensión del Teorema ?? que cubra al movimiento browniano.

DEFINICIÓN. Una **martingala local continua** es un proceso estocástico $M = (M_t, t \geq 0)$ con trayectorias continuas tal que existe una sucesión de tiempos de paro $T_1 \leq T_2 \leq \dots$ tales que $T_n \rightarrow \infty$ casi seguramente y

- (1) M_0 es \mathcal{F}_0 -medible y
- (2) $(M - M_0)^{T_n}$ es una martingala acotada.

Si M es cualquier proceso estocástico con trayectorias continuas y T es un tiempo de paro tal que M_0 es \mathcal{F}_0 -medible y $(M - M_0)^T$ es una martingala acotada, decimos que el tiempo de paro T **reduce** al proceso M . Por ejemplo, el movimiento browniano es una martingala local. De hecho, cualquier martingala con trayectorias continuas es una martingala local, como se puede ver al definir $T_n = \inf \{t \geq 0 : |M_t - M_0| \geq n\}$.

COROLARIO 3. Si M es una martingala local continua, existe un único proceso creciente, nulo en cero, continuo y adaptado $\langle M \rangle$ tal que $M^2 - M$ es una martingala local continua. Además, para cualquier sucesión de particiones Δ_n sin puntos de acumulación cuyo paso tiende a cero, la sucesión T^{Δ_n} converge a $\langle M \rangle$ uniformemente en compactos en probabilidad.

DEFINICIÓN. Una **semimartingala continua** es un proceso estocástico X que se puede descomponer como $M + A$ donde M es una martingala local continua y A es un proceso de variación acotada en compactos.

Un detalle importante es que la descomposición es única (en el sentido de indistinguibilidad) si $M_0 = 0$.

COROLARIO 4. Si $X = M + A$ es una semimartingala continua, entonces $T^\Delta(X)$ converge uniformemente en compactos a $\langle M \rangle$ en probabilidad.

Como corolario adicional, obtenemos que si $X = X_0 + M + A$ y $Y = Y_0 + N + B$ son las descomposiciones canónicas de las semimartingalas X y Y podemos definir a la covariación de X y Y , denotada por $\langle X, Y \rangle$ por medio de la fórmula

$$\langle X, Y \rangle = \langle M, N \rangle$$

y que

$$\sum_{\Delta_n} (X_{t \wedge t_i} - X_{t \wedge t_{i-1}})(Y_{t \wedge t_i} - Y_{t \wedge t_{i-1}}) \rightarrow \langle X, Y \rangle_t$$

si Δ_n es una sucesión de particiones de $[0, \infty)$ sin puntos de acumulación cuya norma tiende a cero y la convergencia es en probabilidad uniformemente en compactos de la variable t .

Ahora pasaremos a la construcción de la covariación entre dos martingalas locales continuas. Para esto, recordemos la *fórmula de polarización* $(a+b)^2 - (a-b)^2 = 4ab$. Sean M y N dos martingalas locales continuas y definamos la **covariación** entre M y N , denotada $\langle M, N \rangle$ por medio de la fórmula

$$\langle M, N \rangle = \frac{\langle M + N \rangle - \langle M - N \rangle}{4}.$$

COROLARIO 5. *Sean M y N martingalas locales continuas. Entonces $\langle M, N \rangle$ es el único proceso continuo, nulo en cero, con trayectorias de variación finita y adaptado tal que $MN - \langle M, N \rangle$ es una martingala continua. Sea Δ_n es una sucesión de particiones de $[0, \infty)$ sin puntos de acumulación cuya norma tiende a cero. Entonces la sucesión de procesos*

$$T_t^{\Delta_n}(M, N) = \sum_{\Delta_n} (M_{t_i \wedge t} - M_{t_{i-1} \wedge t}) (N_{t_i \wedge t} - N_{t_{i-1} \wedge t})$$

converge uniformemente en compactos en probabilidad a (M, N) .

La integral estocástica se define en dos etapas: martingalas locales continuas y luego semimartingalas. Primero se introduce el espacio de integrandos.

DEFINICIÓN. Un proceso estocástico $K = (K_t, t \geq 0)$ se dice progresivamente medible si la aplicación $(s, \omega) \mapsto K_s(\omega)$ de $[0, t] \times \Omega$ en \mathbb{R} es $\mathcal{B}_{[0,t]} \otimes \mathcal{F}_t$ -medible.

Cualquier proceso con trayectorias continuas por la derecha y adaptado es progresivamente medible. Esto se puede consultar por ejemplo en [?, Prop. 1.13 p.5]

DEFINICIÓN. El espacio $L_2^{\text{loc}}(M)$ se define como la clase de procesos estocásticos progresivamente medibles K para los cuales existe una sucesión de tiempos de paro (S_n) tal que $S_n \leq S_{n+1}$, $S_n \rightarrow \infty$ casi seguramente y

$$\mathbb{E} \left(\int_0^{S_n} K_s^2 d\langle M \rangle_s \right) < \infty.$$

Un proceso progresivamente medible K es **localmente acotado** si existe una sucesión de tiempos de paro (S_n) tal que $S_n \leq S_{n+1}$, $S_n \rightarrow \infty$ casi seguramente y K^{S_n} es acotado.

Si M es una martingala local continua y K es localmente acotado entonces $K \in L_2^{\text{loc}}(M)$.

TEOREMA 4.2. *Sea M una martingala local continua y $H \in L_2^{\text{loc}}(M)$. Entonces existe una única martingala local continua que se anula en cero, denotada $H \cdot M$ tal que para cualquier martingala local continua N*

$$\langle H \cdot M, N \rangle = H \cdot \langle M, N \rangle.$$

Al proceso $H \cdot M$ se le conoce como la **integral estocástica (indefinida) de H respecto de M** .

Pasemos ahora al caso de las semimartingalas. Sea X una martingala local continua con descomposición canónica

$$X = X_0 + M + A$$

donde M es una martingala local continua (nula en cero) y A es un proceso continuo de variación acotada en compactos. El espacio adecuado de integrandos lo conformaran los procesos progresivamente medibles y localmente acotados. Si K es un tal proceso, se define la **integral estocástica** de K respecto de X , denotada por $K \cdot X$, como el proceso estocástico dado por

$$(K \cdot X)_t = K \cdot M + K \cdot A.$$

Notemos que $K \cdot X$ es una nueva semimartingala. El siguiente resultado resume las propiedades más importantes de la integral estocástica.

TEOREMA 4.3. *Sean $X = X_0 + M + A$ una semimartingala continua, y H, H_n y K procesos progresivamente medibles localmente acotados. Entonces*

- (1) $K \cdot (H \cdot X) = KH \cdot X$,
- (2) *si T es un tiempo de paro entonces $\mathbf{1}_{[0,T]} H \cdot X = (H \cdot X)^T = H \cdot X^T$,*
- (3) *si H es un proceso elemental, esto es tiene la forma*

$$H = \sum \lambda_i \mathbf{1}_{[t_{i-1}, t_i]}$$

donde λ_i es $\mathcal{F}_{t_{i-1}}$ -medible, entonces

$$(H \cdot X)_t = \sum_i \lambda_i (X_{t_i} - X_{t_{i-1}})$$

- (4) *si $H_n \rightarrow H$ uniformemente en compactos en probabilidad y $|H_n| \leq K$ entonces $H_n \cdot X \rightarrow H \cdot X$ uniformemente en compactos en probabilidad,*
- (5) *y si H es continuo por la derecha y Δ_n es una sucesión de particiones de $[0, t]$ cuya norma tiende a cero entonces*

$$\int_0^t H_s dX_s = \lim_{n \rightarrow \infty} \sum_{\Delta_n} H_{t_{i-1}} (X_{t_i} - X_{t_{i-1}}).$$

en probabilidad.

Si $X = (X^1, \dots, X^d)$ es un proceso estocástico con valores en \mathbb{R}^d tal que cada componente es una semimartingala continua, decimos que X es una **semimartingala vectorial**. Si $F : \mathbb{R}^d \rightarrow \mathbb{R}$ es dos veces diferenciable y $e^i \in \mathbb{R}^d$ denota al i -ésimo vector de la base canónica que tiene todas las entradas iguales a cero salvo la i -ésima igual a 1, denotaremos por $D_i F$ a la derivada de F en la dirección e^i . La notación $D_{i,j} F$ se utilizará para $D_j(D_i F)$, misma que se abreviará como D_i^2 cuando $i = j$. Cuando $d = 1$, se utiliza la notación D y D^2 .

TEOREMA 4.4 (Fórmula de Itô). *Sea $X = (X^1, \dots, X^d)$ una semimartingala vectorial y $F : \mathbb{R}^d \rightarrow \mathbb{R}$ de clase C_2 . Entonces el proceso $F(X) = (F(X_t))_{t \geq 0}$ es una semimartingala real con descomposición*

$$F(X) = F(X_0) + \sum_{i=1}^d D_i F(X_s) \cdot X^i + \frac{1}{2} \sum_{i,j=1}^d D_{i,j} F(X) \cdot \langle X^i, X^j \rangle.$$

Esta descomposición de $F(X)$ se conoce con el nombre de fórmula de Itô y usualmente se escribe de la siguiente manera:

$$F(X_t) = F(X_0) + \sum_{i=1}^d \int_0^t \frac{\partial F}{\partial x_i}(X_s) dX_s^i + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2 F}{\partial x_i \partial x_j}(X_s) d\langle X^i, X^j \rangle_s.$$

Un caso particular especialmente importante es la fórmula de integración por partes. Sean X y Y dos semimartingalas reales. Entonces:

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle X, Y \rangle_t.$$

2. Aplicaciones a la integral estocástica

2.1. La exponencial estocástica. Comencemos con la construcción de la exponencial estocástica de una martingala local continua M .

TEOREMA 4.5. *Existe un único proceso continuo y adaptado $\mathcal{E}(M)$ tal que*

$$\mathcal{E}(M)_t = 1 + \int_0^t \mathcal{E}(M)_s dM_s.$$

Se tiene la fórmula explícita

$$\mathcal{E}(M)_t = e^{M_t - \frac{1}{2} \langle M \rangle_t}.$$

DEMOSTRACIÓN. Sea $\mathcal{E}(M)_t = e^{M_t - \frac{1}{2} \langle M \rangle_t}$. Al aplicar la fórmula de Itô con la función $f(x_1, x_2) = e^{x_1 - y_2/2}$ y con la semimartingala vectorial $X = (M, \langle M \rangle)$, vemos que

$$\mathcal{E}(M)_t = 1 + \int_0^t \mathcal{E}(M)_s dM_s - \int_0^t \frac{1}{2} \mathcal{E}(M)_s d\langle M \rangle_s + \frac{1}{2} \int_0^t \mathcal{E}(M)_s d\langle M \rangle_s,$$

por lo que $\mathcal{E}(M)$ satisface la ecuación diferencial estocástica anunciada.

Por otra parte, notemos que $\mathcal{E}(M) > 0$, por lo que podemos aplicar la fórmula de Itô y concluir que

$$\mathcal{E}(M)_t^{-1} = 1 - \int_0^t \mathcal{E}(M)_s^{-1} dM_s + \int_0^t \mathcal{E}(M)_s^{-1} d\langle M \rangle_s.$$

Supongamos que X es continuo, adaptado y satisface la ecuación diferencial estocástica

$$X_t = 1 + \int_0^t X_s dM_s.$$

Entonces la fórmula de integración por partes nos permite deducir que

$$\begin{aligned} X_t \mathcal{E}(M)_t^{-1} &= 1 + \int_0^t X_s \mathcal{E}(M)_s^{-1} (d\langle M \rangle_s - M_s) \\ &\quad + \int_0^t X_s \mathcal{E}(M)_t^{-1} dM_s - \int_0^t X_s \mathcal{E}(M)_t^{-1} d\langle M \rangle_s \\ &= 1. \end{aligned}$$

Por lo tanto, concluimos que $X = \mathcal{E}(M)$. \square

2.2. El teorema de caracterización de Lévy. En esta sección haremos una primera aplicación de la fórmula de Itô a los procesos estocásticos.

TEOREMA 4.6 (Teorema de caracterización de Lévy). *Sea M una martingala local continua con variación cuadrática $\langle M \rangle_t = t$. Entonces M es un (\mathcal{F}_t) -movimiento browniano.*

DEMOSTRACIÓN. Aplicaremos la versión compleja de la martingala exponencial. Esto es, para $u \in \mathbb{R}$, consideremos a

$$\mathcal{E}(iuM_t) = e^{iuM_t + u^2 t/2},$$

que es una martingala local compleja. Es decir, su parte real y su parte imaginaria son martingalas locales como se puede verificar fácilmente. Por tener trayectorias acotadas en compactos, vemos que $\mathcal{E}(iuM)$ es una martingala compleja (y no sólo local). Por lo tanto, se sigue que para $s < t$:

$$\mathbb{E}\left(e^{iuM_t + u^2 t/2} \mid \mathcal{F}_s\right) = e^{iuM_s + u^2 s/2},$$

por lo cual para todo $A \in \mathcal{F}_s$

$$\mathbb{E}\left(\mathbf{1}_A e^{iu(M_t - M_s)}\right) = \mathbb{E}(\mathbf{1}_A) e^{-u^2(t-s)/2}.$$

Se sigue que si $\mathbb{P}(A) > 0$, entonces bajo la medida $\mathbb{P}(\cdot \mid A)$, $M_t - M_s$ tiene la misma función característica que una variable gaussiana centrada de varianza $t - s$ y por lo tanto la misma distribución. Así, para todo $C \in \mathcal{B}_{\mathbb{R}}$:

$$\mathbb{P}(A, M_t - M_s \in C) = \mathbb{P}(A) \mathbb{P}(B_{t-s} \in C).$$

Notemos que la fórmula anterior sigue siendo válida si $\mathbb{P}(A) = 0$. Se concluye que $M_t - M_s$ es independiente de \mathcal{F}_s y que M es un (\mathcal{F}_t) -movimiento browniano. \square

2.3. Martingalas locales continuas como cambios de tiempo del movimiento browniano. El objetivo de esta sección es probar que toda martingala local continua un movimiento browniano cambiado de tiempo.

TEOREMA 4.7 (Dambis-Dubins-Schwarz, [?, ?]). *Sea M una martingala local continua nula en cero y tal que $\langle M \rangle_\infty = \infty$. Entonces existe un movimiento browniano β tal que $M_t = \beta_{\langle M \rangle_t}$.*

DEMOSTRACIÓN. Puesto que $\langle M \rangle_\infty = \infty$, su inverso continuo por la derecha $\langle M \rangle^{-1}$ es finito casi-seguramente. Sea $\beta = M \circ \langle M \rangle^{-1}$. Puesto que $\langle M \rangle^{-1}$ puede tener saltos, lo mismo podría sucederle a β . Sin embargo, $\langle M \rangle^{-1}$ tiene un salto en t si y sólo si $\langle M \rangle$ es constante en $[\langle M \rangle_{t-}^{-1}, \langle M \rangle_t^{-1}]$. Puesto que los intervalos de constancia para M y $\langle M \rangle$ coinciden, se sigue que β es un proceso estocástico con trayectorias continuas. Veamos ahora que β es una martingala local continua respecto de una filtración que satisface las condiciones habituales. En efecto, note-mos que $\langle M \rangle_t^{-1}$ es un (\mathcal{F}_t) -tiempo de paro. Por lo tanto, el proceso β es adaptado respecto a la filtración cambiada de tiempo $(\mathcal{G}_t, t \geq 0)$ donde $\mathcal{G}_t = \mathcal{F}_{\langle M \rangle_t^{-1}}$. Esta filtración es completa puesto que la filtración original ya lo era. Por otra parte, es fácil verificar que $\mathcal{F}_{\langle M \rangle_t^{-1}+} = \mathcal{F}_{\langle M \rangle_t^{-1}}$ (como se afirma en el ejercicio 4 Cap. 1 de [?]). Sea

$$S_n = \inf \{t \geq 0 : |M_t| \geq n \text{ ó } \langle M \rangle_t \geq n\}.$$

Entonces M^{S_n} y $[M^2 - \langle M \rangle]^{S_n}$ son martingalas acotadas. Consideremos ahora a

$$T_n = \inf \{t \geq 0 : \langle M \rangle_t^{-1} \geq S_n\}.$$

Notemos que

$$\{T_n \leq t\} = \{S_n \leq \langle M \rangle_t^{-1}\} \in \mathcal{F}_{\langle M \rangle_t^{-1}}.$$

Por lo tanto, T_n es un (\mathcal{G}_t) -tiempo de paro. Puesto que M^{S_n} es una martingala acotada, podemos aplicar muestreo opcional para concluir que si $s \leq t$:

$$\mathbb{E}\left(M_{S_n \wedge \langle M \rangle_t^{-1}} \mid \mathcal{G}_s\right) = M_{S_n \wedge \langle M \rangle_s^{-1}}.$$

Notemos ahora que

$$\langle M \rangle_t^{-1} \wedge S_n = \langle M \rangle_{t \wedge T_n}^{-1}.$$

Por lo tanto, vemos que

$$\mathbb{E}\left(\beta_t^{T_n} \mid \mathcal{G}_s\right) = \beta_s^{T_n}$$

y así hemos probado que β es una martingala local continua. Por otra parte, al notar que

$$\langle M \rangle_{\langle M \rangle_s^{-1} \wedge S_n} = s \wedge T_n,$$

podemos aplicar un argumento similar para probar que $(\beta^2 - \text{Id})^{T_n}$ es una martingala acotada y por lo tanto la variación cuadrática de β es la identidad. Por el teorema de caracterización de Lévy, β es un $(\mathcal{G}_t)_{t \geq 0}$ -movimiento browniano. Por

construcción, vemos que $M = \beta \circ \langle M \rangle$ (donde de nuevo se utiliza que los intervalos de constancia de M y de $\langle M \rangle$ coinciden). \square

El teorema anterior es la clave para ver que ciertos procesos de interés son soluciones a ecuaciones diferenciales estocásticas. (Posteriormente, analizaremos a profundidad la ecuación diferencial estocástica satisfecha por la norma al cuadrado del movimiento browniano en dimensión δ .) También hay una versión del teorema anterior que no utiliza la hipótesis de que la variación cuadrática sea infinita. Sin embargo, el movimiento browniano se encuentra entonces en una extensión de nuestro espacio de probabilidad. Además, hay una versión multidimensional del teorema anterior conocido como Teorema de Knight.

TEOREMA 4.8 ([?]). *Sean M^1, \dots, M^n martingalas locales continuas tales que $\langle M^i \rangle_\infty = \infty$ y $\langle M^i, M^j \rangle = 0$. Entonces existe un movimiento browniano n -dimensional $\beta = (\beta^1, \dots, \beta^n)$ tal que $M^i = \beta^i \circ \langle M^i \rangle^{-1}$.*

2.4. La norma del movimiento browniano en \mathbb{R}^d . Sea $B = (B^1, \dots, B^d)$ un movimiento browniano y definamos a

$$Z_t = \|\vec{x} + B_t\|^2 = \sum_{i=1}^d (x_i + B_t^i)^2.$$

Al aplicar la fórmula de Itô con la función $f(\vec{y}) = \|\vec{x} + \vec{y}\|^2$ a la semimartingala vectorial B , vemos que

$$Z_t = f(B_t) = x + \sum_{i=1}^d \int_0^t 2(x_i + B_s^i) dB_s^i + dt$$

donde $x = \|\vec{x}\|^2$. Definamos a

$$M_t = \sum_{i=1}^d \int_0^t 2(x_i + B_s^i) dB_s^i.$$

Entonces M es una martingala local continua con variación cuadrática

$$\sum_{i=1}^d \int_0^t 4Z_s ds.$$

Sea ahora $h : \mathbb{R} \rightarrow \mathbb{R}$ una función de clase C_2 . Entonces

$$\begin{aligned} h(Z_t) &= h(x) + \int_0^t h'(Z_s) dZ_s + \frac{1}{2} \int_0^t h''(Z_s) 4Z_s ds \\ &= h(x) + \int_0^t h'(Z_s) dM_s + \int_0^t h'(Z_s) \delta + h''(Z_s) 2Z_s ds. \end{aligned}$$

Vemos entonces que si

$$2xh''(x) + \delta h'(x) = 0$$

entonces $h(Z)$ será una martingala local continua. Por otra parte, al probar funciones de la forma $h(x) = x^\alpha$, vemos que si $\delta \neq 2$ entonces

$$h(x) = x^{1-\delta/2}$$

satisface la ecuación diferencial anterior mientras que cuando $\delta = 2$, la función

$$h(x) = \log x$$

lo hace. Sean $0 < r < x < R$ y definamos a $T_{r,R}$ como la primera vez que B sale del anillo $\{\vec{x} : r < \|\vec{x}\|^2 < R\}$. En otras palabras, definamos a

$$T_r = \inf \{t \geq 0 : Z_t \leq r\}, \quad T^R = \inf \{t \geq 0 : Z_t \geq R\} \quad \text{y} \quad T_{r,R} = R_r \wedge T^R.$$

Notemos que $T_{r,R} < \infty$ casi seguramente puesto que las variables $\|B_{n+1} - B_n\|^2$ son independientes e idénticamente distribuidas y $\mathbb{P}(\|B_1\|^2 > 2R) > 0$. Por Borel-Cantelli, casi seguramente existe n tal que $\|B_{n+1} - B_n\|^2 > 2R$ y para dicha n forzosamente se sigue que $T_{r,R} \leq n$.

Puesto que $h(Z^{T_{r,R}})$ es una martingala local continua acotada, es una martingala uniformemente integrable y por lo tanto

$$h(x) = \mathbb{E}(h(Z_0)) = \mathbb{E}(h(Z_{T_{r,R}})) = h(r)p + h(R)(1-p) \quad \text{donde} \quad p = \mathbb{P}(T_r < T_R).$$

Se sigue que

$$\mathbb{P}(T_r < T_R) = \begin{cases} \frac{R^{1-\delta/2} - r^{1-\delta/2}}{R^{1-\delta/2} - r^{1-\delta/2}} & \delta \neq 2 \\ \frac{\log R/x}{\log R/r} & \delta = 2 \end{cases}.$$

Puesto que las trayectorias de Z son continuas, se sigue que $T_R \rightarrow \infty$ conforme $R \rightarrow \infty$. Se deduce

$$\mathbb{P}(T_r < \infty) = \begin{cases} 1 & \delta \leq 2 \\ \left(\frac{x}{r}\right)^{1-\delta/2} & \delta > 2 \end{cases}.$$

Por otro lado, puesto que $T_r \rightarrow T_0$ conforme $r \rightarrow 0$ entonces vemos que

$$\mathbb{P}(T_0 < T_R) = \begin{cases} 0 & \delta \geq 2 \\ 1 - \left(\frac{x}{R}\right)^{1-\delta/2} & \delta < 2 \end{cases}.$$

Nótese que se ha utilizado $\delta < 2$ en vez de $\delta = 1$. Esto se sigue de que es posible definir a un proceso que actúe como la norma al cuadrado del movimiento browniano en dimensión δ para cualquier $\delta \geq 0$. En efecto, a continuación utilizaremos el teorema de Dambis-Dubins-Schwarz para verificar que cuando δ es entero no-negativo, entonces Z satisface una ecuación diferencial estocástica (parametrizada por δ). Se utilizará esta ecuación diferencial estocástica para darle sentido a Z cuando δ no es natural. Antes de eso, continuemos con algunas consecuencias de los cálculos que hemos hecho:

COROLARIO 6. *Sea B un movimiento browniano δ -dimensional que parte de cero. Si $\delta \geq 2$, B jamás regresa a cero. Si $\delta \leq 2$ entonces B regresa a cualquier vecindad de cero y el conjunto de puntos en que se encuentra en una vecindad de cero no es acotado. Si $\delta > 2$, B es transitorio.*

DEMOSTRACIÓN. Ya hemos probado que para $\delta \geq 2$ y $x \neq 0$ entonces $x + B$ jamás se anula. Apliquemos lo anterior al proceso $B_{\varepsilon+t}, t \geq 0$ donde $\varepsilon > 0$. Puesto que $B_\varepsilon \neq 0$ casi seguramente, al condicionar por B_ε , vemos que $B_{\varepsilon+t}, t \geq 0$ jamás se anula casi seguramente. Al ser válida esta conclusión para cualquier $\varepsilon > 0$, vemos que $B_t, t > 0$ jamás se anula.

También hemos visto que si $\delta \leq 2$ y $x \neq 0$ entonces $x + B$ regresa a cualquier vecindad (fija) de cero. Al aplicar esto al proceso $B_{t+n}, t \geq 0$, condicionalmente a B_n (que es casi seguramente distinto de cero), vemos que casi seguramente $B_{t+n}, t \geq 0$ regresa a cualquier vecindad fija V de cero. Así, para toda $n \geq 1$ existe $t_n \geq n$ tal que B_{t_n} pertenece a V y por lo tanto, el conjunto de visitas de B a V no es acotado.

Finalmente, si $\delta > 2$ y $\vec{x} \neq 0$ entonces $\|x + B_t\|^{2-\delta}$ es una martingala local no-negativa. Esto implica que se trata de una supermartingala no-negativa. Por lo tanto, converge casi seguramente a un límite finito, digamos ξ . Por el lema de Fatou y la autosimilitud del movimiento browniano vemos que

$$\mathbb{E}(\xi) \leq \liminf_{t \rightarrow \infty} \mathbb{E}\left(\frac{1}{\|x + B_t\|^{2-\delta}}\right) = \lim_{t \rightarrow \infty} \mathbb{E}\left(\frac{1}{\|x + \sqrt{t}B_1\|^{2-\delta}}\right) = 0. \quad \square$$

Sean $\delta \geq 2$ y $\vec{x} \neq 0$. Puesto que $Z \neq 0$ casi seguramente, el proceso $1/2\sqrt{Z}$ es continuo, por lo que podemos definir al proceso β mediante

$$\beta = \frac{1}{2\sqrt{Z}} \cdot M.$$

El teorema de caracterización de Lévy nos dice que β es un movimiento browniano y por construcción

$$(5) \quad Z_t = \|x\|^2 + \int_0^t 2\sqrt{Z_s} d\beta_s + \delta t.$$

Por supuesto la ecuación diferencial anterior tiene sentido aún cuando δ no sea un entero positivo y esta es la manera en la que consideraremos al cuadrado de la norma del browniano δ -dimensional aún cuando δ no sea un entero. El único problema con la ecuación anterior es que no podemos utilizar el teorema de existencia y unicidad para ecuaciones diferenciales estocásticas puesto que el coeficiente de la ecuación no es Lipschitz.

2.5. Movimiento browniano y funciones armónicas. En esta sección, veremos como el movimiento browniano nos permite resolver la ecuación de Poisson. El lector puede consultar un desarrollo más a profundidad de estos temas en [?] y [?].

Sea $\delta \in \mathbb{Z}_+$ y consideremos un abierto $D \subset \mathbb{R}^\delta$ con cerradura \overline{D} y frontera $\partial(D)$. Consideremos además una condición de frontera de Dirichlet $f : \partial(D) \rightarrow \mathbb{R}$ continua y un término que representa la fuente de calor externa $g : D \rightarrow \mathbb{R}$. Una solución a la ecuación de Poisson es una función continua $u : \overline{D} \rightarrow \mathbb{R}$ de clase C_2 en D y tal que

$$\begin{cases} \Delta u(x) = -g(x) & x \in D \\ u(x) = f(x) & x \in \partial(D) \end{cases}.$$

Si $g = 0$, la ecuación de Poisson resultante se denomina ecuación de Laplace. Sea $B = (B^1, \dots, B^\delta)$ un movimiento browniano; utilizaremos a B para dar un resultado de unicidad para la ecuación de Poisson.

TEOREMA 4.9. *Supongamos que D, f y g son acotadas y que u es solución a la ecuación de Poisson. Sea*

$$S = \inf \{t \geq 0 : B_t \notin \overline{D}\}.$$

Entonces

$$u(x) = \mathbb{E}_x \left(f(B_S) + \int_0^S g(B_s) \, ds \right)$$

para toda $x \in D$.

DEMOSTRACIÓN. Sea $M_t = u(B_t^S) + \int_0^{t \wedge S} g(B_s) \, ds$. Al utilizar la fórmula de Itô, vemos que

$$\begin{aligned} M_t &= u(x) + \sum_i \int_0^{t \wedge S} D_i u(B_s) \, dB_s^i + \frac{1}{2} \int_0^{t \wedge S} \Delta u(B_s) \, ds + \int_0^{t \wedge S} g(B_s) \, ds \\ &= u(x) + \sum_i \int_0^{t \wedge S} D_i u(B_s) \, dB_s^i, \end{aligned}$$

donde la última igualdad se deduce pues u satisface la ecuación de Poisson.

Así, M es una martingala local. Puesto que u es continua y \overline{D} es acotado, se sigue que u es acotada. Por lo tanto M es una martingala acotada. Además, al ser D acotado, se sigue que $S < \infty$ casi seguramente y por lo tanto

$$M_t \rightarrow u(B_S) + \int_0^S g(B_s) \, ds = f(B_S) + \int_0^S g(B_s) \, ds,$$

casi seguramente y en L_1 . Al aplicar muestreo opcional, se sigue que

$$\mathbb{E}(M_S) = \mathbb{E}(M_0)$$

lo cual significa que

$$u(x) = \mathbb{E}_x \left(f(B_S) + \int_0^S g(B_s) \, ds \right).$$
□

2.6. La fórmula de Feynman-Kac. La fórmula de Feynman-Kac es introducida por Kac en [?] para calcular la distribución F_t de la variable aleatoria

$$A_t = \int_0^t v(B_s) ds$$

donde $v \geq 0$ satisface ciertas condiciones y B es un movimiento browniano. Un caso particular es cuando $v = \mathbf{1}_{(0,\infty)}$, en cuyo caso la distribución de A_t/t había sido encontrada por Lévy en [?] y coincide con la llamada distribución arcoseno, que es la distribución de una variable Beta de parámetros $1/2$ y $1/2$. Las investigaciones de Kac siguen a unas anteriores de Erdős y Kac publicadas en [?] y [?] en la que consideran teoremas límites para funcionales de caminatas aleatorias y ven que en ciertos casos no dependen de la distribución de salto de la caminata aleatoria. De hecho, el punto de vista de Kac para encontrar la distribución de A_t es discretizar a A_t , encontrar una ecuación en diferencias para calcular la distribución de la aproximación, resolverla y pasar al límite. El nombre de Feynman aparece en la fórmula puesto Kac argumenta que su método está influenciado fuertemente por la derivación de Feynman de la ecuación de Shrödinger. En [?] se pueden consultar aplicaciones de la medida de Wiener a la física cuántica con una discusión sobre la fórmula de Feynman-Kac.

La formulación moderna de la fórmula de Feynman-Kac nos presenta una liga entre ciertas ecuaciones diferenciales parabólicas y ciertas difusiones. En efecto, nos afirma (en el caso unidimensional) que si existe una solución $u(t, x)$ a la ecuación

$$\frac{\partial u}{\partial t} + b \frac{\partial u}{\partial x} + \sigma^2 \frac{\partial^2 u}{\partial x^2} + f = vu$$

para $u : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ donde b, σ, f y v dependen de t y de x y se satisface la condición terminal

$$u(x, T) = \psi(x)$$

entonces u está dada por la fórmula

$$u(t, x) = \mathbb{E} \left(\int_t^T e^{- \int_{t_1}^{t_2} v(X_{t_2}) dt_2} f(t_1, X_{t_1}) dt_1 + e^{- \int_t^T v(X_{t_1}) dt_1} \psi(X_T) \right),$$

donde se asume que X satisface la ecuación diferencial estocástica

$$X_t = x + \int_0^t \sigma(s, X_s) dB_s + \int_0^t b(x, X_s) ds.$$

En particular, lo anterior representa un resultado de unicidad bajo el supuesto probabilístico de existencia débil a la ecuación diferencial estocástica.

Ahora veremos cómo probar dichos resultados, enfocándonos en casos particulares que muestren las ideas principales.

Comencemos con la liga entre el movimiento browniano y la ecuación de calor.

PROPOSICIÓN 4.3. Si u es continua en $[0, \infty) \times \mathbb{R}^\delta$, de clase C_2 en $(0, \infty) \times \mathbb{R}^\delta$ y satisface el problema de Cauchy

$$(6) \quad \begin{cases} \frac{\partial u}{\partial t} - \frac{1}{2} \Delta u = 0 \\ u(0, x) = f(x) \end{cases}$$

para alguna función continua y acotada f , entonces

$$u(t, x) = \mathbb{E}_x(f(B_t)).$$

DEMOSTRACIÓN. Probemos primero, mediante un argumento analítico, que u es acotada. En efecto, se afirma que para toda $\delta > 0$ y $M > 0$,

$$\max_{\delta \leq t \leq T, \|x\| \leq M} u(t, x) \leq \max_{\|x\| \leq M} u(\delta, x).$$

En efecto, sean $\varepsilon > 0$ y $v(t, x) = u(t, x) - \varepsilon t$ y supongamos que v se maximiza en el interior de $[\delta, t] \times \{\|x\| \leq M\}$, digamos en (t^*, x^*) . Notemos primero que

$$\frac{\partial v}{\partial t} - \Delta v = \frac{\partial u}{\partial t} - \varepsilon - \Delta u = -\varepsilon.$$

Por otra parte, puesto que v se maximiza en (t^*, x^*) , vemos que

$$\frac{\partial v}{\partial t}(t^*, x^*) \geq 0 \quad \text{y} \quad \Delta v(t^*, x^*) \leq 0.$$

Esto implica que

$$\frac{\partial v}{\partial t}(t^*, x^*) - \Delta v(t^*, x^*) \geq 0,$$

una contradicción. Por lo tanto v alcanza su máximo en $[\delta, t] \times \{\|x\| \leq M\}$ en la frontera para cualquier $\varepsilon > 0$ y por lo tanto, u también. (Un argumento similar aplica al mínimo.) Al tomar el límite conforme $\delta \rightarrow 0$, vemos que

$$\sup_{t \leq T, \|x\| \leq M} |u(t, x)| \leq \sup_x |f(x)| < \infty$$

(pues supusimos que f es acotada) y al tomar el límite conforme $M \rightarrow \infty$, concluimos que u es acotada.

Sea $\varepsilon \in (0, t)$. Puesto que u es acotada y satisface la ecuación de calor entonces $u(s, B_{t-s})$ es una martingala en $[0, t - \varepsilon]$ y no sólo una martingala local. Por lo tanto

$$\mathbb{E}_x(u(\varepsilon, B_{t-\varepsilon})) = \mathbb{E}_x(u(t, B_0)) = u(t, x).$$

Puesto que u es continua y acotada y $u(0, x) = f(x)$, podemos utilizar el teorema de convergencia acotada para ver que

$$\mathbb{E}_x(f(B_t)) = u(t, x). \quad \square$$

Generalizaremos ahora el razonamiento anterior para obtener la formulación moderna de la fórmula de Feynman-Kac. Como se observa en [?], la fórmula de

Feynman-Kac se comprende muy bien cuando se comienza con el movimiento browniano matado en un tiempo exponencial. En efecto, si T es exponencial de parámetro λ e independiente de B y definimos

$$\tilde{B}_t = \begin{cases} B_t & t < T \\ \Delta & T \geq t \end{cases}$$

(donde Δ se interpreta como el estado cementerio y extendemos a cualquier función real como cero en Δ) entonces para cualquier función continua y acotada se tiene que la función

$$u(t, x) = \mathbb{E}_x \left(f(\tilde{B}_t) \right) = e^{-\lambda t} \mathbb{E}_x(f(B_t))$$

satisface el problema de Cauchy

$$\begin{cases} \frac{\partial u}{\partial t} - \frac{1}{2} \Delta u = \lambda u \\ u(0, x) = f(x) \end{cases}$$

En un caso más general, consideremos a

$$u(t, x) = \mathbb{E}_x \left(f(B_t) e^{- \int_0^t v(B_s) ds} \right).$$

La interpretación es que consideramos la esperanza de un Browniano matado a tasa $v(x)$ cuando se encuentra en el estado x . Si f es continua y acotada y v es no-negativa entonces u es continua y acotada. Al utilizar la propiedad de Markov vemos que

$$\mathbb{E}_x \left(f(B_t) e^{- \int_0^t v(B_s) ds} \mid \mathcal{F}_s \right) = e^{- \int_0^s v(B_r) dr} u(t-s, B_s)$$

para $s \leq t$. Definamos

$$\Pi_t = e^{- \int_0^t v(B_s) ds}.$$

Si u fuera de clase C_2 entonces la fórmula de Itô nos diría que

$$\begin{aligned} \Pi_t u(t-s, B_s) &= u(t, x) + \int_0^s \Pi_r D_2 u(t-r, B_r) dB_r - \int_0^s \Pi_r D_1 u(t-r, B_r) dr \\ &\quad + \frac{1}{2} \int_0^r \Pi_r \Delta u(t-r, B_r) dr - \int_0^s u(t-r, B_r) v(B_r) \Pi_r dr. \end{aligned}$$

Así, vemos que una condición natural para que $\Pi_s u(t-s, B_s)$ sea una martingala local es que u satisfaga la ecuación

$$\begin{cases} \frac{\partial u}{\partial t} - \frac{1}{2} \Delta u = vu \\ u(0, x) = f(x) \end{cases}.$$

Por otro lado, mostremos que hay a lo más una solución acotada para la ecuación anterior. En efecto, si u es una solución continua y acotada a dicha ecuación entonces la fórmula de Itô nos dice que

$$\Pi_s u(t-s, B_s)$$

es una martingala acotada. Por lo tanto

$$u(t, x) = \mathbb{E}_x(\Pi_t u(0, B_t)) = \mathbb{E}_x(\Pi_t f(B_t)).$$

3. El teorema de Girsanov

La fórmula de Itô nos dice que la clase de semimartingalas es invariante ante composición con funciones de clase C_2 . Ahora examinaremos otra propiedad de invariancia de las semimartingalas: la invariancia ante cambios de medida (localmente) absolutamente continuos. Si \mathbb{P} y \mathbb{Q} son medidas de probabilidad absolutamente continuas y X es una semimartingala al utilizar la medida de probabilidad entonces el célebre teorema de Girsanov nos ayudará a encontrar la descomposición de semimartingala de X cuando se utiliza la medida \mathbb{Q} .

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad dotado de una filtración $(\mathcal{F}_t, t \geq 0)$ que satisface las condiciones habituales. Recordemos que una medida de probabilidad \mathbb{Q} en (Ω, \mathcal{F}) es absolutamente continua respecto de \mathbb{P} , denotado $\mathbb{Q} \ll \mathbb{P}$, si para todo $A \in \mathcal{F}$ con $\mathbb{P}(A) = 0$ se tiene que $\mathbb{Q}(A) = 0$.

PROPOSICIÓN 4.4. *Supongamos que $\mathbb{Q} \ll \mathbb{P}$ y sea*

$$\tilde{D}_t = \frac{d\mathbb{P}|\mathcal{F}_t}{d\mathbb{Q}|\mathcal{F}_t}.$$

Entonces \tilde{D} admite una modificación D que es una una martingala càd no-negativa y uniformemente integrable. Para todo T tiempo de paro se tiene:

$$D_T = \frac{d\mathbb{P}|\mathcal{F}_T}{d\mathbb{Q}|\mathcal{F}_T}.$$

Si \mathbb{Q} es equivalente a \mathbb{P} entonces $D_t > 0$ para toda $t \geq 0$ casi seguramente.

DEMOSTRACIÓN. Si $A \in \mathcal{F}_s$ y $s \leq t$ entonces $A \in \mathcal{F}_t$ y por definición de \tilde{D}_s y \tilde{D}_t : $\mathbb{E}(\mathbf{1}_A \tilde{D}_s) = \mathbb{Q}(A) = \mathbb{E}(\mathbf{1}_A \tilde{D}_t)$. Por lo tanto \tilde{D} es una \mathbb{P} -martingala. Puesto que hemos asumido las condiciones habituales para $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ vemos que \tilde{D} admite una modificación càdlàg que también es una martingala y también se satisface la relación

$$D_t = \frac{d\mathbb{P}|\mathcal{F}_t}{d\mathbb{Q}|\mathcal{F}_t}.$$

Notemos que lo anterior vale también para $t = \infty$, por lo que D es uniformemente integrable. Si T es un tiempo de paro y $A \in \mathcal{F}_T$ entonces, al aplicar muestreo opcional, vemos que

$$\mathbb{Q}(A) = \mathbb{E}_{\mathbb{P}}(\mathbf{1}_A D_{\infty}) = \mathbb{E}_{\mathbb{P}}(\mathbf{1}_A D_T).$$

Por lo tanto

$$D_T = \frac{d\mathbb{P}|\mathcal{F}_T}{d\mathbb{Q}|\mathcal{F}_T}.$$

Finalmente, si $S = \inf \{t \geq 0 : D_t = 0\}$ entonces

$$\mathbb{Q}(S < \infty) = \mathbb{E}_{\mathbb{P}}(\mathbf{1}_{S < \infty} D_S) = 0$$

y si \mathbb{Q} es equivalente a \mathbb{P} , esto implica que $\mathbb{P}(S < \infty) = 0$. \square

Así, en el caso en que tengamos dos medidas de probabilidad equivalente, el proceso de derivadas de Radon-Nikodym es una martingala estrictamente positiva. El siguiente resultado nos permitirá expresar a dicha martingala, cuando tenga trayectorias continuas, como una exponencial estocástica.

PROPOSICIÓN 4.5. *Sea D una martingala local continua estrictamente positiva. Existe entonces una única martingala local continua L tal que $D = \mathcal{E}(L)$. Además:*

$$L_t = \log(D_0) + \int_0^t D_s^{-1} dD_s.$$

DEMOSTRACIÓN. Notemos que si $D = \mathcal{E}(L)$ entonces $1/D = \mathcal{E}(-L)$.

Para probar la unicidad, supongamos que $D = \mathcal{E}(L) = \mathcal{E}(\tilde{L})$. Entonces

$$1 = D \frac{1}{D} = \mathcal{E}(L) \frac{1}{\mathcal{E}(\tilde{L})} = e^{L - \tilde{L}}.$$

Para la existencia, utilizamos la fórmula de Itô con la función \log , que es infinitamente diferenciable en $(0, \infty)$ a la martingala local continua estrictamente positiva D . Se tiene entonces que

$$\log(D_t) = \log(D_0) + \int_0^t D_s^{-1} dD_s - \frac{1}{2} \int_0^t D_s^{-2} d\langle D \rangle_s.$$

Si

$$L_t = \log(D_0) + \int_0^t D_s^{-1} dD_s,$$

entonces

$$\langle L \rangle_t = \int_0^t D_s^{-2} d\langle D \rangle_s,$$

por lo que

$$\log(D_t) = L_t - \frac{1}{2} \langle L \rangle_t$$

y por lo tanto

$$D_t = \exp(L_t - \frac{1}{2} \langle L \rangle_t) = \mathcal{E}(L)_t.$$

\square

Ahora podemos enunciar el teorema de Girsanov.

TEOREMA 4.10 (Teorema de Girsanov). *Sea \mathbb{Q} equivalente a \mathbb{P} en \mathcal{F}_∞ . Sea D la versión càdlàg del proceso de derivadas de Radon-Nikodym y supongamos que D es continuo. Sea L una martingala local continua tal que $D = \mathcal{E}(L)$. Si M es cualquier $(\mathcal{F}_t, \mathbb{P})$ -martingala local continua el proceso \tilde{M} dado por*

$$\tilde{M}_t = M_t - \langle M, L \rangle_t$$

es una $(\mathcal{F}_t, \mathbb{Q})$ -martingala local continua.

Notemos que en particular, M es una \mathbb{Q} -semimartingala. Notemos además que la variación cuadrática no depende de la medida que estemos utilizando puesto que los límites en probabilidad coinciden para medidas equivalentes. Así, si M es un movimiento browniano bajo \mathbb{P} , entonces \tilde{M} lo es bajo \mathbb{Q} .

DEMOSTRACIÓN. Mostremos primero que si XD es una \mathbb{P} -martingala local continua entonces X es una \mathbb{Q} -martingala local. En efecto,

$$T_n = \inf \{t \geq 0 : D_t \geq n \text{ ó } |X_t| \geq n\}.$$

Entonces (T_n) es una sucesión creciente de tiempos de paro que convergen a ∞ , $(XD)^{T_n}$ es una martingala acotada y X^{T_n} es un proceso acotado. Por lo tanto, si $A \in \mathcal{F}_s$ y $s \leq t$ entonces

$$\mathbb{E}_{\mathbb{Q}}(X T_n \wedge t \mathbf{1}_A) = \mathbb{E}_{\mathbb{P}}(X_{t \wedge T_n} D_{t \wedge T_n} \mathbf{1}_A) = \mathbb{E}_{\mathbb{P}}(X_{s \wedge T_n} D_{s \wedge T_n} \mathbf{1}_A) = \mathbb{E}_{\mathbb{Q}}(X_{T_n \wedge s} \mathbf{1}_A).$$

Por otra parte, notemos que puesto que $\mathbb{P}(T_n \leq t) \rightarrow 0$ se sigue que $\mathbb{Q}(T_n \leq t) \rightarrow 0$ y que por lo tanto $T_n \rightarrow \infty$ \mathbb{Q} -casi seguramente. Así, X es una \mathbb{Q} martingala local continua.

Ahora aplicaremos la observación anterior. Si M es una \mathbb{P} -martingala local continua y $\tilde{M} = M - \langle M, L \rangle$, podemos aplicar la fórmula de Itô para escribir

$$\begin{aligned} D_t \tilde{M} &= D_0 \tilde{M}_0 + \int_0^t D_s d\tilde{M}_s + \int_0^t \tilde{M}_s dD_s + \langle \tilde{M}, D \rangle \\ &= D_0 \tilde{M}_0 + \int_0^t D_s dM_s + \int_0^t \tilde{M}_s dD_s - \int_0^t D_s d\langle M, L \rangle_s + \langle M, D \rangle \\ &= D_0 \tilde{M}_0 + \int_0^t D_s dM_s + \int_0^t \tilde{M}_s dD_s. \end{aligned}$$

Por lo tanto $D\tilde{M}$ es una \mathbb{P} -martingala local continua y se deduce que entonces \tilde{M} es una \mathbb{Q} -martingala local continua. \square

Uno de los ejemplos típicos de aplicación del teorema de Girsanov es al movimiento browniano con deriva. En efecto, si B es un movimiento browniano bajo \mathbb{P} y

$$\mathbb{Q}_t(A) = \mathbb{E}_{\mathbb{P}}(\mathbf{1}_A e^{\mu B_t - \mu^2 t/2}),$$

para $A \in \mathcal{F}$, entonces \mathbb{Q}_t es una medida de probabilidad absolutamente continua respecto de \mathbb{P} y equivalentemente a ella en \mathcal{F}_t . Por lo tanto, bajo \mathbb{Q}_t el proceso

$B - \mu \text{Id}$ es un browniano en $[0, t]$; equivalentemente, bajo \mathbb{Q}_t , B es un browniano con deriva $-\mu$ en $[0, t]$. Otro ejemplo es que si $T_b(X) = \inf\{t \geq 0 : X_s \geq b\}$ entonces

$$\begin{aligned}\mathbb{E}\left(e^{-\lambda T_b(B+\mu \text{Id})}\right) &= \lim_{t \rightarrow \infty} \mathbb{E}\left(e^{-\lambda T_b(B+\mu \text{Id}) \wedge t}\right) \\ &= \lim_{t \rightarrow \infty} \mathbb{E}\left(e^{-\lambda T_b(B) \wedge t} e^{\mu B_{T_b(B) \wedge t} - \mu^2 T_b(B) \wedge t / 2}\right) \\ &= \mathbb{E}\left(e^{-\lambda T_b(B)} e^{\mu b - \mu^2 T_b(B) / 2}\right) \\ &= e^{\mu b} e^{-|b| \sqrt{2\lambda + \mu^2}}.\end{aligned}$$

En particular, vemos que

$$\mathbb{P}(T_b(B + \mu \text{Id}) < \infty) = \lim_{\lambda \rightarrow 0} \mathbb{E}\left(e^{-\lambda T_b(B+\mu \text{Id})}\right) = e^{\mu b - \mu |b|} = \begin{cases} = 1 & \text{sgn}(b) = \text{sgn } \mu \\ < 1 & \text{otro caso} \end{cases}.$$

Una extensión de la idea anterior permite resolver el siguiente problema.

EJERCICIO 4.1. Sea B un movimiento browniano que comienza en cero y $\gamma \in \mathbb{R}$. Sea

$$T = \inf\{t \geq 0 : |B_t + \gamma t| = 1\}.$$

- (1) Pruebe que si $\gamma = 0$ entonces T y B_T son independientes.
- (2) Al utilizar el teorema de Girsanov muestre la independencia entre T y B_T cuando $\gamma \neq 0$.

Otro ejemplo de aplicación es el siguiente. Notemos que

$$\mathbb{E}\left(f\left(\sup_{s \leq t} B_s + \mu s\right)\right) = \mathbb{E}\left(f\left(\sup_{s \leq t} B_s\right) e^{-\mu B_t}\right).$$

En particular

$$\mathbb{P}\left(\sup_{s \leq t} B_s + \mu s \in dx, B_t \in dy\right) = \mathbb{P}\left(\sup_{s \leq t} B_s \in dx, B_t \in dy\right) e^{\mu y - \mu^2 t / 2}.$$

Este es un resultado no trivial puesto que se conoce explícitamente la densidad conjunta de $(B_t, \sup_{s \leq t} B_s)$:

$$\mathbb{P}\left(\sup_{s \leq t} B_s \in dy, B_t \in dx\right) = \frac{2(2y-x)}{\sqrt{2\pi t^3}} e^{-(2y-x)^2/2t} \mathbf{1}_{y>0, x \leq y}.$$

Una de las aplicaciones del teorema de Girsanov es a la técnica de remoción de deriva.

EJERCICIO 4.2. Considere la ecuación diferencial estocástica

$$(7) \quad dX_t = dB_t + b(X_t) dt \quad X_0 = x$$

donde b es medible y acotada. Suponga que bajo P , X es un movimiento browniano que comienza en x . Utilice el teorema de Girsanov para encontrar una medida de probabilidad $\tilde{\mathbb{P}}$ tal que si definimos a

$$B_t = X_t - \int_0^t b(X_s) \, ds$$

entonces $(B_t)_{t \leq 1}$ sea un movimiento browniano bajo $\tilde{\mathbb{P}}$. Note que X resuelve entonces la ecuación diferencial estocástica (??); esta solución es llamada solución por transformación de deriva.

Bibliografía

- [Dam65] K. È. Dambis, *On decomposition of continuous submartingales*, Teor. Verojatnost. i Primenen. **10** (1965), 438–448. MR 0202179 (34 #2052)
- [Doo84] J. L. Doob, *Classical potential theory and its probabilistic counterpart*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 262, Springer-Verlag, New York, 1984. MR 731258 (85k:31001)
- [DS65] Lester E. Dubins and Gideon Schwarz, *On continuous martingales*, Proc. Nat. Acad. Sci. U.S.A. **53** (1965), 913–916. MR 0178499 (31 #2756)
- [EK46] P. Erdős and M. Kac, *On certain limit theorems of the theory of probability*, Bull. Amer. Math. Soc. **52** (1946), 292–302. MR 0015705 (7,459b)
- [EK47] P. Erdős and M. Kac, *On the number of positive sums of independent random variables*, Bull. Amer. Math. Soc. **53** (1947), 1011–1020. MR 0023011 (9,292g)
- [Itô87] Kiyosi Itô, *Differential equations determining a Markoff process*, Selected papers (Daniel W. Stroock and S.R.S Varadhan, eds.), Springer-Verlag, New York, 1987, Translated from the 1942 Japanese original.
- [Kac49] M. Kac, *On distributions of certain Wiener functionals*, Trans. Amer. Math. Soc. **65** (1949), 1–13. MR 0027960 (10,383b)
- [Kni71] Frank B. Knight, *A reduction of continuous square-integrable martingales to Brownian motion*, Martingales (Rep. Meeting, Oberwolfach, 1970), Springer, Berlin, 1971, pp. 19–31. Lecture Notes in Math., Vol. 190. MR 0370741 (51 #6967)
- [KS91] Ioannis Karatzas and Steven E. Shreve, *Brownian motion and stochastic calculus*, second ed., Graduate Texts in Mathematics, vol. 113, Springer-Verlag, New York, 1991. MR 1121940 (92h:60127)
- [Lév39] Paul Lévy, *Sur certains processus stochastiques homogènes*, Compositio Math. **7** (1939), 283–339. MR 0000919 (1,150a)
- [Pro04] Philip E. Protter, *Stochastic integration and differential equations*, second ed., Applications of Mathematics (New York), vol. 21, Springer-Verlag, Berlin, 2004, Stochastic Modelling and Applied Probability. MR 2020294 (2005k:60008)
- [PS78] Sidney C. Port and Charles J. Stone, *Brownian motion and classical potential theory*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1978, Probability and Mathematical Statistics. MR 0492329 (58 #11459)
- [RY99] Daniel Revuz and Marc Yor, *Continuous martingales and Brownian motion*, third ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293, Springer-Verlag, Berlin, 1999. MR 1725357 (2000h:60050)
- [Sim05] Barry Simon, *Functional integration and quantum physics*, second ed., AMS Chelsea Publishing, Providence, RI, 2005. MR 2105995 (2005f:81003)
- [Ste01] J. Michael Steele, *Stochastic calculus and financial applications*, Applications of Mathematics (New York), vol. 45, Springer-Verlag, New York, 2001. MR 1783083 (2001i:60080)
- [WZ65] Eugene Wong and Moshe Zakai, *On the convergence of ordinary integrals to stochastic integrals*, Ann. Math. Statist. **36** (1965), 1560–1564. MR 0195142 (33 #3345)