UNIVERSIDAD DE VALLADOLID

Facultad de Ciencias

TRABAJO FIN DE GRADO

Grado en Matemáticas

APLICACIONES DEL CÁLCULO ESTOCÁSTICO EN MATEMÁTICA FINANCIERA

Autor: Alfonso Álamo Zapatero

Tutor: Eustasio del Barrio Tellado

Valladolid, Agosto del 2013

TÍTULO: Aplicaciones del cálculo estocástico en matemática financiera.

AUTOR: Alfonso Álamo Zapatero.

TUTOR: Eustasio del Barrio Tellado.

MIEMBROS DEL TRIBUNAL:

1. **PRESIDENTE:** Carlos Matrán Bea.

2. **SECRETARIO:** Luis A. Tristán Vega.

3. **VOCAL**: Luis M. Abia Llera.

FECHA DE LECTURA: Viernes 6 de Septiembre del 2013.

Agradecimientos

El autor pretende dar constancia de la inestimable ayuda recibada por su tutor, Eustasio del Barrio. También ve oportuno dedicar alguna que otra línea a sus padres, por estar siempre cuando se les necesita.

Resumen del proyecto

El principal objetivo de este trabajo es servir de introducción al cálculo estocástico y a la matemática financiera, temas de vital importancia en la actualidad. En lo referente al cálculo estocástico se abordan temas como la integración respecto del movimiento Browniano y la teoría de martingalas. En lo que atañe a la matemática financiera se estudian conceptos como el de mercado viable y la teoría de valoración de Black y Scholes. Para terminar se tratan algunas conexiones entre la matemática financiera y el transporte óptimo.

Abstract

The aim of this proyect is to act as an introduction to stochastic calculus and financial mathematics, being vitual issues nowadays. Regarding stochastic calculus, we shall assess integration with respect to Brownian motion and martingale theory. As for financial mathematics, concepts such as viable market and rational asset pricing of Black-Scholes are examined. Finally, the relationship between financial mathematics and optimal transport will be evaluated.

Palabras clave

Martingala, Movimiento Browniano, Integración estocástica, Activo financiero, Derivado financiero, Mercado viable, Ecuación de Black-Scholes, Transporte Óptimo.

Aplicaciones del cálculo estocástico en matemática financiera

Alfonso Álamo Zapatero

23 de agosto de 2013

Índice general

Parte I

Cálculo estocástico

Introducción

La primera parte del libro está destinada a presentar los pilares básicos del cálculo estocástico: la teoría de martingalas y la integral respecto del movimiento Browniano.
La teoría de martingalas fue desarrollada casi de manera íntegra por J.L. Doob (1910-2004); en el primer capítulo se exponen los conceptos claves de esta teoría, llave que permite extender el cálculo integral a funciones y/o medidas no deterministas.
Después de estudiar las desigualdades y los teoremas de convergencia de Doob se presenta una formulación matemática, como proceso estocástico, del conocido movimiento Browniano: el llamado proceso de Wiener.
Resulta que este proceso es una martingala continua con variación total no acotada; luego si se pretende extender el concepto de integral utilizando enfoques clásicos como el Stieltjes, será necesaria una revisión de los mismos.
Esto es precisamente lo que se trata en el segundo capítulo: en la línea de Stieltjes y apoyándose fuertemente en la teoría de medida, se generaliza esta perspectiva de la integración para cierta clase (lo suficientemente amplia) de procesos estocásticos, esto es, funciones no deterministas.
De todos modos, la construcción que se presenta comparte, como era de esperar, rasgos con la habitual de la integral de Lebesgue: por ejemplo, se razonará de manera jerárquica ampliando sucesivamente el dominio de la integral. Cabe destacar que, en contra de lo que sucede generalmente en la teoría de la medida abstracta, utilizaremos un marco topológico; es más, de espacio vectorial topológico: los espacios normados. Conceptos como el de isometría serán necesarios para completar la construcción de la integral.
Mención a parte merece la llamada fórmula de Îto; este resultado es al cálculo estocástico lo que la regla de la cadena al cálculo ordinario, es más, veremos que admite una versión que la generaliza.

Con todo este bagaje se puede abordar el concepto de ecuación diferencial estocástica; si bien son ecuaciones integrales, se presentan como una generalización de las ecuaciones diferenciales ordinarias.
Mención a parte merece el llamado teorema de Girsanov (al final del capítulo primero se aborda una versión lo suficientemente general del mismo). Este profundo resultado viene a describir cómo se comportan los procesos estocásticos al sustituir la probabilidad base por otra equivalente; su utilidad se hará patente en la parte de teoría económica, en el paradigma de Black-Scholes.

Capítulo 1

Martingalas y movimiento Browniano

1.1. Definiciones previas

El concepto de martingala tal y como hoy lo conocemos fue introducido por J. L. Doob a mediados del siglo XX, sin embargo, desde la época de B. Pascal (s. XVII) ya se empleaba tal terminología para referirse a ciertas estrategias en juegos de apuestas. La actual teoría de martingalas es de vital importancia para el modelado de gran cantidad de fenómenos que se presentan día a día.

A partir de ahora admitiremos que trabajamos en un espacio de probabilidad $(\Omega, \mathcal{F}, P)$. Sea $\mathcal{T}$ un subconjunto no vacío de $[0, \infty)$, generalmente $\mathcal{T}$ será el propio $[0, \infty)$, $\mathbb{Z}_{\geq 0}$ o un intervalo $[0, L]$. Tal conjunto desempeñará el papel de índice temporal. Comenzaremos introduciendo dos conceptos básicos en el cálculo estocástico: martingala y filtración.

Una filtración es una colección de σ-álgebras cuya pretensión es representar el historial de un proceso: el elemento t-ésimo albergará, hablando de manera informal y sin ningún tipo de rigor, la información hasta el instante t. Es por tanto natural exigir la condición de que formen una sucesión creciente.

Las martingalas son un tipo especial de procesos estocásticos diseñados en primera instancia para representar las ganancias de un juego de apuestas justo. No obstante han resultado ser una pieza clave en el análisis estocástico tanto por sus numerosas aplicaciones prácticas como por su papel preponderante en la justificación teórica de los pilares del propio cálculo estocástico. En palabras de M. Loève: *"Martingales, Markov dependence and stationarity are the only three dependence concepts so far isolated which are sufficiently general and sufficiently amenable to investigation, yet with a great number of deep properties"*.

Formalizando los dos conceptos anteriores:

Definición 1.1.1 (Filtración) *Por filtración entederemos cualquier sucesión expansiva de sigma álgebras* $(\mathcal{F}_t)_{t\in\mathcal{T}}$. *Además:*

- *En el caso de que* $\mathcal{T} = [0, L]$ *con* $L > 0$ *o* $\mathcal{T} = [0, \infty)$, *se dice que la filtración es continua por la derecha si*

$$\bigcap_{s>t} \mathcal{F}_s = \mathcal{F}_t$$

- *Se dice que la filtración es completa si cada conjunto de medida nula para* $\mathcal{F}$ *está en* $\mathcal{F}_0$ *(luego estará en cualquier otro* $\mathcal{F}_t$ *por ser la sucesión creciente).*

A partir de ahora, siempre que la indexación se haga sobre un intervalo de $[0, \infty)$ daremos por sentado que la filtración es completa y continua por la derecha

Definición 1.1.2 (Martingala) *Se dice que un proceso* $X = (X(t))_{t\in\mathcal{T}}$ *es una martingala respecto de la filtración* $(\mathcal{F}_t)_{t\in\mathcal{T}}$ *si es integrable, adaptado a la filtración y se tiene que para cada par de instantes de* $\mathcal{T}$ $0 \leq s < t$

$$E(X(t)\|\mathcal{F}_s) = X(s) \text{ } P\text{-a.s. (1)}$$

En caso de tenerse que $E(X(t)\|\mathcal{F}_s) \leq X(s)$ *(resp.*$E(X(t)\|\mathcal{F}_s) \geq X(s)$*), en un conjunto de probabilidad 1, se diría que* $X = (X(t))_{t\in\mathcal{T}}$ *es una supermartingala (submartingala respectivamente).*

La condición de martingala viene a representar que si $s < t$ son dos instantes de tiempo, con la información que se dispone hasta instante s, la mejor predicción posible para el valor del instante posterior $X(t)$ es $X(s)$.

Tres observaciones:

a) Es inmediato comprobar que la condición (1) es equivalente a

$$E(X(t) - X(s) \| \mathcal{F}_s) = 0$$

para todos $t > s$ índices temporales, esto es: podemos pensar que los incrementos de una martingala son en esencia ruidos aleatorios.

b) Si la indexación se hace sobre $\mathbb{Z}_{\geq 0}$, por la propiedad de torre de la esperanza condicionada la condición (1) es equivalente a que, para cada índice n se tenga que:

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

la cual resulta mucho más sencilla de comprobar. Es obvio que se tienen extensiones análogas para las condiciones pertinentes de los conceptos de sub/super-martingala.

c) Sea X una variable aleatoria integrable y una filtración $(\mathcal{F}_n)_{n\geq 1}$; de nuevo por la propiedad de torre de la esperanza condicionada se tiene que la sucesión $X_n = E(X\|\mathcal{F}_n)$ es una martingala. Este proceso recibe el nombre de ***martingala de la esperanza condicionada de*** X.

1.1.1. Integración estocástica discreta

Introducimos someramente la integración estocástica. Sólo estamos en codiciones de abordar el caso discreto (esto es, $\mathcal{T}$ es numerable: supondremos sin pérdida de generalidad que $\mathcal{T} = \{0, 1, \ldots, N\}$ donde N puede ser infinito). El siguiente capítulo se dedica por completo al caso continuo. Primeramente definiremos un nuevo tipo de procesos que a la postre resultarán ser los integrandos más habituales: los procesos predecibles.

Definición 1.1.3 (Proceso predecible) *Se dice que un proceso* $\phi = (\phi_n)_{n\geq 0}$ *es un proceso predecible respecto de una filtración* $(\mathcal{F}_n)_{n\geq 0}$ *si* ϕ_{n+1} *es* $\mathcal{F}_n$ *medible para cada* $n \in \mathcal{T}$

Informalmente, tal concepto viene a representar la idea de que los valores que puede tomar el proceso a tiempo $n + 1$ se pueden determinar a través de la historia hasta el tiempo n.

Notemos que en virtud de las propiedades de la esperanza condicionada, cualquier martingala predecible es constante casi seguro. En efecto: por ser martingala $E(X_{n+1}\|\mathcal{F}_n) = X_n$ y por ser predecible $E(X_{n+1}\|\mathcal{F}_n) = X_{n+1}$, ambas condiciones con probabilidad 1. Es entonces obvio que X_n es constante en n casi seguro.

Definición 1.1.4 (Integración estocástica discreta) *Si $u = (u_n)_{n\geq 0}$ es un proceso predecible y $M = (M_n)_{n\geq 0}$ es una martingala, ambos respecto de la misma filtración, se define la integral estocástica discreta de u respecto de M como el proceso:*

$$I_0(u, M) = 0$$

$$I_n(u, M) = u_1\Delta M_1 + \ldots + u_n\Delta M_n$$

donde $\Delta M_n = M_n - M_{n-1}$.

Veamos una sencilla propiedad que asegura que, bajo la acotación del proceso u, la integral estocástica conserva el carácter de martingala de M. Es por ello que a tal característica se la conoce con un nombre propio: la propiedad de estabilidad.

Proposición 1.1.5 (Propiedad de estabilidad) *Bajo las mismas hipótesis y notaciones que en la definición anterior, si M es una martingala (resp. sub-super martingala) y u es acotado, entonces el proceso $I(u, M)$ es una martingala (resp. sub-super martingala).*

Demostración:
Si $u = (u_n)_n$ es acotado y predecible entonces $u_{n+1}\Delta M_{n+1}$ sigue siendo integrable. Entonces, teniendo en cuenta que $\Delta I_{n+1}(u, M) = u_{n+1}\Delta M_{n+1}$, para cada $n \geq 0$ se tiene que:

$$E(\Delta I_{n+1}(u, M)\|\mathcal{F}_n) = E(u_{n+1}\Delta M_{n+1}\|\mathcal{F}_n) = u_{n+1}E(\Delta M_{n+1}\|\mathcal{F}_n) = 0$$

tal y como queríamos ver.
□

Esta propiedad proporciona una caracterización de las martingalas:

Proposición 1.1.6 *Un proceso $M = (M_n)_{n\in\mathcal{T}}$ es una martingala respecto de una filtración $(\mathcal{F}_n)_{n\geq 0}$ si y solo si para cada proceso predecible respecto de tal filtración y acotado $u = (u_n)_{n\geq 0}$ se tiene que*

$$E(I_n(u, M)) = 0 \text{ para cada } n \in \mathcal{T}$$

Demostración:
Si M es una martingala, entonces el proceso $(I_n(u, M))_n$ también lo será, luego todos sus elementos compartirán esperanza; como $I_0(u, M) = 0$, cada $I_n(u, M)$ tiene esperanza nula, que es justo lo que queríamos probar.

Supongamos ahora que $E(I_n(u, M)) = 0$ para cada $n \geq 0$ para cada proceso predecible respecto de tal filtración y acotado u.
Tomemos un elemento m de $\mathcal{T}$ y un conjunto A que sea $\mathcal{F}_m$-medible; definimos entonces el proceso v como $v_l = 1_A$ si $l = m + 1$ y $v_l = 0$ en otro caso. De esta forma, para cada $p > m$ se tiene que:

$$0 = E(I_p(u, M)) = E(1_A(M_{m+1} - M_m))$$

Como el conjunto A era arbitrario se tiene que $E(M_{m+1} - M_m)\|\mathcal{F}_m) = 0$ y esto último es válido para cada índice m.
□

1.1.2. Teoremas de Kolmogorov

Enunciaremos dos resultados debidos a Kolmogorov de suma importancia en la teoría de procesos estocásticos: el teorema de existencia y el criterio de continuidad.
El teorema de existencia (también conocido como teorema de extensión) nos permite asegurar la existencia de un proceso dada una familia de distribuciones de probabilidad en dimensión finita (que a la postre serán sus distribuciones finito dimensionales) siempre y cuando satisfagan una condición de compatibilidad.

Teorema 1.1.7 (Teorema de extensión de Kolmogorov) *Consideremos la familia de probabilidades de Borel*

$$\{P_{t_0, t_1 \ldots t_n}\}_n$$

con $0 \leq t_0 < t_1 < \ldots < t_n$ *verificando la siguiente condición de compatibilidad:*
Si $\{0 \leq s_0 < s_1 < \ldots < s_m\} \subseteq \{0 \leq t_0 < t_1 < \ldots < t_n\}$, *entonces:*

$$P_{t_0, t_1 \ldots t_n} \circ \pi^{-1} = P_{s_0, s_1 \ldots s_m}$$

donde $\pi : \mathbb{R}^n \to \mathbb{R}^m$ *es la proyección canónica asociadad a los dos conjuntos de índices. Entonces, existe un proceso* X *cuyas distribuciones finito dimensionales son precisamente las que componen la familia* $\{P_{t_0, t_1 \ldots t_n}\}_n$.

El criterio de continuidad aporta una condición suficiente para que un proceso admita una versión con trayectorias continuas. La idea estriba en ser capaces de alcanzar cierto control sobre los incrementos de las trayectorias. Es una herramienta de uso frecuente, pues no suele ser difícil obtener la desigualdad que plantea.

Teorema 1.1.8 (Criterio de continuidad de Kolmogorov) *Si el proceso estocástico $X = (X(t))_{t\geq 0}$ verifica que*

$$E(|\Delta X(t)|^p) \leq M|\Delta t|^k$$

para todo tiempo t y donde $p > 0$ y $k > 1$, entonces existe una versión del proceso con trayectorias continuas.

Veamos una aplicación conjunta de ambos teoremas:
Supongamos que $x(t)$ representa la trayectoria (aleatoria) de una cierta partícula a lo largo del tiempo t. Asumimos $x(t)$ tiene incrementos independientes y estacionarios; es más $\Delta x(t)$ sigue una ley normal centrada de dispersión Δt. Por supuesto las trayectorias son continuas y además, sin pérdida de generalidad podemos suponer que la partícula siempre parte del origen. Surge la pregunta ¿matemáticamente existe tal proceso? Se comprueba, apoyándose en los dos teoremas anteriores que la respuesta es afirmativa.

Para ello notemos que:

- Caso de existir tal proceso, por las condiciones de normalidad e independencia, se tendría que:

$$E(|\Delta x(t)|^{2k}) = \frac{(2k)!}{2^k k!}|\Delta t|^k$$

$$Cov(x(t), x(s)) = min\{s, t\}$$

- Caso de existir, tal proceso sería gaussiano, en efecto: el vector $(x(t_1),\ x(t_2), \ldots, x(t_n))$ es una transformación lineal del vector $(x(t_1), x(t_2) - x(t_1), \ldots,\ x(t_n) - x(t_{n-1}))$, que sigue una distribución normal multivariante debido a que sus componentes son normales e independientes.

- En virtud de los dos apartados anteriores se tiene que cualquier distribución finito dimensional, a tiempos $t_1, t_2, \ldots, t_n$, sigue una ley normal multivariante $N(0, \Sigma)$, donde

$$\Sigma = \begin{pmatrix} t_1 & t_1 & \cdots & t_1 \\ t_1 & t_2 & & \\ \vdots & & \ddots & \\ t_1 & & & t_n \end{pmatrix}$$

Se comprueba además que se cumplen las condiciones de compatibilidad requeridas, luego en virtud del teorema de extensión existe tal proceso.

Para la continuidad de las trayectorias basta observar que por hipótesis

$$E(|\Delta x(t)|^4) = \frac{(4)!}{2^3}(\Delta t)^2$$

luego también se verifican las condiciones del criterio de continuidad.

Tal proceso recibe el nombre de movimiento Browniano. Aunque su origen es fruto de la necesidad de explicar las trayectorias observadas de ciertas partículas (granos de pólen en suspensión en un medio líquido), el movimiento Browniano ha resultado ser una construcción matemática fundamental tanto en el análisis clásico como en el estocástico.

1.2. Tiempos de parada

La interpretación habitual de un tiempo aleatorio resulta obvia: representar el instante temporal en que se presenta un determinado evento. En esta sección se estudia a grandes rasgos un clase particular de tiempos aleatorios: los tiempos de parada. La idea central subyace en que un tiempo aleatorio no tiene por qué comportarse bien en lo que a nociones de medibilidad se refiere. Para subsanar tal dificultad técnica se introduce el concepto de tiempo de parada. Además, la nueva condición requerida para que un tiempo aleatorio sea de parada, adquiere una clara interpretación práctica en la línea que venimos siguiendo: el hecho de que el tiempo aleatorio tome valores hasta el instante t sólo depende del desarrollo del proceso hasta tal instante. Las líneas siguientes formalizarán y aclararán tales ideas.

Definición 1.2.1 (Tiempo de parada) *Se dice que una variable aleatoria τ es un tiempo de parada respecto de una filtración $(\mathcal{F}_t)_{t \in \mathcal{T}}$ si se verifica que $(\tau \leq t) \in \mathcal{F}_t$ para cada $t \in \mathcal{T}$.*

Existe un concepto íntimamente relacionado con el de tiempo de parada para indexaciones continuas: tiempo opcional. La diferencia es que en vez de exigir que $(\tau \leq t) \in \mathcal{F}_t$ para cada $t \in \mathcal{T}$ se impone que $(\tau < T) \in \mathcal{F}_t$ para cada $t \in \mathcal{T}$. Si la filtración es continua por la derecha no es difícil comprobar que ambos conceptos coinciden. Como se comentó en la sección anterior, siempre que la indexación lo permita, trabajaremos con filtraciones continuas por la derecha. Como se puede intuir, esta condición hace que la filtración presente un comportamiento «razonable» como se pone de manifiesto en la observación anterior.

Fijándonos en el caso discreto conviene resaltar que la condición $(\tau \leq n) \in \mathcal{F}_n$ para cada $n \geq 1$ es trivialmente equivalente a que $(\tau = n) \in \mathcal{F}_n$ (basta sólo recordar que las σ-álgebra conforman una sucesión expansiva).

Además, esta última condición suele ser mucho más asequible de comprobar.

Se tienen las siguientes propiedades, sobre las cuales nos apoyaremos frecuentemente:

Proposición 1.2.2 *El conjunto de tiempos de parada respecto de una filtración es cerrado para máximos, mínimos y sumas.*

Demostración:
El que es cerrado para máximos y mínimos es trivial. Veámoslo para sumas:

$(T+S > t) = (T = 0, S > t) \cup (0 < T < t, T+S > t) \cup (T > t, S = 0) \cup (T \geq t, S > 0)$

El primer, tercer y cuarto conjunto son, evidentemente, elementos de $\mathcal{F}_t$. Para el segundo conjunto notemos que podemos reescribirlo como

$$\bigcup_{0<r<t, r\in\mathbb{Q}_{\geq 0}} (t > T > r, S > t - r)$$

es esto, como una unión numerable de conjuntos de $\mathcal{F}_t$.
□

Para probar el siguiente resultado basta observar que:

$$(\sup_{n\geq 1} T_n \leq t) = \bigcap_{n\geq 1} (T_n \leq t)$$

y que

$$(\inf_{n\geq 1} T_n \geq t) = \bigcap_{n\geq 1} (T_n \geq t)$$

Proposición 1.2.3 *Si* $(\tau_n)_{n\geq 1}$ *es un sucesión de tiempos de parada, entonces tanto* $\sup_n \tau_n$ *como* $\text{ínf}_n \tau_n$ *son tiempos de parada.*

Con argumentos similares se prueban los tres siguientes resultados:

Proposición 1.2.4 *El tiempo de llegada de un proceso a un conjunto boreliano es un tiempo de parada respecto de la filtración natural.*

Proposición 1.2.5 *Si* $(X_n)_{n\in\mathcal{T}}$ *es un proceso adaptado a* $(\mathcal{F}_n)_{n\in\mathcal{T}}$ *y* τ *es un tiempo de parada, entonces* $(X_{min\{\tau,n\}})_{n\in\mathcal{T}}$ *vuelve a ser un proceso adaptado a* $(\mathcal{F}_n)_{n\in\mathcal{T}}$.

Proposición 1.2.6 *Si* $(X_n)_{n\in\mathcal{T}}$ *es una(sub-super) martingala adaptada a* $(\mathcal{F}_n)_{n\in\mathcal{T}}$*y* τ *es un tiempo de parada, entonces* $(X_{min\{\tau,n\}})_{n\in\mathcal{T}}$ *vuelve a ser una (sub-super)martingala respecto de* $(\mathcal{F}_n)_{n\in\mathcal{T}}$.

Introducimos ahora un concepto cuya pretensión es ser capaz de recoger y medir la información acumulada hasta el instante que determina un tiempo de parada: σ-álgebra de parada

Definición 1.2.7 (σ-álgebra de parada) *Si* τ *es un tiempo aleatorio respecto de una filtración* $(\mathcal{F}_t)_{t\in\mathcal{T}}$*, se define la sigma-álgebra de parada como aquella que está formada por los sucesos medibles* A *tales que*

$$A \cap \{\tau \leq t\} \in \mathcal{F}_t$$

para cada $t \in \mathcal{T}$. *La denotaremos por* $\mathcal{F}_\tau$

Se comprueba que $\mathcal{F}_\tau$ es efectivamente una σ-álgebra; es más, a partir de la definición no es difícil comprobar que si $\sigma \leq \tau$ entonces $\mathcal{F}_\tau \subseteq \mathcal{F}_\sigma$ y que $\mathcal{F}_{min\{\tau,\sigma\}} = \mathcal{F}_\tau \cap \mathcal{F}_\sigma$.

El resultado que se presenta a continuación es una útil herramienta a nivel teórico. En cierto sentido formaliza la idea intuitiva de que las martingalas, al menos en lo que a propiedades se refiere, han de comportarse correctamente con tiempos de parada. No obstante, hay que exigir cierta regularidad en tales tiempos, como por ejemplo la acotación. Por otra parte, asumiendo ciertas hipótesis sobre la familia de variables aleatorias (como por ejemplo la integrabilidad uniforme) se pueden relajar o incluso suprimir algunas restricciones impuestas sobre los tiempos aleatorios. Volveremos **en la sección 4** sobre ello.

Teorema 1.2.8 (Teorema de muestreo opcional) *Si $(X_n)_{n\in\mathcal{T}}$ es una supermartingala y τ, σ son tiempos de parada acotados, verificando que $\sigma \leq \tau$ a.s. entonces:*

$$E(X_\tau \| \mathcal{F}_\sigma) \leq X_\sigma$$

con probabilidad 1. En particular, si $(X_n)_{n\in\mathcal{T}}$ es una martingala entonces $E(X_\tau \| \mathcal{F}_\sigma) = X_\sigma$ casi seguro.

Demostración:
Como siempre, supondremos sin pérdida de generalidad que $\mathcal{T} = \{0, 1, \ldots, N\}$ para algún N, que puede ser ∞. Consideremos el proceso $u = (u_n)_n$ tal que $u_n = 1_{\{\sigma < n \leq \tau\}}$.

Como $(\sigma < n \leq \tau) = (\sigma < n) \cap (\Omega - (\tau < n))$ se tiene que la variable aleatoria u_n es $\mathcal{F}_{n-1}$ medible, por ello u es predecible y no negativo. Como τ es acotado, existirá un k de suerte que $\tau \leq k$, luego:

$$|I_n(u, X)| \leq |X_0| + |X_1| + \ldots + |X_k|$$

uniformemente en n, luego $I_n(u, X)$ es integrable. De esta forma, $I(u, X)$ es una submartingala con valor inicial nulo y tal que $I_k(u, X) = X_\tau - X_\sigma$; por ende se tendrá que:

$$0 = E(I_0(u, X)) \geq E(I_k(u, X)) = E(X_\tau - X_\sigma)$$

Consideremos ahora $A \in \mathcal{F}_\sigma$ y apliquemos la anterior ecuación a los tiempos acotados Λ y Σ, donde Λ coindice con τ en A y toma el valor k en $\Omega - A$ (de manera análoga queda definido Σ).

Razonando como antes se tiene que

$$\int_A X_\tau dP \leq \int_A X_\sigma dP$$

Como el conjunto A era arbitrario se tiene el resultado.
□

Como aplicación se obteniene el siguiente resultado, también de uso extendido:

Teorema 1.2.9 (Teorema de parada opcional) *Si $(X_n)_{n\geq 1}$ es una (sub-super) martingala y σ es un tiempo de parada acotado, entonces X^σ sigue siendo una (sub-super)martingala*

Demostración:
Probaremos el caso en que X es una supemartingala y que $\mathcal{T} = \{0, 1, \ldots, N\}$ para algún N, pudiendo ser infinito. Para cada $n \geq 1$ se tiene que

$$X_{\text{mín}\{n,\sigma\}} = X_0 + \sum_{m\leq n} u_m \Delta X_m$$

donde $u_m = 1_{\{m\leq\sigma\}}$, el cual es predecible.
De esta forma se tiene que X^σ es adaptado a la filtración y como $u_m \geq 0$ se tiene que X^σ es una super-martingala.
□

1.3. Desigualdades

Las desigualdades eran consideradas por Kolmogorov como la piedra angular del razonamiento matemático. La teoría de martingalas aporta numerosas desigualdades útiles en muchos aspectos, pues gran cantidad de familias de variables aleatorias admiten transformaciones que las convierten en (sub-super)martingalas. Presentamos aquí cinco de ellas de capital importancia en dicha teoría, obtenidas todas ellas por Doob.

Nos centraremos en primera instancia en el caso discreto, la extensión de los resultados al caso continuo no será difícil.

1.3.1. Desigualdades maximales

Las llamadas desigualdades maximales que presentamos a continuación resultan ser herramientas extremadamentes útiles, pues permiten acotar uniformemente la submartingala tanto en valor como en promedio.

Proposición 1.3.1 (Desigualdades maximales; caso discreto) *Si $(X_n)_{n\geq 1}$ es una submartingala no negativa, entonces:*

a) *Para cada $\lambda \geq 0$, $\lambda P(\max_{1\leq j\leq n} X_j > \lambda) \leq E(X_n)$, es más, para cualquier submartingala se tiene que $\lambda P(\max_{1\leq j\leq n} X_j > \lambda) \leq E(X_n^+)$*

b) *Para* $p > 1$, $E(X_n^p) \geq (\frac{p-1}{p})^p E(\max_{1\leq j\leq n} X_j)^p$

Demostración:

a) Para la primera probaremos sólo el primer caso:
Fijado un natural n, definimos el tiempo de paro

$$\tau(n) = \tau = \min\{n, \min\{1 \leq k \leq n \text{ tales que } X_k \geq \lambda\}\}$$

Esto es, τ es el primer instante hasta tiempo n en que el proceso alcanza la cota λ; si nunca la alcanza entonces $\tau = n$. Como el proceso es una submartingala y $1 \leq \tau \leq n$ se tiene que $E(X_n) \geq E(X_\tau)$.
Notemos por una parte que

$$(\max_{1\leq j\leq n} X_j \leq \lambda) \subset (X_\tau \geq \lambda)$$

y por otra que

$$(\max_{1\leq j\leq n} X_j < \lambda) \subset (\tau = n)$$

De esta forma, denotando a $\max_{1\leq j\leq n} X_j$ por $X_{(n)}$, se tiene que:

$$E(X_n) \geq E(X_\tau) = E(X_\tau 1_{(X_{(n)}\geq\lambda)}) + E(X_\tau 1_{(X_{(n)}<\lambda)}) \geq \lambda P(X_{(n)} \geq \lambda) +$$

$$+ E(X_n 1_{(X_{(n)}<\lambda)})$$

Reordenando,

$$\lambda P(X_{(n)} \geq \lambda) \leq E(X_n) - E(X_n 1_{(X_{(n)}<\lambda)}) = E(X_n 1_{(X_{(n)}\geq\lambda)})$$

y la conclusión es ahora inmediata.

b) El tener una acotación para el comportamiento de las colas es extremadamente útil; recordando que para una variable aleatoria Z no negativa, su momento de orden p se expresa mediante la siguiente integral

$$E(Z^p) = p \int_0^\infty t^{p-1} P(Z > t) dt$$

y apoyándonos en la desigualdad que proporciona el apartado anterior, se tiene que, para natural n:

$$E(X_{(n)}^p) = p\int_0^\infty t^{p-1}P(X_{(n)} > t)dt \leq p\int_0^\infty t^{p-2}E(X_n 1_{(X_{(n)}\geq t)})dt =$$

$$= \int_0^\infty t^{p-2}(\int_\Omega X_n 1_{(X_{(n)}\geq t)}dP)dt$$

En base al teorema de Fubini, la integral anterior se puede expresar como

$$\frac{p}{p-1}E(X_n X_{(n)}^{p-1})$$

Si q denota al exponente conjugado de p *(i.e. $p+q = pq$ con ambas cantidades mayores que* 1*)* entonces, en virtud de la desigualdad de Hölder, la anterior expresión es menor que

$$qE(X_n^p)^{\frac{1}{p}}E(X_{(n)}^p)^{\frac{p-1}{p}}$$

luego

$$E(X_{(n)}^p) \leq qE(X_n^p)^{\frac{1}{p}}E(X_{(n)}^p)^{\frac{p-1}{p}}$$

Observando que $q = \frac{p}{p-1}$, al despejar $E(X_n^p)$ en la expresión anterior se obtiene que

$$E(X_n^p) \geq (\frac{p-1}{p})^p E(X_{(n)}^p) = (\frac{p-1}{p})^p E(\max_{1\leq j\leq n} X_j)^p$$

tal y como queríamos ver.
□

Al hilo de lo que comentábamos al comienzo de la sección: sea $(X_n)_{n\geq 1}$ una sucesión de variables aleatorias centradas e independientes; definimos $S_n = X_1 + \ldots + X_n$. No es difícil comprobar que la sucesión S_n forma una martingala, luego en virtud de la desigualdad de Jensen S_n^2 es una submartingala. El resultado de aplicar la primera de las desigualdades anteriores a tal sucesión conduce directamente a la desigualdad maximal de Kolmogorov, hecha a medida para probar la ley fuerte de los grandes números. Es más, tal resultado resulta un corolario trivial de la aplicación de los teoremas de convergencia que se expondrán en la siguiente sección, a las denominadas martingalas inversas. No profundizaremos más en ello.

Centrémonos en el caso continuo. Citemos de forma somera el hecho de que toda martingala admite una versión con trayectorias càdlàg (se dice que una función en el sentido ordinario es càdlàg si en cada punto es continua por la derecha y tiene límites por la izquierda). Además se tiene que una submartingala admite una versión con trayectorias càdlàg si y solamente si $E(X(t))$ es una función continua de t.

Veamos que si una submartingala $X(t)$ tiene trayectorias continuas por la derecha en el intervalo $[0, L]$ (L bien podría ser ∞), entonces las desigualdades maximales se siguen verificando, por supuesto con las pertinentes adaptaciones.

Tomemos $[a, b] \subseteq [0, L]$. Sea $(r_n)_{n\geq 1}$ una numeración de los racionales de $[a, b]$ junto con los extremos $\{a\}$ y $\{b\}$ del intervalo. Definamos $F_n = \{r_m \colon m \leq n\} \cup \{a, b\}$, los cuales forman una sucesión monótona de conjuntos creciente hacia $([a, b] \cap \mathbb{Q}) \cup \{a, b\} = F$.

Sean $\varphi_n = \sup\limits_{t\in F_n} X(t)$ y $\varphi = \sup\limits_{t\in F} X(t)$; como el proceso X restringido a cada F_n es una martingala se tiene que

$$\lambda P(\varphi_n > \lambda) \leq E(X(b)^+)$$

para cada $\lambda \geq 0$. Ahora, como la sucesión de conjuntos $(\varphi_n > \lambda)$ crece monótonamente hacia $(\varphi > \lambda)$, en virtud de la continuidad secuencial de la probabilidad se tendrá que

$$P(\varphi > \lambda) = \lim_n P(\varphi_n > \lambda)$$

y por tanto

$$\lambda P(\varphi > \lambda) = \lambda P(\sup_{t\in F} X(t) > \lambda) \leq E(X(b)^+)$$

Ahora entra en danza la continuidad por la derecha de las trayectorias, pues gracias a ella se puede asegurar que

$$\sup_{t\in F} X(t) = \sup_{t\in[a,b]} X(t)$$

Un razonamiento análogo permite transladar la otra desigualdad al caso continuo. Se tiene así el siguiente resultado:

Proposición 1.3.2 (Desigualdades maximales; caso continuo) *Si* $(X(t))_{t\geq 0}$ *es una submartingala respecto de la filtración* $(\mathcal{F}_t)_{t\geq 0}$ *con trayectorias continuas por la derecha, entonces:*

- *Para cada* $\lambda \geq 0$, $\lambda P(\sup_{0\leq t\leq L} X(t) > \lambda) \leq E(X(L)^+)$
- *Si además la martingala es no negativa , para cada* $p > 1$, $E(X(L)^p) \geq (\frac{p-1}{p})^p E(\sup_{0\leq t\leq L} X(t))^p$

Quizás, la desigualdad que se presenta en un mayor número de razonamientos sea la siguiente:

Teorema 1.3.3 (Desigualdad de Doob) *Si* $(X(t))_{t\geq 0}$ *es una martingala respecto de la filtración* $(\mathcal{F}_t)_{t\geq 0}$ *con trayectorias continuas por la derechay* $p > 1$ *entonces, para cada* $\lambda \geq 0$ *y para cada* $L \geq 0$*, se tiene que :*

$$P(\sup_{0\leq t\leq L} |X(t)| \geq \lambda) \leq \frac{1}{\lambda^p} E(|X(L)|^p)$$

Notemos que esta desigualdad supone una generalización, en cierto sentido, de la desigualdad maximal de Kolmogorov.

1.3.2. Desigualdades up/down-crossing

Tomemos $a < b$. Nos interesaremos por el número de veces que una martingala supera la altura b partiendo desde debajo de a en los n primeros instantes de tiempo. A esta cantidad la llamaremos el número de upcrossing's de la martingala en $[a, b]$ hasta tiempo n y la denotaremos por $U_n[a, b]$. Es natural preguntarse por el número de veces que la martingala queda, bajo las mismas condiciones, por debajo de la altura a. A esta cantidad la llamaremos el número de downcrossing's de la martingala en $[a, b]$ hasta tiempo n y la denotaremos por $D_n[a, b]$.

Las desigualdades de este apartado viene encaminadas a presentar ciertas cotas para la estimación de tales valores.

Fijados $a < b$ y n natural, consideremos la sucesión de tiempos de parada siguientes:

$$
\begin{aligned}
\tau_1 &= \min\{k \geq 0 | X_k \leq a\} \\
\tau_2 &= \min\{k \geq \tau_1 | X_k \geq b\} \\
&\vdots \\
\tau_{2l} &= \min\{k \geq \tau_{2l-1} | X_k \geq b\} \\
\tau_{2l+1} &= \min\{k \geq \tau_{2l} | X_k \leq a\}
\end{aligned}
$$

Si alguno de los mínimos anteriores no está definido (porque el conjunto en cuestión es vacío), convenimos en que el valor de su elemento asociado en la sucesión será n.
Por otra parte, es fácil comprobar que, $U_n[a,b] = \max\{k | \tau_{2k} \leq n\}$ y que $D_n[a,b] = \max\{k | \tau_{2k+1} \leq n\}$. Enunciamos ya las desigualdades buscandas:

Proposición 1.3.4 (Desigualdades up/down-crossing) *Si $(X_n)_{n\geq 1}$ es una submartingala, entonces:*

- $E(U_n[a,b]) \leq \frac{E(X_n^+)+|a|}{b-a}$
- $E(D_n[a,b]) \leq \frac{E(X_n-a)^+}{b-a}$

Un razonamiento similar al empleado en el apartado anterior para extender las desigualdades maximales del caso discreto al continuo sirve ahora como llave para obtener las desigualdades up/down-crossing a tiempo continuo:

- $E(U_{[0,L]}[a,b]) \leq \frac{E(X^+(L))+|a|}{b-a}$
- $E(D_{[0,L]}[a,b]) \leq \frac{E(X(L)-a)^+}{b-a}$

1.4. Convergencia y representación de martingalas

Pasamos ahora a abordar teoremas de convergencia, todos ellos debido también a Doob. El primero de ellos podría interpretarse como el análogo al caso real de que toda sucesión monótona y acotada es convergente. En el segundo de ellos mostraremos lo útil que resulta ser el que una martingala sea uniformemente integrable, pues con tan solo esa suposición se puede probar un profundo teorema de convergencia de numerosas aplicaciones.

Llamemos $D[a,b]$ al límite puntual de $D_n[a,b]$, que está bien definido en virtud de la monotonía de la sucesión. El límite de una sucesión existe si y solamente si el límite superior y el inferior coinciden, los cuales resultan ser el superior y el inferior respectivamente de los valores de adherencia. Se tiene que $D[a,b] = \infty$ si y solamente existen infinitos puntos burlando las barreras de a y b, esto es, subsucesiones que mayoran y minoran tanto a a como a b, por ende los límites inferiores y superiores son distintos, luego no hay convergencia. Acabamos de probar un sencillo resultado que resulta ser extremadamente útil para obtener convergencias debido a las desigualdades up/down cross.

Proposición 1.4.1 (Criterio general de convergencia) *Sea sucesión $(X_n)_n$ de variables aleatorias, entonces:*

$$(X_n(\omega))_n \textit{es convergente si y solamente si } D[a,b](\omega) < \infty$$

para todos los reales $a < b$

Por tanto, si se encuentra una forma de conseguir que $\lim_n D_n[a,b] = D[a,b] < \infty$ uniformemente en $a < b$ con probabilidad 1, se tendrá la convergencia casi seguro de la sucesión. Tentativo es obtener una condición más fuerte, como por ejemplo la integrabilidad : $D[a,b] \in L^1(\Omega, P)$. Del teorema de la convergencia monótona de Lebesgue se tiene que $E(D[a,b]) = \lim_n E(D_n[a,b])$, y de la desigualdad down-crossing $E(D[a,b]) = \lim_n E(D_n[a,b]) \leq \lim_n \frac{E(X_n-a)^+}{b-a} \leq \lim_n \frac{\sup_n E(|X_n|+|a|)}{b-a}$ Vemos por tanto que para la integrabilidad de $D[a,b]$ bastaría la acotación uniforme en la norma de $L^1(\Omega, P)$. De esta forma hemos probado el siguiente resultado de importancia capital:

Teorema 1.4.2 (Teorema de convergencia de submartingalas de Doob) *Sea $(X_n)_n$ una submartingala uniformemente acotada en $L^1(\Omega, P)$, entonces la sucesión converge casi seguro hacia una variable aleatoria X integrable.*

La integrabilidad es consecuencia directa del lema de Fatou; en efecto:

$$E|X| = E(\liminf |X_n|) \leq \liminf E|X_n| \leq \sup_n E(|X_n|) < \infty$$

1.4.1. Integrabilidad uniforme

Examinemos ahora una interesante propiedad que puede presentar una familia de variables aletorias estrictamente más fuerte que la simple integrabilidad de cada uno de sus miembros: la integrabilidad uniforme.

Definición 1.4.3 (Integrabilidad uniforme) *Se dice que una familia $(X_i)_{i\in I}$ es uniformemente integrable si*

$$\sup_{i\in I}\{\int_{|X_i|>t} X_i\, dP\} \to_{t\to\infty} 0$$

Veamos el por qué del adjetivo uniforme:

Decir que una variable aleatoria X es integrable es tanto como decir que para todo $\epsilon > 0$ existe un $M = M(\epsilon) > 0$ tal que

$$\int_{|X|>M} X\, dP \ < \epsilon$$

Si se considera una familia integrable $(X_i)_{i\in I}$ y se fija una tolerancia ϵ, por hipótesis existirán cantidades $M = M(\epsilon, i)$ tales que

$$\int_{|X_i|>M(\epsilon,i)} X_i\, dP \ < \epsilon$$

La integrabilidad uniforme significa exactamente que la cantidad $M(\epsilon, i)$ es independiente de i.

Las siguientes observaciones son necesarias:

- No es difícil encontrar sucesiones de variable aletorias integrables pero no uniformemente integrables. Basta tomar $\Omega = [0, 1]$, como σ-álgebra la de Lebesgue y como probabilidad la medida también de Lebesgue, ambas por supuesto restringidas al intervalo en cuestión. Si definimos $X_n = n1_{(0,\frac{1}{n})}$, entonces:

  $$\int_{|X_n|>M} X_n\, dP \ = 1 \text{ si } n > M$$

  En caso de que $n \leq M$ la integral anterior es nula. Es obvio que la familia $(X_n)_n$ no puede ser uniformemente integrable.

- Si X es una variable aleatoria integrable y $(X_n)_n$ representa la martingala de la esperanza condicionada, entonces tal proceso es uniformemente integrable.
- Si $X_1, X_2, \ldots, X_n, \ldots$ es una sucesión de variables aleatorias integrables convergente en media, entonces es una familia uniformemente integrable.
- A modo de recíproco del resultado anterior se tiene que toda submartingala uniformemente integrable es convergente en media.

Veamos ahora el principal resultado de la sección: el teorema de representación de martingalas.
Lo que establece es que toda martingala uniformemente integrable se puede expresar como la martingala de la esperanza condicionada de cierta variable aleatoria integrable.

Teorema 1.4.4 (Representación de martingalas) *Sea $(X_n)_{n\geq 0}$ una martingala uniformemente integrable, con filtración natural $(\mathcal{F}_n)_{n\geq 0}$. Si X es su límite en media, entonces, para cada $n \geq 0$:*

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

Demostración:

Tomemos dos naturales $m > n$, entonces, por ser $(X_n)_{n\geq 0}$ martingala se tiene que :

$$\int_A X_m dP = \int_A X_n dP$$

para cualquier conjunto $A \in \mathcal{F}_n$, de esta forma:

$$|\int_A X_n - X dP| = |\int_A X_m - X dP| \leq \int_A |X_m - X| dP \leq \int_\Omega |X_m - X| dP$$
$$= \|X_m - X\|_{L^1(\Omega,P)}$$

y el último sumando tiende a cero cuando $m \to \infty$, luego

$$\int_A (X_n - X) dP = 0$$

para cada $A \in \mathcal{F}_n$, luego con probabilidad uno se tendrá que

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

$\square$

Los resultados anteriores muestran que el hecho de que una martingala sea una familia uniformemente integrable atañe un buen comportamiento en el marco asintótico. Por tanto, no es descabellado pensar que las hipótesis de los teoremas de parado y muestreo opcional puedan relajarse ante la condición de la integrabilidad uniforme.

La clave de los siguientes resultados es que si X es una martingala uniformemente integrable, X_∞ es su límite en media y τ es un tiempo de parada, entonces $X^\tau = E(X_\infty \| \mathcal{F}_\tau)$

A partir de este hecho no resulta difícil comprobar el siguiente resultado:

Teorema 1.4.5 *Si $(X_n)_{n\geq 0}$ es una martingala uniformemente integrable y τ, σ son tiempos de parada monótonos (i.e.: verificando $\tau \leq \sigma$ a.s.) no necesariamente acotados, entonces:*

a) $E(X_\sigma \| \mathcal{F}_\tau) = X_\tau$ ***(teorema de muestreo opcional)***

b) X^τ *es una martingala uniformemente integrable* ***(teorema de parada opcional)***

Además estamos en condiciones de extender estos hechos al caso continuo. La clave es estriba en discretizar los tiempos de parada. Las dificultades habituales que puede presentar el pertinente paso al límite se solventan con el hecho, ya mencionado, de que toda martingala admite una versión con trayectorias càdlàg. De esta forma llegamos al siguiente resultado:

Teorema 1.4.6 *Si $(X_t)_{t\geq 0}$ es una martingala càdlàg uniformemente integrable y τ, σ son tiempos de parada monótonos (i.e.: verificando $\tau \leq \sigma$ a.s.) no necesariamente acotados, entonces:*

a) $E(X_\sigma \| \mathcal{F}_\tau) = X_\tau$ ***(teorema de muestreo opcional)***

b) X^τ *es una martingala uniformemente integrable* ***(teorema de parada opcional)***

1.5. Movimiento Browniano

Existen ciertas clases partículas microscópicas que al introducirlas en un fluido como el agua no se hunden gradualmente como predice la teoría de la gravedad; experimentan cambios bruscos de dirección, sentido y celeridad continua y erráticamente. El ejemplo más popular lo consituyen células sexuales masculinas del polen de plantas. El biólogo R. Brown achacó en primera instancia dicho comportamiento a cierto movimiento vital de tales células, rectificando posteriormente al hallar un movimiento análogo en partículas ajenas a cualquier tipo de vida. En 1905, de manera independiente (y de hecho siguiendo caminos distintos) A. Einstein y M. Smoluchowski dieron una descripción satisfactoria de tal comportamiento utilizando postulados de la teoría atómica. En añadidura: tal éxito supuso el golpe de gracia de la teoría atomista a las denominadas corrientes energicistas, que rechazaban la idea de que la materia estuviese constituida por átomos. Mencionemos que allá por 1900 Louis Bachelier, en su tésis doctoral no sólo examinó el problema de la valoración de ciertos activos financieros sino que además introdujo una versión primitiva del movimiento Browniano. Sin embargo, fue Norbert Wiener quien a través de artículos publicados entre 1920 y 1923 sentó las bases del citado proceso tal y como hoy lo conocemos. En los años sucesivos a los pioneros trabajos de Lèvy, Kac, Doob, Kakutani... especialmente en la década de los años cincuenta, se empieza a tomar conciencia de la importancia del movimiento Browniano en el análisis clásico, especialmente de su conexión con las funciones armónicas, analíticas y con ciertas EDP's. De esta forma el movimiento Browniano deja de ser un mero modelo para la representanción de ciertas trayectorias de pólen y se convierte en una construcción matemática preponderante.

Para terminar una leve reflexión propiciada por el abuso de lenguaje: conviene tener en mente que una cosa es el movimiento observado en la naturaleza y otra es la descripción matemática. Por ello, en ciertos contextos al primero se le conoce como movimiento Browniano y al segundo como proceso de Wiener. Nosotros no haremos tal distinción.

Trataremos únicamente el movimiento Browniano unidimensional:

Definición 1.5.1 *Se dice que un proceso $B = (B(t))_{t\geq 0}$ es un movimiento Browniano si:*

a) $B(0) = 0$

b) *Tiene trayectorias continuas.*

c) *Tiene incrementos independientes.*

d) $B(t) - B(s)$ *sigue una ley* $N(0, t-s)$*, por ende, tiene incrementos estacionarios.*

Según vimos en la **sección 1** tal proceso existe y además comprobamos que:

- $E(|\Delta B(t)|^{2k}) = \frac{(2k)!}{2^k k!}|\Delta t|^k$
- La función de covarianza es $k(s,t) = Cov(B(t), B(s)) = min\{s,t\}$
- Es un proceso gaussiano; el vector $(B(t_1), B(t_2), \ldots, B(t_n))$ sigue una ley normal multivariante con vector nulo de medias y matriz de covarianza

$$\Sigma = \begin{pmatrix} t_1 & t_1 & \cdots & t_1 \\ t_1 & t_2 & & \\ \vdots & & \ddots & \\ t_1 & & & t_n \end{pmatrix}$$

- Aunque no lo hicimos entonces, de la fórmula del Jacobiano se deduce que la densidad de $(B(t_1), B(t_2), \ldots, B(t_n))$ es

$$f_{t_1,\ldots,t_n}(x_1, \ldots, x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} exp(\frac{-(x_i - x_{i-1})^2}{2(t_i - t_{i-1})})$$

Existe una bella construcción explícita, totalmente ajena al teorema de extensión de Kolmogorov y a su criterio de continuidad. La clave estriba en definir el proceso en mallas equiespaciadas de anchura 2^{-n} sobre el intervalo $[0,1]$. Por interpolación lineal se obtienen los restantes valores y justificando varios pormenores (apoyándose en las hipótesis y en resultados clásicos de la teoría de probabilidad como los lemas de Borel-Cantelli) se construye el proceso de manera íntegra sobre $[0,1]$.

Nota 1.5.2 *Varias observaciones:*

- *Un proceso con incrementos independientes y estacionarios recibe el nombre de proceso de Lévy. De* (3) *y* (4) *se tiene que el movimiento Browniano es un proceso de Lévy.*
- *El movimiento Browniano desempeña un papel fundamental en la teoría de procesos estocásticos; podría pensarse incluso como "la distribución normal estándar" para procesos.*

- *El razonamiento seguido por Einstein, consiste en probar mediante argumentos heurísticos, que la densidad ρ de la posición $x(t)$, asumiendo ciertas hipótesis de independencia, verifica la ecuación en derivadas parciales $\frac{\partial \rho}{\partial t} = \frac{1}{2}\frac{\partial^2 \rho}{\partial x^2}$. Esta última relación es una ecuación de difusión, cuya única solución es una campana gaussiana: la conexión con la distribución normal es ahora evidente. De todos modos, añadiremos que en física se prefiere utilizar el proceso de Wiener para modelar la velocidad antes que la posición de una partícula sometida a un movimiento Browniano.*

1.5.1. Propiedades como proceso

Tomemos dos instantes temporales $0 \leq s < t$; entonces, si $(\mathcal{F}_s)_s$ denota la filtración natural del movimiento Brownniano, se tiene que:

$$E(B(t)\|\mathcal{F}_s) = E(B(t)-B(s)+B(s)\|\mathcal{F}_s) = E(B(t)-B(s)\|\mathcal{F}_s)+E(B(s)\|\mathcal{F}_s)$$

Debido a la independecia de los incrementos, $E(B(t)-B(s)\|\mathcal{F}_s) = E(B(t)-B(s))$, luego

$$E(B(t)\|\mathcal{F}_s) = B(s)$$

Es decir, el movimiento Browniano es una martingala (continua).

La función $\rho(t,x,y) = \frac{1}{\sqrt{2\pi t}}e^{\frac{-(y-x)^2}{2t}}$ recibe el nombre de núcleo de transición del movimiento Browniano; tal función es de capital importancia como ponen de manifiesto los siguientes resultados, los cuales, si bien no son difíciles de probar, conllevan cálculos bastante tediosos:

- $P(B(t+s) \in A\|B(s) = x) = \int_A \rho(t,x,y)dy$
- La densidad de $(B(t_1), B(t_2), \ldots, B(t_n))$ se expresa de la siguiente forma en función de ρ :

$$f_{t_1,\ldots,t_n}(x_1,\ldots,x_n) = \rho(t_1,0,x_1)\rho(t_2-t_1,x_1,x_2)\ldots\rho(t_n-t_{n-1},x_{n-1},x_n)$$

- $\rho(t+s,x,y) = \int_{\mathbb{R}} \rho(t,x,u)\rho(s,u,y)du$, esto es, ρ verifica la llamada **ecuación de Chapman-Kolmogorov**.
- ρ verifica la **ecuación del calor**: $\frac{\partial \rho}{\partial t} = \frac{1}{2}\frac{\partial^2 \rho}{\partial^2 x}$

Recordemos que se dice que un proceso sobre $\mathbb{R}$, $X = (X(t))_t$, tiene la **propiedad de Markov**, si para cualesquiera tiempos $0 \leq t_1 < t_2 < \ldots <$

$t_n < t_{n+1}$, puntos $x_1, x_2, \ldots, x_n$ y para cualquier conjunto medible borel $A \subset \mathbb{R}$ se tiene que:

$$P(X(t_{n+1}) \in A \| X(t_n) = x_n, X(t_{n-1}) = x_{n-1}, \ldots, X(t_1) = x_1) =$$

$$P(X(t_{n+1}) \in A \| X(t_n) = x_n)$$

Para el movimiento Browniano, denotando $\rho(t,x,A) = \int_A \rho(t,x,y)dy$, en virtud de lo anterior, se tiene que:

$$P(B(t_{n+1}) \in A \| B(t_n) = x_n, B(t_{n-1}) = x_{n-1}, \ldots, B(t_1) = x_1) =$$

$$= \frac{f_{t_1,\ldots,t_n}(x_1, \ldots, x_n)\rho(t_{n+1} - t_n, x_n, A)}{f_{t_1,\ldots,t_n}(x_1, \ldots, x_n)}$$

$$= \rho(t_{n+1} - t_n, x_n, A) = \rho(t_{n+1} - t_n, x_n, A)\frac{\rho(t_n, 0, x_n)}{\rho(t_n, 0, x_n)} =$$

$$P(B(t_{n+1} \in A \| B(t_n) = x_n)$$

esto es, el movimiento Browniano es un proceso de Markov.

El siguiente resultado recoge varias propiedades fundamentales como proceso que ya hemos ido comprobando:

Teorema 1.5.3 *El movimiento Browniano verifica las siguientes propiedades:*

1. *Es una martingala continua.*
2. *Es un proceso de Markov.*
3. *Es un proceso de Lévy.*
4. *Es un proceso Gaussiano.*

Exponemos a continuación, de manera anecdótica, un resultado debido a Paul Lévy que caracteriza completamente al movimiento Browniano:

Teorema 1.5.4 (Caracterización de Paul Lévy) *Sea $X(t)$ un proceso que comienza en 0 y tiene trayectorias continuas, entonces el proceso es un movimiento Browniano si y solo si $X(t)$ y $X^2(t) - t$ son martingalas.*

1.5.2. Propiedades de las trayectorias

Estudiaremos ahora varias propiedades de las trayectorias, las cuales hacen que el movimiento Browniano sea tanto difícil de tratar como verdaderamente interesante.

Definición 1.5.5 *Sea $[a,b]$ un intervalo compacto de $\mathbb{R}$, f una aplicación definida en $[a,b]$. Para cada $n \geq 1$ definimos:*

$$V_n^2(f,[a,b]) = \sum_{j=1}^{n} (f(t_j) - f(t_{j-1}))^2$$

donde $t_j = a + j\frac{b-a}{n}$ Caso de existir, el límite de las sumas anteriormente definidas recibe el nombre de variación cuadrática de f en $[a,b]$.

Definición 1.5.6 *Si $X = X(t)$ es un proceso estocástico, su variación cuadrática es el proceso $[X]$, de forma que $[X]_t$ es el límite en probabilidad, caso de existir, de $V_n^2(X,[0,t])$*

Se comprueba que el movimiento Browniano tiene variación cuadrática finita:

Proposición 1.5.7 *Si B es un movimiento Browniano estándar, entonces $[B]_t = t$.*

Demostración:

Por una parte:

$$E(\sum_{i=1}^{n} (B(t_i) - B(t_{i-1}))^2) = \sum_{i=1}^{n} (t_i - t_{i-1}) = t$$

ya que los incrementos son normales centrados con varianza $t_i - t_{i-1}$
Por otra:

$$Var(\sum_{i=1}^{n} (B(t_i) - B(t_{i-1}))^2) = \sum_{i=1}^{n} Var(B(t_i) - B(t_{i-1}))^2 =$$

$$= \sum_{i=1}^{n} (t_i - t_{i-1})^2 E(Z^2-1)^2 \leq \frac{tE(Z^2-1)^2}{n}$$

donde Z representa una variable aleatoria normal $(0,1)$; la desigualdad de Chebychev nos permite concluir.
□

Definición 1.5.8 *Sea $[a,b]$ un intervalo compacto de $\mathbb{R}$, f una aplicación definida en $[a,b]$. Se define la variación de f en $[a,b]$ como*

$$V(f,[a,b]) = \sup\{\sum_{t_j \in P} |f(t_j) - f(t_{j-1})|, \text{ donde } P \text{ recorre las particiones de } [a,b]\}$$

Es fácil comprobar, apoyándose en la proposición anterior, que el movimiento Browniano tiene variación total infinita en cada compacto.

Proposición 1.5.9 *Se verifica P- casi seguro, que*

$$V(B,[a,b]) = \infty$$

Demostración:

Sin pérdida de generalidad podemos suponer que $[a,b] = [0,t]$.
De la proposición anterior se deduce que existen sucesiones de particiones del intervalo $0 = t_0^n < t_1^n < \ldots < t_n^n$ de suerte que su diámetro δ_n converge a 0 y tales que

$$\sum_{i=1}^{n} (B(t_i^n) - B(t_{i-1}^n))^2 \to t$$

cuando $n \to \infty$ con probabilidad 1. En efecto: basta tomar cualquier sucesión de particiones cuyo diamátro tienda a cero; por la proposición anterior se tiene que

$$\sum_{i=1}^{n} (B(t_i^n) - B(t_{i-1}^n))^2 \to t$$

en probabilidad, luego, tomando subsucesiones se llega a lo que hemos afirmado.
Ahora bien, se tiene que el conjunto de convergencia anterior está contenido en $(V(B,[0,t]) = \infty)$, y por tanto tiene probabilidad 1.
Caso contrario, debido a la continuidad de las trayectorias del proceso B, se tendría que:

$$\sum_{i=1}^{n} (B(t_i^n) - B(t_{i-1}^n))^2 \leq \max_{1 \leq i \leq n} |B(t_i^n) - B(t_{i-1}^n)| \sum_{i=1}^{n} |B(t_i^n) - B(t_{i-1}^n)| \leq$$

$$\leq V(B,[0,t]) \max_{1 \leq i \leq n} |B(t_i^n) - B(t_{i-1}^n)| \to 0$$

□

Con un poco más de trabajo se puede probar el siguiente resultado sorprendente:

Proposición 1.5.10 *Para cada $t_0 \geq 0$, el movimiento Browniano no es diferenciable en t_0 con probabilidad 1; es más, con probabilidad 1, no es diferenciable en ningún punto.*

1.5.3. Otras propiedades del movimiento Browniano

Consideremos el movimiento de Brown restringido a $[0,1]$ como una aplicación $B : \Omega \to \mathcal{C}([0,1])$ de un espacio de probabilidad al espacio de las funciones continuas en $[0,1]$. Si $(\varphi_n)_n$ es un sistema ortonormal completo en $L^2[0,1]$, entonces, existirán coeficientes $a_n(\omega)$ de suerte que:

$$B(\omega) = \sum_{n\geq 0} a_n(\omega)\varphi_n$$

de donde $a_n(\omega) = (B(\omega), \varphi_n)$ (el paréntesis indica el producto interior usual en $L^2[0,1]$). Consideremos el operador integral de $L^2[0,1]$ en $L^2[0,1]$ definido por el núcleo integral $k(s,t) = \text{mín}\{s,t\}$: la función de covarianza asociada al proceso de Wiener.
Apoyándonos en resultados de análisis funcional, se comprueba que tal operador es compacto y autoadjunto; en virtud del teorema de descomposición espectral existe una sucesión de autofunciones γ_n que conforman un sistema ortonormal completo en $L^2[0,1]$. No es difícil comprobar que

$$\gamma_n(t) = 2^{\frac{1}{2}} \sin(\frac{(2n+1)\pi t}{2})$$

Por otra parte, como $a_n(\omega) = (B(\omega), \varphi_n) = \int_0^1 B(\omega)(s)\varphi_n(s)\,ds$, y $B(s)$ sigue una ley normal, y, además, la integral $\int_0^1 B(s)\varphi_n(s)\,ds$ se puede aproximar por integrales de funciones simples, que serán a la postre combinaciones lineales de $B(t_i)$ para ciertos i (luego normales), los coeficientes a_n resultan ser límite de normales, y por ello variables aleatorias normales. Para caracterizarlas basta por tanto calcular su esperanza y su varianza.
Se puede probar que la utilización del teorema de Fubini en la siguiente situación está justificada, así que:

$$E(a_n) = E(\int_0^1 B(s)\varphi_n(s)\,ds) = \int_0^1 E(B(s))\varphi_n(s)\,ds = 0$$

En lo que a la varianza se refiere :

$$Var(a_n) = E(\int_0^1 B(s)\varphi_n(s)\,ds)^2 = E(\int_0^1 \int_0^1 B(t)B(s)\varphi_n(s)\varphi_n(t)\,dsdt)$$

Como antes, el uso del teorema de Fubini está justificado, luego $Var(a_n) = \int_0^1 \int_0^1 k(s,t)\varphi_n(s)\varphi_n(t)\, dsdt)$. Si tomamos $\varphi_n = \gamma_n$, entonces $Var(a_n) = \lambda_n$ donde λ_n denota al autovalor n-ésimo. Tenemos por tanto que $a_n =^L N(0,\lambda_n)$, luego $\lambda_n^{-\frac{1}{2}} a_n =^L N(0,1)$ Se comprueba que $\lambda_n = \frac{2}{(2n+1)\pi)}$, luego existen $c_1, c_2, \ldots, c_n, \ldots$ variables aleatorias independientes, normales estandar de suerte que:

$$B(t) = \sqrt{2} \sum_{n=0}^{\infty} c_n \frac{\sin(\frac{(2n+1)\pi t}{2})}{\frac{(2n+1)\pi}{2}}$$

para todo $t \in [0,1]$.
Ésta es la llamada representación de Paley-Wiener, que se encuandra en un marco más general como expone el llamado teorema de Karhunen-Loeve.

El siguiente resultado, cuya sencilla interpretación geométrica le da nombre, resulta ser útil a la hora de simplificar ciertos cálculos. Por otra parte, evidencia el carácter simétrico del movimiento Browniano.

Proposición 1.5.11 (Principio de reflexión) *Sea $B^a(t)$ un movimiento Browniano que comienza en a y sea $b \geq a$. Para cada $t \geq 0$:*

$$P(B^a(s) \geq b \text{ para algún } s \in [0,t]) = 2P(B^a(t) \geq b)$$

1.6. El Teorema de Girsanov

El teorema de Girsanov es un importante resultado en la teoría de procesos estocásticos. Ayuda a entender cómo son los cambios que experimenta un proceso cuando se sustituye la probabilidad base por otra equivalente. De capital importancia en la teoría económica debido al concepto de EMM (*equivalent martingale measure*), esto es, una medida de probabilidad equivalente a la probabilidad base bajo la cual los precios actualizados de un activo forman una martingala. Cuando se aborden temas económicos, se comprobará la vital importancia de tales medidas de martingala y de paso se hará patente que el teorema de Girsanov es una herramienta indispensable. Los trabajos de Girsanov datan de los años 60, no obstante ya se tenía resultados en esta línea de los años 40. En 1977 Lenglart extendió el resultado a clases más amplias de procesos. Nosotros trataremos una versión lo suficientemente general para nuestros propósitos en matemática financiera.

Se llama movimiento Browniano con deriva a cualquier proceso $D = D(t) = B(t) + \lambda t$ donde λ es una constante real no nula.

El teorema de Girsanov afirma que existe una probabilidad equivalente Q a la probabilidad base P bajo la cual el proceso D es un movimiento Browniano.

Antes de entrar directamente en la materia es conveniente considerar la siguiente situación, pues el planteamiento y la solución no distan mucho de la dinámica que seguiremos en el teorema de Girsanov.

Sea X una variable aleatoria normal centrada, de dispersión σ; entonces, si $m \neq 0$ la variable $Y = X+m$ deja de ser centrada: es normal de media no nula m y varianza σ. Es natural preguntarse si existirá alguna ley de probabilidad equivalente (recordemos que se dice que dos medidas de probabilidad son equivalentes si comparten todos sus conjuntos de medida nula) a la dada, bajo la cual Y sea normal centrada. Consideremos la variable

$$Z = e^{-\frac{m}{\sigma^2}X-\frac{m^2}{2\sigma^2}}$$

Tomemos ahora una medida Q tal que

$$\frac{dQ}{dP} = Z$$

donde la derivada ha de entenderse en el sentido de Radon-Nikodym. Se tiene por tanto que Q es absolutamente continua respecto de P; ahora bien, Z es estrictamente positiva, luego $P << Q$ y por tanto ambas medidas son equivalente. Resta comprobar que Y se distribuye bajo Q de forma normal pero esto es puro cálculo operativo si se utilizan las herramientas que proporciona el teorema de Radon - Nikodym:

$$Q(Y \in A) = \int_{(Y\in A)} \frac{dQ}{dP}dP = \int_{(X+m\in A)} e^{-\frac{m}{\sigma^2}X-\frac{m^2}{2\sigma^2}}dP =$$
$$= \int_{(A-m)} \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{m}{\sigma^2}x-\frac{m^2}{2\sigma^2}}e^{\frac{x^2}{2\sigma^2}}dx$$

Haciendo el cambio $y = x + m$ se tiene que la integral anterior es igual a

$$\int_A \frac{1}{\sqrt{2\pi}\sigma}e^{\frac{y^2}{2\sigma^2}}dx$$

justo lo que queríamos ver.

Abordamos ya el teorema de Girsanov:
Consideremos la función $f_\lambda(t,x) = exp(-\lambda x - \frac{\lambda^2}{2}t)$, la cual verifica que

$$\frac{\partial f_\lambda}{\partial t} + \frac{1}{2}\frac{\partial^2 f_\lambda}{\partial x^2} = 0$$

,

luego en virtud del lema de Îto, que veremos en el segundo capítulo, el proceso $Z_\lambda(t) = f_\lambda(t, B(t))$ es una martingala. Consideremos la probabilidad Q_λ cuya densidad respecto de P es $Z_\lambda(L)$. De la propiedad de martingala se comprueba que en el espacio $(\Omega, \mathcal{F}_t, P)$, la probabilidad Q tiene como densidad $Z_\lambda(t)$. Dicho esto, ya se está en condiciones de formular el teorema:

Teorema 1.6.1 (Teorema de Girsanov) *Si $D = D(t)$ es un movimiento Browniano con deriva λ en $[0, L]$, i.e $D(t) = B(t) + \lambda t$ con $t \in [0, L]$, entonces existe una probabilidad Q equivalente a P bajo la cual D es un movimiento Browniano sin deriva. Es más, se tiene que Q es la probabilidad Q_λ anteriormente definida:*

$$\frac{dQ_\lambda}{dP} = exp(-\lambda B(L) - \frac{\lambda^2}{2} L)$$

Volveremos más adelante sobre este resultado y comprobaremos su importancia en teoría económica.

Capítulo 2

Integración estocástica

2.1. Construcción de la integral

En este capítulo se aborda la formalización de una integral respecto del movimiento Browniano. La principal dificultad que se presenta es que tal proceso es de variación no acotada en cualquier intervalo compacto de la recta real, imposibilitando de pleno la perspectiva de la integral de Stieltjes. No obstante, con ciertas restricciones sobre el integrando se puede construir una teoría satisfactoria en lo que a aplicaciones se refiere. Como comentario: la integral estocástica respecto del movimiento Browniano se puede generalizar a integrales respecto de las llamadas semimartingalas, las cuales viene a resultar un instrumento mucho más flexible.

En lo que sigue tomaremos $L > 0$ y trabajaremos en el intervalo compacto $[0, L]$ de la recta real. Supondremos también que $(\mathcal{F}_t)_{t\in[0,L]}$ es una filtración respecto de la cual el movimiento de Brown es adaptado (por supuesto completa y continua por la derecha).

Definición 2.1.1 *Sean $\mathcal{P} = \{0 = t_0 \leq t_1 \leq \ldots \leq t_n = L\}$ una partición de $[0, L]$ y $H_1, H_2, \ldots, H_n$ variables aleatorias cuadrado integrables con H_i $\mathcal{F}_{t_{i-1}}$ medible. Entonces, al proceso $\phi(t)$ definido en $[0, L]$ como*

$$\phi(t) = \sum_{1\leq l\leq n} H_l I_{[t_{l-1},t_l]}(t)$$

si $t \in [0, L]$, se le llama proceso simple.

Al conjunto de los procesos simples en $[0, L]$ lo denotaremos por $\mathcal{S}[0, L]$. Es inmediato comprobar que tiene estructura de $\mathbb{R}$-espacio vectorial.
Denotaremos por $\mathcal{M}_2[0, L]$ a las martingalas de cuadrado integrable en $[0, L]$.

Este espacio tendrá un papel relevante en la teoría que sigue veces, pues como se verá, en gran cantidad de ocasiones la integral estocástica será una martingala cuadrado integrable.
Definiremos a continuación espacios de procesos que a la postre resultarán ser los dominios más generales de la integral estocástica.

Definición 2.1.2 *Llamaremos $\hat{\mathcal{H}}[0,L]$ al conjunto*

$$\{\textit{Procesos progresivamente medibles } \phi \textit{ tales que } \int_0^L \phi^2(t)\,dt \in L^1(\Omega,P)\}$$

Por otra parte, la notación $\mathcal{H}[0,L]$ representará al conjunto

$$\{\textit{Procesos progresivamente medibles } \phi \textit{ tales que } \int_0^L \phi^2(t)\,dt \textit{ es finita } P\textit{-casi seguro}\}$$

Como comentario: se tiene que el movimiento Browniano $B = B(t) \in \hat{\mathcal{H}}[0,L]$ para cualquier $L > 0$. En efecto: en virtud del teorema de Fubini

$$\int_\Omega \int_0^L B^2(t)\,dtdP = \int_0^L \int_\Omega B^2(t)\,dPdt = \int_0^L Var(B(t))\,dt = \int_0^L t\,dt = \frac{L}{2}$$

Notemos que la condición $\int_0^L x^2(t)\,dm \in L^1(\Omega,P)$ pretende asegurar cierta finitud sobre la variable aleatoria que representa tal integral; una restricción más fuerte que el simple hecho de que sea finita P casi seguro. Al hilo de lo que comentamos se tiene el siguiente hecho nada difícil de comprobar:

Proposición 2.1.3 *Se tienen las siguientes contenciones:*

$$\mathcal{S}[0,L] \subset \hat{\mathcal{H}}[0,L] \subset \mathcal{H}[0,L]$$

La demostración es muy sencilla. La segunda contención es clara; para la primera basta una simple inducción en el número de sumandos del proceso y la aplicación posterior de la desigualdad de Cauchy-Schwartz.

Observemos que tanto a $\hat{\mathcal{H}}[0,L]$ como a $\mathcal{M}_2[0,L]$ se les puede dotar de estructura de espacio normado:

1. En $\hat{\mathcal{H}}[0,L]$: $\|\phi\|^2 = E(\int_0^L \phi^2(t)\,dm)$
2. En $\mathcal{M}_2[0,L]$: $\|X\|^2 = E(X(L)^2)$

El comprobar que ambas aplicaciones son efectivamente normas no es difícil, no obstante, a título de ejemplo veamos la propiedad de anulación para la norma de $\mathcal{M}_2[0,L]$:
La condición $\|X\|^2 = 0$ es justamente que $E(X(L)^2) = 0$, lo que se traduce, en virtud de la desigualdad L^2 de Doob en que $E(\sup_{0\le t\le L} |X(t)|)^2 = 0$, esto es: X es indistinguible del proceso 0.
Sin entrar en detalles, la norma en $\mathcal{M}_2[0,L]$ proviene de un producto interno y la topología métrica que induce es completa, luego $\mathcal{M}_2[0,L]$ es un espacio de Hilbert.

A partir de ahora consideraremos siempre que dichos espacios están equipados con tales normas.

Veamos ahora un resultado topológico de vital importancia aunque aún no estemos en condiciones de apreciarla:

Teorema 2.1.4 (Densidad de S(0,L)) *El conjunto $\mathcal{S}[0,L]$ es denso en $\hat{\mathcal{H}}[0,L]$*

Demostración

Fijado $H \in \hat{\mathcal{H}}$ veamos que la sucesión $H_n(t) = n\int_{\frac{(k-1)L}{n}}^{\frac{kL}{n}} H(s)ds$ si $t \in (\frac{(k-1)L}{n}, \frac{kL}{n}]$ con $k = 1, 2\ldots, n$ y $H_n(t) = 0$ en $[1, \frac{L}{n}]$ es una sucesión de procesos simples convergente hacia H en $\hat{\mathcal{H}}$. Comprobemos primeramente que son procesos simples, para lo cual, por construcción, basta ver que:

$$\int_0^L H_n^2(t)dt = n\sum_{k=1}^{n-1}(\int_{\frac{(k-1)L}{n}}^{\frac{kL}{n}} H(s)ds)^2 \le n\sum_{k=1}^{n-1}\frac{1}{n}\int_{\frac{(k-1)L}{n}}^{\frac{kL}{n}} H^2(s)ds$$

$$\le \int_0^L H^2(s)ds$$

Admitiendo que

$$\int_0^L (H_n(t) - H(t))^2 ds \to 0$$

en probabilidad, unido a la desigualdad anterior, el teorema de la convergencia dominada permitiría probar que

$$E(\int_0^L (H_n(t) - H(t))^2 ds) \to 0$$

Veamos que efectivamente se tiene la anterior convergencia en probabilidad:
Fijado un $\varepsilon > 0$, existirá una función escalonada H_ε de suerte que

$$\|H - H_\varepsilon\| = \|H - H_\varepsilon\|_{L^2([0,L],m)} < \varepsilon$$

puesto que con probabilidad 1, $H(t)$ es una función de $L^2([0,L],m)$ y las escalonadas son densas en él.
Para una función escalonada r de $L^2([0,L],m)$ definimos la sucesión r_n como antes; así pues, se tiene que $r_n \to r$ puntualmente y en base a la acotación anterior se puede usar el teorema de la convergencia dominada, luego r_n converge a r en $L^2([0,L],m)$.
Notando que $H_n - (H_\varepsilon)_n = (H - H_\varepsilon)_n$ se obtiene que:

$$\begin{aligned}\|H_n - H\| &= \|H_n - (H_\varepsilon)_n + (H_\varepsilon)_n - H_\varepsilon + H_\varepsilon - H\| \\ &\leq 2\|H_\varepsilon - H\| + \|(H_\varepsilon)_n - H_\varepsilon\| \to 0\end{aligned}$$

con probabilidad 1.
□

Con estas ideas en mente podemos abordar ya la construcción de la integral estocástica. Comenzaremos con la integral indefinida para posteriormente definir el proceso integral asociado (análogo al caso determinista)

2.1.1. La integral indefinida

La construcción se llevará a cabo de forma jerárquica; comenzando por las funciones simples para las que los resultados son elementales, se amplía a clases cada vez más extensas.

La integral para procesos simples será análoga a la de Stieltjes para funciones escalonadas, precisando más:

Definición 2.1.5 *Si $H(t) = \sum_{1\leq l\leq n} H_l I_{[t_{l-1},t_l]}(t)$ es un proceso simple, definimos su integral respecto del movimiento Browniano, y la denotamos por $\int_0^L H(t)\,dB$, a la variable aleatoria*

$$\sum_{1\leq l\leq n} H_l(B_{t_l} - B_{t_l-1})$$

Veamos dos importantes propiedades de dicha integral: conserva las normas y es lineal sobre $\mathbb{R}$.

Proposición 2.1.6 *Si $X(t)$ y $Y(t)$ son un procesos simples y $a, b \in \mathbb{R}$ entonces:*

- $E(\int_0^L X(t)\, dB)^2 = E(\int_0^L X^2(t)\, dt)$, *es decir* $\|X\|_{\hat{H}} = \| \int_0^L X(t)\, dB(t)\|_{L^2(\Omega,P)}$
- $\int_0^L aX(t) + bY(t)\, dB = a \int_0^L X(t)\, dB + b \int_0^L Y(t)\, dB$

Por tanto, la aplicación

$$\begin{array}{rcl} I : \mathcal{S}[0, L] & \longrightarrow & L^2(P) \\ \phi & \mapsto & \int_0^L \phi(t)\, dB \end{array}$$

es una isometría $\mathbb{R}$-lineal.

Una vez definida la integral en $\mathcal{S}[0, L]$ no es difícil comprobar que se extiende de manera unívoca a $\hat{\mathcal{H}}[0, L]$. La clave de este hecho estriba en las tres propiedades siguientes: $L^2(P, \Omega)$ es completo, $\mathcal{S}[0, L]$ es denso en $\hat{\mathcal{H}}[0, L]$ y la integral definida para procesos simples es una isometría.
Sea $X = X(t) \in \hat{\mathcal{H}}[0, L]$ un proceso, definamos su integral Browniana de la manera siguiente:
Por densidad existe una sucesión de procesos simples $(S_n)_n$ convergentes a X, luego $(S_n)_n$ es de Cauchy para la norma de $\hat{\mathcal{H}}[0, L]$. Denotando por $I(S_n)$ la integral Browniana del proceso S_n se tendrá, en virtud de la propiedad isométrica de la integral, que $(I(S_n))_n$ es de Cauchy en $L^2(P, \Omega)$, luego convergente. Definimos la integral estocástica de X como la variable límite de la anterior sucesión.
Veamos que tal definición es consistente: si $(R_n)_n$ y $(Q_n)_n$ son sucesiones de procesos simples convergentes hacia X también lo hará la sucesión $(S_n)_n$ definida por $S_{2n} = R_n$ y $S_{2n+1} = Q_n$, luego $(I(S_n))_n$ será convergente en $L^2(P, \Omega)$,y por ende cualquier subsucesión suya es convergente compartiendo todas el mismo límite. Basta observar que $(I(R_n))_n$ e $(I(Q_n))_n$ son subsucesiones suyas. En la misma línea del razonamiento anterior se comprueba que la extensión de la integral a $\hat{\mathcal{H}}[0, L]$ es también una isometría.

Se tiene por tanto el siguiente resultado:

Proposición 2.1.7 *La aplicación*

$$\begin{array}{rcl} I : \hat{\mathcal{H}}[0, L] & \longrightarrow & L^2(P) \\ \phi & \mapsto & \int_0^L \phi(t)\, dB \end{array}$$

es una isometría $\mathbb{R}$-lineal.

2.1.2. La integral definida

Construida la integral en todo $[0, L]$ resulta natural proceder en analogía con el caso clásico e introducir la noción integral definida:

Definición 2.1.8 (Integral definida) *Dado un proceso $X \in \hat{\mathcal{H}}[0, L]$, se dice que el proceso F definido de forma puntual como*

$$F(t) = \int_0^t X(s)\, dB(s)$$

sobre $[0, L]$ es la integral definida de X. También recibe el nombre de proceso integral asociado a X.

Conviene entender la aplicación que asocia a cada proceso de $\hat{\mathcal{H}}[0, L]$ su integral respecto del movimiento Browniano como una regularización en muchos sentidos (al igual que ocurre en el caso determinista). Precisando más, el proceso integral asociado resulta ser una martingala, y además, a tenor de los resultados de la sección anterior, de cuadrado integrable al igual que el movimiento Browniano. Es más, siguiendo esta línea de similitudes se comprueba de manera inmediata que F es un proceso centrado y que admite una versión con trayectorias continuas (probar esto último requiere algo más trabajo). Pero ahí no terminan las semejanzas, a pesar de que las trayectorias sean continuas, son irregulares y tienen variación cuadrática finita.

Formalizamos los hechos anteriores en las tres siguientes proposiciones:

Proposición 2.1.9 *Se verifican las siguientes propiedades:*

1. *La aplicación*
$$\begin{array}{rcl} I : \hat{\mathcal{H}}\,[0, L] & \longrightarrow & M_2[0, L] \\ u & \mapsto & I(u) \end{array}$$
con $I(u)(t) = \int_0^t u\, dB$ es una isometría $\mathbb{R}$-lineal

2. *Si $F(t) = \{I(u)(t)\}_{0 \leq t \leq L}$, entonces $F(t)$ es un proceso centrado ; i.e. $E(F(t)) = 0$ para todo t de $[0, L]$.*

3. $P(\sup_{0 \leq t \leq L} |F(t)| \geq \lambda) \leq \frac{1}{\lambda^2} E(\int_0^L u^2(t)\, dt)$

La prueba de (1) no es difícil: primero se prueba el resultado para procesos simples y luego se extiende a $\hat{\mathcal{H}}[0, L]$ utilizando las propiedades de densidad, isometría y completitud de $L^2(P)$. La parte (2) es inmediata y (3) es consecuencia de la desigualdad maximal de Doob.

Además, de la propiedad de isometría enunciada en la proposición anterior, se deduce que $Var(F(t)) = E((F(t)^2) = \int_0^t u^2(s)\, ds$

Las otras dos proposiciones que mencionábamos son las que siguen:

Proposición 2.1.10 *Si u es un proceso de $\hat{\mathcal{H}}[0, L]$ y $F(t)$ representa su integral Browniana, entonces existe un proceso G, con trayectorias continuas, tal que para todo t de [0,L] $F(t) = G(t)$ P-casi seguro. Es decir, G es una versión de F.*

Demostración

Consideremos una sucesión de procesos simples $(u_n)_{n\geq 1}$ convergente a u en $\hat{\mathcal{H}}[0, L]$, es decir:

$$\lim_n E(\int_0^L |u(t) - u_n(t)|^2 dt) \to 0$$

Llamando $I_n(t) = \int_0^t u_n(s)dB(s)$, se tiene, por lo anteriormente visto que $I_n - I_m$ es una martingala.
En virtud de la desigualdad de Doob:

$$P(\sup_{0\leq t\leq L} |I_n(t) - I_m(t)| > \lambda) \leq \frac{1}{\lambda^2} E(|I_n(t) - I_m(t)|^2) =$$

$$= \frac{1}{\lambda^2} E(\int_0^L |u_n(t) - u_m(t)|^2 dt)$$

Ahora bien, por ser la sucesión $(u_n)_{n\geq 1}$ convergente en $\hat{\mathcal{H}}[0, L]$ será de Cauchy en $\hat{\mathcal{H}}[0, L]$, luego

$$E(\int_0^L |u_n(t) - u_m(t)|^2 dt) \to 0$$

cuando $n, m \to \infty$ y por tanto podemos elegir una subsucesión $(u_{n_k})_{k\geq 1}$ de suerte que

$$P(\sup_{0\leq t\leq L} |I_{n_{k+1}}(t) - I_{n_k}(t)| > 2^{-k}) \leq 2^{-k}$$

De esta forma, llamando $A_k = (\sup_{0\leq t\leq L} |I_{n_{k+1}}(t) - I_{n_k}(t)| > 2^{-k})$, se tiene que

$$\sum_{k\geq 1} P(A_k) < \infty$$

luego, en virtud de los lemas de Borel-Cantelli,

$$P(\limsup_k A_k) = 0$$

Todo lo anterior nos sirve para afirmar la existencia de un subconjunto $A \subset \Omega$ de probabilidad nula, tal que para todo $\omega \in \Omega - A$ se pueden encontrar índices $k_1(\omega)$ de suerte que para cada $k \geq k_1(\omega)$:

$$\sup_{0 \leq t \leq L} |I_{n_{k+1}}(t,\omega) - I_{n_k}(t,\omega)| \leq 2^{-k}$$

Esto es, si $\omega \notin A$, la sucesión $I_{n_k}(t,\omega)$ es uniformemente convergente en $[0, L]$. Además, por ser la integral Browniana de un proceso simple una combinación lineal de movimientos Brownianos tendrá trayectorias continuas, luego la sucesión anterior converge uniformemente en $\mathcal{C}[0, L]$, y por ende, su límite es una función continua, digamos $G(t, \omega)$.
Por otra parte, como $I_n(t)$ converge a $\int_0^t u(s)dB(s)$ en media cuadrática, también lo hará la subsucesión $I_{n_k}(t)$, luego con probabilidad 1

$$G(t,\omega) = \int_0^t u(s)dB(s)$$

para cada $t \in [0, L]$.

$\square$

Proposición 2.1.11 *Si u es un proceso de $\hat{\mathcal{H}}[0, L]$ y $F(t)$ representa su integral Browniana, entonces, para cualquier t comprendido entre 0 y L :*

$$V_n^2(F, [0,t]) \to \int_0^t u^2(s)\, ds$$

donde la convergencia ha de entenderse en $L^1(P)$. Por tanto:

$$[F]_t = \int_0^t u^2(s)\, ds$$

La condición de trabajar en $[0, L]$ resulta ciertamente restrictiva; no es difícil extender la integral a intervalos compactos genéricos de $\mathbb{R}$:

Definición 2.1.12 *Si $[a, b]$ es un intervalo compacto de $\mathbb{R}$ se define la integral de un proceso $u \in \hat{\mathcal{H}}[a, b]$ respecto del movimiento Browniano como sigue:*

$$\int_a^b u(t)\, dB(t) := \int_0^b u(t)\, dB(t) - \int_0^a u(t)\, dB(t)$$

Notemos que por construcción se verifica la identidad de Chasles.

Presentamos ahora un resultado debido de H. Îto que establece, en esencia, que la integral estocástica anteriormente definida es en realidad una biyección entre $\hat{\mathcal{H}}[0, L]$ y el subconjunto de $\mathcal{M}_2[0, L]$ formado por los procesos centrados. De hecho el que el proceso sea o no centrado será irrelevante en cierto sentido; precisando más:

Teorema 2.1.13 (Teorema de representación de Îto) *Para cada martingala $M = M(t)$ de $\mathcal{M}_2[0, L]$ existe un único proceso $u = u(t)$ perteneciente a $\hat{\mathcal{H}}[0, L]$ de suerte que:*

$$M(t) = \int_0^t u(s)\, dB + E(M(t))$$

para todo $t \in [0, L]$

Más tarde (en la parte de matemática financiera) se comproborá la importancia capital de este resultado.

Una vez construida la integral para procesos u tales que $\int_0^L u^2(t)dt$ sea una variable aleatoria integrable, es natural preguntarse si se pueden seguir debilitando restricciones acerca del integrando. La respuesta es parcialmente afirmativa. Esto es, bien se puede relajar la condición anterior a que $\int_0^L u^2(t)dt$ sea finita casi seguro pero se perderá el carácter de martingala: se tendrá una semimartingala. También se difuminarán otras propiedades como puede ser el carácter isométrico. Para llevar a cabo esta extensión se sigue un argumento de localización; el concepto de tiempo parada cobra ahora una importancia fundamental, pues si partimos el intervalo mediante una sucesión de tiempos de parada apropiados, se mantendrán las propiedades anteriormente vistas siempre que nos centremos en la partición que determinan.

El concepto de martingala local fue introducido también por H. Îto para englobar un nuevo conjunto de procesos que aún no siendo martingalas, presentaban un compartamiento "razonable" a la hora de integrar.

Definición 2.1.14 (Martingala local) *Se dice que un proceso $\phi(t)$ es una martingala local respecto de la filtración $(\mathcal{F}_t)_t$ si existe una sucesión creciente de tiempos de parada $(\tau_n)_n$ convergente hacia L, tales que el proceso detenido en cada τ_n, ϕ^{τ_n}, es una martingala respecto de $(\mathcal{F}_t)_{t\in[0,L]}$. Denotaremos por $\mathcal{M}_{2,loc}[0, L]$ al conjunto de las martingalas locales en $[0, L]$.*

Tomemos un proceso $H \in \mathcal{H}$ y definimos la siguiente sucesión creciente de tiempos de parada:

$$\tau_n = \text{ínf}\{0 \leq s \leq L \text{ tal que } \int_0^s H^2(s)ds = n\}$$

siempre que el conjunto anterior sea no vacío, en caso contrario tomaremos $\tau_n = L$.
Definimos ahora la siguiente sucesión de procesos de $\mathcal{H}$:

$$H_n(s) = H(s)1_{\{s\leq\tau_n\}}$$

Teniendo en cuenta que para cada $m > n$ $H_n(s) = H_m(s)1_{\{s\leq\tau_n\}}$ y admitiendo que $\int_0^\tau H(s)dB(s) = \int_0^L H(s)1_{\{s\leq\tau\}}dB(s)$ si τ un tiempo de parada, se llega a que:

$$\int_0^t H_n(s)dB(s) = \int_0^t H_{n+1}(s)1_{\{s\leq\tau_n\}}dB(s) = \int_0^{\text{mín}\{t,\tau_n\}} H_{n+1}(s)dB(s)$$

Notemos que las integrales anteriores están bien definidas para cada $t \in [0, L]$, pues

$$(\int_0^t H_n^2(s)ds \leq n) = \Omega$$

luego $\int_0^t H_n^2(s)ds \in L^\infty(P) \subset L^1(P)$
Ahora definimos la siguiente sucesión de subconjuntos de Ω

$$A_n = (\int_0^L H^2(s)ds < n)$$

Por hipótesis se tiene que

$$A = \bigcup_n A_n = (\int_0^L H^2(s)ds < \infty)$$

es un conjunto de probabilidad uno, y por construcción que

$$A_n \subset (\int_0^t H_n(s)dB(s) = \int_0^t H_m(s)dB(s))$$

Ya estamos en condiciones de extender la integral respecto del movimiento Browniano al proceso $H \in \mathcal{H}$, a la cual denotaremos por $\hat{I}(H)$:
Fijado $t \in [0, L]$, para cada $\omega \in A$ definimos

$$\hat{I}(H)_t = I(H_n)_t$$

supuesto que $\omega \in A_n$. Fuera del conjunto A lo definimos como queramos (notemos que $\Omega - A$ es nulo para P). Por todo lo anterior se tiene que la defínción es consistente, que este nuevo proceso tiene trayectorias continuas con probabilidad 1 y que extiende a la integral en $\hat{\mathcal{H}}$. Además se puede comprobar que

$$\hat{I}(H)_{\text{mín}\{t,\tau_n\}} = I(H_n)_t$$

esto es, $\hat{I}(H)$ es una martingala local. Es más, se puede probar el siguiente teorema:

Teorema 2.1.15 *Existe una única extesión lineal de I (la anteriormente construida)*

$$\begin{array}{rcl} \hat{I} : H[0,L] & \longrightarrow & M_{2,loc}[0,L] \\ u & \mapsto & \hat{I}(u) \end{array}$$

verificando

$$\sup_{0 \le t \le L} |\hat{I}(H_n)_t| \to 0$$

en probabilidad si $H_n \in \mathcal{H}[0,L]$ verifican $\int_0^L H_n^2(t)\,dt \to 0$ en probabilidad.

Demostración:

Admitiendo la existencia la unicidad es clara, pues si tomamos un proceso $H \in \mathcal{H}$, considerando la sucesión de procesos H_n construida anteriormente, se tiene que

$$\int_0^T (H_n(s) - H(s))^2 ds \to 0$$

en probabilidad, luego cualquier otra extensión de la integral J verificando que

$$\sup_{0 \le t \le L} |J(H_n)_t| \to 0$$

en probabilidad coincidirá con ella.

Veamos ahora que la extensión que hemos construido antes verifica la propiedad de convergencia del teorema:

Tomemos $\varepsilon > 0$ y N natural, entonces:

$$P(\sup_{0\le t\le L} |\hat{I}(H_n)_t| > \varepsilon) \le$$

$$\le P(\sup_{0\le t\le L} |\hat{I}(H_n)_t| > \varepsilon, \int_0^L H_n^2(s)ds < \frac{1}{N}) + P(\int_0^L H_n^2(s)ds \ge \frac{1}{N})$$

Se puede probar que $P(\sup_{0\le t\le L} |\hat{I}(H_n)_t| \ge \varepsilon, \int_0^L H_n^2(s)ds < \frac{1}{N}) = O(\frac{1}{N})$; la idea estriba en emular el razonamiento empleado en la extensión de la integral, de manera esquemática:
Se define τ_N como el primer instante en que la integral $\int_0^s H_n^2(s)ds$ es mayor o igual que N, considerando el proceso $G(s) = H_n(s)1_{\{s\le\tau_N\}}$, se comprueba que $\hat{I}(H_n)_t = I(G)_t$ siempre que $\omega \in (\int_0^L H_n^2(s)ds < \frac{1}{N})$, luego

$$P(\sup_{0\le t\le L} |\hat{I}(H_n)_t| > \varepsilon, \int_0^L H_n^2(s)ds < \frac{1}{N}) \le P(\sup_{0\le t\le L} |I(G)_t| \ge \varepsilon)$$

Debido, primeramente a la desigualdad de Markov, y posteriormente a la desigualdad cuadrática de Doob, se llega a que

$$P(\sup_{0\le t\le L} |I(G)_t| \ge \varepsilon) \le \frac{1}{\varepsilon^2} E(\sup_{0\le t\le L} |I(G)_t|^2) \le \frac{4}{N\varepsilon^2} = O(\frac{1}{N})$$

Una vez visto esto la conclusión es trivial, pues:

$$P(\sup_{0\le t\le L} |\hat{I}(H_n)_t| \ge \varepsilon) \le P(\int_0^L H_n^2(s)ds \ge \frac{1}{N}) + O(\frac{1}{N})$$

□

2.1.3. La fórmula de Îto

La fórmula que vamos a presentar a continuación es una de las herramientas más útiles del análisis estocástico. Suele presentarse como el análogo para procesos de la regla de la cadena para funciones diferenciables (deterministas).

Teorema 2.1.16 (Lema de Îto) *Si $f \in \mathcal{C}^2[0, L]$, entonces:*

$$f(B(t)) - f(0) = \int_0^t f^{'}(B(s))\, dB + \frac{1}{2}\int_0^t f^{''}(B(s))\, ds \qquad (2.1)$$

Nota 2.1.17 *Las siguientes observaciones son necesarias:*

1. *La prueba se basa en el terorema de Taylor, en efecto: como* $f(B(t)) - f(0) = \sum_{i=1}^{n} f(B(t_i)) - f(B(t_{i-1}))$ *y* $f(B(t_i)) - f(B(t_{i-1})) = f^{'}(B(t_{i-1}))\Delta t_i + \frac{1}{2} f^{''}(B(\tau_i))\Delta t_i^2$ *donde* τ_i *es un punto entre* t_{i-1} *y* t_i*, basta probar que*

$$\sum_{i=1}^{n} f^{'}(B(t_{i-1}))\Delta t_i \to \int_0^t f^{'}(B(s))dB$$

y que

$$\sum_{i=1}^{n} f^{''}(B(\tau_i))\Delta t_i^2 \to \int_0^t f^{''}(B(s))\, ds$$

ambas convergencias entendidas en probabilidad.

2. *A partir de (1) es inmediato comprobar que el proceso definido puntualmente por* $B(t)^2 - t$ *es una martingala pues, según del lema de Îto aplicado al polinomio* $f(x) = x^2$ *se tiene que*

$$B(t)^2 - t = 2\int_0^t B(s)\, dB$$

siendo el miembro de la derecha una martingala por las propiedades de la integral estocástica.

3. *Frecuentemente la fórmula de Îto se expresa en la llamada notación diferencial:*

$$df(B(t)) = f^{'}(B(t))dB(t) + \frac{1}{2} f^{''}(B(t))dt$$

Recalquemos que la expresión anterior es sólo una notación efectiva cuya única pretensión es expresar que el proceso $f(B(t))$ *verifica (1). Por otra parte, dicha notación conlleva ventajas mnemotécnicas y permite una interpretacaión heurística del lema de Îto. Un incremento «infinitesimal» de* $f(B(t))$ *se traduce en otro incremento «infinitesimal» de* $f^{'}(B(t))$ *ponderado por la medida aleatoria a la que da lugar el movimiento Browniano más un incremento «infinitesimal» de* $\frac{f^{''}(B(t))}{2}$ *ponderado por la medida de Lebesgue.*

Observamos que sea cual sea $f \in \mathcal{C}^2[0, L]$ entonces $f(B(t))$ es la suma de una integral estocástica más una integral ordinaria. Esto nos lleva a considerar una nueva clase de procesos, los procesos de Îto, que serán en esencia la suma de una integral estocástica y de una integral ordinaria.

Definición 2.1.18 (Procesos de Îto) *Se dice que un proceso $X(t)$ es un proceso de Îto si existe una representación de la forma*

$$X(t) = X(0) + \int_0^t H(s)\, dB + \int_0^t K(s)\, ds$$

donde $H \in \mathcal{H}[0, L]$ y $K(\omega) \in L^1([0, L], m)$ $P - a.s.$
Equivalentemente, en notación diferencial:

$$dX(t) = H(t)dB(t) + K(t)dt$$

Se tiene los siguientes resultados fundamentales:

- La representación anterior es única salvo indistinguibilidad; precisando más, si

  $$dX(t) = H(t)dB(t) + K(t)dt$$

  y

  $$dX(t) = A(t)dB(t) + B(t)dt$$

  entonces $A = H$ y $K = B$ casi seguro para $m \otimes P$.

  De aquí se deduce que si X es una martingala entonces $K = 0$.

- Se tiene que $d[X]_t = H(t)^2 dt$ donde $[X]$ representa la variación cuadrática del proceso X.

Resulta que el concepto de proceso de Îto proporciona el marco adecuado para enteder la llamada fórmula de Îto; el siguiente resultado aclara esta afirmación:

Teorema 2.1.19 (Fórmula de Îto, primera versión) *Si f es una función de clase $\mathcal{C}^2$ y X es un proceso de Îto, entonces $f(X)$ es de nuevo un proceso de Îto; es más, si X tiene la representación;*

$$dX_t = K_t dt + H_t dB(t)$$

entonces

$$df(X_t) = f^{'}(X_t)dX_t + \frac{1}{2}f^{''}(X_t)d[X]_t$$

$$= (f^{'}(X_t)K(t) + \frac{1}{2}f^{''}(X_t)H^2(t))dt + f^{'}(X_t)H(t)dB(t)$$

Las siguientes observaciones son necesarias:

- Si g es una función ordinaria suficientemente regular, entonces es claro que podemos interpretarla como un proceso de Îto, para el cual, siguiendo la notación anterior, el proceso H es idénticamente nulo, luego $[f]_t = 0$. De esta forma, la fórmula de Îto coincide con la regla de la cadena convencional:

$$df(g(t)) = f^{'}(g(s))dg(s)$$

El término $\frac{1}{2}f^{''}(X_t)d[X]_t$, el cual no es necesariamente nulo, es conocido como "correción de Îto".

- Notemos que si f es una función afín entonces la fórmula de Îto se reduce a la regla de la cadena (determinista). Es más, aunque no hemos hablado del movimiento Browniano multidimensional, existe una fórmula de Îto en $\mathbb{R}^n$, que se reduce a las reglas oportunas del análisis clásico si la función f es armónica.

Teorema 2.1.20 (Fórmula de Îto, segunda versión) *Sea $X(t)$ un proceso de Îto y f una función definida en un abierto de $\mathbb{R}^2$, de variables x y t, dos veces derivable respecto de x y una vez respecto de t. Entonces, el proceso $Y(t) = f(t, X(t))$ es un proceso de Îto con la representación*

$$Y(t) = f(0, X_0) + \int_0^t \frac{\partial f}{\partial s}(s, X(s))\, ds + \int_0^t \frac{\partial f}{\partial x}(s, X(s))u(s)\, dB$$

$$+ \int_0^t \frac{\partial f}{\partial x}(s, X(s))v(s)\, ds + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s, X(s))u^2(s)\, ds$$

Escrito en notación diferencial:

$$df(t, X(t)) = \frac{\partial f}{\partial t}(s, X(s))dt + \frac{\partial f}{\partial x}(s, X(s))dX(t)$$

$$+\frac{1}{2}\frac{\partial^2 f}{\partial x^2}(s, X(s))d[X]_t$$

Notemos que si $f(t, x)$ es cualquier función definida en un subconjunto apropiado de $\mathbb{R}^2$, dos veces derivable en x y una vez en t, verificando que

$$\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2} = 0$$

entonces el proceso $f(t, B(t))$ es una martingala respecto de la filtración natural del movimiento Browniano. Observemos que $x^2 - t$, $\exp(ax - \frac{a^2 t}{2})$ verifican la EDP anterior. Esta observación será muy interesante en un futuro.

Como comentario: en este capítulo se ha construido la llamada integral de Îto, no obstante hay otras construcciones para la integral estocástica distintas, como por ejemplo la de Stratonovich. El por qué de esta elección reside en que la construcción de Îto, entendida como un paso al límite en sumas análogas a las de Riemann, pondera el valor del proceso en el extremo izquierdo, lo que puede interpretarse como un tratamiento del proceso de manera predecible. La integral de Stratonovich, por ejemplo, considera el promedio de las evaluciones en ambos extremos. En matemática financiera asumir que conocemos cierta información futura de los procesos involucrados (por ejemplo, admitir que se conoce el comportamiento de los precios de un activo en el futuro) es, generalmente, descabellado. La integral de Stratonovich presenta notables diferencias con la de Îto como el hecho de que la integral definida de un proceso de $\hat{\mathcal{H}}$ deja de ser una martingala; sin embargo presenta varias ventajas operacionales (consecuencias todas ellas de su propia construcción) que aquí no entraremos a valorar.

2.2. Ecuaciones diferenciales estocásticas

Por EDO (ecuación diferencial ordinaria) se suele entender al par de condiciones bien definidas:

$$[3]\begin{cases} x^{'}(t) = f(t, x(t)) & t_0 \leq t \leq t_1 \\ x(t_0) = x_0 \end{cases}$$

La incógnita es la función $x(t)$ definida en $[t_0, t_1]$, solución de [3]. También es frecuente encontrar la condición [3] formulada en notación diferencial:

$$[4]\begin{cases} dx(t) = f(t, x(t))dt & t_0 \leq t \leq t_1 \\ x(t_0) = x_0 \end{cases}$$

Conveniremos en que una función es solución de [4] si y solamente si lo es de [3]. El no tener definida propiamente una derivada (a priori) para procesos estocástico nos lleva a utilizar la notación diferencial.

Definición 2.2.1 *Una ecuación diferencial estocástica (EDE) es un par de condiciones:*

$$[5]\begin{cases} dX(t) = f(t, X(t))dt + \sigma(t, X(t))dB(t) & t_0 \leq t \leq t_1 \\ X(t_0) = X_0 \end{cases}$$

Diremos que un proceso $X(t)$ definido en $[t_0, t_1]$ es solución de [5] *si verifica que*

$$X(t) = X(t_0) + \int_{t_0}^{t} f(s, X(s))\, ds + \int_{t_0}^{t} \sigma(s, X(s))\, dB(s)$$

A las funciones f y σ se las denomida deriva y volatilidad respectivamente. Un proceso que sea solución de [5] *recibe el nombre de proceso de difusión.*

Convenimos por tanto que las ecuaciones diferenciales estocásticas han de interprestarse como ecuaciones integrales. No obstante, la notación diferencial aporta, como ya se dijo, ventajas tanto mnemotécnicas como heurísticas. Esto es, podemos pensar en que el proceso solución $X(t)$ es un función (generalmente no determinista) que verifica que un incremento infinitesimal de su valor es equivalente a un incremento infinitesimal de la deriva ponderado por la medida de Lebesgue (medida determinista) más un incremento de la volatilidad ponderado por una medida aleatoria (lo que podemos interpretar como un pulso aleatorio, de vital importancia en el modelado de fenónemos físicos).

Es natural preguntarse si cualquier ecuación del tipo de [5] posee solución y si existe, por su unicidad. Si la volatilidad es nula y la condción inicial es constante, el problema se convierte en una ecuación del tipo [4], para las cuales sabemos que no siempre hay solución, y que caso de existir no tiene por qué ser única. A continuación se presenta un resultado al uso que proporciona condiciones suficientes para la existencia y unicidad de la solución del problema. El lector familiarizado con nociones de EDO's encontrará en él la extensión natural del resultado clásico de existencia y unicidad en tal contexto.

Teorema 2.2.2 (Existencia y unicidad) *Consideremos el problema* [5] *en el intervalo $[0, L]$ y supongamos que se verifica que:*

$$\begin{aligned} \|f(t,x) - f(t,y)\| &\le M_1 \|x - y\| \\ \|\sigma(t,x) - \sigma(t,y)\| &\le M_2 \|x - y\| \\ \|f(t,x)\| &\le C_1 (1 + \|x\|) \\ \|\sigma(t,x)\| &\le C_2 (1 + \|x\|) \end{aligned}$$

para todos x, y reales y $t \in [0, L]$ y que además X_0 es una variable de cuadrado integrable, entonces [5] *tiene solución y además: es única salvo indistinguibilidad, tiene trayectorias continuas y es uniformemente acotada en $L^2(P)$*

La prueba es semejante al caso determinista: se define un funcional auxiliar en cierto espacio de procesos, y se comprueba utilizando las condiciones impuestas que es una contracción. Su único punto fijo resulta ser por construcción la solución del problema.

Consideremos los siguientes ejemplos:

Movimiento Browniano geométrico

El movimiento Browniano geométrico (o exponencial) es el proceso

$$X(t) = x_0 exp((\mu - \frac{1}{2}\sigma^2)t + \sigma B(t))$$

donde σ y μ son constante positivas.

Este proceso desempeña un papel relevante en el modelo de Black-Scholes sobre la valoración de opciones europeas. Tal planteamiento asume que el crecimiento del activo en el instante t digamos $X(t)$ (aunque es frecuente denotarlo por $S(t)$ en virtud de la nomeclatura inglesa: spot price), es proporcional al propio activo en dicho instante ponderado por una combinación lineal de coeficientes constantes de la medida de Lebesgue y de la medida aleatoria que genera el movimiento Browniano.

Precisando más:

$$dX(t) = X(t)(\mu dt + \sigma dB(t))$$

Se tiene por tanto la siguiente EDE:

$$[a] \begin{cases} dX(t) = \mu X(t)dt + \sigma dB(t) & \mathrm{t}_0 \leq t \leq t_1 \\ X(0) = x_0 \end{cases}$$

Aplicando la fórmula de Îto a $X(t) = f(t, B(t))$ e igualando coeficientes con la ecuación anterior, se llega a que:

$$\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2} = \mu f$$

$$\frac{\partial f}{\partial x} = \sigma f$$

Despejando en la segunda ecuación se obtiene que $f(t,x) = exp(\sigma x + g(t))$; substituyendo en la primera se concluye que $g(t) = (\mu - \frac{1}{2}\sigma^2)t$. Por tanto, la solución a $[a]$ modelo es exactamente el movimiento Browniano geométrico.

Proposición 2.2.3 (Propiedades del movimiento Browniano geométrico) *El mov. Browniano geométrico verifica que:*

1. $E(X(t)) = x_0 e^{\mu t}$
2. $Var(X(t)) = x_0^2 e^{2\mu t}(e^{\sigma^2 t} - 1)$
3. $Cov(X(t), X(s)) = x_0^2 e^{\mu(t+s)}(e^{\sigma^2 s} - 1)$

Proceso de Ornstein-Uhlenbeck

Modelo propuesto por Ornstein y Uhlenbeck para modelar la velocidad que experimenta una partícula sometida a un movimiento de difusión en pequeños intervalos de tiempo.

Explícitamente, tal velocidad $X(t)$ verificaría $[b]$

$$[b]\begin{cases} dX(t) = -\alpha X(t)dt + \sigma dB(t) \\ X(0) = x_0 \end{cases}$$

donde α y σ son constantes positivas.

Resolviendo la EDE obtenemos una expresión explícita para la velocidad:

$$X(t) = x_0 e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-s)}\, dB$$

La diferencial $dX(t)$ que heurísticamente corresponde a un incremento infinitesimal en velocidad de la partícula, resulta ser directamente proporcional a la propia velocidad. Ahora bien, tal término es negativo; esto es, representa una fuerza de fricción. Por otra parte, notemos que se añade un ruido aleatorio: $\sigma dB(t)$.

Para resolver $[b]$ emplearemos un argumento similar a la clásica de variación de las constantes; consideremos una solución de la forma:

$$X(t) = a(t)[x_0 + \int_0^t b(s)\, dB]$$

Por tanto:

$$dX(t) = a^{'}(t)[x_0 + \int_0^t b(s)\, dB]dt + a(t)b(t)dB =$$

$$= \frac{a^{'}(t)}{a(t)} X(t)dt + a(t)b(t)dB$$

Comparando con $[b]$ se tiene que:

$$\frac{a^{'}(t)}{a(t)} = -\alpha$$

$$a(t)b(t) = \sigma$$

Por tanto, si $a(0) = 1$ se concluye que $a(t) = e^{-\alpha t}$ y $b(t) = \sigma e^{\alpha t}$

Proposición 2.2.4 (Propiedades del proceso de Ornstein-Uhlenbeck) *El proceso de Ornstein-Uhlenbeck verifica que:*

1. $E(X(t)) = x_0 e^{-\alpha t}$
2. $Var(X(t)) = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t})$
3. $Cov(X(t), X(s)) = \frac{\sigma^2}{2\alpha}(e^{-\alpha(t-s)} - e^{-\alpha(t+s)})$

Puente Browniano

El considerar un movimiento Browniano en el intervalo $[0,1]$ de forma que en los extremos se anule es de suma utilidad en muchas aplicaciones. Sorprendentemente, tal proceso se presenta también en ciertos razonamientos abstractos, como puede ser el test de Kolmogorov-Smirnov. Hay muchas formas equivalentes de definirlo, por ejemplo:

- $X(t) = (1-t)B(t)$
- $X(t) = B(1-t) - (1-t)B(1)$
- $X(t) = (1-t)\int_0^t \frac{1}{1-s}\, dB$
- Como el proceso solución de la ecuación diferencial estocástica

$$[c]\begin{cases} dX(t) = -\frac{X(t)}{1-t}dt + dB(t) & 0 < \text{t} < 1 \\ X(0) = 0 \end{cases}$$

Para resolver $[c]$ se procede de la misma forma que para $[b]$:
Consideremos una solución de la forma:

$$X(t) = a(t)[x_0 + \int_0^t b(s)\, dB]$$

Por tanto:

$$dX(t) = a^{'}(t)[x_0 + \int_0^t b(s)\, dB]dt + a(t)b(t)dB =$$

$$= \frac{a^{'}(t)}{a(t)}X(t)dt + a(t)b(t)dB$$

Comparando con $[c]$ se tiene que:

$$\frac{a^{'}(t)}{a(t)} = -\frac{1}{1-t} \text{ y que } a(t)b(t) = 1$$

Si $a(0) = 1$ se tiene que $a(t) = 1 - t$ y $b(t) = \frac{1}{1-t}$, luego

$$X(t) = (1-t)\int_0^t \frac{1}{1-s}\,dB$$

No es difícil comprobar que que efectivamente $\lim_{t\to 1} X(t) = 0$ P-a.s.
Veamos ahora ciertas propiedades:

Proposición 2.2.5 (Propiedades del proceso del puente Browniano) *El puente Browniano verifica que:*

1. $E(X(t)) = 0$
2. $Var(X(t)) = t(1-t)$
3. $Cov(X(t), X(s)) = s(1-t)$

Notemos que es un proceso centrado, y que la varianza viene descrita por una parábola, de eje vertical, simétrica respecto de $\frac{1}{2}$ cuyo máximo valor se alcanza precisamente en tal punto, el punto medio del intervalo. Además es nula en los extremos.

Capítulo 3

Referencias bibliográficas

Se han escrito muchos textos sobre cálculo estocástico, movimiento Browniano y teoría de martingalas. Señalaremos unos pocos:

- Una excelente referencia para la teoría de martingalas es [1]. En el capítulo 6, *"Derivates and conditional probability"*, uno puede encontrar una exposición de la teoría de martingalas mucho más profunda que la que hemos hecho. Se aborda el interesante concepto de martingala inversa, el cual hemos tenido que omitir.
  El capítulo 7: *"Stochastic processes"* comienza con el teorema de existencia de Kolmogorov, para luego introducir el movimiento Browniano. En él se estudian ciertas propiedades de este proceso que nosotros hemos pasado por alto, como por ejemplo la propiedad fuerte de Markov.
- El libro de Karatzas y de Shreve ([2]) es una magnífica referencia general para los tres capítulos; todas las demostraciones que hemos omitido se pueden encontrar en él.
- Otra excelente referencia para los tres capítulos, que también contiene todas las demostraciones omitidas es [4]. Aunque este libro constituirá el pilar fundamental sobre el que se apoya la siguiente parte, contiene una sólida y sencilla, en lo que a lectura se refiere, exposición del cálculo estocástico.
  En el capítulo 2: *"Martingales measures"*, se introduce el concepto de martingala a tiempo discreto y el de integración estocástica discreta; en el capítulo 5: *"Discrete-time American options"*, se exponen las desigualdades y los teoremas de convergencia de Doob.
  El capítulo 6 (*"Continuous-time Stochastic calculus"*) aborda la teoría de martingalas a tiempo continuo y la construcción de la integral Browniana, terminando con una somera introducción a las ecuaciones dife-

renciales estocásticas.
El teorema de Girsanov se estudia en el capítulo 7: *“ Continuous-time european options”.*

- Las referencias [3] y [4] son también muy interesantes. El libro de Nualart, aunque aborda temas que escapan a los pretextos de estas notas, contiene una exposición del cálculo de Îto. Por otra parte, [3] es una obligada referencia general.

Parte II

Aplicaciones a la matemática financiera

Introducción

En las líneas que siguen a esta introducción se pretende, por una parte, ilustrar someramente el problema de la valoración racional de derivados financieros y, por otra, ejemplificarlo con un caso particular: el afamado modelo de Black-Scholes, que resuelve de manera explícita la valoración racional de las denominadas opciones europeas bajo ciertas condiciones generales.
Se comienza presentando ciertos conceptos básicos e ideas generales hasta que se esté preparado para comprender los mecanismos subyacentes en la valoración de derivados, con especial énfasis en las opciones. Posteriormente se introduce el modelo de mercado binomial; el modelo más simple posible. El que sea sencillo no quiere decir que sea yermo, todo lo contrario: presenta numerosas características que sirven para poner de manifiesto ciertas peculiaridades de los modelos generales de mercado discreto. A continuación se estudian dos condiciones sobre los mercados que son lo suficientemente generales como para poder reproducir en mayor o en menor medida gran parte de los comportamientos bursátiles y lo suficientemente restricitivas como para evitar comportamientos anómalos: la viabilidad y la completitud. El otro interés que presentan tales hipótesis es que permiten dar una solución cerrada (al menos de manera teórica) a nuestro problema. En el caso continuo, con las anteriores ideas en mente, se presenta el modelo de Black-Scholes que da una solución explícita al problema de valoración de opciones europeas bajo ciertas restricciones no excesivamentes fuertes pero tampoco realistas.
Incidiendo en las plausibles restricciones que presenta el pardigma de Black-Scholes, en el último capítulo se introducen, a título de ejemplo, métodos alternativos de valoración relacionados con el problema del transporte óptimo.

Capítulo 4

Dinámica discreta y generalidades

El intercambio de bienes ha sido, es y será una constante a lo largo de la historia del ser humano. Desde tiempos prehistóricos la especialización en una o en otra actividad condicionaba el tipo de beneficios que se adquirían; el trueque se convertía por tanto en una de las piedras angulares para la supervivencia. El no tener un criterio universal y absoluto para la valoración de los productos se resolvió en la creación del dinero. Si bien no se solventaban los anteriores problemas, totalmente subjetivos por cierto, servía como marco de referencia para la valoración. El dinero vino acompañado en primera instancia de préstamos, prestamistas y prestatarios, respondiendo todos ellos a la idea abstracta de «alquilar dinero». Esto resultó ser un negocio extremadamente eficaz; con él surgieron los bancos y posteriormente todo un tejido financiero que daría lugar a las primeras bolsas de inversión. Éstas tenían como principales objetivos poner en contacto empresas y entidades del estado con los principales ahorradores y canalizar el ahorro hacia la inversión. Con el paso de los siglos, este entramado ha ido evolucionado en mayor o menor medida hasta la década de los 70, fecha en que surgen los llamados derivados financieros. A partir de entonces, los mercados han experimentado un incomensurable crecimiento, y con ellos la matemática financiera.

4.1. Generalidades

Centrémonos ya en los mercados actuales.
Dos de los principios fundamentales que se asumen en economía son el de **comportamiento racional**, esto es, los inversores prefieren siempre tener más que menos y **el riesgo neutral**: a priori, nada indica que los bienes en un mercado ideal se van a desplomar o a revalorizar, se les presupone una tendencia de comportamiento neutral (estable).

Es claro que el objetivo fundamental que se persigue inviertiendo en bolsa es ganar dinero; la dinámica en una primera aproximación (ingenua) puede resultar sencilla: comprar cierta cantidad de un activo financiero y esperar que al cabo de un tiempo se haya revalorizado para ganar dinero. En resumidas cuentas: **especulación**. No obstante, tal actitud ante el mercado no es la única. Encontramos otras dos posiciones bien delimitadas dentro de los negociantes: **coberturistas** y **arbitrajistas**. La filosofía de los coberturistas es la reducción del riesgo de sus operaciones a través de otros movimientos financieros. Los arbitrajistas se centran en varios mercados de forma simultánea intentando explotar desiquilibrios puntuales.

Comenzaremos introduciendo varios conceptos al uso de capital importancia:

- Por **activo financiero con riesgo** entenderemos cualquier clase de activo (acciones, deuda, divisas...) cuya valoración el tiempo no es conocida a priori. Por **activo financiero sin riesgo** o simplemente **bono** nos referimos a cualquier capital cuya evolución viene determinada mediante una tasa de cotización conocida; por ejemplo, un capital depositado en una cuenta bancaria.
- Una **cartera de inversión** (o *portfolio* en inglés) es un conjunto de estrategias de inversión en el mercado diseñadas para una colección de activos financieros. Suele contener activos de renta fija y de renta variable. La diversificación de una cartera es una herramienta fundamental para reducir el riesgo natural que atañe. Entre los activos más frecuentes encontramos acciones, bonos, monedas internacionales, fondos de inversión, etc.
- Un **contrato de futuro** es un acuerdo para comprar o vender cierto activo financiero por un precio prefijado en una fecha futura prefijada.
- Un **derivado** de un activo con riesgo es un producto financiero cuyo valor se basa en el precio del activo. Destacan espcialmente los *swaps*, *forwards* y las opciones (europeas, americanas, exóticas, asiáticas, etc).
- Un derivado **swap** o permuta financiera es contrato entre dos partes consistente en el intercambio de capital en ciertas fechas temporales.
- Una **opción europea de compra** o european call option, es un derecho pero no una obligación, que se otorga al negociante para comprar a un precio fijo cierta cantidad de un activo financiero en una fecha fijada.
- Una **opción europea de venta** o european put option, es un derecho pero no una obligación, que se otorga al negociante para vender a un

precio fijo cierta cantidad de un activo financiero en una fecha fijada.

- El **precio de ejercicio** o strike-price es el precio que se ha de pagar en los contratos de futuro o en las opciones. Análogamente se define la **fecha de ejercicio** o de vencimiento.
- Una **opción americana** (de venta o compra) es a todos los efectos una opción europea (de venta o de compra) salvo que puede ejercerse en cualquier instante de tiempo anterior a la fecha de ejercicio.
- Una **posición corta** ante un contrato consiste en desempeñar el rol de vendedor, por contra, una **posición larga** hace alusión al papel de comprador.
- **Capitalizar** consiste en llevar a valor (invertir) los réditos obtenidos en cierta operación.

4.2. Modelos de mercado finitos y opciones europeas

De ahora en adelante tomaremos como marco de trabajo un modelo de mercado donde los cambios de valores de los activos se producen a tiempo discreto. Por simplicidad trabajaremos con un único activo financiero, cuyo precio al contado (**spot price**) vendrá representado por el proceso estocástico $(S_n)_{n\geq 0}$. A priori, tenemos por tanto un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ que supondremos finito, esto es, sólo existe una cantidad finita de puntos $\omega \in \Omega$ tales que $P(\omega) > 0$. En este caso la σ-álgebra $\mathcal{F}$ será $P(\Omega)$: el conjunto de partes de Ω.

El modelo también recoge al propio proceso $(S_n)_{n\geq 0}$, una filtración $(\mathcal{F}_n)_{n\geq 0}$ que representa heurísticamente la historia del proceso de precios (por supuesto $(S_n)_{n\geq 0}$ es adaptado a la filtración) y una colección de instantes de tiempo $\{0, 1, 2, ...N\}$ donde evolucionan los precios del activo. Podría darse el caso extremo de que $N = \infty$.

Asumimos que no hay ni costes de transacción, ni límites de posiciones (límites de contratos que un inversor puede mantener) ni límites monetarios en los préstamos. Bajo estas suposiciones se dice que el mercado es **libre de fricción**. En otras palabras, es un modelo de **mercado perfecto**.

Veamos como algunos de los conceptos anteriormente definidos admiten una reformulación matemática en el lenguaje de la probabilidad en términos de procesos estocásticos.

Comencemos por el bono, el activo financiero sin riesgo. Supongamos C_n representa el capital en el instante n del bono. Si C_0 es el montante inicial, al cabo del primer período de tiempo se tendrá que $C_1 = (1 + r_1)C_0$ donde r_1 es el tipo de interés (discreto) en el intervalo $[0, 1)$. A la cantidad $(1 + r_1)$ se la llama R_1. Se define de manera análoga $R_n = R_{n-1}(1+r_n)$ donde $r_n \geq 0$ representa el tipo de interés en el intervalo $[n-1, n)$. Como las cantidades R_n arrastran la recursión de los r_i se tiene que $C_n = R_n C_0$
Con estas ideas en mente podemos abordar la formulación matemática de cartera de inversión en nuestro modelo:

Definición 4.2.1 (Cartera de inversión) *Una cartera de inversión (o estrategia de inversión) es un par de procesos predecibles (A_n, B_n). El primero de ellos representa la cantidad de bonos que se poseen a tiempo n y el segundo la cantidad del activo financiero. El proceso de valor asociado a la cartera es el proceso estocástico V definido como sigue:*

$$V_n = A_n R_n + B_n S_n$$

Nota 4.2.2 *1. Salvo que digamos lo contrario, cualquier cartera de inversión será autofinanciada, esto es, la inversión de capital que se llevará a cabo en el instante $n+1$ viene dada por una reordenanción de las ganancias o pérdidas a instante n. Dicho de otro modo, los únicos cambios de valor que experimenta el proceso V_n son debidos a ganancias o pérdidas de la propia cartera (de ahí la denominación: no hay ni inyecciones de capital exterior ni costes de penalización adicionales). Matemáticamente esto se expresa como sigue:*

$$V_n = A_{n+1} R_n + B_{n+1} S_n$$

con $n \geq 0$. Notemos que la ecuación anterior es equivalente a

$$0 = (A_{n+1} - A_n)R_n + (B_{n+1} - B_n)S_n$$

es decir

$$0 = \Delta A_n R_n + \Delta B_n S_n$$

Observemos que esto último es una relación de ortogonalidad. Tal característica será muy útil en el desarrollo posterior de la materia.

2. *Llamaremos factor de descuento a tiempo n y lo denotaremos por β_n a la cantidad de dinero que necesitamos invertir en bonos a tiempo 0 para tener una unidad monetaria a tiempo n. Siempre que nos referimas a una cantidad actualizada a tiempo n nos estaremos refiriendo al valor de dicho montante multiplicado por el factor de descuento.Denotaremos a estas cantidad mediante el acento circunflejo francés. Así pues, para el modelo que manejamos de cartera de inversión se tendría que $\beta_n = R_n^{-1}$ y que $\hat{V}_n = A_n + R_n^{-1} B_n S_n$.*

El proceso de valor asociado a una cartera suele ser un criterio para su clasificación. Introducimos un tipo especial de carteras que serán muy útiles en la práctica.

Definición 4.2.3 (Carteras admisibles) *Se dice que una cartera de inversión es admisible si el proceso de valor asociado es siempre no negativo, es decir*

$$P(\bigcap_{n \geq 1} (V_n \geq 0) = 1$$

Es decir, si en cualquier instante la estrategia de inversión no genera pérdidas. Notemos que efectivamente se trata de una condición razonable para muchos propósitos.

Hasta la década de los 70, los mercados financieros se centraban principalmente en cuatro tipos de activos: acciones, divisas, deuda y futuros; sin embargo, el advenimiento de los derivados financieros supuso una completa reestructuración del mercado a todos los niveles.
Dentro de los derivados más importantes encontramos las opciones.
La teoría de valoración de opciones cuenta desde 1973 con una gran fama, pues los celebrados trabajos de Myron Scholes y Fisher Black dieron fórmuladas cerradas para el llamado precio racional de las opciones europeas. Esto no quiere decir que tales predicciones sean siempre las correctas; nada más lejos de la realidad. Las hipótesis sobre las que se basa su construcción matemática no eran del todo reales, aunque si representaban un modelo en cierta medida coherente.

Comencemos con una primera aproximación ingenua al problema de la valoración de una **opción europea** (de compra) en el modelo mercado en

que trabajamos.
Comúnmente se suele denotar por K al precio de ejercicio; de esta forma las ganancias que revierte una opción europea de compra en una posición larga son $H = (S_N - K)^+$.

Como estamos adquiriendo un derecho y no una obligación es lógico que se exija un precio a pagar. Éste recibe el nombre de **prima de la opción** y generalmente se denota mediante la letra C. El problema central de estas notas consiste en crear una teoría matemática que permita asignar un precio justo (fair-price) a la prima de la opción (**option pricing**).
A priori a uno se le ocurren dos estrategias bien distintas:

- Construir una cartera de inversión admisible cuyo valor a instante N sea el mismo que el de la opción, i.e. $H = V_N$. Un buen cantidato a prima de la opción sería el coste inicial de la cartera V_0.
- Utilizar argumentos probabilísticos y hallar el valor esperado de la variable aleatoria H bajo cierta ley asociada a S_N, a priori totalmente desconocida. Una posición ingenua consistiría en pensar que se puede estimar tal ley a partir de observaciones en el mercado. Nada más lejos de la realidad: ciertos resultados de teoría económica afirman que la ley de S_N poco o nada tiene que ver con la situación del mercado en los primeros instantes.

Las observaciones anteriores aportan ideas interesantes. Formalizamos un par de conceptos implícitos en ellas: Sea H una variable aleatoria no negativa (que representará, de forma general, la parte positiva de los réditos de algún contrato financiero), una cartera de cobertura para H es una cartera de inversión admisible tal que $V_N = H$ siendo V el proceso de valor asociado.

Definición 4.2.4 (Variable replicable) *Sea H una variable aleatoria no negativa. Se dice que es replicable si existe una cartera de cobertura para ella.*

El que una variable sea replicable es una herramiento teórica muy útil. Los mercados completos son aquellos en los que todas las variables que representen los beneficios de una opción europea son replicables.

El restringirse a opciones europeas no aporta (en primera instancia) ninguna ventaja ni supone ninguna simplificación desde el punto de vista teórico en el desarrollo del problema. En el caso general, el beneficio H resulta ser una función f dependiente del valor del activo a tiempo N, S_N, y del precio de ejercicio K.

Hemos visto que para una opción europea de compra se tiene que $f(S_N, K) = (S_N - K)^+$; además, es inmediato comprobar que para una opción europea de venta f es justamente $(K - S_N)^+$. Daremos más ejemplos cuando se trabaje en la dinámica continua.

Volviendo al problema de la valoración; se puede probar que bajo ciertas hipótesis que aseguren el «buen» comportamiento del mercado (viabilidad) **tanto el argumento de la cartera de cobertura como la valoración mediante la esperanza conducen al mismo resultado**. Antes de abordar este profundo hecho estudiaremos un modelo sencillo (el modelo binomial) que aportará interesantes ideas y servirá para comprobar tanto la veracidad como la practicidad de los resultados teóricos que se expondrán en las siguientes secciones.

4.2.1. Modelo binomial multiperíodo

El modelo binomial de N pasos es un sencillo modelo de mercado, no excesivamente realista, pero de vital importancia para desarrollos posteriores. En él se anticipan varios conceptos que luego desarrollaremos en profundidad, como el medida de martingala equivalente, esto es, una medida de probabilidad equivalente a la dada bajo la cual el proceso de precios es una martingala. Se verá que cualquier cartera que cubra un derivado comparte el proceso de valor asociado, cuyo valor inicial es exactamente la esperanza bajo la probabilidad sin riesgo (que también veremos que es única) del contigente del derivado: H.

Tomemos dos sucesiones $(a_n)_n$ y $(b_n)_n$, indexadas de 1 a N tales que $0 < a_n < b_n$. El adjetivo "binomial" hace referencia a que el valor de los precios (S_n) solo puede tomar dos valores por instante de tiempo; suponemos por tanto, sin pérdida de generalidad, que:

$$S_n \in \{a_n S_{n-1}, b_n S_{n-1}\}$$

Supongamos que se tiene la siguiente situación:

$$P(S_n = a_n S_{n-1} \| \mathcal{F}_{n-1}) = 1 - p_n$$
$$P(S_n = b_n S_{n-1} \| \mathcal{F}_{n-1}) = p_n$$

donde asumimos que no conocemos las cantidades (p_n) pero si las $(a_n)_n$ y $(b_n)_n$ (aunque esto no es nada realista).

Por tanto, la distribución de S_n condicionalmente dado S_0 queda totalmente caracterizada por el vector $(p_1, p_2, \ldots, p_n)$
Consideremos ahora la sucesión de intereses $(r_n)_n$ indexadas evidentemente de 1 a N . Para evitar arbitrajes es claro que ha de tenerse que $a_n < 1+r_n < b_n$ para todo n. En efecto, si se tuviese que $a_n < b_n < 1+r_n$ según la hipótesis de mercado racional cualquier corredor invertiría en bonos. Caso contrario: $1 + r_n < a_n < b_n$, todo el mundo invertiría en el activo financiero por la misma razón.
Supongamos ahora que se desea valorar una opción europea de beneficio $H = (S_N - K)^+$. Para ello nos planteamos las dos estrategias anteriores: la construcción de una cartera de cobertura o buscar una ley de probabilidad adecuada bajo la cual hallar la esperanza de H. El siguiente resultado que resuelve de pleno el problema, sirve para anticipar lo anteriormente comentado: bajo ciertas condiciones sobre el mercado ambos caminos coinciden.

Teorema 4.2.5 (Comportamiento del modelo binomial) *Considerando las tres sucesiones anteriores, sus relaciones y la sucesión de precios S_n definidas todas ellas como antes, entonces, existe una única cartera de cobertura cuyo valor inicial es*

$$V_0 = E_Q(R_N^{-1} H)$$

donde $Q = (q_1, \ldots, q_N)$ es la única medida de probabilidad tal que $\hat{S}_n$ dado S_0 es una martingala. Explícitamente, si $q_n = P(S_n = b_n S_{n-1} \| \mathcal{F}_{n-1})$, se tiene que:

$$q_n = \frac{1 + r_n - a_n}{b_n - a_n}$$

Demostración:
Por hipótesis:

$$E(\hat{S}_n \| \mathcal{F}_{n-1}) = R_n^{-1}((1 - q_n)a_n + q_n b_n)S_{n-1}$$

que es igual a $\hat{S}_{n-1}$ si y solamente si

$$(1 - q_n)a_n + q_n b_n = \frac{R_n}{R_{n-1}} = 1 + r_n$$

equivalentemente:

$$p_n = \frac{1 + r_n - a_n}{b_n - a_n}$$

Por hipótesis tal valor es mayor que 0 y menor que 1, luego define una probabilidad, y además se observa que es única, luego existe una única probabilidad (a la que llamaremos Q) tal que $\hat{S}_n$ dado S_0 es una martingala.

Obviamente el proceso $\hat{V}_n = E_Q(R_N^{-1}H\|\mathcal{F}_n)$ es una Q- martingala. Dado $\mathcal{F}_n$, las variables $\hat{V}_n - \hat{V}_{n-1}$ y $\hat{S}_n - \hat{S}_{n-1}$ son ambas funciones del cociente $\frac{S_n}{S_{n-1}}$ y por dando están soportadas en dos puntos. Al ser diferencias de martingalas, su esperanza condicionada bajo q_n nula. Todo esto implica que existe una variable $\mathcal{F}_{n-1}$ medible, digamos B_n, tal que

$$\hat{V}_n - \hat{V}_{n-1} = B_n(\hat{S}_n - \hat{S}_{n-1})$$

Dado el proceso B, definimos el proceso A como sigue:

$$A_nR_{n-1} + B_nS_{n-1} = R_{n-1}\hat{V}_{n-1}$$

De esta forma se tiene que tanto A como B son procesos predecibles, además:

$$A_n + B_n\hat{S}_{n-1} = \hat{V}_{n-1}$$

luego, de la definición de B_n se tiene que

$$A_n + B_n\hat{S}_n = \hat{V}_{n-1} + B_n(\hat{S}_n - \hat{S}_{n-1}) = \hat{V}_n$$

multiplicando por R_n:

$$R_nA_n + B_nS_n = R_n\hat{V}_n$$

Evaluando la anterior línea en $n-1$ en vez de en n y comparando con que $A_nR_{n-1} + B_nS_{n-1} = R_{n-1}\hat{V}_{n-1}$ se comprueba que el par (A, B) define una cartera autofinanciada.

Además, de lo anterior también se deduce que el valor a tiempo n de la cartera es

$$V_n = R_n\hat{V}_n = R_nE_Q(R_n^{-1}H\|\mathcal{F}_{n-1})$$

luego $V_N = H$ y $V_0 = R_0E_Q(R_n^{-1}H)$. El que la cartera sea admisible es obvio, pues $\hat{V}$ es la martingala de la esperanza condicionada de una variable aleatoria no negativa, luego es no negativa y por ende también lo será el proceso V. Esto termina la prueba de la existencia de la cartera pedida, resta ver que es única.
Sea (A, B) una cartera autofinanciada y admisible que replique a H. Sea $V_n = A_nR_n + B_nS_n$ su proceso de valor asociado. Como ya se ha comprobado existe una única medida de probabilidad Q de suerte que que $\hat{S}$ es una martingala; por ser la cartera autofinanciada se tendrá que:

$$\hat{V}_{n-1} = A_n + B_n\hat{S}_{n-1} = A_{n-1} + B_{n-1}\hat{S}_{n-1}$$

luego, operando se llega a que:

$$\hat{V}_n - \hat{V}_{n-1} = B_n(\hat{S}_n - \hat{S}_{n-1})$$

Al ser $\hat{S}$ una Q-martingala y B predecible se tendrá que $\hat{V}$ es una Q-martingala, luego $\hat{V}_n = E_Q(\hat{V}_N \| \mathcal{F}_n)$. Para concluir basta recordar que la cartera replicaba a H, esto es: $V_N = H$.
$\square$

4.3. Mercados viables

4.3.1. Oportunidades de arbitraje

El concepto de arbitraje responde a la idea de la posible existencia de estrategias de inversión que proporcionen de forma segura un beneficio. En un mercado ideal las oportunidadades de arbitraje no deberían existir, en efecto: en virtud del principio racional serían las estrategias de inversión más deseadas y por tanto las más solicitadas creándose un desiquilibrio en el mercado.

A la hora de definir una oportunidad de arbitraje en el sentido matemático haremos dos distinciones: fuerte y débil.

Definición 4.3.1 (Oportunidad de arbitraje fuerte) *Se dice que una cartera admisible es un arbitraje en el mercado (o una oportunidada de arbitrage) si su proceso de valor asociado verifica:*

$$V_0 = 0$$

$$E(V_N) > 0$$

La restricción de que la cartera sea admisible no parece ser esencial si se pretende reflejar el concepto de arbitraje en base al valor que presenta una cartera, es más, puede llegar a ser ligeramente restrictivo. Para solventar tal obstáculo se introduce el concepto de arbitraje débil, de gran utilidad en desarrollos teóricos:

Definición 4.3.2 (Oportunidad de arbitraje débil) *Se dice que una cartera de inversión es una oportunidad de arbitraje débil si su proceso de valor verifica que:*

1. $V_0 = 0$
2. *La variable aleatoria es* $V_N \geq 0$ *con probabilidad 1 y* $P(V_N > 0) > 0$*, es decir,* $P(V_N) > 0$

Nota 4.3.3 *Las siguientes observaciones son necesarias:*

- *Si no se especifica lo contrario nos referiremos siempre al arbitraje fuerte.*
- *Notemos que si la variable* V_N *es no negativa, la condición* $E(V_N) > 0$ *es trivialmente equivalente a que* $P(V_N > 0) > 0$
- *El arbitraje es un concepto íntimamente ligado al equilibrio de un mercado. En los mercados reales sí que hay arbitrajes; los negociadores que los llevan a cabo se conocen como arbitrajistas. Entre las causas principales del arbitraje podemos encontrar: un activo que no cotiza con el mismo valor a dos o más mercados distintos, dos activos que producen el mismo flujo de efectivo pero no cotizan al mismo precio y las diferencias en la valoración de divisas.*
- *Una diferencia importante entre un arbitrajista y un especulador es que el primero solo poseé el bien durante un corto período de tiempo, mientras que el especulador lo mantiene con la esperanza de que se revalorice y pueda tomar ventaja de tal diferencia de precio.*

La clave del arbitraje débil es que resulta ser equivalente al artibraje fuerte. Es decir, para que se produzca un arbitraje en el mercado no es necesario que la estrategia seguida sea admisible. Esta debilitación en las hipótesis resulta de gran ayuda en la prueba de varios resultados en las secciones siguientes.

Recogemos tal equivalencia en el siguiente teorema, del que omitimos su demostración:

Teorema 4.3.4 (Equivalencia de arbitrajes) *En un mercado finito el arbitraje débil es equivalente al arbitraje fuerte.*

Como se ha comentado antes, en un mercado ideal no deberían existir arbitrajes. Parece lógico imponer a cualquier modelo de mercado que pretenda ser realista la condición anterior.

Definición 4.3.5 (Mercado viable) *Se dice que un modelo de mercado es viable si no se puede construir en él ningún arbitraje.*

En breve se comprobará que el hecho de que un modelo de mercado posea tal propiedad es equivalente a que exista una probabilidad sin riesgo, esto es, una probabilidad equivalente a la probabilidad base respecto de la cual los precios actualizados son una martingala.

4.3.2. Medidas equivalentes de martingala

A parte de evitar arbitrajes, en un modelo de mercado ideal regido por principios como el de racionalidad, leyes de valoración únicas y riesgo neutro, parece lógico asumir que los precios actualizados de un bien se distribuyan de la forma más cercana posible a un comportamiento neutral. Por tanto, no es descabellado contemplar la hipótesis de la existencia de una probabilidad respecto de la cual los precios al contado sean una martingala. En las siguientes líneas se profundiza en estas ideas y se abre la puerta al llamado teorema de equivalencia, que culmina la sección.

Una probabilidad con las características anteriores se la conoce como probabilidad sin riesgo, precisando más:

Definición 4.3.6 (Probabilidad sin riesgo) *Se dice que una probabilidad sobre $(\Omega, \mathcal{F})$ es una probabilidad sin riesgo para S si es equivalente a P y sobre ella los precios actualizados $(\hat{S}_n)_n$ forman una martingala.*

Sea $\phi = (A_n, B_n)$ una cartera de inversión. Al suponer que es autofinanciada se tendrá que

$$V_n = A_{n+1}R_n + B_{n+1}S_n$$

luego

$$\hat{V}_n = A_{n+1} + B_{n+1}\hat{S}_n$$

por tanto

$$\hat{V}_{n+1} - \hat{V}_n = \Delta\hat{V}_n = A_{n+1} + B_{n+1}\hat{S}_{n+1} - A_{n+1} + B_{n+1}\hat{S}_n = B_{n+1}\Delta\hat{S}_n$$

Por ser el proceso $(B_n)_n$ acotado y $(\Delta\hat{S}_n)_n$ integrable se tiene que $(\Delta\hat{V}_n)_n$ es también integrable.

Si suponemos que Q es una probabilidad sin riesgo entonces:

$$E_Q(\Delta\hat{V}_n \| \mathcal{F}_n) = E_Q(B_{n+1}\Delta\hat{S}_n \| \mathcal{F}_n) = B_{n+1}E_Q(\Delta\hat{S}_n \| \mathcal{F}_n) = 0$$

Q-casi

seguro (o lo que es lo mismo, P- casi seguro). Se tiene por tanto el siguiente resultado:

Proposición 4.3.7 *Sea Q una probabilidad sin riesgo, entonces, para cualquier cartera de inversión, su proceso de valor actualizado* $(\hat{V}_n)_n$ *es una Q martingala.*

Consideremos una cartera de inversión cuyo valor inicial sea nulo. Si $(\hat{V}_n)_n$ (su valor asociado actualizado) fuese una martingala respecto de alguna probabilidad Q entonces:

$$E_Q(\hat{V}_N) = E_Q(\hat{V}_0) = E_Q(V_0) = 0$$

luego si $\hat{V}_N \geq 0$ entonces $\hat{V}_N = 0$ Q-a.s.
Si suponemos además que Q es una probabilidad equivalente a P se tendría entonces que $\hat{V}_N = 0$ P-a.s. evitando así que la cartera sea una oportunidad de arbitraje. Del resultado anterior se sigue que **si el modelo de mercado admite una probabilidad sin riesgo entonces no puede haber ningún arbitraje, esto es: sería un modelo de mercado viable.**

Podría pensarse que la condición de que en un mercado exista una probabilidad sin riesgo es una estrictamente más fuerte que la restricción sobre la no existencias de arbitrajes y de esta forma intentar buscar alguna caracterización equivalente a esta última hipótesis. Sorprendentemente esto no es así, es decir: ambas son equivalentes.

4.3.3. El teorema de equivalencia

El teorema de equivalencia recoge el hecho de que al menos en un modelo de mercado finito, suponer la no existencia de arbitrajes es tanto como suponer la existencia de una probabilidad sin riesgo.

Teorema 4.3.8 (Teorema de equivalencia) *Un modelo de mercado finito (para un bien S) es viable si y solo si existe una probabilidad sin riesgo para S*

Notemos que estamos trabajando en el caso en que sólo disponemos de un único bien. Todos los resultados anteriores se pueden generalizar al caso de m bienes $S^1, S^2, \ldots, S^m$:

- Una **cartera de inversión** será una $m+1$-upla de procesos predecibles $(A_n, B_n^1, B_n^2 \ldots, B_n^m)$. El primero de ellos representa la cantidad de bonos que se poseen a tiempo n y el proceso i-ésimo la cantidad del activo financiero i-ésimo. El de valor asociado a la cartera es el proceso estocástico V definido como sigue:

$$V_n = A_n R_n + B_n^1 S_n^1 + B_n^2 S_n^2 + \ldots + B_n^m S_n^m$$

Conceptos tales como la autofinanciación, carteras admisibles, oportunidades de arbitraje, etc... se extienden sin mayor dificultad.

- En cuanto a la probabilidad sin riesgo: se dice que una probabilidad Q es una EMM para la colección de bienes (vector de bienes) $\mathbf{S} = (S^1, S^2, \dots, S^m)$ si lo es para cada una de las componentes.
- El teorema de equivalencia se extiende forma natural : «*Un modelo de mercado finito (para un vector de bienes* $\mathbf{S}$*) es viable si y solo si existe una probabilidad sin riesgo para* $\mathbf{S}$».

  La parte restante que queda por probar presenta una dificultad sustancialmente mayor. Requiere, bajo nuestras suposiciones sobre el modelo de mercado, una versión del teoremea de Hanh-Banach. Ésta puede probarse sin necesidad de recurrir al axioma de elección, sin embargo, en contextos más generales, la demostración del teorema de equivalencia sí que necesita la versión clásica del teorema, la cual se apoya en el mencionado axioma.

Proposición 4.3.9 (Separación por hiperplanos) *Sea L un subespacio lineal de $\mathbb{R}^n$ y sea $K \subset \mathbb{R}^n$ un subconjunto compacto y convexo disjunto de L. Entonces existe un funcional $\phi : \mathbb{R}^n \to \mathbb{R}$ tal que $L \subset Ker(\phi)$ y $\phi(x) > 0$ para cada $x \in K$; esto es, existe un hiperplano, $Ker(\phi)$, que separa K y L conteniendo a este último.*

La prueba también se apoya en el siguiente lema técnico:

Proposición 4.3.10 *Dada la variable aleatoria V_0 y una familia finita de procesos predecibles $B^1, \dots, B^d$, entonces existe un único proceso predecible A tal que la $d+1$-úpla $(A, B^1, \dots, B^d)$ es una cartera de inversión autofinanciada de valor inicial V_0 .*

Notemos que podemos aplicar el teorema de separación anterior al $\mathbb{R}$-espacio vectorial $\mathbb{R}^\Omega$ (variables aleatorias de $\Omega \to \mathbb{R}$, pues al ser la σ-álgebra trivial toda apliación será medible); basta recordar que el espacio de probabilidad $(\Omega, \mathcal{F}, P)$ era discreto y finito, es decir, existe sólo una cantidad finita de puntos $\omega \in \Omega$ con probabilidad positiva, digamos $\{\omega_1, \dots, \omega_n\}$, luego si identificamos todas las variables aleatorias iguales casi siempre, existe una biyección entre $\mathbb{R}^\Omega$ y $\mathbb{R}^n$: la aplicación que asigna a cada variable X el vector

$$(X(\omega_1), X(\omega_2), \dots, X(\omega_n))$$

Denotaremos por C al cono *(subconjunto cerrado para suma de vectores y multiplicación por escalares no negativos)* de $\mathbb{R}^n$:

$$C = \{u \in \mathbb{R}^n : u_i \geq 0 \text{ para cada } i \text{ y tal que existe al menos}$$
$$\text{un índice } \ j \ \text{ tal que } \ u_j > 0\}$$

Definimos el proceso de ganancias de una cartera de inversión θ, $G_n(\theta)$, como:

$$G_n(\theta) = V_n(\theta) - V_0(\theta)$$

De esta forma, la viabilidad del mercado se traduce en que para cualquier cartera admisible ha de tenerse que

$$G_n(\theta) = V_n(\theta) \notin C$$

siempre que $V_0(\theta) = 0$
Notemos que de la **proposición 4.3.10** se deduce que cualquier cartera $\theta = (\theta_0, \ldots, \theta_d)$ queda determinada por sus d últimas componentes: $\hat{\theta} = (\theta_1, \ldots, \theta_d)$ y por su valor inicial; tomemos una cartera de valor inicial nulo y sea $\hat{\theta} = (\theta_1, \ldots, \theta_d)$ el conjunto de sus d-útimas componentes, siguiendo la notación habitual (recordemos que β representaba el factor de descuento) se tiene que:

$$V_N(\theta) = \beta_N^{-1}\hat{V}_N(\theta) = \beta_N^{-1}(V_0(\theta) + \hat{G}_N(\theta)) = \beta_N^{-1}\hat{G}_N(\hat{\theta})$$

luego el que $\hat{G}_N(\hat{\theta}) \in C$ implica que $V_N(\theta)$ sea no negativo con probabilidad 1 y estrictamente no negativo con probabilidad positiva, luego θ es un arbitraje débil. De esta forma hemos probado el siguiente resultado sobre el que de nuevo nos apoyaremos en la prueba del teorema de equivalencia:

Proposición 4.3.11 *En un modelo de mercado finito y viable, el proceso de ganancias actualizadas asociado a cualquier d-upla de procesos predecibles* $\hat{\theta} = (B^1, \ldots, B^d)$ *no puede estar en el cono* C*.(Recordemos que identificábamos las variables aleatorias de* Ω *con elementos de* $\mathbb{R}^n$*).*

Demostración (del teorema de equivalencia):

Una aplicación estaba ya demostrada, probemos la otra.

En aras de la brevedad, a partir de ahora, no haremos distinción entre elementos $\mathbb{R}^\Omega$ y $\mathbb{R}^n$ vía la identificación que venimos considerando. Supongamos

que el modelo de mercado es viable; construyamos una probabilidad Q equivalente a P bajo la cual el proceso de precios actualizado sea una martingala (respecto de la filtración del modelo, por supuesto).
El cono C visto desde Ω, es el subconjunto de variables aleatorias (iguales casi siempre) X tales que $X \geq 0$ y tales que existe un índice j de suerte que $X(\omega_j) > 0$, donde suponemos sin pérdida de generalidad, que $\Omega = \{\omega_1, \ldots, \omega_n\}$ y que $P(\omega_i) = p_i > 0$

Tomemos el subconjunto de $\mathcal{L} \subset \mathbb{R}^n$

$\mathcal{L} = \{G_n(\hat{\theta}) : 1 \leq n \leq N, \hat{\theta} = (\theta_1, \ldots, \theta_d)$ y cada componente es un proceso predecible $\}$

Es inmediato comprobar que es un subespacio lineal; además también es trivial comprobar que el conjunto

$$\mathcal{K} = \{X \in C : E_P(X) = 1\}$$

es cerrado y acotado (luego compacto en $\mathbb{R}^n$) y disjunto con $\mathcal{L}$.
Del teorema de separación por hiperplanos *(teorema de Hanh-Banach)* existirá un funcional $f : \mathbb{R}^n \to \mathbb{R}$ que separa $\mathcal{L}$ y $\mathcal{K}$, esto es, se anula en $\mathcal{L}$y estrictamente positivo en $\mathcal{K}$. Por el teorema de representación de Riesz sabemos que f es un producto escalar, es decir:

$$f(x_1, \ldots, x_n) = (x_1, \ldots, x_n) \circ (r_1, \ldots, r_n)$$

Veamos que cada r_i es estrictamente positivo.
Consideremos ahora los vectores $\epsilon_i = \frac{1}{p_i} e_i$ donde con e_i se representa al vector i-ésimo de la base canónica de $\mathbb{R}^n$. Es inmediato que estas variables aleatorias tienen media 1 bajo P, luego son elementos de $\mathcal{K}$ y por ende $f(\epsilon_i) = \frac{r_i}{p_i} > 0$, de donde se deduce que cada r_i es mayor que 0.

Sea $\alpha = \sum_i r_i > 0$, definimos el funcional $g = \frac{f}{\alpha}$; si $\frac{r_i}{\alpha} = q_i$, es obvio que

$$g(x_1, \ldots, x_n) = (x_1, \ldots, x_n) \circ (q_1, \ldots, q_n)$$

Notemos que $\sum_i q_i = 1$, luego la n-úpla $(q_1, \ldots, q_n)$ induce una probabilidad Q en Ω: $Q(\omega_i) = q_i > 0$; además, por construcción es obvio que Q es equivalente a P.

Para ver que Q es una medida de martingala resta ver que los precios actualiazdos bajo Q son una martingala:

En primer lugar notemos que si $\mathbf{x} \in \mathcal{L}$, entonces $g(\mathbf{x}) = \frac{f(\mathbf{X})}{\alpha} = 0$, luego $E_Q(G_n(\hat{\theta})) = g(G_n(\hat{\theta})) = 0$ para cualquier d-upla $\hat{\theta} = (\theta_1, \ldots, \theta_d)$ de procesos predecibles, en particular, para una cartera $\theta = (\theta_0, \ldots, \theta_d)$ cuyo valor inicial sea 0 se tendrá que $E_Q(G_n(\hat{\theta})) = 0$. Con esta información, de las consideraciones anteriores a la prueba del teorema se deduce que $E_Q(V_N(\hat{\theta})) = 0$ para cualquier cartera de la forma anterior y de la **Proposición 4.3.10** se tiene el hecho de que cualquiera de esas carteras se puede generar a partir de familias finitas de procesos predecibles n dimensionales, en particular de las de la forma $(0, \ldots, \theta_i, \ldots, 0)$ donde 0 denota al proceso indénticamente nul0, que es por supuesto predecible. De todo los anterior se tiene que

$$E_Q(\sum_{1 \leq n \leq N} \theta_n^i \Delta \hat{S}_n^i) = 0$$

para cualquier proceso predecible y acotado θ^i; la **Proposición 1.1.6** concerniente a la integración estocástica discreta y a la propiedad de estabilidad completa el resultado.
□

4.4. Mercados completos

El concepto de mercado completo surge de la necesidad formal de justificar la existencia de carteras de cobertura.
Las carteras de cobertura resultan ser excelentes formas de valoración para las opciones europeas (y para los derivados en general): basta con construir una y fijarse en su costo inicial. Si la prima de la opción es más cara que el costo se tendría una oportunidad de arbitraje, en efecto: el corredor que se encuentra en la posición corta, con el dinero recibido de la prima construiría una cartera de cobertura y por ende a tiempo N habría ganado tanto como la diferencia de precios entre la prima y el costo inicial. Si por contra, el costo es más caro inviertiendo los papeles entre el inversor de posición corta y el de larga se llegaría de manera análoga a un arbitraje.

El problema estriba en que nada nos garantiza a priori que el beneficio de la opción, H, sea replicable, así que sin una condición que nos asegure que esto es posible nada del razonamiento anterior serviría. Para paliar este obstáculo se introduce el concepto de mercado completo:

Definición 4.4.1 (Mercado completo) *Se dice que un mercado es completo si toda variable que represente el precio de una opción europea es replicable.*

Nota 4.4.2 *Las siguientes observaciones son necesarias:*

- *Es claro que ni la viabilidad implica la completitud ni viceversa.*

- *Si bien la hipótesis de viabilidad en un mercado permite obtener numerosos resultados y le dota de cierto grado de semejanza con los mercados reales, la completitud por sí sola no conlleva ninguna virtud destacada. Es por ello que se exige frecuentemente que los mercados completos sean viables.*

El siguiente resultado cuya prueba no es nada trivial caracteriza completamente, y de un modo práctico, los mercados viables que resultan ser completos.

Teorema 4.4.3 (Caracterización de la completitud en mercados viables) *Dado un modelo de mercado finito y viable, entonces es completo si y solamente si posee una única medida de martingala equivalente.*

La prueba se basa en un lema técnico que ofrece otra caracterización de los mercados viables y finitos que resultan ser completos:

Proposición 4.4.4 *Consideremos un modelo de mercado finito y viable y sea Q una EMM. Entonces el modelo es completo si y solo cada Q-martingala $M = M(n)$ puede ser representada como sigue:*

$$M(n) = M(0) + \sum_{i=1}^{n} H(i)\Delta\hat{S}_i$$

siendo $H = H(i)$ un proceso predecible.

4.5. Valoración a través de martingalas

Ya estamos en condiciones de dar carpetazo al problema de valoración de opciones europeas en mercados viables y completos al menos de manera teórica.

Como se ha venido señalando, parece haber dos estrategias fundamentales a la hora de valorar: construir carteras de cobertura o calcular la esperanza de H según cierta ley de probabilidad (presumiblemente la única EMM que tiene el mercado). Notemos que es razonable pensar que la primera sí que pueda llevarse a la práctica, al menos en ciertas ocasiones.
Los conceptos de mercado viable y mercado completo introducidos con el fin de acercar el modelo de mercado a la realidad, resultan ser además el marco adecuado para el desarrollo de la teoría de valoración a través de martingalas. En esta sección se comprueba que el costo inicial de cualquier cartera de cobertura en un modelo de mercado finito, completo y viable es exactamente $E_Q(\hat{H})$ siendo Q la única probabilidad sin riesgo del modelo. De esta forma coinciden los dos procedimientos anteriores, hecho que está en concordancia con la ley de valoración única (principio económico que asume que el valor de un activo financiero ha de ser único con el fin de evitar arbitrajes).
La **esencia del razonamiento de Black y de Scholes** es que en ausencia de arbitrajes los inversores pueden construir carteras (que se ajusten continuamente debido a las variaciones del mercado) con el fin de compensar el riesgo. El capital que se necesita para construir estas carteras será exactamente la prima de la opción en virtud de la ausencia de arbitrajes.
Las líneas siguientes vienen a justificar tales asertos, al menos en el caso discreto. Sus trabajos de 1973, que se enmarcan en el caso continuo, persiguen, a grosso modo, alcanzar valores explícitos de las primas de opciones europeas. En el siguiente capítulo abordaremos esta cuestión, la cual requiere numerosas herramientas del cálculo estocástico. No obstante, tras exponer los siguientes resultados, mostraremos someramente como del modelo (discreto) binomial se obtienen por paso al límite dichas estimaciones sin necesidad de entrar en temas delicados de la teoría de probabilidad como puede ser la integral de Îto.

Teorema 4.5.1 (Carteras de cobertura en mercados viables) *Sea H una variable replicable en modelo de mercado finito y viable. Entonces se tiene que:*

1. *$\hat{H}$ es Q-integrable*
2. *El proceso de valor actualizado respecto de cualquier cartera de cobertura para H es exactamente la martingala de la esperanza condicionada de $\hat{H}$ sobre Q. Es decir:*

$$\hat{V}_n = E_Q(\hat{H} \| \mathcal{F}_n)$$

Nota 4.5.2 *Del teorema anterior se obtienen las siguientes conclusiones:*

- *Dos cualesquiera carteras de cobertura para H presentan las mismas ganancias en cada instante n.*

- *Se comprueba que realmente las carteras de cobertura son carteras admisibles, pues* $\hat{H}$ *es no negativa.*
- *El costo* V_0 *es exactamente* $E_Q(\hat{H})$*, en efecto:*

$$V_0 = E_Q(V_0) = E_Q(\hat{V}_0) = E_Q(\hat{V}_N) = E_Q(\hat{H})$$

Si estamos en un mercado viable y completo y H representa la ganancia de una opción europea de compra, entonces H es replicable y estamos bajo las hipótesis del teorema anterior. Se tiene por tanto que **el precio racional para una opción europea de compra** en un mercado viable y completo es $E_Q(\hat{H}) = \int_\Omega \hat{H}\, dQ$ donde $H = f(K, S_N) = (S_N - K)^+$

4.6. Del modelo binomial a la fórmula de Black-Scholes

Detengámonos un momento antes de entrar en la dinámica continua.
Uno podría considerar el modelo binomial como una discretización de un modelo a tiempo continuo. Con esta idea en mente no parece descabellado pensar que cuanto más refinada sea la discretización, más se aproximará tal modelo a la realidad. Mostraremos someramente y sin entrar en detalles cómo con una discretización particular se alcanza la fórmula de valoración de Black-Scholes.

Fijemos un horizonte temporal $T > 0$ y un entero positivo N. Como instantes temporales en la discretización tomaremos los nodos equiespaciados de la malla de extremos 0 y T y de anchura $\frac{T}{N}$, esto es: $0, \frac{T}{N}, \frac{2T}{N}, \ldots, T$. Denotaremos por S_k al precio del activo en el nodo k-ésimo. Siguiendo las mismas notaciones que las empleadas en el modelo binomial, consideremos los siguientes escalados para las sucesiones $(a_n)_n$, $(b_n)_n$,$(r_n)_n$:

$$a_{n,N} = e^{\frac{\mu T}{N} - \sigma\sqrt{\frac{T}{N}}}$$

$$b_{n,N} = e^{\frac{\mu T}{N} + \sigma\sqrt{\frac{T}{N}}}$$

$$r_{n,N} = e^{\frac{rT}{N}} - 1$$

No entraremos en el por qué de este escalado, simplemente nos apoyaremos en él y comprobaremos que conduce a la citada fórmula de valoración.

Utilizando las ecuaciones siguientes del modelo binomial

$$P(S_{n,N} = a_{n,N}S_{n-1,N}\|\mathcal{F}_{n-1,N}) = 1 - p_{n,N}$$
$$P(S_{n,N} = b_{n,N}S_{n-1,N}\|\mathcal{F}_{n-1,N}) = p_{n,N}$$

se obtiene que

$$S_N = S_0 e^{\mu T + \sigma\sqrt{T}\frac{2X_N - N}{\sqrt{N}}}$$

donde X_N es el número de veces que el precio del activo sube. Lo más lógico y realista es asumir que el precio sube y baja con igual probabilidad, esto es, X_N se distribuye siguiendo una ley binomial de tamaño N y parámetro $\frac{1}{2}$. De esta forma, en virtud del teorema central del límite se tiene que

$$\log\frac{S_N}{S_0} = \mu T + \sigma\sqrt{T}\frac{X_N - N/2}{\sqrt{N}/2}$$

converge en ley a una normal de media μT y varianza $\sigma^2 T$

Ahora bien, vimos que la verdadera distribución de los precios es irrelevante a la hora de valorar, lo que importaba era trabajar con la probabilidad sin riesgo $Q_N = (q_{1,N}, \ldots, q_{N,N})$. Se comprueba que bajo esta medida X_N es binomial de tamaño N y parámetro $q_{N,N}$.

$$q_{n,N} = \frac{1 + r_{n,N} - a_{n,N}}{b_{n,N} - a_{n,N}}$$

Operando y utilizando desarrollos de Taylor (se omiten las operaciones algebraicas) se tiene que

$$q_{N,N} = \frac{1}{2} - \frac{1}{2}\sqrt{\frac{T}{N}}(\frac{\mu + \frac{1}{2}\sigma^2 - r}{\sigma}) + O(\frac{1}{N})$$

De aquí se deduce inmediatamente que $q_{N,N}(1 - q_{N,N}) \to \frac{1}{4}$ y que

$$\log\frac{S_N}{S_0} = \mu T + \sigma\sqrt{T}\frac{X_N - NqN,N}{\sqrt{N}/2} - \sqrt{N}\frac{\mu + \frac{1}{2}\sigma^2 - r}{\sigma} + O(\frac{1}{\sqrt{N}})$$

converge en ley a una normal de media $(r - \frac{\sigma^2}{2})T$ y varianza $\sigma^2 T$

Como queremos valorar una opción europea de comprar únicamente restar hallar la esperanza de

$$e^{-rT}(S_T - K)^+$$

pero sabemos que $\frac{S_N}{S_0}$ es log-normal de parámetros $(r - \frac{\sigma^2}{2})T$ y $\sigma^2 T$. De nuevo, haciendo las cuentas se obtiene que:

$$S_0\Phi(\frac{\log{(S_0/K)}+(r+\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}})-Ke^{-rT}\Phi(\frac{log(S_0/K)+(r-\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}})$$

donde Φ representa la función de distribución de una ley normal $(0,1)$.
Ésta es la llamada fórmula de Black- Scholes; en el siguiente capítulo llegaremos a ella utilizando directamente un modelo a tiempo continuo.

Capítulo 5

Dinámica continua

5.1. Introducción al caso continuo

La comprensión adecuada de la estructura de los modelos de mercados a tiempo discreto es fundamental para entender, a grosso modo, características y comportamientos de los mercados reales. No obstante tales modelos presentan serias deficiencias; para intentar paliarlas se introducen los modelos de mercado a tiempo continuo.
La forma de proceder es, en esencia, indéntica. La mayor parte de los conceptos se transcriben de forma obvia al lenguaje continuo.

El esquema que seguiremos será análogo al caso discreto:
En principio nos centraremos un único activo financiero, representado su precio en el instante t por $S(t)$. Consideraremos un espacio ambiente de probabilidad $(\Omega, \mathcal{F}, P)$, un intervalo de la recta real, digamos $[0, T]$, que representará el tiempo en el que evolucionan los precios del activo y una filtración $(\mathcal{F}_t)_t$, continua por la derecha, que como siempre viene a representar de forma heurística la historia del proceso de precios; $B = B(t)$ representará el movimiento Browniano estándar. Supondremos también que tanto $S(t)$ como $B(t)$ son adaptados a $(\mathcal{F}_t)_t$

En este caso el **bono** presenta un comportamiento ligeramente distinto: consideramos un tipo de interés **continuo** y constante, esto es, si un bono posee un capital inicial C, al cabo de un tiempo t se habrá visto incrementado hasta Ce^{rt}.
Con mayor generalidad: si $M(t)$ representa una cierta cantidad de dinero a tiempo t, entonces decir que la cantidad M se invierte en un depósito bancario (o se comporta como un bono) a interés continuo y constante r es decir

que $M(t)$ verifica la siguiente ecuación diferencial:

$$M'(t) = rM(t)$$

Reescrita en notación diferencial: $dM(t) = rM(t)dt$, que con condición inicial $M(0) = C$ tiene por solución

$$M(t) = Ce^{rt}$$

En lo que al **comportamiento de los precios** $S(t)$ se refiere seguiremos el planteamiento de Samuelson (1965) y de Black-Scholes (1973): supondremos que el proceso S verifica la siguiente ecuación diferencial estocástica:

$$dS(t) = S(t)(\mu dt + \sigma dB(t))$$

Es decir, tanto la deriva (coeficiente de dt en la combinación lineal $\mu dt + \sigma dB(t)$) como la volatilidad (coeficiente de $dB(t)$) son constantes. Sabemos que su solución es el llamado movimiento Browniano geométrico o exponencial:

$$S(t) = S_0 exp((\mu - \frac{1}{2}\sigma^2)t + \sigma B(t))$$

que ya fue estudiado en el capítulo de cálculo estocástico. Veamos que esta forma de valorar los activos **goza de importantes propiedades matemáticas no necesariamente realistas.**
Tomando logaritmos en la ecuación anterior se tiene que:

$$\log(S(t)) = \log(S_0) + (\mu - \frac{1}{2}\sigma^2)t + \sigma B(t)$$

Es decir, $\log(S(t)) - \log(S(0))$ es un movimiento Browniano con volatilidad σ y con deriva $\mu - \frac{1}{2}\sigma^2$, de donde se sigue que $S(t)$ es log-normal.
También se observa que $S = S(t)$ tiene trayectorias continuas y que $\log(S(t))$ tiene incrementos estacionarios e independientes.
Además, para los precios actualizados se tendrá que $\hat{S}(t) = S_0 exp((\mu - r - \frac{1}{2}\sigma^2)t + \sigma B(t))$, escrito en notación diferencial:

$$d\hat{S}(t) = (\mu - r)\hat{S}(t)dt + \hat{S}(t)dB(t)$$

junto con $\hat{S}(0) = S_0$

Conviene tener en mente las siguientes útiles ideas:
Tal y como se vio en el bloque de cálculo estocástico, el proceso definido puntualmente por $M(t) = exp(\sigma B(t) - \frac{1}{2}\sigma^2 t)$ es una martingala (basta notar que la función $f = exp(\sigma x - \frac{\sigma^2 t}{2})$ verifica que $f_t + \frac{1}{2} f_{xx} = 0$).
Como $S(t) = e^{\mu t} M(t)$ entonces, para $u < t$ se tiene que:

$$E(S(t) \| \mathcal{F}_u) = E(e^{\mu t} M(t) \| \mathcal{F}_u) = e^{\mu t} E(M(t) \| \mathcal{F}_u) = e^{\mu t} M(u) = e^{\mu(t-u)} S(u)$$

De las propiedades de la exponencial se tiene que el proceso de precios $S = S(t)$ es una martingala respecto de la filtracón ambiente si y solamente si $\mu = 0$. Si $\mu > 0$ es una submartingala. En el caso de que $\mu < 0$ es una supermartingala.
En esta línea se tiene el siguiente resultado análogo: el proceso de precios $\hat{S} = \hat{S}(t)$ es una martingala respecto de la filtracón ambiente si y solamente si $\mu = r$. Si $\mu > r$ es una submartingala. En el caso de que $\mu < r$ es una supermartingala.
Pasemos a analizar el concepto de cartera de inversión.

Definición 5.1.1 (Cartera de inversión, caso continuo) *Una cartera de inversión (o estrategia de inversión) es un par de procesos progresivamente medibles a tiempo continuo $(A(t), G(t))$. El primero de ellos representa la cantidad de bono que se poseen a tiempo t y el segundo la cantidad del activo financiero. Se supone además que:*

- $\int_0^T |A(t)|\, dt < \infty$ *P-a.s. , es decir* $A(\omega) \in L^1([0,T], m)$ *P-a.s*
- $\int_0^T G^2(t)\, dt < \infty$ *P-a.s. , es decir,* $G \in \mathcal{H}[0,T]$

El de valor asociado a la cartera es el proceso estocástico V definido como sigue:

$$V(t) = A(t)e^{rt} + G(t)S(t)$$

Nota 5.1.2 *Las siguientes observaciones son necesarias:*

1. *Diremos que una cartera es autofinanciada si la diferencial de su proceso de valor asociado es un proceso de Îto con representación:*

$$dV(t) = A(t)de^{rt} + G(t)dS(t) = rA(t)e^{rt}dt + G(t)dS(t)$$

equivalentemente:

$$V(t) = V(0) + \int_0^t A(u)\, de^{ru} + \int_0^t G(u)\, dS(u)$$

A priori esta condición puede parecer alejada de la autofinanciación discreta: nada más lejos de la realidad, pues para el caso discreto, tal y como vimos, se tiene que

$$\Delta V_n = A_{n+1}\Delta R_n + B_{n+1}\Delta S_n$$

2. *El proceso de valor actualizado será:* $\hat{V}(t) = A(t) + e^{-rt}G(t)S(t)$ *es decir* $\hat{V}(t) = A(t) + G(t)\hat{S}(t)$
3. *Si* V *es el proceso de valor asociado a una estrategia autofinanciada, en virtud de la fórmula de integración por partes de Îto y de que* $\hat{V}(t) = e^{-rt}V(t)$ *se tiene que* $d\hat{V}(t) = G(t)d\hat{S}(t)$.

Se tiene un resultado que caracteriza las carteras de inversión:

Proposición 5.1.3 (Caracterización de las carteras de inversión) *Si el par* $(A(t), G(t))$ *lo forman dos procesos progresivamente medibles a tiempo continuo verificando que*

- $\int_0^T |A(t)|\, dt < \infty$ *P-a.s.*
- $\int_0^T G^2(t)\, dt < \infty$ *P-a.s.*

Entonces el par $(A(t), G(t))$ *es una cartera de inversión autofinanciada si y solamente si*

$$\hat{V}(t) = V(0) + \int_0^t G(x)d\hat{S}(x)$$

para todo $t \in [0, T]$.

5.2. Propiedades de los modelos de mercado continuos

5.2.1. Características generales

Nada más comenzar hemos asumido por una parte que los precios de los activos obedecen a un modelo Browniano exponencial con deriva y volatilidad constantes y por otra que el tipo de interés también es constante. Como ya se ha comentado ésto no siempre es así; cada modelo tiene sus propias peculiaridades que le hacen más o menos útil, más o menos realista.
Un modelo de mercado a tiempo continuo que asume los dos supuestos anteriores recibe el nombre de **modelo de Black-Scholes**. Los conceptos de cartera de inversión y de cartera de inversión autofinanciadas se generalizan

de manera trivial a modelos continuos genéricos.

Veamos ahora cómo se extienden los demás conceptos fundamentales estudiados en los modelos de mercado discretos a un modelo continuo arbitrario:

- El concepto de EMM desempeñaba un papel primordial en los modelos de mercado a tiempo discreto; su extensión razonable a tiempo continuo sigue ocupando un lugar central en la teoría: se dice que una probabilidad Q sobre $(\Omega, \mathcal{F})$ equivalente a P es una **probabilidad sin riesgo** o una medida de martingala equivalente si el proceso de precios actualizados $\hat{S}$ es una Q- martingala local.

  Tomemos una cartera autofinanciada. Por definción su proceso de valor asociado verifica que $d\hat{V}(t) = G(t)d\hat{S}(t)$, es decir $\hat{V}(t) = \hat{V}(0) + \int_0^t G(t)d\hat{S}(t)$. Por lo tanto, decir que Q es una probabilidad sin riesgo es tanto como decir que el proceso de valor asociado a cualquier cartera autofinanciada es una martigala local bajo Q.

- Sea $H \in L^2(\Omega, Q)$ una variable aleatoria no negativa y $\mathcal{F}_T$ -medible. Una cartera autofinanciada que permita obtener un beneficio a tiempo T igual a H se dice que es una **cartera de cobertura** para H (o que cubre el derivado que representa H). Además, se dice que H es **replicable** si admite alguna cartera de cobertura.
- Se dice que un modelo de mercado es **viable** si existe una probabilidad sin riesgo.
- Se dice que un modelo de mercado es **completo** si toda variable de cuadrado integrable respecto de la probabilidad sin riesgo es replicable.
- Se dice que una **cartera de inversión es admisible** si es autofinanciada y su proceso de valor actualizado es siempre no negativo y P-integrable.
- Una **oportunidad de arbitraje** es una cartera autofinanciada con valor inicial nulo y tal que $V(T) \geq 0$ y $P(V(T) > 0) > 0$.
- El **teorema de equivalencia** sigue siendo válido, esto es: el que un modelo de mercado admita una probabilidad sin riesgo es equivalente a la ausencia de arbitrajes.
- La valoración racional de derivados en mercados viables es análoga: se determina a través de una cartera de cobertura. Es más, si H representa el beneficio del derivado, asumiendo que tal variable es

integrable, el valor de cualquier cartera de cobertura en el instante $t \leq T$ es $E_Q(\hat{H} \| \mathcal{F}_t)$. En efecto: $E_Q(\hat{H} \| \mathcal{F}_t) = E_Q(\hat{V}(T) \| \mathcal{F}_t) = \hat{V}(t)$ en virtud de las propiedades de martingala.
De aquí se deduce que el precio racional para el activo será $\hat{V}(0) = V(0) = E_Q(\hat{H}) = \int_\Omega \hat{H}\, dQ$ de donde $H = f(S(T), K)$.

5.2.2. Completitud y viabilidad del modelo de Black-Scholes

Comprobemos que el modelo de Black-Scholes es viable y completo.

- **Viabilidad:**

En virtud de las caracterizaciones anteriores basta construir una medida de martingala equivalente. Los precios actualizados son $\hat{S}(t) = S_0 exp((\mu - r - \frac{1}{2}\sigma^2)t + \sigma B(t))$, luego si W es el proceso definido puntualmente como $W(t) = B(t) + \frac{\mu - r}{\sigma} t$ se tendrá que $\hat{S}(t) = S_0 exp(\sigma W(t) - \frac{\sigma^2 t}{2})$.

Si W fuese un movimiento Browniano $\hat{S}(t)$ sería una martingala (ver la observación posterior al **teorema 2.1.20** de cálculo estocástico), ahora bien, en virtud del teorema de Girsanov **(teorema 1.61.** de cálculo estocástico) la probabilidad Q definida por $\frac{dQ}{dP} = exp((\frac{r-\mu}{\sigma})W(T) - \frac{1}{2}\frac{(r-\mu)^2}{\sigma^2})T)$ (recordemos que la indexación temporal se hacía en $[0, T]$) es una probabilidad equivalente a P respecto de la cual el proceso W se comporta como un movimiento Browniano, esto es, Q es una probabilidad sin riesgo.

Afinando un poco más el razonamiento anterior (utilizando toda la información que aporta el teorema de Girsanov) se obtiene el siguiente resultado:

Proposición 5.2.1 *1. Existe una única probabilidad Q sin riesgo para el modelo de Black-Scholes. Explícitamente viene dada por:*

$$\frac{dQ}{dP} = exp((\frac{r-\mu}{\sigma})W(T) - \frac{1}{2}\frac{(r-\mu)^2}{\sigma^2})T)$$

2. El precio actualizado del activo bajo tal probabilidad verifica que:

$$d\hat{S}(t) = \sigma \hat{S}(t) d\hat{W}(t)$$

- **Completitud:**

Veamos de forma constructiva, aunque sin entrar en pormenores, la completitud del modelo de Black-Scholes.

Tomemos $H \in L^2(Q, \Omega)$; se puede probar que el proceso $M(t) = E_Q(\hat{H} \| \mathcal{F}_t)$ es un elemento de $\mathcal{M}_2[0, L]$ (martingalas de cuadrado integrable sobre $[0, T]$). Por el teorema de representación de Îto (teorema ?? de cálculo estocástico) existe un proceso $X = X(t)$ en el espacio $\hat{\mathcal{H}}$ de suerte que

$$M(t) = E_Q(M(0)) + \int_0^t X(s) dW(s)$$

Notemos que $M(0) = E_Q(M(0))$ ya que $M(0) = E_Q(\hat{H})$.

Siguiendo las notaciones habituales, se define la cartera de componentes

$$A(t) = M(t) - \frac{X(t)}{\sigma}$$
$$G(t) = \frac{X(t)}{\sigma \hat{S}(t)}$$

Se comprueba que está bien definida y que es autofinanciada.

Su valor acutalizado será $\hat{V}(t) = A(t) + \hat{S}(t)G(t) = M(t) - \frac{X(t)}{\sigma} + \frac{X(t)}{\sigma} = M(t)$, luego

$$V(T) = e^{rT} M(T) = e^{rT} E_Q(e^{-rT} H \| \mathcal{F}_T) = E_Q(H \| \mathcal{F}_T)$$

es decir, la cartera replica a H, tal y como queríamos ver.

5.3. Valoración racional de Black-Scholes

Consideremos un derivado financiero sobre un activo en el modelo de Black-Scholes. Sabemos que los réditos que proporciona serán de la forma $H = f(S(T), K)$ siendo como es habitual K el precio de ejercicio y T la fecha de ejercicio o *maturity*. Si el derivado es una opción europea habíamos comprobado que f es tan solo la diferencia entre ambas cantidades orientada (dependiendo de si es de compra o de venta). De forma más general, para los principales tipos de opciones se tienen las siguientes expresiones para f:

- Opción europea de compra: $f = (S_T - K)^+$
- Opción europea de venta: $f = (K - S_T)^+$
- Opción asiática de compra: $f = (\int_0^T S(t)\, dt)^+$
- Opción *lookback* de compra: $f = S_T - \min\limits_{0 \leq t \leq T} S(t)$
- Opción *down and out barrier*: $f = (S_T - K)^+ 1_{\min\limits_{0 \leq t \leq T} S(t) \geq \lambda}$

En una opción de tipo *lookback* la liquidación depende del mínimo y/o del máximo valor del precio del activo subyacente; en una opción asiática las ganancias dependen del promedio de los valores del activo a lo largo de su tiempo de vida. Las opciones de tipo barrera (*barrier*) son aquellas cuyo valor depende del hecho de que el precio del activo ligado alcance o no (por arriba o por debajo) una cierta barrera λ.

Notemos que un tipo especialmente frecuente de opciones, las americanas (tanto de compra como de venta), no están recogidas en la lista. En cuanto a las de venta: debido a que por su propia naturaleza pueden ejercerse en cualquier instante de tiempo, su valoración depende de un proceso de optimización sobre su tiempo de ejercicio, algo que suscita una mayor dificultad. Para las de compra: puede probarse el curioso hecho de que el que se ejercite una opción de compra antes de su fecha de ejercicio no supone ninguna ventaja, luego son equivalentes a las opciones europeas de compra.

Hay tantos tipos de opciones como uno pueda imaginar. A las más comunes, como pueden ser la europea o la americana se las conoce como *vanilla options*. Otros tipos como pueden ser la asiática, las opciones de barrera, opciones rusas, opciones *basket*, etc, componen las llamadas opciones exóticas.

Supongamos que estamos bajo las hipótesis del modelo de Black-Scholes. La notación que emplearemos será la introducida a lo largo de las ecciones anteriores.
Trabajaremos en el espacio $(\Omega, Q, \mathcal{F})$. El proceso de precios en función de W será

$$S(t) = S_0 exp(\sigma W(t) + (r - \frac{\sigma^2}{2})t)$$

Observemos que la deriva μ no aparece, es más, será irrelevante en la valoración.
Por otra parte, es inmediato comprobar que $\log \frac{S(t)}{S(0)}$ sigue una ley normal de parámetros $\{(r - \frac{\sigma^2}{2})t, \sigma^2 t\}$, pues W es un movimiento Browniano estándar. Si Z es una variable aleatoria normal $(0, 1)$ entonces, en virtud de los resultados ya vistos, la prima para una opción cuya liquidez H sea replicable y venga dada por $H = f(S(T), K)$ es exactamente

$$e^{-rT} E(f(S_0 e^{\sigma\sqrt{T}Z + (r - \frac{\sigma^2}{2})T)}, K)$$

En el caso de que el derivado sea una opción europea (por ejemplo de compra) las cosas se simplifican notoriamente. En primer lugar $f = (S_T - K)^+$, luego

H es cuadrado integrable y por ende replicable. Además, la fórmula anterior adquiere el siguiente aspecto:

$$e^{-rT}E((S_0e^{\sigma\sqrt{T}Z+(r-\frac{\sigma^2}{2})T}-K)^+)$$

Por procedimientos de cálculo uno puede obtener una expresión explícita para tal valor:

$$S_0\Phi(\frac{\log(S_0/K)+(r+\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}})-Ke^{-rT}\Phi(\frac{log(S_0/K)+(r-\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}})$$

donde Φ representa la función de distribución de una ley normal $(0,1)$.
De todos modos, se puede llegar más lejos y obtener un valor para cualquier $t\in[0,T]$:
Consideremos una cartera de cobertura (autofinanciada se sobrentiende) para H cuyo valor asociado sea el proceso $V=V(t)$. Supongamos además que $V(t)=F(t,S(t))$.
Si logramos obtener una expresión explícita para F habremos acabado.

Se puede probar que bajo hipótesis muy generales sobre la función $g(x)=f(x,K)$, como por ejemplo que sea continua y derivable a trozos, se tiene que la función $F(t,x)$ es al menos dos veces derivable respecto de x y al menos una vez respecto de t. En nuestro caso, como $g(x)=(x-K)^+$ es continua y derivable a trozos así que se tienen las condiciones de regularidad citadas para F.
Teniendo en cuenta que $dS(t)=S(t)(\mu dt+\sigma dB(t))$, aplicando la fórmula de Îto al proceso V se llega a que:

$$dV(t)=D_1F(t,S(t))dt+D_2F(t,S(t))dS(t)+\frac{1}{2}D_{2,2}F(t,S(t))\sigma^2S(t)^2dt$$

Por ser autofinanciada:

$$dV(t)=A(t)de^{rt}+G(t)dS(t)=(V(t)-G(t)S(t))rdt+G(t)dS(t)$$

Ahora bien, en virtud de la unicidad de la representación de los procesos de Îto se obtienen las siguientes igualdades:

$$D_1F(t,S(t))dt+\frac{1}{2}D_{2,2}F(t,S(t))\sigma^2S(t)^2=(V(t)-G(t)S(t))r$$
$$D_2F(t,S(t))=G(t)$$

Realizando las sustituciones oportunas se llega a llamada **EDP de Black-Scholes**:

$$\frac{\partial F}{\partial t} + \frac{1}{2}\frac{\partial^2 F}{\partial s^2}\sigma^2 s^2 + sr\frac{\partial F}{\partial s} = rF$$

sujeta a que $F(T,s) = f(s,K)$

Si la condición frontera es sencilla, como en el caso de las opciones europeas se puede resolver explícitamente tal ecuación. En otros casos, hay que recurrir forzosamente a métodos numéricos.

En el caso de una opción europea de compra se tiene explícitamente que

$$F(t,s) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{\frac{-y^2}{2}}(xe^{\sigma\sqrt{u}y - \frac{\sigma^2}{2}u} - Ke^{-ru})^+ dy = x\Phi(d_+) - Ke^{-r(T-t)}\Phi(d_-)$$

donde

$$u = T - t$$

$$d_+ = \frac{\log{(x/K)} + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

y

$$d_- = \frac{\log{(x/K)} + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

Capítulo 6

Métodos de valoración basados en el transporte óptimo

6.1. Carencias del modelo de Black-Scholes.

El principal problema que se achaca al paradigma de Black-Scholes es que presupone que el comportamiento de los precios es log-normal. Este hecho es algo realmente rígido y muy poco realista; es por ello que inmediatamente después de los trabajos de Fisher Black y Myron Scholes surgieron diversas alternativas y generalizaciones a este paradigma: los modelos volatilidad local, de volatilidad estocástica, de salto y difusión, etc.

Centrémonos en un derivado financiaro del activo con riesgo $\mathbf{S}$; consideraremos un modelo a tiempo discreto soportado en los instantes $t_1 < t_2 < \ldots < t_N$. Como es habitual denotaremos por S_i el valor del activo a tiempo t_i.
Supongamos que las ganancias que produce el derivado se calculan a través de una función $\Phi = \Phi(S_1, \ldots, S_N)$.
De los capítulos anteriores se deduce que, en un modelo de mercado viable, la prima del derivado será:

$$C = E_Q(\hat{\Phi}(S_1, \ldots, S_N)) = \int_\Omega \hat{\Phi}(S_1, \ldots, S_N) dQ$$

siendo Q una probabilidad sin riesgo, la cual, imponiendo la hipótesis de completitud al mercado, se puede suponer única.

Como se ha visto, la suposición clave del modelo de Black-Scholes es que es conocida la distribución del proceso de precios (salvo dos parámetros, constantes por cierto, la volatilidad y la deriva). A partir de aquí no es difícil encontrar una probabilidad respecto de la cual éstos son una martingala y que además sea equivalente a la probabilidad base, luego, al menos de manera teórica, se conocen todos los elementos de la integral anterior y por ende puede llevarse a cabo su cálculo. Sin embargo, en un modelo de mercado viable y completo genérico, sin suposiciones adicionales es, generalmente, imposible encontrar explícitamente la probabilidad de martingala, luego no es posible hallar la prima del derivado. No obstante, aunque uno no conozca la prima, sí puede preguntarse por cotas inferiores, es decir, **valores tales que la cantidad razonable a pagar para por un derivado quede por encima de ellos.** En concreto, nos interesaremos por cantidades similares a

$$\inf_{\lambda}\{\int_{\Omega}\hat{\Phi}(S_1,\dots,S_N)d\lambda\}$$

donde λ varía en las probabilidades de martingala para el proceso S

- Notemos que en virtud del teorema del cambio de variable, la integral

  $$\int_{\Omega}\hat{\Phi}(S_1,\dots,S_N)dQ$$

  se puede expresar como una integral sobre $\mathbb{R}^N$:

  $$\int_{\mathbb{R}^N}\phi(s_1,\dots,s_N)d(Q\#S)$$

  donde $\phi=\hat{\Phi}$ y la notación $\mu\#f$ siendo f una función y μ una medida se toma directamente de C. Villani; representa la medida inducida por f a partir de μ, esto es: $\mu\#f(A)=\mu(f\in A)$. De esta forma, ahora el inferior se calcularía sobre el conjunto de **todas las probabilidades de martingala y de Borel sobre $\mathbb{R}^N$.**

- Imponemos que nuestro modelo **quede calibrado con las primas de opciones europeas de compra**, esto es, para cada instante t_i, asumimos que conocemos la prima de una opción europea de compra para cualquier precio de ejercicio K (esto no es del todo ficticio, pues pueden estimarse a partir de observaciones de mercado). Recordemos que, denotando por $C(t_i,K)$ a la prima de la acción, se tenía que:

  $$C(t_i,K)=\int_{\Omega}(S_i-K)^+dQ=\int_{\mathbb{R}^+}(s-K)^+d(S_i\#Q)$$

Observemos que estamos obviando el trabajar con los precios descontados: para el propósito que perseguimos en este apartado, es inmediato comprobar que esta elipsis está justificada (y además simplifica las operaciones) siempre que el tiempo de interés sea determinista.
La clave de esta suposición es que permite obtener explícitamente la expresión de la función de distribución de cada $S_i\#Q$, a las cuales, de ahora en adelante, las denotaremos por μ_i. De esta forma, teniendo en cuenta que las marginales de $S\#Q$ son exactamentes las probabilidades $S_i\#Q = \mu_i$, **podemos admitir que conocemos las marginales de la medida de probabilidad $S\#Q$.**

Veámos cómo obtener la ley F de una variable aleatoria X a partir de la función $C(k) = E(X-K)^+$:

Por definición:

$$E(X-K)^+ = \int_{\mathbb{R}} (y-k)^+ dF(y) = \int_k^\infty (y-k)dF(y)$$

En virtud del teorema de Fubini, la última integral se expresa de la siguiente forma:

$$\int_{(k<x<y<\infty)} dxdF(y) = \int_k^\infty (1-F(x))dx$$

es decir:

$$C(k) = \int_k^\infty (1-F(x))dx$$

Puede comprobarse que C es derivable en casi todo punto, por tanto, $C^{'}(k) = -(1-F(k))$ siempre que la derivada esté definida; teniendo en cuenta que F ha de ser monótona, se determina completamente F.

Así pues, nuestro objetivo es hallar el **inferior de las integrales**

$$\int_{\mathbb{R}^N} \phi(s_1,\ldots,s_N)d(Q\#S)$$

sobre las probabilidades de $\mathbb{R}^N$ con marginales $Q\#S_i = \mu_i$ y tales que el proceso S sea una martingala respecto de su propia filtración, esto es: $E(S_{i+1}\|S_i) = S_i$.

Denotaremos a este conjunto $\mathcal{M}(\mu_1, \ldots, \mu_N)$ y al estimador de C por $\hat{C}$; en resumidas cuentas, de todo lo anterior se tiene que:

$$\hat{C} = \inf_{\lambda \in \mathcal{M}(\mu_1, \ldots, \mu_N)} \int_{\mathbb{R}^N} \phi(s_1, \ldots, s_N) d\lambda$$

Veamos ahora, que en el caso en que tengamos dos instantes temporales (para fijar ideas: información de hoy y de mañana), nuestro problema está íntimamente relacionada con el llamado problema del transporte óptimo y por ende con la optimización lineal.

6.2. El transporte óptimo de Kantorovich

Imaginemos que tenemos una montaña de arena y un agujero, ambos con igual volumen, y que deseamos deseamos rellenar este último completamente con la arena.
Es claro que transportar la arena hasta el agujero conlleva cierto trabajo, luego debemos ser cuidadosos con la distribución elegida para el transporte
El **problema del transporte óptimo** consiste, a grosso modo, en hallar cuál es la forma de transportar la arena que minimiza el trabajo.
Por supuesto se puede generalizar a otras muchas situaciones como por ejemplo a modelos de difusión de gases, problemas de optimización lineal, modelos de información, etc.

Matemáticamente el problema del transporte óptimo admite dos formulaciones: la más moderna, la de Kantorovich, y la clásica, debida a Monge.

Tomemos dos conjuntos X e Y, el primero, siguiendo la motivación anterior, representa la región del espacio donde está la arena, y el segundo la región que contiene al agujero. Supondremos además que son espacios de medida; la medida de cada uno de ellos representará, bien la densidad de la disposición de la arena (μ), o bien la densidad de la profundidad del agujero (ν). Normalizando las medidas anteriores (que supondremos finitas) se puede suponer que son espacios de probabilidad.
Denoteremos por $c : X \times Y \to \mathbb{R}$ al coste de transportar el punto $x \in X$ al punto $y \in Y$; generalmente se tendrá que $c(x, y) = h(x - y)$ para alguna función convexa h.

- **La formulación de Kantorovich** está basada en el concepto de **planes de transporte**: un plan de transporte λ es una medida de probabilidad en el espacio producto $X \times Y$ de forma que el peso que asigna a un conjunto $A \times B$ representa la masa de A que se trasnporta a B. Si queremos que se respeten las hipótesis, obviamente ha de tenerse que la marginal de λ sobre X es exactamente μ y que la otra marginal es la probabilidad ν. Denotaremos al conjunto de planes de transporte de marginales μ y ν, siguiendo la notación de C. Villani por $\Pi(\mu, \nu)$. Con esta notación el problema consiste en hallar

$$\mathcal{K} = \inf_{\lambda \in \Pi(\mu,\nu)} \int_{X \times Y} c(x,y) d\lambda$$

- **La formulación de Monge** data de comienzos del s. XIX y centra su atención en funciones en vez de en medidas. En la notación actual consiste en hallar una **función transporte**, esto es, una función $T : X \to Y$ (por supuesto medible) de suerte que $T\#\mu = \nu$ que minimice el coste $c(x,y)$:

$$\mathcal{M} = \inf_{T \text{ t.q. } T\#\mu=\nu} \int_X c(x, T(x)) d\mu$$

Notemos que, por ser $\nu = T\#\mu$ una probabilidad, en virtud del teorema de Fubini (el coste es acotado en $X \times Y$, luego integrable para cualquier medida finita) se tiene que

$$\int_X c(x,T(x))d\mu = \int_Y \int_X c(x,T(x))d\mu d\nu = \int_{X\times Y} c(x,T(x))d(\mu \otimes \nu) =$$

$$\int_{X \times Y} c(x,y) d(\mu \otimes \nu)$$

pues para cualquier conjunto $A \subset Y$ tal que $A \cap T(X) = \emptyset$ se tiene que $\nu(A) = 0$ por defición de ν.
De esta forma se tiene que $\mathcal{M} \geq \mathcal{K}$.

Es claro que si ampliamos el conjunto $\mathcal{M}(\mu_1, \mu_2)$ al conjunto $\Pi(\mu_1, \mu_2)$ estamos en las condiciones del problema de Kantorovich tomando $X = Y = \mathbb{R}$ y $\mu_1 = \mu, \mu_2 = \nu$ y $c(x,y) = \phi(x,y)$. Las siguientes líneas están dedicadas a estudiar ciertas propiedades del nuevo estimador

$$\tilde{C} = \inf_{\lambda \in \Pi(\mu,\nu)} \int_{\mathbb{R}^2} \phi(s_1, s_2) d\lambda$$

y analizar, en algún caso, cómo se pueden generalizar tales argumentos al estudio de $\hat{C}$

1. Veamos en primer lugar que si suponemos que nuestra función ϕ es continua y acotada (lo cual se tiene en casi la totalidad de los casos), entonces el inferior se alcanza:

   - Primeramente consideraremos la topología débil en $\Pi(\mu,\nu)$; esto es, la topología inducida por la convergencia de las integrales
   $$\int_{\mathbb{R}^2} f d\lambda_n \to \int_{\mathbb{R}^2} f d\lambda$$
   donde el conjunto de funciones test f es el de funciones continuas y acotadas: $\mathcal{C}_B(\mathbb{R}^2)$ *(Esta noción de convergencia es la habitual cuando X e Y son espacios de probabilidad polacos).*
   Veamos que $\Pi(\mu,\nu)$ es secuencialmente compacto; para ello, en virtud del teorema de Prohorov basta ver que ajustado (en inglés *«tight»*), lo cual es muy sencillo.
   Se dice que $(\lambda_i)_{i\in I}$ es una familia ajustada si para cada $\varepsilon > 0$ existe un compacto K de suerte que $\lambda_i(K^c) < \varepsilon$ para cada índice i.

   Sea $\varepsilon > 0$, por ser μ y ν probabilidades existirán K y L compactos de $\mathbb{R}$ de suerte que:
   $$\mu(K^c) < \frac{\varepsilon}{2}$$
   y
   $$\nu(L^c) < \frac{\varepsilon}{2}$$
   Veamos que basta tomar el compacto $K \times L$: si λ es una probabilidad de $\Pi(\mu,\nu)$, entonces:
   $$\lambda((K\times L)^c) = \lambda((K^c\times Y)\cup(X\times L^c)) \leq \lambda((K^c\times Y))+\lambda(X\times L^c) =$$
   $$\mu(K^c) + \nu(L^c)$$
   siendo el último sumando menor que ε.
   Un argumento similar prueba que $\mathcal{M}(\mu_1,\mu_2)$ **también es secuencialmente compacto.**
   - Comprobemos ahora que en nuestro caso se alcanza el inferior: Sea (ε_n) una sucesión de números reales monótona decreciente hacia 0. Por la definición de inferior existirá una sucesión de probabilidades $(\lambda_n)_n \subset \Pi(\mu,\nu)$ tales que
   $$I_n = \int_{\mathbb{R}^2} c(x,y)d\lambda_n < \inf_{\lambda\in\Pi(\mu,\nu)} \int c(x,y)d\lambda + \varepsilon_n$$

Por ser $\Pi(\mu,\nu)$ secuencialmente compacto existirá una subsucesión $(\lambda_{n_j})_j$ convergente débilmente hacia λ_0. Para no recargar la notación supondremos que tal subsucesión es la propia (λ_n). Entonces, por ser el coste continuo y acotado se tiene que

$$I_n \to \int_{\mathbb{R}^2} c(x,y)d\lambda_0$$

luego basta ver que $\lambda_0 \in \Pi(\mu,\nu)$, esto es, que λ_0 tiene como marginales a μ y ν, pero , en virtud del teorema de Portmanteau se tiene que

$$\mu(A) = \lambda_n(A \times Y) \to \lambda_0(A \times Y)$$

$$\nu(B) = \lambda_n(X \times B) \to \lambda_0(X \times B)$$

tal y como queríamos ver.
Por tanto:

$$\inf_{\lambda\in\Pi(\mu,\nu)} \int c(x,y)d\lambda = \min_{\lambda\in\Pi(\mu,\nu)} \int c(x,y)d\lambda = \int_{\mathbb{R}^2} c(x,y)d\lambda_0$$

Si admitimos que $\mathcal{M}(\mu_1,\mu_2)$ es no vacío, la misma demostración permite probar que se alcanza el mínimo; ahora bien, la suposición anterior es un tanto delicada. Claramente $\Pi(\mu_1,\mu_2)$ es siempre no vacío, pues contiene a la probabilidad $\mu_1 \otimes \mu_2$, sin embargo, a priori, nada nos asegura que no lo sea $\mathcal{M}(\mu_1,\mu_2)$.
Conviene introducir la siguiente noción:
Se dice que la probabilidad μ_1 *es menor que* μ_2 *en el orden convexo* si para cada par de variables aleatorias X e Y cuyas leyes sean μ_1 y μ_2 respectivamente, se tiene que

$$E(\phi(X)) \leq E(\phi(Y))$$

para cada función convexa ϕ, o lo que es lo mismo :

$$\int \phi d\mu_1 \leq \int \phi d\mu_2$$

Un teorema debido a Strassen afirma que $\mathcal{M}(\mu_1,\mu_2)$ es no vacío si y solamente si μ_1 es menor que μ_2 en el orden convexo.

La condición necesaria es evidente en virtud de la desigualdad de Jensen; sean $\pi \in \mathcal{M}(\mu_1, \mu_2) \neq \emptyset$ y ϕ una función convexa, entonces:

$$\int \phi(y) d\mu_2(y) = \int \phi(y) d\pi(x, y) =$$

$$= \int \int \phi(y) d\pi_y(y) d\mu_1(x) \geq \int \phi(x) d\mu_1(x)$$

donde π_y representa la marginal de la probabilidad π sobre su segunda componente; la última desigualdad es la desigualdad de Jensen.
Para ver la otra implicación, mucho más complicada, es recomendable seguir la referencia [10]

- En el caso de que $\phi(s_1, s_2) = |s_1 - s_2|^p$ puede comprobarse que:

$$\min_{\lambda \in \Pi(\mu,\nu)} \int_{\mathbb{R}^2} \phi(x, y) d\lambda = \int_0^1 |\mu^{-1}(t) - \nu^{-1}(t)|^p dt$$

siendo μ^{-1} y ν^{-1} las funciones cuantiles de μ y ν respectivamente. No entraremos en detalles.

6.3. Programación lineal

Veamos cómo obtener explícitamente y de manera sencilla el valor $\hat{C}$ en el caso en que las marginales $\mu = \mu_1$ y $\nu = \mu_2$ sean probabilidades discretas sobre $\mathbb{R}$ utilizando métodos de programación lineal.

Supondremos que μ está concetrada en $X = \{x_1, x_2, \ldots, x_I\} \subset \mathbb{R}$, que a cada punto x_i le da peso p_i y, que ν está concetrada en el conjunto $Y = \{y_1, y_2, \ldots, y_J\} \subset \mathbb{R}$ y asigna al punto y_j el peso q_j.
Sea λ una probabilidad sobre $\mathbb{R}^2$ centrada en $X \times Y$; llamaremos $\lambda_{i,j}$ al peso del punto (x_i, y_j).
Por ser λ una probabilidad se tendrá que:

$$\sum_{I \times J} \lambda_{i,j} = 1$$

y que $\lambda_{i,j} \geq 0$ para cada par de índices $(i, j) \in I \times J$. Imponiendo que μ y ν sean sus marginales se tiene que:

$$\sum_{i\in I} \lambda_{i,j} = q_j$$

y que

$$\sum_{j\in J} \lambda_{i,j} = p_i$$

para todos los índices i y j.

La condición de martingala se traduce en que:

$$\sum_{j\in J} y_j P(S_2 = y_j \| S_1 = x_i) = \sum_{j\in J} y_j \frac{\lambda_{i,j}}{p_i} = x_i$$

es decir:

$$\sum_{j\in J} y_j \lambda_{i,j} = p_i x_i$$

para cada índice $i \in I$.

De esta forma se tiene que hallar

$$\hat{C} = \inf_{\lambda\in\mathcal{M}(\mu_1,\ldots,\mu_N)} \int_{\mathbb{R}^N} \phi(s_1,\ldots,s_N) d\lambda$$

consiste exactamente en minimizar

$$\sum_{I\times J} \phi(x_i, y_j)\lambda_{i,j}$$

sujeto a :

- $\sum_{I\times J} \lambda_{i,j} = 1$
- $\lambda_{i,j} \geq 0$ con $(i,j) \in I \times J$
- $\sum_{i\in I} \lambda_{i,j} = q_j$ para cada $j \in J$
- $\sum_{j\in J} \lambda_{i,j} = p_i$ para cada $i \in I$
- $\sum_{j\in J} y_j \lambda_{i,j} = p_i x_i$

Como se presuponen conocidos tanto la función ϕ como los x_i y los y_j, el problema anterior es entonces un problema clásico de programación lineal que puede resolverse de la forma habitual: utilizando el algoritmo símplex.

No obstante, la formulación del problema anterior cuenta con un total de $I \times J$ variables, esto es, sin entrar en un análisis del coste computacional del algoritmo, algo totalmente inaccesible para su implementación en la mayor parte de los casos.
De esta forma, salvo para valores realmente pequeños de I y J (provenientes de una situación en general nada realista), la utilización del símplex es totalmente ineficaz; aún así, las matrices que intervienen en el problema presentan una estructura extremadamente dispersa, lo cual abre la puerta a la impletación de otros algoritmos.

Capítulo 7

Referencias bibliográficas

Al igual que ocurría en la sección de cálculo estocástico, hay una elevada cantidad de libros escritos sobre estos temas; citaremos tres:

- El libro de Elliot y Kopp ([6]) es una excelente referencia general para los dos primeros capítulos.
  En el capítulo primero (*"Pricing by arbitrage"*), se exponen los conceptos previos necesarios y varias reflexiones generales que sirven de motivación; además, se aborda el modelo de mercado binomial.
  En el segundo capítulo (*"Martingales measures"*), se presenta la teoría de valoración a través de martingalas en modelos de mercado finitos; el teorema de equivalencia se expone en el capítulo tercero: *"The first fundamental theorem"*. Las nociones fundamentales de nuestro capítulo de dinámica continua también se plasman en este libro, en el capítulo 7: *"Continuous-time european options"*. Todas las demostraciones del capítulo 5 de estas notas se encuentran en este libro.
- El libro de Musiela y Rutkowski ([8]), es una amplia referencia para muchos temas de matemática financiera; su segundo capítulo (*"Discrete-time security markets"*) aborda modelos de mercado finitos. El modelo de Black-Scholes se encuentra en el tercer capítulo: *"Benchmark models in cotinuous time"*. Las demostraciones omitidas del capítulo 6 se encuentran en esa sección.
- En lo que al transporte óptimo se refiere: el libro de Villani [9], es una referencia obligada.

Bibliografía

[1] Billingsley, P. (1986). *Probability and measure.* Second edition. Ed. John Wiley sons.

[2] Karatzas, I. y Shreve, S.E. (1991). *Brownian Motion and Stochastic Calculus.* Second edition. Ed. Springer

[3] Chung, K.L. y Williams, R.J. (1990). *Introduction to stochastic integration.* Second edition. Ed. Birkhauser, London.

[4] Elliot, R.J. (1982). *Stochastic calculus and applications.* Ed. Springer.

[5] Nualart, D. (2006) *The Malliavin calculus and related topics.* Second edition. Ed. Springer.

[6] Elliot, R.J. y Kopp, P.E. (2004). *Mathematics of financial markets.* Second edition. Ed. Springer.

[7] Van der Vaart, A. W. (2012) *Martingales, diffusions and financial mathematics.* Versión electrónica. http://www.math.vu.nl/sto/onderwijs/fw/stochint.pdf

[8] Musiela, M. y Rutkowski, M. (2004). *Martingale methods in financial modelling.* Second edition. Ed. Springer.

[9] Villani, C. (2003). *Topics in Optimal Transportation.* First edition. American Mathematical Society.

[10] Beiglbock, M. y Juillet, N. (2012). *On a problem of optimal transport under marginal martingale constraints.* arXiv.