UNIVERSIDAD TECNOLÓGICA DE LA MIXTECA

“NUEVAS SOLUCIONES PARA ALGUNOS MODELOS FINANCIEROS CON VOLATILIDAD NO CONSTANTE”

TESIS

PARA OBTENER EL TÍTULO DE:

LICENCIADO EN MATEMÁTICAS APLICADAS

PRESENTA:

NELSON GARCÍA PÉREZ

DIRECTOR:

M.C. JUAN CARLOS MENDOZA SANTOS

HUAJUAPAN DE LEÓN, OAXACA, JULIO DE 2010

Nuevas Soluciones para Algunos Modelos Financieros con Volatilidad no Constante

Nelson García Pérez

Julio de 2009

A mis padres, Hipolito y Victorina,

a mis hermanos, Arturo, Cielo, Omar y Marisa.

Agradecimientos

Primeramente, manifiesto mi más sincero agradecimiento a mis padres y hermanos, cuyo ejemplo y apoyo incondicional fueron factores principales para no desistir ante tropiezos.

La aportación de este trabajo no sería posible sin el apoyo y las sugerencias de mi asesor de tesis, el M.C. Juan Carlos Mendoza Santos, por lo que le doy gracias por su confianza y dedicación.

Agradezco a mis sinodales, el M.C. Adolfo Maceda Méndez, el M.C. Jose del Carmen Jiménez Hernández y el M.C. Vulfrano Tochihuitl Bueno, por su tiempo y por sus siempre útiles comentarios al respecto.

Expreso mi agradecimiento al Dr. Francisco Solano Tajonar Sanabria, por sus valiosas sugerencias en la revisión y recopilación del material bibliográfico, las cuales me permitieron dar estructura al presente trabajo. Así mismo, agradezco al Dr. Erick Treviño Aguilar, a quien también debo mi gran interés por el formalismo de este trabajo.

A todas aquellas personas e instituciones que han contribuido, directa o indirectamente, en el desarrollo del presente trabajo, mi más sincero agradecimiento.

Índice general

Introducción

En el afán de lograr una mayor comprensión del mundo actual, las técnicas matemáticas y estadísticas han mostrado ser de suma utilidad. La modelización de la realidad económica no ha sido ajena a esa incursión de técnicas, pero esta matematización está mucho más presente en el campo de la economía financiera. Por ejemplo, como caso muy particular, lo anterior está reflejado en las operaciones realizadas en el Mercado Mexicano de Derivados, S.A. de C.V.

Aunque la estadística y los mercados financieros, si se tiene en cuenta las manifestaciones de Andrew W. Lo[1], han estado unidos de manera indisoluble desde la publicación en 1565 de *"Liber de Ludo Aleae"* por Cardano, generalmente se prefiere remontarse a una fecha mucho más reciente como la de 1900. En dicho año, Louis Bachelier presenta su tesis doctoral *"Théorie de la Spéculation"* en donde establece, en terminología actual, que los precios financieros en un instante t están gobernados por un proceso estocástico de incrementos independientes y que se distribuyen de acuerdo a un modelo gaussiano o normal. Este trabajo de Bachelier contenía ideas cuya maduración y consolidación teórica tuvo que esperar a los trabajos fundamentales de matemáticos como Norbert Wiener, Paul Lévy, Andrey Nikolaevich Kolmogorov, Joseph Leo Doob, Kiyosi Itô y Paul-André Meyer, en un periodo que va desde los años veinte hasta los años sesenta del pasado siglo, trabajos de una elevada dificultad matemática, lo que también ha contribuido a su difusión retardada entre los economistas y financieros.

Otro acontecimiento importante se da alrededor de 1960, que es cuando Paul Samuelson descubre la tesis de Bachelier. Gracias a este hecho, en 1965 Samuelson[2] introduce el movimiento Browniano geométrico, y es hasta esta fecha que comienza el periodo heroico de la Matemática Financiera.

El estudio de los mercados financieros, cada vez más desarrollados y con productos cada vez más complejos, exige el desarrollo de una metodología para la determinación de los

[1] Lo, A. W. (2000). " Finance: A selective survey". JASA, vol 95. No. 450, p. 629.

[2] Samuelson, P. (1965). "Proof that properly anticipated prices fluctuate randomly". Industrial Management Review 6, p. 41-50.

precios de los activos correspondientes y para el análisis de los mercados. El modelo de Bachelier, como las primeras incursiones en un territorio inhóspito, dista mucho de ser definitivo.

Es importante resaltar que el movimiento Browniano, así como sus transformaciones y generalizaciones que se obtienen mediante la manipulación de la herramienta matemática del cálculo estocástico, forman la piedra angular de la modelización estocástica de los mercados financieros de la actualidad. Haciendo uso de estas nuevas metodologías, F. Black y M. Scholes obtienen a principios de los años 70 una fórmula para la determinación de los precios de un tipo de opciones financieras. La fórmula fundamental que proponen Black y Scholes[3] corresponde a la solución de una ecuación diferencial parcial de segundo orden con coeficientes constantes, parabólica y lineal, ecuación que ellos mismos determinan y resuelven.

Desde la propuesta de la ecuación del modelo de Black-Scholes, han surgido una gran variedad de modelos de valuación de opciones. Algunos de ellos plantean extensiones del modelo original con el propósito de adecuarlo a la valuación de opciones distintas a las contempladas inicialmente (por ejemplo, opciones exóticas), y otros más se construyen con el fin de obtener modelos más cercanos a la realidad.

Como caso particular, los analistas financieros han notado que los coeficientes que aparecen en la ecuación obtenida por Black y Scholes no deben permanecer constantes como se propone originalmente. Bajo este contexto, tanto la aplicación como las propuestas de métodos que permitan obtener soluciones analíticas de ecuaciones diferenciales parciales adquiere especial interés por parte de los inversionistas, economistas y financieros.

En esta tesis de licenciatura se está interesado en cumplir dos objetivos centrales. El primer objetivo es presentar, desde un enfoque clásico, los fundamentos matemáticos de los derivados financieros, permitiendo de este modo que en trabajos futuros los interesados se centren casi directamente en la aplicación. Para cumplir tal objetivo, los primeros cuatro capítulos de la tesis están organizados como a continuación se describe.

En el Capítulo 1 se introducen los conceptos y resultados fundamentales de la teoría formal de la Probabilidad. Éste capítulo está estructurado de tal manera que permita al lector interesarse en los fundamentos rigurosos de la Probabilidad Clásica.

En el Capítulo 2 se exponen las características básicas del movimiento Browniano, se introduce la construcción de la integral estocástica y se presenta la fórmula fundamental del cálculo estocástico: la fórmula de Itô. En el presente se resalta el carácter complejo de estos temas, así como su importancia en la modelización del mercado financiero.

[3]Black, F. and M. Scholes (1973). " The pricing of options and corporate liabilities". Journal of Political Economy, **81**, pp. 637-654.

En el Capítulo 3 se realiza una breve descripción de los aspectos teóricos de las ecuaciones diferenciales estocásticas, necesarios para la comprensión de los modelos que describen el comportamiento de los mercados financieros.

Para concretar la aplicación de los resultados previamente revisados, en el Capítulo 4 se analiza la estructura teórica de los modelos financieros. Debido a la complejidad, primero se resalta el análisis de los modelos de un sólo periodo y luego se da una introducción a los modelos en tiempo continuo. Este capítulo concluye con el análisis detallado del modelo propuesto por Black y Scholes desde el enfoque de las medidas de riesgo neutral.

Respecto al segundo objetivo central, se generaliza la aplicación del método propuesto por J. F. Harper en [7] que permite resolver la ecuación de Black y Scholes, y se proponen dos soluciones generales para la obtención de soluciones analíticas de ecuaciones diferenciales parciales con coeficientes dependientes del tiempo que tienen la forma de la ecuación de Black y Scholes. Además, utilizando las soluciones que aquí se proponen, se resuelven de manera explícita dos ecuaciones diferenciales parciales conocidas. Todo lo anterior está incluido en el Capítulo 5.

Con el propósito de que el contenido básico de la tesis sea abordado sin la necesidad de depender de muchos otros trabajos, se incluye un conjunto de apéndices en donde se mencionan los principales resultados utilizados y que no son definidos en el desarrollo de la tesis. Así mismo, en esta última parte se incluye las notaciones estándar utilizadas, principalmente, en los primeros tres capítulos.

Capítulo 1
Probabilidad

Existen escritos que muestran que el término *probabilidad* ya se empleaba desde el siglo XVI. Publicaciones sobre oportunidades en juegos datan a G. Cardano (1501-1576). El término probabilidad también aparece en los trabajos de J. Kepler (1571-1630) y de G. Galilei (1564-1642).

Los historiadores de las matemáticas están de acuerdo en considerar como el origen de la teoría de la probabilidad a los trabajos de B. Pascal (1623-1662) y de P. Fermat (1571-1630). Sus primeros estudios se relacionan con algunos problemas implícitos en los juegos de azar. Fue el matemático holandés C. Huygens (1629-1695) quien recopiló las ideas de Pascal y Fermat y escribió un libro que tuvo gran influencia en sus lectores. En el año 1685, Jacques Bernoulli (1654-1705) propuso problemas de probabilidad para los cuales fue necesario desarrollar una teoría que permitiera resolverlos. Después de los trabajos de J. Bernoulli y de A. De Moivre (1667-1754), muchos matemáticos de aquella época empezaron a trabajar sobre los problemas de probabilidad, incluyendo a D. Bernoulli (1700-1782), L. Euler (1707-1803), C. F. Gauss (1777-1855), y P. S. Laplace (1749-1827).

En el año 1900, durante el Segundo Congreso Internacional de Matemáticas realizado en París, David Hilbert (1826-1943) señaló como uno de los problemas matemáticos más importantes, *la fundamentación del cálculo de probabilidades.* Después de algunos años, el matemático soviético Andrei Nikolaevich Kolmogorov (1903-1987) encontró la conexión entre las ideas de Borel, Lebesgue y la probabilidad, y en 1933 publicó su teoría axiomática del cálculo de probabilidades.

La *teoría de la probabilidad* es la rama de la matemática cuyo objetivo es describir e investigar modelos matemáticos de fenómenos aleatorios. Esta teoría está provocando gran interés para su estudio debido a sus exitosas aplicaciones a muchas áreas de la física, biología, ciencias sociales e ingenierías.

Un ejemplo en donde el uso de la probabilidad resulta provechoso es en el área de Economía, en particular en la teoría de las Finanzas. Aquí, las tasas de interés y el valor de los activos (tales como acciones, bonos, divisas) son afectados por cambios aleatorios, pero sujetos a leyes de probabilidad específicas. Usar estos modelos permite proveer productos seguros a los inversionistas.

En esta primera parte presentamos los conceptos básicos de la teoría de la Probabilidad que permitirán comprender los siguientes capítulos. Además, se hace notar la necesidad de estudiar los conceptos básicos de la teoría de la Medida y otras herramientas matemáticas que permiten un análisis riguroso.
Para ejemplos, se recomienda al lector revisar la bibliografía básica introductoria de la probabilidad que más le agrade.

1.1. Eventos y sus probabilidades

El concepto básico sobre el que se desarrolla toda la teoría de probabilidad es el de experimento aleatorio.

Definición 1.1.1. *Un **experimento aleatorio** es un experimento que presenta las siguientes propiedades:*

a) *Tiene al menos dos posibles resultados.*

b) *El conjunto de posibles resultados se conoce antes de que el experimento se realice.*

c) *Puede repetirse esencialmente bajo las mismas condiciones.*

De especial importancia en el cálculo de probabilidades es el conjunto de todos los posibles resultados.

Definición 1.1.2. *El conjunto de todos los posibles resultados de un experimento aleatorio es llamado* **espacio muestral***, y será denotado por* Ω*.*

Ahora se pretende caracterizar a los posibles resultados del experimento que contienen información relevante del experimento, para ello, se denotará por $\mathcal{F}$ a tal colección de subconjuntos de Ω. Para elementos de $\mathcal{F}$ se define nuevos elementos generados por las operaciones usuales de conjuntos (unión, intersección y complemento). Es razonable solicitar que $\mathcal{F}$ sea cerrado bajo operaciones finitas de elementos en $\mathcal{F}$, en este caso se dice que la colección $\mathcal{F}$ es una ***álgebra*** de Ω.

Si Ω es finito, toda la información relevante del experimento pertenece a la álgebra $\mathcal{F} = \mathcal{P}(\Omega)$. Sin embargo, cuando Ω es infinito se necesita extender las operaciones a una colección numerable de elementos en $\mathcal{F}$.

Para cubrir ambos casos, se pide que la colección $\mathcal{F}$ sea cerrada no sólo bajo operaciones finitas, sino también bajo operaciones numerables. A ésta nueva colección se le conoce como *σ-álgebra*, siendo éste concepto otro de los pilares de la teoría matemática de la probabilidad.

El cálculo de la probabilidad de algún elemento de $\mathcal{F}$ requiere, en general, del conteo de los resultados que sean favorables a la ocurrencia de dicho elemento. Por tal razón, el concepto de σ-*álgebra* requiere de un procedimiento para contar, el cual está disponible en el concepto de sucesión.

Debido a los objetivos que se persiguen, aquí la σ-*álgebra* será definida para espacios muestrales, pero la definición es válida para conjuntos arbitrarios.

Definición 1.1.3. *Sean Ω un espacio muestral no vacío y $\mathcal{F} \subseteq \mathcal{P}(\Omega)$. La colección $\mathcal{F}$ es llamada una* σ***-álgebra*** *(σ-field) de Ω si cumple las siguientes condiciones:*

a) $\Omega \in \mathcal{F}$.

b) *Si $A \in \mathcal{F}$, entonces $A^c \in \mathcal{F}$.*

c) *Si (A_n) es una sucesión numerable de conjuntos en $\mathcal{F}$, entonces*

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}.$$

A la pareja $(\Omega, \mathcal{F})$ se le llama ***espacio medible*** *y a los elementos de $\mathcal{F}$ se les llama* ***eventos*** *(o* ***conjuntos medibles****).*

La pareja $(\Omega, \mathcal{F})$ es llamada *espacio medible* debido a que es posible definir una medida de probabilidad sobre dicho espacio. El conjunto vacío $\emptyset$ es llamado ***evento imposible,*** el conjunto $\{\omega\}$ que contiene un sólo posible resultado ω de Ω es conocido como ***evento elemental***, y Ω es el ***evento seguro***.

Las σ-álgebras son utilizadas para representar estructuras de información. Por ejemplo, si $\mathcal{F} = \{\emptyset, \Omega\}$ se carece completamente de información, si $\mathcal{F} = \mathcal{P}(\Omega)$ se cuenta con información completa. En general, no hay experimentos que carezcan completamente de información, ni que lo sepan todo, así que una σ-álgebra intermedia podría representar lo poco o mucho de la información disponible.

Algunas propiedades de las σ-álgebras de Ω se mencionan a continuación:

1. $\emptyset \in \mathcal{F}$.
2. Si $A, B \in \mathcal{F}$, entonces $A - B = A \cap B^c \in \mathcal{F}$.
3. Si (A_n) es una sucesión de eventos en $\mathcal{F}$, entonces
$$\bigcap_{n=1}^{\infty} A_n \in \mathcal{F}.$$
4. La intersección arbitraria de σ-álgebras de Ω es nuevamente una σ-álgebra de Ω.

La demostración de la propiedad 4 es análoga al caso de dos σ-álgebras, por tal razón, aquí sólo se demostrará este caso.

Teorema 1.1.1. *La intersección de dos σ-álgebras de Ω es una σ-álgebra de Ω.*

Demostración.
Sean $\mathcal{F}_1$ y $\mathcal{F}_2$ dos σ-álgebras de Ω. Se demostrará que $\mathcal{F} = \mathcal{F}_1 \cap \mathcal{F}_2$ es una σ-álgebra de Ω.

a) Puesto que $\mathcal{F}_1$ y $\mathcal{F}_2$ son σ-álgebras de Ω, entonces $\Omega \in \mathcal{F}_1$ y $\Omega \in \mathcal{F}_2$. Así, $\Omega \in \mathcal{F}_1 \cap \mathcal{F}_2 = \mathcal{F}$.

b) Sea $A \in \mathcal{F}$, entonces $A \in \mathcal{F}_1$ y $A \in \mathcal{F}_2$. Luego, $A^c \in \mathcal{F}_1$ y $A^c \in \mathcal{F}_2$. Por lo tanto, $A^c = \mathcal{F}_1 \cap \mathcal{F}_2 = \mathcal{F}$.

c) Sea (A_n) una sucesión de elementos en $\mathcal{F}$, entonces $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}_1$ y $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}_2$. En consecuencia,
$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}_1 \cap \mathcal{F}_2 = \mathcal{F}.$$

Por lo tanto, $\mathcal{F}$ es una σ-álgebra de Ω. ■

El resultado anterior garantiza que la siguiente definición tenga sentido.

Definición 1.1.4. *Sea $\mathcal{C}$ una colección de subconjuntos de Ω. La* ***mínima σ-álgebra generada por*** *$\mathcal{C}$, denotada por $\sigma(\mathcal{C})$, es la colección*

$$\sigma(\mathcal{C}) = \bigcap \{\mathcal{F} : \mathcal{F} \text{ es } \sigma\text{-álgebra de } \Omega \text{ y } \mathcal{C} \subseteq \mathcal{F}\}.$$

Un ejemplo particular de mínima σ-álgebra que contiene todos los intervalos abiertos, cerrados, semiabiertos de $\mathbb{R}$ es la ***σ-álgebra de Borel***, definida como

$$\mathcal{B}(\mathbb{R}) = \sigma\{(a,b) \subseteq \mathbb{R} : a < b\} = \sigma\{(-\infty, b], b \in \mathbb{R}\} = \sigma\{(-\infty, b), b \in \mathbb{R}\}.$$

A los elementos de $\mathcal{B}(\mathbb{R})$ se les llama *conjuntos de Borel, Borelianos o conjuntos Borel medibles.*

Toca turno definir las probabilidades que son asignadas a los eventos. Sólo se tienen que asignar probabilidades a los elementos de la σ-álgebra, de otra manera se estaría simplemente trabajando de más.

Definición 1.1.5. *Una* **medida de probabilidad** *en un espacio medible $(\Omega, \mathcal{F})$ es una función $\mathbb{P} : \mathcal{F} \to [0,1]$ que satisface:*

a) $\mathbb{P}(\Omega) = 1$.

b) *Para cada sucesión (A_n) de elementos de $\mathcal{F}$, disjuntos a pares (esto es, $A_n \cap A_m = \emptyset$ si $n \neq m$), se cumple*
$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

La terna $(\Omega, \mathcal{F}, \mathbb{P})$ es llamada **espacio de probabilidad** *y $\mathbb{P}(A)$ se llama* **probabilidad de** *A.*

La medida $\mathbb{P}(A)$ cuantifica la creencia acerca de la ocurrencia del evento $A \in \mathcal{F}$. Si $\mathbb{P}(A) = 1$, diremos que "A ocurre con probabilidad 1" o "A ocurre casi seguramente (c.s.)".

Combinando las dos propiedades anteriores se obtiene que $\mathbb{P}(\emptyset) = 0$. La propiedad **b)** se conoce como ***σ-aditiva*** o ***aditividad numerable***. Ésta propiedad no puede extenderse a una colección no numerable. Si se hiciera, considérese por ejemplo $\Omega = [0, 1]$ y $\mathcal{F} = \mathcal{P}(\Omega)$, entonces se esperaría que

$$\mathbb{P}([0,1]) = \sum_{x\in\,[0,1]} \mathbb{P}(\{x\}),$$

pero como se verá en secciones posteriores, ésto es falso.

Lo anterior expresa que no es posible definir una medida de probabilidad para todos los posibles subconjuntos de un conjunto no numerable Ω. La demostración de la existencia de conjuntos en donde no es posible definir una medida de probabilidad involucra el *axioma de elección* (véase Apéndice A.4), y no es posible definir tales conjuntos de forma explícita.

A continuación se mencionan algunas propiedades simples pero importantes que cumplen las medidas de probabilidad.

Lema 1.1.1. *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y $A, B \in \mathcal{F}$. Entonces*

a) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

b) $\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A - B)$.

c) *(Propiedad monótona) Si $A \subseteq B$, entonces* $\mathbb{P}(A) = \mathbb{P}(B) - \mathbb{P}(B - A) \leq \mathbb{P}(B)$.

d) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

e) *(Aditividad finita) Si $A_1, \ldots, A_n \in \mathcal{F}$ son disjuntos a pares, entonces*

$$\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mathbb{P}(A_k).$$

f) *(Desigualdad de Boole) Para cualquier sucesión de eventos (A_n),*

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

Demostración.

Sólo se demostrá la propiedad **d)**.

d) Para A, B eventos, el evento

$$A \cup B = (A - B) \cup (A \cap B) \cup (B - A) = (A - A \cap B) \cup (A \cap B) \cup (B - A \cap B)$$

está formado por la unión de tres eventos disjuntos a pares.

Por la propiedad b),

$$\begin{aligned} \mathbb{P}(A \cup B) &= \mathbb{P}(A - A \cap B) + \mathbb{P}(A \cap B) + \mathbb{P}(B - A \cap B) \\ &= \mathbb{P}(A) - \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B) + \mathbb{P}(B) - \mathbb{P}(A \cap B) \\ &= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B). \end{aligned}$$

Por lo tanto,

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

■

A la desigualdad de Boole también se le conoce como propiedad **σ-subaditiva**. Puesto que la medida de probabilidad es no negativa, nótese que de la propiedad **d)** se tiene que

$$\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B),$$

siendo la anterior un caso particular de la desigualdad de Boole.

Cualquier evento A que tiene probabilidad cero, esto es $\mathbb{P}(A) = 0$, es llamado ***evento nulo***. Parece razonable suponer que cualquier subconjunto B de un evento nulo A será por si mismo nulo, pero esto puede no tener sentido, puesto que B puede no ser un evento, y entonces $\mathbb{P}(B)$ puede no estar definido.

Definición 1.1.6. *Un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ es **completo** si todos los subconjuntos de los eventos nulos son eventos. En otro caso, se dice que el espacio es **incompleto**.*

Cualquier espacio incompleto puede ser completado. Los detalles de este importante hecho requiere de conceptos sofisticados de la teoría de la Medida. En esencia, si $\mathcal{N}$ es la colección de todos los subconjuntos de los eventos nulos en $\mathcal{F}$, se considera $\mathcal{G} = \sigma(\mathcal{F} \cup \mathcal{N})$. Luego, para cualquier evento G en $\mathcal{G}$,

$$\mathbb{P}^*(G) := \text{ínf}\{\mathbb{P}(F) : F \in \mathcal{F}, G \subseteq F\} \tag{1.1}$$

define una medida de probabilidad en $(\Omega, \mathcal{G})$, siendo éste una extención de $\mathbb{P}$, ya que $\mathbb{P}^*(G) = \mathbb{P}(G)$ para todo evento $G \in \mathcal{F} \subseteq \mathcal{G}$ y $\mathbb{P}^*(B) = 0$ para todo subconjunto B de eventos nulos en $\mathcal{F}$. Al espacio de probabilidad $(\Omega, \mathcal{G}, \mathbb{P}^*)$ se le conoce como la **completación de** $(\Omega, \mathcal{F}, \mathbb{P})$.

Ejemplo 1.1. *Sean $\Omega = \{a, b, c\}$, $\mathcal{F} = \Big\{\emptyset, \{a\}, \{b, c\}, \Omega\Big\}$. Para $A \in \mathcal{F}$ arbitrario, se define*

$$\mathbb{P}(A) = \begin{cases} 1 & si\ a \in A, \\ 0 & si\ a \notin A. \end{cases}$$

Para este ejemplo en particular, $\emptyset$ y $\{b, c\}$ son los conjuntos nulos. Así, $\mathcal{N} = \Big\{\emptyset, \{b\}, \{c\}, \{b, c\}\Big\}$.
Luego,

$$\begin{aligned} \mathcal{G} = \sigma(\mathcal{F} \cup \mathcal{N}) &= \Big\{\{a, b\}, \{a, c\}\Big\} \cup \mathcal{N} \cup \mathcal{F} \\ &= \Big\{\emptyset, \{a\}, \{b\}, \{c\}, \{b, c\}, \{c, a\}, \{a, b\}, \Omega\Big\}. \end{aligned}$$

Utilizando (1.1), no es complicado definir explícitamente la nueva medida de probabilidad $\mathbb{P}^*$ *en* $(\Omega, \mathcal{G})$*, aquí se omiten los detalles.* $\square$

Ejemplo 1.2. *Si* Ω *es a lo más numerable y* $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ *es un espacio de probabilidad, entonces el espacio de probabilidad es completo.* $\square$

En lo sucesivo, por *probabilidad* nos referiremos a una *medida de probabilidad;* a menos de que se especifique lo contrario, $(\Omega, \mathcal{F}, \mathbb{P})$ representará un espacio de probabilidad completo.

Ahora se pretende mostrar las propiedades que permitan el cálculo de probabilidades en procedimientos de límites. El siguiente resultado se verifica para dos tipos de sucesiones particulares, aquellas que son monótonas crecientes o decrecientes, tales propiedades se conocen como *propiedades de continuidad secuencial.*

Teorema 1.1.2. *Sea* (A_n) *una sucesión de eventos.*

a) *Si* $A_n \uparrow A$*, entonces*

$$\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \mathbb{P}(A).$$

b) *Si* $A_n \downarrow A$*, entonces*

$$\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \mathbb{P}(A).$$

Demostración.

a) Por la propiedad monótona de la probabilidad, la sucesión numérica $\{\mathbb{P}(A_n) : n \in \mathbb{N}\}$ es no dreciente y acotada superiormente por uno. Entonces, el límite de esta sucesión existe.

Se define $B_1 = A_1$ y $B_k = A_k - A_{k-1}$. Por construcción, la colección formada por B_k son disjuntos a pares; además,

$$A = \bigcup_{k=1}^{\infty} B_k, \quad A_n = \bigcup_{k=1}^{n} B_k.$$

Por la propiedad de aditividad finita y numerable,

$$\mathbb{P}(A) = \sum_{k=1}^{\infty} \mathbb{P}(B_k) = \lim_{n\to\infty} \sum_{k=1}^{n} \mathbb{P}(B_k) = \lim_{n\to\infty} \mathbb{P}(A_n).$$

b) Si (A_n) es una sucesión no creciente de eventos tal que $A = \bigcap_{n\in\mathbb{N}} A_n$, entonces se cumple que $A_n^c \uparrow A^c$. Luego entonces,

$$\begin{aligned}
\mathbb{P}(A) = 1 - \mathbb{P}(A^c) &= 1 - \lim_{n\to\infty} \mathbb{P}(A_n^c) \\
&= \lim_{n\to\infty} (1 - \mathbb{P}(A_n^c)) \\
&= \lim_{n\to\infty} \mathbb{P}(A_n).
\end{aligned}$$

Por lo tanto,

$$\lim_{n\to\infty} P(A_n) = P(A) = \mathbb{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right).$$

■

La siguiente definición es fundamental para demostrar la propiedad de continuidad en el caso de sucesiones arbitrarias.

Definición 1.1.7. *Sea (A_n) una sucesión de eventos.*
*El **límite superior** de (A_n) es el conjunto dado por*

$$[A_n\, i.o.] := \limsup_n A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k, \tag{1.2}$$

donde $\omega \in [A_n\, i.o.]$ si y sólo si para cada n existe algún $k \geq n$ para el cual $\omega \in A_k$, i.e., $\omega \in [A_n\, i.o.]$ si y sólo si está en una infinidad de A_n.

*El **límite inferior** de (A_n) es el conjunto dado por*

$$[A_n\, a.a.] := \liminf_n A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k, \tag{1.3}$$

donde $\omega \in [A_n\, a.a.]$ si y sólo si está en todo A_n excepto un número finito de ellas.

Si los límites en (1.2) y (1.3) existen y coinciden con A, entonces se dice que A_n tiene como límite a A, y se escribe $A_n \to A$ o $\lim_{n\to\infty} A_n = A$.

Teorema 1.1.3. *Sea (A_n) una sucesión de eventos.*

a) *Entonces,*

$$\begin{aligned} \mathbb{P}\Big(\liminf_n A_n\Big) &\leq \liminf_n \mathbb{P}(A_n) \\ &\leq \limsup_n \mathbb{P}(A_n) \leq \mathbb{P}\Big(\limsup_n A_n\Big). \end{aligned}$$

b) *Si $A_n \to A$, entonces $\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A)$.*

Demostración.

a) Se define $B_n = \bigcap_{k=n}^{\infty} A_k$, $C_n = \bigcup_{k=n}^{\infty} A_k$. Entonces

$$B_n \uparrow \liminf_n A_n \quad \text{y} \quad C_n \downarrow \limsup_n A_n.$$

Además,

$$\mathbb{P}(B_n) \leq \inf\{\mathbb{P}(A_n), \mathbb{P}(A_{n+1}), \ldots\}, \quad \mathbb{P}(C_n) \geq \sup\{\mathbb{P}(A_n), \mathbb{P}(A_{n+1}), \ldots\}.$$

Luego,

$$\begin{aligned} \mathbb{P}\Big(\liminf_n A_n\Big) &= \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} B_n\right) \\ &= \lim_{n\to\infty} \mathbb{P}(B_n) \\ &\leq \lim_{n\to\infty} (\inf\{\mathbb{P}(A_n), \mathbb{P}(A_{n+1}), \ldots\}) \\ &= \liminf_n \mathbb{P}(A_n), \end{aligned} \tag{1.4}$$

en donde (1.4) se debe al Teorema 1.1.2.

De forma análoga,

$$\limsup_n \mathbb{P}(A_n) \leq \mathbb{P}\Big(\limsup_n A_n\Big).$$

Puesto que $\liminf_n \mathbb{P}(A_n) \leq \limsup_n \mathbb{P}(A_n)$, entonces la cadena de desigualdad queda demostrada.

b) Supóngase ahora que $A_n \to A$, entonces

$$\limsup_n A_n = \liminf_n A_n = A.$$

Por el inciso anterior,

$$\begin{aligned} \mathbb{P}(A) = \mathbb{P}\Big(\liminf_n A_n\Big) &\leq \liminf_n \mathbb{P}(A_n) \\ &\leq \limsup_n \mathbb{P}(A_n) \leq \mathbb{P}\Big(\limsup_n A_n\Big) = \mathbb{P}(A), \end{aligned}$$

de donde

$$\liminf_n \mathbb{P}(A_n) = \limsup_n \mathbb{P}(A_n) = \mathbb{P}(A).$$

Por lo tanto,

$$\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(A).$$

■

Cuando Ω es a lo más numerable (esto es, Ω es finito o numerable), la σ-álgebra que se considera es el conjunto potencia de Ω, $\mathcal{F} = \mathcal{P}(\Omega)$. Para espacios muestrales más generales, no siempre puede tomarse esta estructura tan grande.

Teorema 1.1.4. *Sea Ω a lo más numerable.*

a) *Una probabilidad en $(\Omega, \mathbb{P})$ está caracterizada por sus valores en los eventos elementales:*

$$p_\omega = \mathbb{P}(\{\omega\}), \quad \omega \in \Omega.$$

b) *Sea $(p_\omega)_{\omega\in\Omega}$ una familia de números reales indexadas por Ω. Entonces existe una única probabilidad $\mathbb{P}$ tal que $\mathbb{P}(\omega) = p_\omega$ si y sólo si*

- $p_\omega \geq 0$.
- $\sum_{\omega\in\Omega} p_\omega = 1$.

Demostración.

a) Sea $A \in \mathcal{F}$ arbitrario, entonces $A = \bigcup_{\omega\in A}\{\omega\}$ o es la unión finita o es la unión numerable de eventos elementales disjuntos a pares. Si $\mathbb{P}$ es una probabilidad, por la aditividad numerable (o aditividad finita) se tiene que

$$\mathbb{P}(A) = \sum_{\omega\in A} \mathbb{P}(\{\omega\}) = \sum_{\omega\in A} p_\omega.$$

Por lo tanto, la probabilidad está caracterizada por los valores p_ω.

b) Suponga que $\mathbb{P}(\omega) = p_\omega$. Por definición, $p_\omega \geq 0$, además

$$1 = \mathbb{P}(\Omega) = \mathbb{P}\left(\bigcup_{\omega \in \Omega} \{\omega\} \right) = \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}) = \sum_{\omega \in \Omega} p_\omega.$$

Para lo inverso, si los valores p_ω satisfacen $p_\omega \geq 0$ y $\sum_{\omega \in \Omega} p_\omega = 1$, se define una probabilidad $\mathbb{P}$ por

$$\mathbb{P}(A) = \sum_{\omega \in A} p_\omega, \quad \text{para } A \in \mathcal{F},$$

con la convención de que una suma "vacía" es igual a 0.

Entonces,

$$\mathbb{P}(\emptyset) = 0, \quad \mathbb{P}(\Omega) = \sum_{\omega \in \Omega} p_\omega = 1.$$

Verificar la propiedad de la aditividad numerable no presenta complicaciones cuando Ω es finito. Supóngase que Ω es numerable.

Si $(A_i)_{i \in I}$ es una colección numerable de elementos de $\mathcal{F}$, disjuntos a pares, entonces

$$\begin{aligned} \mathbb{P}\left(\bigcup_{i \in I} A_i \right) &= \sum_{\omega \in \bigcup_{i \in I} A_i} p_\omega \\ &= \sum_{i \in I} \sum_{\omega \in A_i} p_\omega \\ &= \sum_{i \in I} \mathbb{P}(A_i). \end{aligned}$$

Con todo lo anterior, el teorema queda demostrado.

■

Si Ω es finito, cualquier familia de términos no negativos que sumen uno es un ejemplo de una probabilidad en Ω. Entre todas ellas, la siguiente es particularmente importante.

Definición 1.1.8. *Una probabilidad $\mathbb{P}$, en un conjunto finito Ω, se llama* **uniforme** *si $p_\omega = \mathbb{P}(\{\omega\})$ no depende de ω.*

En este caso, es inmediato que

$$\mathbb{P}(A) = \frac{\#(A)}{\#(\Omega)},$$

donde $\mathbb{P}$ es llamada *probabilidad clásica.*

En algunos casos, determinar la probabilidad de un evento B depende de que haya o no ocurrido un evento A, a esta probabilidad se le conoce como *probabilidad condicional.*

Definición 1.1.9. *Sean A y B dos eventos tal que $\mathbb{P}(A) > 0$. Entonces, la* **probabilidad condicional de** B **dado** A*, denotada por $\mathbb{P}(B|A)$, está definida por*

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

La ocurrencia de algún evento A cambia la probabilidad de ocurrencia de otro evento B, la probabilidad original $\mathbb{P}(B)$ es reemplazada por $\mathbb{P}(B|A)$. Si la probabilidad no cambia, es decir $\mathbb{P}(B|A) = \mathbb{P}(B)$, entonces diremos que A y B son *independientes.*

Definición 1.1.10. *Los eventos A y B son* ***independientes*** *si* $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Más general, $(A_i)_{i \in I}$ *es una colección de eventos independientes si para cada subconjunto finito J de I, se cumple*

$$\mathbb{P}\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} \mathbb{P}(A_i).$$

1.2. Variables aleatorias

La situación estándar en el modelado de fenómenos aleatorios es que las cantidades de interés, en lugar del espacio de probabilidad subyacente, son las funciones del espacio de probabilidad sobre otro espacio medible. Estas funciones se llaman *variables aleatorias.* Hablando estrictamente, se usa el término variable aleatoria cuando esas funciones van de un espacio de probabilidad al espacio medible $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Definición 1.2.1.

a) *Una* **variable aleatoria** *(v.a.) X es una función del espacio muestral Ω a $\mathbb{R}$, esto es,*

$$X : \Omega \to \mathbb{R},$$

y que cumple que la imagen inversa de cualquier Boreliano pertenece a $\mathcal{F}$, i.e.,

$$X^{-1}(A) = \{\omega : X(\omega) \in A\} \in \mathcal{F}, \qquad \text{para cada } A \in \mathcal{B}(\mathbb{R}). \tag{1.5}$$

b) *X se llama v.a.* ***simple*** *si, para algún $n \in \mathbb{N}$,*

$$X = \sum_{k=1}^{n} x_k I_{A_k},$$

donde $x_k \in \mathbb{R}, 1 \leq k \leq n$, I_{A_k} es la función indicadora (o característica), $A_k \in \mathcal{F}$ y $\{A_k, 1 \leq k \leq n\}$ es una partición de Ω (esto es, $A_k \neq \emptyset, A_i \cap A_j = \emptyset$ si $i \neq j$, $\bigcup_{k=1}^{n} A_k = \Omega$).

c) *X se llama v.a.* ***elemental*** *siempre que*

$$X = \sum_{n=1}^{\infty} x_n I_{A_n},$$

donde $x_n \in \mathbb{R}, n \geq 1$, $A_k \in \mathcal{F}$ y la colección $\{A_n, n \geq 1\}$ es una partición de Ω.

Si una función X cumple (1.5), simplemente se dice que X ***es $\mathcal{F}$-medible.*** En particular, se dice que una función $g : \mathbb{R} \to \mathbb{R}$ es ***Borel medible*** si g es $\mathcal{B}(\mathbb{R})$-medible.

Como es de esperarse, suma y producto de v.a.'s definen nuevamente v.a.'s (véase e.g., Gut[6]).

No es complicado verificar que la colección de imágenes inversas de una v.a. forma una σ-álgebra de Ω. Por tanto, tiene sentido definir

$$\sigma(X) = \{X^{-1}(A) : A \in \mathcal{B}(\mathbb{R})\},$$

como la mínima σ-álgebra generada por la v.a. X, esto es , $\sigma(X)$ es la mínima σ-álgebra que hace que X sea v.a. $\mathcal{F}$-medible.

El siguiente resultado es un caso especial del resultado conocido como el *lema de Doob-Dynkin* y su demostración puede encontrarse en Rao[17] (véase Proposición 3, p. 8).

Teorema 1.2.1. *Sean X, Y v.a.'s. Entonces Y es $\sigma(X)$-medible si y sólo si existe una función Borel medible $g : \mathbb{R} \to \mathbb{R}$ tal que $Y = g(X)$.*

Ejemplo 1.3. *No es complicado demostrar que toda función real continua, o función continua por pedazos, es Borel medible.*
En el primer caso, la demostración se basa principalmente en utilizar el hecho de que la imagen inversa de cualquier conjunto abierto es abierto, y por tanto medible. Para el segundo caso, la función original se puede escribir como una suma de funciones continuas.

Como caso particular, si $f(x) = x^k$ para $k \in \mathbb{N}$, entonces f es Borel medible. De aquí que, si X es una v.a., entonces X^k para $k \in \mathbb{N}$ también lo es. Más aún, por el Teorema 1.2.1, X^k es $\sigma(X)$-medible. □

La razón técnica por la cual se le pide a una función que sea medible es para poder trasladar la probabilidad $\mathbb{P}$ al espacio medible $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Así, a cada v.a. X se le asocia una ***medida de probabilidad inducida*** $\mathbb{P}_X$, dada por la relación

$$\mathbb{P}_X(A) := \mathbb{P}(X^{-1}(A)) = \mathbb{P}(\{\omega : X(\omega) \in A\}) = \mathbb{P}(X \in A), \tag{1.6}$$

para cada $A \in \mathcal{B}(\mathbb{R})$.
A $\mathbb{P}_X$ se le conoce también con el nombre de ***distribución de probabilidad*** (o *ley de probabilidad*) de X. De esta forma se construye el espacio de probabilidad $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_X)$, conocido como *espacio de probabilidad inducido* por la v.a. X.

Teorema 1.2.2. *El espacio inducido $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_X)$, con $\mathbb{P}_X$ definido por (3.22), es un espacio de probabilidad.*

Demostración.
La función $\mathbb{P}_X$ cumple la definición de medida de probabilidad. En efecto,

a) $\mathbb{P}_X(A) = \mathbb{P}(\{\omega : X(\omega) \in A\}) \geq 0$, para cualquier $A \in \mathcal{B}(\mathbb{R})$.

b) $\mathbb{P}_X(\mathbb{R}) = \mathbb{P}(\{\omega : X(\omega) \in \mathbb{R}\}) = 1$.

c) Sea (A_n) una sucesión de eventos Borelianos disjuntos a pares, entonces

$$\begin{aligned}
\mathbb{P}_X\left(\bigcup_{n=1}^{\infty} A_n\right) &= \mathbb{P}\left(\left\{\omega : X(\omega) \in \bigcup_{n=1}^{\infty} A_n\right\}\right) \\
&= \mathbb{P}\left(\bigcup_{n=1}^{\infty} \{\omega : X(\omega) \in A_n\}\right) \\
&= \sum_{n=1}^{\infty} \mathbb{P}(\{\omega : X(\omega) \in A_n\}) \\
&= \sum_{n=1}^{\infty} \mathbb{P}_X(A_n).
\end{aligned}$$

Por lo tanto, el teorema es válido. ■

Existen al menos dos formas de interpretar la igualdad entre v.a.'s.

Definición 1.2.2. *Sean X, Y v.a.'s.*

a) *X y Y son **iguales en distribución,** y se escribe $X \stackrel{d}{=} Y$, si son gobernadas por la misma medida de probabilidad, i.e.,*

$$X \stackrel{d}{=} Y \text{ si y sólo si } \mathbb{P}(X \in A) = \mathbb{P}(Y \in A), \quad \text{para todo } A \in \mathcal{B}.$$

b) *X y Y son **iguales casi seguramente** (o **puntualmente**) si concuerdan para* casi todos *los eventos elementales, excepto posiblemente en conjuntos nulos, i.e.,*

$$X = Y,\ \mathbb{P}\text{-c.s.} \quad \text{si y sólo si } \mathbb{P}(\{\omega : X(\omega) = Y(\omega)\}) = 1,$$

o equivalentemente,

$$X = Y,\ \mathbb{P}\text{-c.s.} \quad \text{si y sólo si } \mathbb{P}(\{\omega : X(\omega) \neq Y(\omega)\}) = 0.$$

En el inciso **a)** de la definición anterior, es posible que las v.a.'s estén definidas en diferentes espacios de probabilidad. Cabe resaltar que la igualdad casi segura es más fuerte que la igualdad en distribución.

1.2.1. Función de distribución

La descripción completa de la distribución de probabilidad de una v.a. X requiere conocer $\mathbb{P}(X \in A)$, para todos los conjuntos $A \in \mathcal{B}(\mathbb{R})$.

Como se mencionará más adelante, toda v.a. tiene asociada una función, llamada *función de distribución*. Entonces, la descripción de una v.a. requiere analizar su función de distribución.

A continuación se define esta importante función.

Definición 1.2.3. *Sea X una v.a.. La* ***función de distribución*** *de X es una función $F : \mathbb{R} \to [0, 1]$, definida como*

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(\{\omega : X(\omega) \leq x\}), \quad x \in \mathbb{R}.$$

A la función $F(x)$ también se le conoce con el nombre de ***función de distribución acumulada*** (fda). Las funciones de distribución contienen toda la información de la v.a., incluyendo su correspondiente medida de probabilidad.

Teorema 1.2.3. *Sea $F(x)$ la función de distribución de una v.a.. Entonces*

a) $\lim_{x \to +\infty} F(x) = 1, \quad \lim_{x \to -\infty} F(x) = 0.$

b) *Si $x \leq y$, entonces $F(x) \leq F(y)$.*

c) *$F(x)$ es continua por la derecha, es decir, $F(x+) = F(x)$.*

Demostración.

a) Se sigue de $\{X \leq x\} \to \Omega$ cuando $x \to +\infty$ y $\{X \leq x\} \to \emptyset$ cuando $x \to -\infty$, junto con la propiedad de continuidad de $\mathbb{P}$.

b) Sea $x \leq y$, entonces $\{X \leq x\} \subseteq \{X \leq y\}$. Luego,

$$F(x) = \mathbb{P}(X \leq x) \leq \mathbb{P}(X \leq y) = F(y).$$

c) Sea $x_1, x_2, \ldots$ una sucesión de números reales no negativos y decrecientes a cero. Entonces

$$F(x + x_n) = F(x) + \mathbb{P}(x < X \leq x + x_n),$$

en donde $A_n = (x < X \leq x + x_n)$ es una sucesión de eventos decrecientes al conjunto vacío. Por lo tanto, $F(x+) = \lim_{n \to \infty} F(x + x_n) = F(x)$.

■

La siguiente definición de función de distribución es más general, pues no hace referencia a v.a.'s ni a espacios de probabilidad particulares.

Definición 1.2.4. *Una función $F : \mathbb{R} \to [0, 1]$ es llamada* ***función de distribución*** *si cumple las tres propiedades del Teorema 3.23.*

Cuando se estudia las distribuciones de probabilidad, en la mayoría de los casos no se especifica el espacio de probabilidad, a pesar de ser este el elemento base en todas las consideraciones. Existe un resultado que garantiza la existencia de un espacio de probabilidad y una v.a. definida sobre él para una función de distribución dada, siendo ésta la principal razón del por qué se considera importante el uso de la función de distribución. Para la demostración del siguiente resultado véase Billingsley[1], Teorema 14.1.

Teorema 1.2.4. *Sea $F(x)$ una función de distribución. Entonces existe un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ y una variable aleatoria X cuya función de distribución es $F(x)$.*

Las siguientes propiedades establecen la forma de calcular probabilidades usando la función de distribución.

Teorema 1.2.5. *Sea X una variable aleatoria con función de distribución $F(x)$. Para cualesquiera números reales $a < b$, se cumple*

a) $\mathbb{P}(X < a) = F(a-)$.

b) $\mathbb{P}(X = a) = F(a) - F(a-)$.

c) $\mathbb{P}(a < X \leq b) = F(b) - F(a)$.

d) $\mathbb{P}(a \leq X \leq b) = F(b) - F(a-)$.

e) $\mathbb{P}(a < X < b) = F(b-) - F(a)$.

f) $\mathbb{P}(a \leq X < b) = F(b-) - F(a-)$.

Demostración.

a) Sea $x_1, x_2, \ldots$ una sucesión de números reales positivos y decrecientes a cero. Sea el evento

$$A_n := (X \leq a - x_n).$$

Entonces $(A_n)_{n \in \mathbb{N}}$, es una sucesión de eventos decrecientes al evento $(X < a)$. Por la propiedad de continuidad de $\mathbb{P}$,

$$\mathbb{P}(X < a) = \lim_{n \to \infty} \mathbb{P}(A_n) = \lim_{n \to \infty} F(a - x_n) = F(a-).$$

b) $\mathbb{P}(X = a) = \mathbb{P}(X \leq a) - \mathbb{P}(X < a) = F(a) - F(a-)$.

c) - f). Se sigue de las anteriores.

■

Nótese que como $F(x)$ es una función no decreciente y continua por la derecha, la probabilidad $\mathbb{P}(X = x)$ representa el tamaño de salto de la función de distribución en el punto x. Cuando $F(x)$ es continua, incluir o excluir los extremos de un intervalo no afecta el valor de la probabilidad de dicho intervalo. Así, para cualquier número real x, la probabilidad del evento $(X = x)$ es cero.

Otro interesante resultado que involucra a la función de distribución es la siguiente.

Teorema 1.2.6. *Toda función de distribución $F(x)$ tiene un conjunto de discontinuidades a lo más numerable.*

Demostración.
Sea D el conjunto de puntos en donde $F(x)$ es discontinua. Para cada número natural n se define los subconjuntos D_n, por

$$D_n = \left\{ x \in D : \frac{1}{n+1} < F(x) - F(x-) \leq \frac{1}{n} \right\}.$$

Ya que $F(x)$ es no decreciente y continua por la derecha, cada conjunto D_n tiene a lo más n elementos. Como $D = \bigcup_{n=1}^{\infty} D_n$, se concluye que D es a lo más numerable. ■

1.2.2. Tipos de variables aleatorias

Las v.a.'s se clasifican en varios tipos, dependiendo de las características de su respectiva función de distribución. Al menos existen tres tipos: *discretas, continuas y mixtas* (mezclas de las dos anteriores).

Definición 1.2.5. *Sea X una v.a. con función de distribución $F(x)$.*

a) *Si $F(x)$ es constante a trozos en un conjunto a lo más numerable de intervalos, entonces se dice que X es una v.a.* ***discreta****.*
En este caso, se dice que $F(x)$ es una ***función de distribución discreta****.*
Si $x_1, x_2, \ldots$ son los puntos de discontinuidad de $F(x)$, se define la función $f(x)$, llamada ***función de probabilidad*** *(o función de masa de probabilidad) de X, dada por*

$$f(x) = \begin{cases} \mathbb{P}(X = x) = F(x) - F(x-) > 0 & si\ x = x_1, x_2, \ldots, \\ 0 & si\ x \neq x_1, x_2, \ldots. \end{cases} \tag{1.7}$$

La función $F(x)$ se reconstruye de la siguiente forma:

$$F(x) = \sum_{u \leq x} f(u).$$

b) *Si $F(x)$ es continua para todo x, entonces se dice que X es una v.a.* ***continua.***

Recíprocamente, a toda función $f(x)$ de la forma (1.7) que cumpla $\sum_i f(x_i) = 1$ se le llama *función de probabilidad*, sin que haya necesariamente una v.a. de por medio.

Las distribuciones continuas se clasifican a su vez en distribuciones *absolutamente continuas* y distribuciones *singulares.*

Definición 1.2.6. *Sea X una v.a. continua con función de distribución $F(x)$. Si existe una función no negativa y Lebesgue integrable f, tal que*

$$F(b) - F(a) = \int_a^b f(x)dx, \quad para\ todo\ a < b, \tag{1.8}$$

entonces se dice que X es una v.a. ***absolutamente continua****.*
A la función f se le llama ***función de densidad*** *de X.*

La función de densidad, también conocida como *función de densidad de probabilidad*, es no negativa y su integral sobre todo $\mathbb{R}$ es uno. Recíprocamente, toda función $f(x)$ no negativa tal que el valor de su integral sobre $\mathbb{R}$ sea uno se llama *función de densidad.*

Si X es una v.a. absolutamente continua con función de distribución $F(x)$ y función de densidad continua $f(x)$, entonces por el *teorema fundamental del cálculo* $F'(x) = f(x)$, donde quiera que exista $F'(x)$.

Las v.a.'s continuas para las cuales no existe una función f no negativa e integrable que cumpla (1.8) se llaman singulares, un ejemplo es la *distribución de Cantor* (véase e.g., Gut[6]). Aquí no se enfatizará sobre las distribuciones de v.a.'s singulares, por tal razón, por v.a.'s continuas nos estaremos refiriendo a las v.a.'s absolutamente continuas, tal y como se encuentra en la mayoría de las bibligrafías que introducen a la probabilidad.

Algunas distribuciones discretas de probabilidad son: *distribución uniforme discreta, distribución binomial, distribución geométrica, distribución Poisson, distribución hipergeométrica.* En el caso de distribuciones continuas, las comunes son: *distribución uniforme continua, distribución exponencial, distribución gama, distribución beta, distribución normal.*

1.2.3. Características numéricas

En muchas aplicaciones, el tratar de estudiar las distribuciones se vuelve muy complicado. En estos casos conviene calcular otros parámetros que nos permiten describir las distribuciones en términos breves.

Suponga que $\Omega = \{1, 2, \ldots, n\}, \mathcal{F} = \mathcal{P}(\Omega)$, y $\mathbb{P}$ asigna valores a_i a $\{i\}$, donde $\sum_{i=1}^n a_i = 1$. Entonces, la *esperanza* matemática $E[X]$ de una v.a. X en Ω es definida por

$$E[X] = \sum_{i=1}^{n} X(i)a_i = \sum_{i=1}^{n} X(i)\mathbb{P}(\{i\}).$$

Cuando $a_i = \frac{1}{n}$ para cada i, esta esperanza es el *promedio* usual de los valores de X. Heurísticamente, este número representa lo que se espera de las "observaciones" de X.

Si X es una v.a. no negativa, entonces $E[X]$ también puede ser interpretada como el "área" bajo la gráfica de X: para cada i uno puede imaginar un rectángulo de altura $X(i)$ y anchura $\mathbb{P}(\{i\})$, entonces $\sum_{i=1}^n X(i)\mathbb{P}(\{i\})$ es la suma de las "áreas" de los rectángulos que forman la parte inferior de la gráfica.

El utilizar el término "área" y por la forma intuitiva de cómo calcular las esperanzas, permiten intentar introducir el concepto de *integración*, siendo la teoría de la *integral de Lebesgue* la que encaje de manera sorprendente.

Para lograr definir la esperanza de v.a.'s arbitrarias, existen tres pasos principales:

- Se define la esperanza para el caso de v.a.'s simples.
- Se define la esperanza de v.a.'s no negativas; para ello, se utiliza un resultado que garantiza que dada una v.a. no negativa siempre es posible encontrar una sucesión creciente de v.a.'s simples que convergen a ella (véase e.g., Gut[6], Lema 1.1., p. 29).
- Finalmente, dada una v.a. arbitraria X, éste se descompone como una diferencia de dos v.a.'s no negativas, $X = X^+ - X^-$, donde

$$X^+(\omega) = \max\{X(\omega), 0\}, \quad X^-(\omega) = \max\{-X(\omega), 0\} = -\min\{X(\omega), 0\}.$$

  Entonces, se define la esperanza de X como la diferencia de las esperanzas de dos v.a.'s no negativas, siempre y cuando ambas sean finitas, o sólo una de ellas sea infinita. Bajo éstas condiciones se asegura utilizar la operación suma del campo $\mathbb{R}$, o en caso extremo utilizar la operación suma del sistema extendido de los número reales que por conveniencia se acostumbra:

$$a + \infty = \infty, \quad a - \infty = -\infty, \qquad \text{para } a \in \mathbb{R}.$$

Nótese que el proceso de cómo definir la esperanza es análogo a la construcción de la integral de Lebesgue, y esto no es una coincidencia, pues la esperanza $E[X]$ es precisamente la integral de Lebesgue de la función X con respecto a la medida de probabilidad $\mathbb{P}$.

Evidentemente, los detalles del proceso para definir $E[X]$ necesita de resultados rigurosos que involucran a la integral de Lebesgue y que están fuera del alcance de este trabajo (véase e.g., Gut[6], Jacod and Protter[9]).

Se define la esperanza para el caso simple.

Definición 1.2.7. *Sea $X = \sum_{k=1}^{n} x_k I_{A_k}$ una v.a. simple. La* ***esperanza*** *(*valor esperado, media o integral*) de X, denotada por $E[X]$, es definida como*

$$E[X] = \int_\Omega X \, d\mathbb{P} = \sum_{k=1}^{n} x_k \mathbb{P}(A_k).$$

Como caso particular, si $\#(\Omega) = n$, A_k es un evento elemental, y $\mathbb{P}$ es la probabilidad clásica, entonces la esperanza se reduce a un promedio aritmético,

$$E[X] = \frac{1}{n} \sum_{k=1}^{n} x_k.$$

En la Definición 1.2.1 no se resaltó que la representación de una v.a. simple no es única, pero ahora se aprovecha el siguiente enunciado para hacerlo notar. Además de que con este resultado se muestra que la esperanza de una v.a. simple es independiente de la representación.

Lema 1.2.1. *Sean* $\{A_k, 1 \leq k \leq n\}$, $\{B_j, 1 \leq j \leq m\}$ *particiones de* Ω, *tales que*

$$X = \sum_{k=1}^{n} x_k I_{A_k}, \qquad X = \sum_{j=1}^{m} y_j I_{B_j}.$$

Entonces

$$E[X] = \sum_{k=1}^{n} x_k \mathbb{P}(A_k) = \sum_{j=1}^{m} y_j \mathbb{P}(B_j).$$

Demostración.
Si $\{A_k, 1 \leq k \leq n\}$ y $\{B_j, 1 \leq j \leq m\}$ son particiones de Ω, entonces se cumple

$$\mathbb{P}(A_k) = \sum_{j=1}^{m} \mathbb{P}(A_k \cap B_j), \quad \mathbb{P}(B_j) = \sum_{k=1}^{n} \mathbb{P}(A_k \cap B_j).$$

Luego,

$$\sum_{k=1}^{n} x_k \mathbb{P}(A_k) = \sum_{k=1}^{n}\sum_{j=1}^{m} x_k \mathbb{P}(A_k \cap B_j),$$
$$\sum_{j=1}^{m} y_j \mathbb{P}(B_j) = \sum_{j=1}^{m}\sum_{k=1}^{n} y_j \mathbb{P}(A_k \cap B_j).$$

Puesto que la colección $\{A_k \cap B_j \neq \emptyset, 1 \leq k \leq n, 1 \leq j \leq m\}$ forma una nueva partición de Ω, entonces $x_k = y_j$ siempre que $A_k \cap B_j \neq \emptyset$, y así el lema queda demostrado. ■

El Lema 1.2.1 muestra que la Definición 1.2.7 es correcta, pues no depende de los eventos A_k. Ahora se define la esperanza para v.a.'s no negativas, y en este caso también existen resultados que garantizan que la esperanza es independiente de la sucesión de v.a.'s que la definen (véase e.g., Gut[6], Teorema 4.2.).

Definición 1.2.8. *Sean X una v.a. no negativa, y (X_n) una sucesión creciente de v.a.'s simples tales que $X = \lim_{n\to\infty} X_n$. Se define la **esperanza de** X o la **integral de** X como*

$$E[X] = \int_{\Omega} X d\,\mathbb{P} = \lim_{n\to\infty} \int_{\Omega} X_n d\,\mathbb{P}. \tag{1.9}$$

Algunos autores definen la integral $\int_{\Omega} X d\,\mathbb{P}$ como

$$E[X] = \sup\{E[Y] : Y \text{ es v.a. simple con } 0 \leq Y \leq X\},$$

la cual es equivalente a la definida en (1.9).

Nótese que $E[X] \geq 0$; se puede tener $E[X] = \infty$, aún cuando X nunca toma el valor ∞.

Finalmente, se define la esperanza de una v.a. arbitraria.

Definición 1.2.9. *Sea X una v.a. arbitraria tal que $X = X^+ - X^-$, donde*

$$X^+(\omega) = \max\{X(\omega), 0\}, \quad X^-(\omega) = -\min\{X(\omega), 0\}.$$

*Si $E[X^+]$ y $E[X^-]$ son finitas, entonces se dice que X tiene **esperanza finita** (o es **integrable.**) En este caso, la* esperanza de X *(o la* integral de X*) se define por $E[X] = E[X^+] - E[X^-]$, o también escrita como,*

$$\int_\Omega X d\mathbb{P} = \int_\Omega X^+ d\mathbb{P} - \int_\Omega X^- d\mathbb{P}. \tag{1.10}$$

*Si $E[X^+]$ y $E[X^-]$ no son ambas iguales a ∞, entonces se dice que X **admite una esperanza,** y su esperanza sigue siendo dada por (1.10), con la convención de que $\infty + a = \infty$, $-\infty + a = -\infty$, para $a \in \mathbb{R}$.*

Nótese que si X admite una esperanza, entonces $E[X] \in [-\infty, +\infty]$; y X es integrable si sólo si su esperanza es finita.

Como caso particular, sea

$$\Omega = \mathbb{N}, \quad \mathcal{F} = \mathcal{P}(\Omega), \quad \text{y} \ \ \mathbb{P}(\{n\}) = p_n,$$

para todo n tal que $\sum_{n=0}^{\infty} p_n = 1$. Se cumple que una v.a. X definida en $(\Omega, \mathcal{F}, \mathbb{P})$ es integrable si y sólo si la serie $\sum_{n=1}^{\infty} X(n)p_n$ es absolutamente convergente. En tal caso,

$$E[X] = \sum_{n=1}^{\infty} X(n)p_n.$$

El conjunto de todas las v.a.'s integrables serán denotados por $\mathcal{L}^1$ (o $\mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ en caso de ambigüedad).

El siguiente teorema contiene algunas de las propiedades más importantes de la esperanza. En caso de que el lector esté interesado en sus respectivas demostraciones, se recomienda revisar Gut[6], Teorema 4.1., p. 49; y Taylor[23], Proposición 2.1.33., p. 42.

Teorema 1.2.7.

a) *$\mathcal{L}^1$ es un espacio vectorial y la esperanza es un funcional lineal. Además, este funcional es positivo (esto es, si $X \geq 0$, entonces $E[X] \geq 0$).*

b) *Si $0 \leq X \leq Y$, $Y \in \mathcal{L}^1$, entonces $X \in \mathcal{L}^1$ y $E[X] \leq E[Y]$.*

c) *$X \in \mathcal{L}^1$ si y sólo si $|X| := X^+ + X^- \in \mathcal{L}^1$, en este caso $|E[X]| \leq E[|X|]$.*
En particular, cualquier v.a. acotada es integrable.

d) *Si $X = Y$, $\mathbb{P}$-c.s., entonces $E[X] = E[Y]$.*

La igualdad $\mathbb{P}$-c.s. entre v.a.'s es una relación de equivalencia. Se puede definir un espacio L^1 al considerar "$\mathcal{L}^1$ módulo la igualdad $\mathbb{P}$-c.s". Así, un elemento de L^1 es una clase de equivalencia, conformada por todas las v.a.'s en $\mathcal{L}^1$ tales que a pares son iguales, $\mathbb{P}$-c.s..

En vista del inciso **d)** del teorema anterior, tiene sentido hablar de la “esperanza” de las clases de equivalencia. Además, como la suma de v.a.’s o el producto de una v.a. por una constante preserva la igualdad, $\mathbb{P}$-c.s., entonces el conjunto L^1 es también un espacio vectorial.

Si $1 \leq p < \infty$, se define $\mathcal{L}^p$ como el espacio de las v.a.’s tales que $|X|^p \in \mathcal{L}^1$; L^p es definido de forma análoga a L^1.

Teorema 1.2.8. *Si $X, Y \in L^2$, entonces $XY \in L^1$ y se cumple la **desigualdad de Cauchy-Schwarz:***

$$|E[XY]| \leq \sqrt{E[X^2]E[Y^2]}.$$

Demostración.
Para $a, b \in \mathbb{R}$, la relación $(|a| - |b|)^2 \geq 0$ siempre es válida.
Luego,

$$|XY| \leq \frac{|X|^2}{2} + \frac{|Y|^2}{2}.$$

De aquí que $X, Y \in L^2$ implica que $XY \in L^1$.

Por otro lado, sean $Z = \frac{|X|}{\sqrt{E[X^2]}}, W = \frac{|Y|}{\sqrt{E[Y^2]}}$. Entonces

$$0 \leq E[(Z - W)^2] = E[Z^2 + W^2 - 2ZW] = 1 + 1 - 2E[ZW],$$

de aquí que $E[ZW] \leq 1$.
Por lo tanto,

$$|E[XY]| \leq \sqrt{E[X^2]E[Y^2]}.$$

■

En terminología matemática uno integra sobre conjuntos. En el caso de integrales que involucran a la probabilidad, se da por hecho que X es una v.a. integrable y se consideran expresiones de la forma

$$E[XI_A] = \int_A X d\mathbb{P} = \int_\Omega XI_A d\mathbb{P}, \quad \text{para} \quad A \in \mathcal{F}.$$

El cálculo de esperanzas implica efectuar integrales de Lebesgue con respecto a la medida de probabilidad $\mathbb{P}$, así que todos los resultados de la teoría de integración son aplicables.

En la mayoría de las situaciones es inconveniente calcular $E[X]$ integrando sobre Ω; el siguiente resultado expresa $E[X]$ como una integral con respecto a la medida de probabilidad $\mathbb{P}_X$, o de forma equivalente, respecto a la función de distribución F_X, pues la medida $\mathbb{P}_X$ está totalmente determinada por F_X.

Teorema 1.2.9 (**Teorema del Cambio de Variable**)**.** *Sea X una v.a., y suponga que g es una función Borel medible tal que $g(X)$ es integrable. Entonces*

$$E[g(X)] = \int_\Omega g(X) d\,\mathbb{P} = \int_\mathbb{R} g(x) dF_X(x) \Big(= \int_\mathbb{R} g d\mathbb{P}_X \Big).$$

Demostración.
Para verificar el resultado se empezará por los casos simples para luego generalizar a funciones integrables.

Sea $g(x) = I_B(x)$ una función indicadora, donde $B \in \mathcal{B}(\mathbb{R})$. Puesto que,

$$\{\omega : g(X(\omega)) = 1\} = \{\omega : X(\omega) \in B\},$$

entonces

$$E[g(X)] = \mathbb{P}(X \in B) = \int_B dF_X(x) = \int_{\mathbb{R}} g(x)d\,F_X(x),$$

así que $E[g(X)]$ y $\int_{\mathbb{R}} gdF_x$ existen y son iguales.

Ahora sea g una función simple, supóngase que $g = \sum_{j=1}^n c_j I_{B_j}$ donde $c_j \in \mathbb{R}$, $B_j \in \mathcal{B}(\mathbb{R})$. La linealidad, tanto del valor esperado como de la integral, implican

$$\begin{aligned} E[g(X)] &= \sum_{j=1}^n c_j E[I_{B_j}(X)] \\ &= \sum_{j=1}^n c_j \int_{\mathbb{R}} I_{B_j}(x) dF_X \\ &= \int_{\mathbb{R}} \Big(\sum_{j=1}^n c_j I_{B_j}(x)\Big) dF_X \\ &= \int_{\mathbb{R}} g(x) dF_X. \end{aligned}$$

Nuevamente, ambas integrales existen y son iguales.

Si g es una función Borel medible no negativa, sea $(g_n)_{n\in\mathbb{N}}$ una sucesión de funciones simples tales que $g_n \uparrow g$. Ya se ha probado que

$$E[g_n(X)] = \int_{\mathbb{R}} g_n dF_X;$$

luego, por el *Teorema de Convergencia Monótona* (véase Apéndice A.4)

$$E[g(X)] = \int_{\mathbb{R}} g dF_X,$$

y otra vez ambas integrales existen y son iguales.

Finalmente, supóngase que $g(x) = g(x)^+ - g(x)^-$ es una función Borel medible tal que $E[|g(X)|] < \infty$. La condición de ser integrable garantiza que $E[g(X)^+]$ y $E[g(X)^-]$ son finitos. Entonces, usando el resultado para funciones no negativas y la linealidad tanto del valor esperado como de la integral,

$$\begin{aligned} E[g(X)] = E[g(X)^+] - E[g(X)^-] &= \int_{\mathbb{R}} g(x)^+ dF_X - \int_{\mathbb{R}} g(x)^- dF_X \\ &= \int_{\mathbb{R}} g(x) dF_X. \end{aligned}$$

Si $E[g(X)]$ existe, supóngase que $E[g(X)^+]$ es finito, entonces $\int_{\mathbb{R}} g(x)^+ dF_X$ es finito, de aquí que $\int_{\mathbb{R}} g(x) dF_X$ existe; por la misma razón, la existencia de $\int_{\mathbb{R}} g dF_X$ implica la existencia de $E[g(X)]$; i.e, ambas integrales existen y son iguales. ■

Como caso particular del teorema anterior, si g es la función identidad, entonces obtenemos la fórmula para calcular la esperanza de la v.a. X.

El teorema anterior expresa que el valor esperado de la v.a. $g(X)$ con respecto a la probabilidad $\mathbb{P}$ en Ω es igual al valor esperado de la función g con respecto a la probabilidad $\mathbb{P}_X$ en $\mathbb{R}$, en notación $E_{\mathbb{P}}[f(X)] = E_{\mathbb{P}_X}[f]$.
Nótese que para calcular la esperanza de g sólo se necesita conocer la distribución de X.

El siguiente corolario presenta dos casos particulares del resultado anterior.

Corolario 1.2.10. *Sean X una v.a. y g una función Borel medible tal que $E[|g(X)|] < \infty$.*

a) *Si X es discreta con función de probabilidad $f_X(x)$, entonces*

$$E[g(X)] = \int_{\Omega} g(X) d\,\mathbb{P} = \sum_{k=1}^{\infty} g(x_k) f_X(x_k) = \sum_{k=1}^{\infty} g(x_k) \mathbb{P}(X = x_k).$$

b) *Si X es absolutamente continua, con función de densidad $f_X(x)$, entonces*

$$E[g(X)] = \int_{\Omega} g(X) d\,\mathbb{P} = \int_{-\infty}^{\infty} g(x) f_X(x) dx.$$

Demostración.

a) Usar la descomposición $A_j = \{X = x_j\}, j = 1, 2, \ldots,$ y $A_0 = \left(\bigcup_{n=1}^{\infty} A_k\right)^c$, donde $\mathbb{P}(A_0) = 0$.

b) Si $g(x) = I_B(x)$ es una función indicadora, con $B \in \mathcal{B}(\mathbb{R})$, entonces

$$\begin{aligned} E[g(X)] &= \int_{\Omega} I_B(x) d\,\mathbb{P} = \mathbb{P}(X \in B) = \int_B f_X(x) dx \\ &= \int_{-\infty}^{\infty} I_B(x) f_X(x) dx \\ &= \int_{-\infty}^{\infty} g(x) f_X(x) dx. \end{aligned}$$

Una vez probado el caso base, el resto de la prueba se desarrolla como en la demostración del teorema anterior.

■

Existen otros parámetros que ayudan a describir una distribución en términos breves, tales como los *momentos*.

Definición 1.2.10. *Sea X una v.a. que tiene esperanza finita y $n \in \mathbb{N}$. Cuando existe, se dice que el valor:*

a) $E[X^n]$ *es el* n***-ésimo momento*** *de* X.

En particular, con $n = 1$*, el* ***primer momento*** *coincide con la* esperanza *de* X.

b) $E[|X|^n]$ *es el* n***-ésimo momento absoluto*** *de* X.

c) $E[(X - E[X])^n]$ *es el* n***-ésimo momento central*** *de* X.

En particular, si $n = 2, E[(X - E[X])^2]$ *se llama la* ***varianza*** *de* X*, y se denota por* $Var(X)$*. El valor* $Var(X)$ *mide el grado de dispersión de los valores particulares de la v.a.* X.

d) $E[|X - E[X]|^n]$ *es el* n***-ésimo momento central absoluto*** *de* X.

e) $E[X(X-1)\dots(X-(n-1))]$ *es el* n***-ésimo momento factorial*** *de* X.

Cuando X es v.a. discreta con esperanza finita μ y función de probabilidad $f(x)$, la varianza de X, cuando existe, se calcula como

$$Var(X) = \sum_x (x - \mu)^2 f(x).$$

Si X es continua, con esperanza finita μ y función de densidad $f(x)$, entonces la varianza de X, cuando existe, se calcula como

$$Var(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx.$$

Cuando la varianza no sea finita se dice que la v.a. no tiene varianza.

La raíz cuadrada positiva de $Var(X)$ se llama ***desviación estándar***, y se le denota por σ_X (o solamente σ cuando no provoque confusión).

En el estudio de los números reales, el concepto de convergencia de una sucesión es único. Sin embargo, en el caso de sucesiones de v.a.'s la situación es diferente, ya que existen al menos cuatro posibilidades de convergencia. Para el siguiente capítulo es de particular interés la que involucra el uso de la esperanza.

Definición 1.2.11. *Sea* (X_n) *una sucesión de v.a.'s.*

a) (X_n) *converge* ***casi seguramente*** *(o* con probabilidad 1*) a la v.a.* $X \left(X_n \xrightarrow{\mathbb{P}\text{-}c.s.} X \right)$ *si*

$$\mathbb{P}(\{\omega : X_n(\omega) \to X(\omega), \ \textit{cuando } n \to \infty\}) = 1.$$

b) (X_n) *converge* ***en probabilidad*** *a la v.a.* X $\left(X_n \xrightarrow{\mathbb{P}} X \right)$ *si, para cada* $\epsilon > 0$,

$$\mathbb{P}(|X_n - X| > \epsilon) \to 0, \ \textit{cuando } n \to \infty.$$

c) *Sea* $p > 0$*.* (X_n) *converge* ***en*** L^p *(o* ***en*** p***-ésima media****) a la v.a.* X $\left(X_n \xrightarrow{L^p} X \right)$ *si* $E[|X_n| + |X|] < \infty$ *para todo* n *y*

$$E[|X_n - X|^p] \to 0, \ \textit{cuando } n \to \infty.$$

En particular, si $p = 2$ *se dice que* (X_n) ***converge en media cuadrada a*** X.

d) *Sea* $C(F_X) = \{x : F_X(x)$ *es continua en* $x\}$*.* (X_n) *converge* ***en distribución*** *a la v.a.* X $\left(X_n \xrightarrow{d} X\right)$ *si*

$$F_{X_n}(x) \to F_X(x), \text{ cuando } n \to \infty, \text{ para todo } x \in C(F_X).$$

El siguiente resultado garantiza que la convergencia es única, en el sentido de que si $X_n \to X$ y $X_n \to Y$, en cualquiera de las tres primeras posibilidades de convergencia, entonces $X = Y$, $\mathbb{P}$-c.s.. Para el caso de convergencia en distribución, la unicidad se refiere a que $F_X(x) = F_Y(x)$ para todo x, esto es, $X \stackrel{d}{=} Y$.

Teorema 1.2.11. *Sea* (X_n) *una sucesión de v.a.'s. Si* (X_n) *converge casi seguramente, en probabilidad, en* L^p*, o en distribución, cuando* $n \to \infty$*, entonces la v.a. límite (distribución) es única.*

Para el caso particular de la convergencia en L^p, la demostración se sigue de la c_p*-desigualdad:*

$$E[|X+Y|^p] \leq c_p(E[|X|^p] + E[|Y|^p]),$$

donde $c_p = 1$ cuando $p \leq 1$, y $c_p = 2^{r-1}$ cuando $p \geq 1$.

A continuación se enuncia las relaciones que existen entre los conceptos de convergencia.

Teorema 1.2.12. *Sean* X *una v.a. y* (X_n) *una sucesión de v.a.'s.*

a) *Si* $X_n \xrightarrow{\mathbb{P}} X$*, entonces* $X_n \xrightarrow{d} X$*.*

b) *Si* $X_n \xrightarrow{\mathbb{P}\text{-}c.s.} X$*, entonces* $X_n \xrightarrow{\mathbb{P}} X$*.*

c) *Si* $X_n \xrightarrow{L^p} X$*, entonces* $X_n \xrightarrow{\mathbb{P}} X$*.*

En particular, la demostración de **c)** se sigue de la *desigualdad de Markov:*

$$\mathbb{P}(|X_n - X| > \epsilon) \leq \frac{E[|X_n - X|^p]}{\epsilon^p} \to 0, \text{ cuando } n \to \infty, \text{ para cada } \epsilon > 0.$$

Si el lector está interesado en las demostraciones de los teoremas anteriores, se recomienda revisar Gut[6], Teorema 2.1., p. 208, y Teorema 3.1., p. 209, pues los resultados que se necesitan en la prueba también son abordados por el autor.

1.3. Vectores aleatorios

Ahora considérese situaciones que involucran más de una v.a. asociadas con el mismo experimento aleatorio. Gran parte de lo desarrollado en la sección anterior se puede generalizar al estudio de los vectores aleatorios, por tal razón, en algunos casos se sigue conservando las denominaciones de los elementos involucrados.

Recuérdese que se ha supuesto que se tiene como elemento base un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. En estos casos, la σ-álgebra Boreliana que se considera está dada por

$$\begin{aligned} \mathcal{B}(\mathbb{R}^n) &= \sigma\{(a_1, b_1) \times (a_2, b_2) \times \ldots \times (a_n, b_n) \subseteq \mathbb{R}^n, a_i < b_i\} \\ &= \sigma\{(-\infty, b_1] \times \ldots \times (-\infty, b_n], b_i \in \mathbb{R}\}. \end{aligned}$$

Definición 1.3.1. *Un* ***vector aleatorio*** *n-dimensional $X = (X_1, \ldots, X_n)$ es una función medible del espacio muestral Ω a $\mathbb{R}^n$, i.e.,*

$$X : \Omega \to \mathbb{R}^n,$$

tal que

$$X^{-1}(A) = \{\omega : X(\omega) \in A\} \in \mathcal{F}, \qquad \textit{para cada } A \in \mathcal{B}(\mathbb{R}^n).$$

Se dice que el vector X es discreto si cada coordenada es una v.a. discreta, y X es vector aleatorio continuo en caso de que cada coordenada lo sea.

Un vector aleatorio n-dimensional X genera el espacio de probabilidad $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mathbb{P}_X)$, en donde $\mathcal{B}(\mathbb{R}^n)$ es la σ-álgebra Boreliana de $\mathbb{R}^n$, y $\mathbb{P}_X$ es la probabilidad inducida por X :

$$\mathbb{P}_X(A) = \mathbb{P}\Big(X^{-1}(A)\Big) := \mathbb{P}(X \in A), \qquad \text{para cada } A \in \mathcal{B}(\mathbb{R}^n).$$

Análogo al caso unidimensional, la probabilidad puede estudiarse mediante la función de distribución conjunta.

Definición 1.3.2. *La* ***función de distribución conjunta*** *(o distribución multivariada) del vector X es una función $F_X : \mathbb{R}^n \to [0, 1]$ definida por*

$$F_X(x) = F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \mathbb{P}(X_1 \leq x_1, \ldots, X_n \leq x_n) = \mathbb{P}(X \leq x),$$

para $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$.

Por ejemplo, si $n = 2$, el valor $F_{X_1,X_2}(x_1, x_2)$ denota la probabilidad de que el vector aleatorio tome algún valor en el rectángulo infinito $(-\infty, x_1] \times (-\infty, x_2]$, esto es, $F_{X_1,X_2}(x_1, x_2)$ es la probabilidad de que $X_1 \leq x_1$ y al mismo tiempo $X_2 \leq x_2$, o simplemente, es la probabilidad del evento $(X_1 \leq x_1) \cap (X_2 \leq x_2)$. En general, el evento $\{X \leq x\}$ se interpreta como $\bigcap_{k=1}^{n} \{X_k \leq x_k\}$.

Las funciones de distribución conjunta satisfacen propiedades semejantes al caso unidimensional, por ejemplo, para el caso bidimensional, la función $F_{X,Y}(x, y) := F(x, y)$ satisface:

1. $\lim_{x,y\to\infty} F(x,y) = 1$.
2. $\lim_{x,y\to-\infty} F(x,y) = 0$.
3. $F(x,y)$ es no decreciente en cada variable.
4. $F(x,y)$ es continua por la derecha en cada variable.
5. Si $a_1 < b_1$ y $a_2 < b_2$, entonces

$$F(b_1,b_2) - F(a_1,b_2) - F(b_1,a_2) + F(a_1,a_2) \geq 0.$$

La propiedad 5 corresponde a la probabilidad del evento $(a_1 < X \leq b_1, a_2 < Y \leq b_2)$. Cuando se cumple ésta propiedad, se garantiza que la probabilidad de que el vector (X,Y) tome valores en el rectángulo $(a_1,b_1] \times (a_2,b_2]$ sea no negativa.

Para el caso n-dimensional, la propiedad 5 merece un tratamiento especial que aquí no se abordará.

Al igual que el caso unidimensional, existe un resultado que garantiza que dada una función de distribución conjunta, existe un espacio de probabilidad y un vector aleatorio definido en él.

Como en el caso unidimensional, los vectores aleatorios tienen asociada una función de probabilidad o de densidad, según sea el caso.

Definición 1.3.3. *Sea X un vector aleatorio.*

a) *Si X es discreta, su **función de probabilidad conjunta** es la función $f_X : \mathbb{R}^n \to [0,1]$ definida por*

$$f_X(x) = \mathbb{P}(X = x), \quad x \in \mathbb{R}^n.$$

b) *Si X es continua, con función de distribución F_X, se dice que X es absolutamente continua si existe una función no negativa y Lebesgue integrable $f_X : \mathbb{R}^n \to [0,\infty)$, tal que*

$$F_X(x) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_X(u_1,\ldots,u_n)du_n \ldots du_1, \quad \textit{para todo } x \in \mathbb{R}^n.$$

*A la función $f_X(x)$ se le llama **función de densidad conjunta** de X.*

En particular, si la función de densidad conjunta es continua , entonces

$$f_X(x) = \frac{\partial F_X(x)}{\partial x_1 \ldots \partial x_n}, \quad x \in \mathbb{R}^n.$$

1.4. Independencia

Anteriormente se ha hablado sobre la independencia de dos eventos, ahora se considerará la independencia sobre variables aleatorias. Intuitivamente, independencia de $X_1, \ldots, X_n$ significa que una afirmación sobre una o mas v.a.'s no afecta ni proporciona ventaja sobre el resto. Una afirmación sobre X_i corresponde a una afirmación sobre eventos de la forma $A_i = \{X_i \in B_i\}$. La definición formal es como sigue.

Definición 1.4.1. *Sean $X_1, \ldots, X_n$ v.a.'s. Si para cualesquiera conjuntos $A_1, \ldots, A_n \in \mathcal{B}(\mathbb{R})$ se cumple*

$$\mathbb{P}(X_1 \in A_1, \ldots, X_n \in A_n) = \mathbb{P}(X_1 \in A_1) \ldots \mathbb{P}(X_n \in A_n),$$

*entonces se dice que las v.a.'s son **independientes**.*

Existen resultados que garantizan que la condición de independencia es equivalente a solicitar que se cumpla la igualdad

$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = F_{X_1}(x_1) \ldots F_{X_n}(x_n), \qquad \text{para cualesquiera } (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

En términos de la función de densidad, cuando ésta exista y salvo un conjunto de medida cero, la condición es

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = f_{X_1}(x_1) \ldots f_{X_n}(x_n).$$

La definición de independencia de v.a.'s puede extenderse a un número finito de vectores aleatorios, no necesariamente todos de la misma dimensión. Aquí se define el caso de dos vectores.

Definición 1.4.2. *Se dice que los vectores $X = (X_1, \ldots, X_n)$, $Y = (Y_1, \ldots, Y_m)$ son independientes, si para cada $A \in \mathcal{B}(\mathbb{R}^n)$, y cada $B \in \mathcal{B}(\mathbb{R}^m)$, se cumple*

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B).$$

Si (X, Y) es un vector aleatorio y $\varphi : \mathbb{R}^2 \to \mathbb{R}$ es una función Borel medible, el cálculo de la esperanza de $\varphi(X, Y)$ primero requiere encontrar la distribución de $\varphi(X, Y)$, lo cual puede ser difícil en muchos casos. El siguiente resultado establece una forma alternativa de calcular la esperanza de funciones compuestas con un vector aleatorio, pero conociendo la distribución del vector aleatorio.

Teorema 1.4.1. *Sean X un vector aleatorio n-dimensional, Y un vector aleatorio m-dimensional, $g : \mathbb{R}^n \to \mathbb{R}$ y $h : \mathbb{R}^m \to \mathbb{R}$ funciones Borel medibles tales que $g(X)$ y $h(Y)$ tienen esperanza finita. Entonces*

a) $E[g(X)] = \int_{\mathbb{R}^n} g(x)dF_X$, *para todo* $x \in \mathbb{R}^n$.

b) *Si* X *y* Y *son independientes, entonces*

$$E[g(X)h(Y)] = E[g(X)]E[h(Y)].$$

En particular, $E[XY] = E[X]E[Y]$.

En muchas aplicaciones, la independencia entre v.a.'s no se cumple. Cuando esto ocurre, entonces se dice que las v.a.'s están correlacionadas.

Definición 1.4.3. *Sean* X *y* Y *v.a.'s , cada una con varianza finita.*

a) *La* ***covarianza*** *de* X *y* Y *es definida como*

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y].$$

b) *El* ***coeficiente de correlación*** *de* X *y* Y*, denotado por* $\rho(X,Y)$*, es el valor real que mide el grado de dependencia lineal que existe entre ellas, y está definido por*

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}.$$

En la definición anterior, nótese que $E[XY]$ existe, pues X y Y tienen varianza finita, y por el Teorema 1.2.8, $XY \in L^1$.

A continuación se mencionan algunas propiedades de la covarianza y el coeficiente de corelación.

1. $Cov(X,Y) = Cov(Y,X)$.
2. $Cov(X,X) = Var(X)$.
3. $Cov(c,Y) = 0$, donde c es una constante.
4. $Cov(cX_1 + X_2, Y) = cCov(X_1,Y) + Cov(X_2,Y)$, donde c es una constante.
5. $-1 \leq \rho(X,Y) \leq 1$.
6. $|\rho(X,Y)| = 1$ si y sólo si existen constante a y b tales que $Y = aX + b$, $\mathbb{P}$-c.s..

Por las propiedades de la covarianza, no es difícil probar que

$$Cov(X,Y) = Cov(X - E[X], Y - E[Y]). \tag{1.11}$$

Luego, aplicando la desigualdad de Cauchy-Schwarz a (1.11),

$$|Cov(X,Y)| = |E[(X - E[X])(Y - E[Y])]| \leq \sqrt{E[(X - E[X])^2]E[(Y - E[Y])^2]}.$$

Por lo tanto, de la desigualdad anterior se concluye que

$$-1 \leq \rho(X,Y) \leq 1.$$

El siguiente resultado se obtiene de manera directa al aplicar el Teorema 1.4.1.

Teorema 1.4.2. *Si X y Y son v.a.'s independientes, entonces*

$$Cov(X, Y) = 0 = \rho(X, Y).$$

Cuando $\rho(X, Y) = 0$ se dice que X y Y son **no correlacionadas**. De la propiedad 6, cuando $|\rho(X, Y)| = 1$ se dice que X y Y están **perfectamente correlacionadas positiva o negativamente**, de acuerdo al signo de $\rho(X, Y)$. Más aún, si $\rho(X, Y) = 1$, entonces $a > 0$; si $\rho(X, Y) = -1$, entonces $a < 0$.

En general, la condición $\rho(X, Y) = 0$ no es suficiente para afirmar que X y Y sean independientes, excepto en el caso de que las v.a.'s tengan distribución normal.

Finalmente, se define la esperanza y la varianza para el caso de vectores aleatorios.

Definición 1.4.4. *Sea $X = (X_1, \ldots, X_n)$ un vector aleatorio.*

a) *Si cada coordenada del vector tiene esperanza finita, se define la esperanza de X como el vector*

$$E[X] = (E[X_1], \ldots, E[X_n]).$$

b) *Si cada coordenada tiene segundo momento finito, entonces la varianza de X se define como la matriz cuadrada Σ_X, llamada* ***matriz de covarianza,*** *cuyos elementos están dados por*

$$\Sigma_{X\,i,j} = Cov(X_i, X_j), \quad i, j = 1, 2, \ldots, n.$$

c) *Se define la* ***matriz de coeficientes de correlación*** *del vector X como la matriz cuadrada, cuyos elementos están dados por*

$$\rho(X_i, X_j), \quad i, j = 1, 2, \ldots, n.$$

La varianza de un vector X puede expresarse como

$$E\Big[(X - E[X])^t (X - E[X])\Big],$$

en donde X^t significa la transpuesta del vector X. Obsérvese que el producto $(X - E[X])^t (X - E[X])$ da como resultado una matriz de dimensión $n \times n$, cuya entrada (i, j) es $E[(X_i - E[X_i])(X_j - E[X_j])] = Cov(X_i, X_j)$. Ésta matriz de covarianza, o también llamada *matriz de varianzas y covarianzas*, tiene la propiedad de ser simétrica y positiva definida, y por tanto, siempre se cumple que es invertible.

1.5. Distribución normal

La distribución normal posiblemente es la distribución de probabilidad de mayor importancia, al menos para los objetivos que persiguen en este trabajo.

Definición 1.5.1. *Sea X una v.a. continua. Se dice que X tiene una* ***distribución normal*** *(o gaussiana) si su función de densidad está dada por*

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big(- \frac{(x-\mu)^2}{2\sigma^2}\Big),$$

en donde $\mu \in \mathbb{R}$ y $\sigma > 0$ son dos parámetros. En este caso, se escribe $X \sim N(\mu, \sigma^2)$.

No es difícil demostrar que $E[X] = \mu$, y $Var(X) = \sigma^2$. En particular, si $\mu = 0$ y $\sigma^2 = 1$, entonces se dice que X tiene una distribución ***normal estándar***. En este caso, la función de densidad se reduce a

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\Big(- \frac{x^2}{2}\Big).$$

Siempre es posible transformar una variable aleatoria normal no estándar X en una estándar Z, a esta operación se le llama *estandarización;*

$$X \sim N(\mu, \sigma^2) \quad \text{si y sólo si} \quad Z = \frac{X-\mu}{\sigma} \sim N(0,1).$$

Si $X \sim N(\mu.\sigma^2)$, entonces la variable $Y = e^X$ tiene distribución $\log normal(\mu, \sigma^2)$, y su función de densidad es

$$f(y) = \begin{cases} \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\Big(- \frac{(\ln y - \mu)^2}{2\sigma^2}\Big) & \text{si } y > 0 \\ 0 & \text{si } y \leq 0. \end{cases}$$

Se puede demostrar que

$$E[Y] = \exp\Big(\mu + \frac{\sigma^2}{2}\Big), \quad Var(Y) = \exp(2\mu + 2\sigma^2) - \exp(2\mu + \sigma^2).$$

Un vector aleatorio X n-dimensional tiene distribución normal o gaussiana si su función de densidad está dada por

$$f_X(x) = \frac{1}{(2\pi)^{\frac{n}{2}} (\det \Sigma)^{\frac{1}{2}}} \exp\Big(- \frac{1}{2}(x-\mu)\Sigma^{-1}(x-\mu)^t\Big),$$

donde z^t nuevamente denota la transpuesta del vector z, $x \in \mathbb{R}^n$, $\mu \in \mathbb{R}^n$ denota la esperanza de X y Σ su respectiva matriz de covarianza.

Para el caso de v.a.'s normales, el valor del coeficiente de correlación permite determinar si las v.a.'s son independientes. Formalmente, ésto último está expresado en el siguiente resultado (véase e.g., Rincón[18], p. 177).

Teorema 1.5.1. *Si (X_1, X_2) es un vector con distribución normal bivariada, tal que $\rho(X_1, X_2) = 0$, entonces X_1 y X_2 son independientes.*

Capítulo 2

Cálculo Estocástico

Observaciones sobre los precios de las acciones, las tasas de interés, los tipos de cambios del peso frente al dólar, posiciones de una partícula de difusión y muchos otros sistemas que evolucionan en el tiempo, ya sea discreto (por ejemplo, $t = 0, 1, 2, \ldots$) o continuo (por ejemplo, $t \geq 0$), son frecuentemente modelados por los procesos estocásticos.

En la práctica, los cambios de precios en el mercado financiero son tan frecuentes que un modelo en tiempo discreto apenas puede modelar los movimientos. Un modelo apropiado es aquel en el cual los precios puedan cambiar en un tiempo arbitrario, estos modelos son comúnmente conocidos como modelos en tiempo continuo. Además, los modelos en tiempo continuo permiten cálculos explícitos, aún si en algunas ocasiones es necesario el uso de métodos numéricos. De hecho, el modelo en tiempo continuo más conocido es el modelo de Black-Scholes, el cual nos proporciona una fórmula extremadamente simple.

La conexión entre procesos estocásticos y finanzas no es reciente. En 1900, en la tesis titulada *Théorie de la Spéculation*, L. Bachelier (1870-1946) ("Pádre de las Matemáticas Financieras Modernas") introduce el modelado de la dinámica de los precios de las acciones de la bolsa de París a través del movimiento Browniano, así como la primera fórmula de valuación de un contrato de opción.

El movimiento Browniano es inadecuado como un modelo de mercado, ya que éste predice precios negativos para el activo. Sin embargo, considerando *funciones* del movimiento Browniano se puede formar una amplia clase de modelos potenciales. El modelo básico subyacente en la teoría de valuación de Black-Scholes (el movimiento Browniano geométrico) surge precisamente de esta forma.

Para analizar adecuadamente los modelos en tiempo continuo, se necesita desarrollar un cálculo basado en el movimiento Browniano. Una de las ramas de los procesos estocásticos que es de gran utilidad para estos fines es el cálculo estocástico. El objetivo principal de este capítulo es familiarizar al lector con esta sofisticada herramienta matemática.

2.1. Procesos estocásticos

Un *proceso estocástico* es un modelo u objeto matemático que representa la evolución de un fenómeno aleatorio a lo largo del tiempo. El comportamiento del fenómeno en un tiempo arbitrario t, es capturada en un *espacio medible* $(\Omega, \mathcal{F}_t)$, en general, la aleatoriedad del fenómeno se almacena en un espacio medible $(\Omega, \mathcal{F})$.

Análogo al capítulo anterior, el primer interés que se tiene es definir todos los posibles eventos del sistema aleatorio. Cuando se trabaja con procesos estocásticos, esto se traduce a especificar la información que está disponible en cada tiempo t. Por ejemplo, en los modelos financieros se requiere que los precios, presentes y pasados, de los activos sean conocidos para producir pronósticos, y conforme el tiempo transcurre la información pasada no se olvida. Esta idea es formalizada en el concepto de *filtración.*

Definición 2.1.1. *Sea $(\Omega, \mathcal{F})$ un espacio medible. Una **filtración** en $(\Omega, \mathcal{F})$ es una colección $(\mathcal{F}_t)_{t\geq 0}$ de σ-álgebras de Ω tal que*

$$\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F},$$

para cualesquiera $0 \leq s < t$.

La definición anterior también es válida para tiempos discretos, y no es difícil notar los respectivos cambios en la notación.

Como caso particular, $\mathcal{F}_0$ contiene eventos de probabilidad cero o uno.

Mientras que a la terna $(\Omega, \mathcal{F}, \mathbb{P})$ se le conoce como espacio de probabilidad, a $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ se le conoce como *espacio de probabilidad filtrado.*

Una filtración puede ser pensada como una estructura de información dinámica. El hecho de que esté aumentando significa que hay más y más información conocida conforme el tiempo transcurre y que la información pasada no se olvida, tal y como se espera que funcionen los modelos financieros.

Sin pérdida de generalidad, se asumirá que para t arbitrario, el espacio de probabilidad $(\Omega, \mathcal{F}_t, \mathbb{P})$ es completo. Así por ejemplo, si $X = Y$, $\mathbb{P}$-c.s., y Y es $\mathcal{F}_t$-medible, la suposición de contar con un espacio completo permite concluir que X es también $\mathcal{F}_t$-medible.

Definición 2.1.2. *Un **proceso estocástico** unidimensional en $(\Omega, \mathcal{F}, \mathbb{P})$ es una colección de variables aleatorias $(X_t)_{t\in T}$, indexadas por un conjunto T, llamado **espacio parametral**, y con valores en un espacio medible $(E, \mathcal{B}(E))$, llamado **conjunto de estados**.*

Los procesos estocásticos pueden ser clasificados de acuerdo a las características de las v.a.'s o de acuerdo a la numerabilidad de T y E. Respecto a este último, los procesos se clasifican como sigue: cadena (T y E numerables), proceso puntual (T no numerable y E

numerable), sucesión de v.a.'s (T numerable y E no numerable), proceso continuo (T y E no numerables, y cuando no cause confusión con la definición de continuidad). Otra clasificación sólo hace referencia a la numerabilidad de T, en este caso, cuando T es numerable se dice que es un ***proceso estocástico a tiempo discreto***, y cuando T es no numerable se dice que es un ***proceso estocástico a tiempo continuo***.

Debido al propósito que aquí se persigue, se considerará $E = \mathbb{R}$ y $\mathcal{B}(\mathbb{R})$ será la respectiva σ-álgebra Boreliana.

Puesto que toda actividad económica se lleva a cabo durante un intervalo acotado, a partir de ahora, a menos de que se especifique de forma explícita lo contrario, $T := [0, T]$ representará un intervalo cerrado y al mismo tiempo representará la cota superior del intervalo.

En lo sucesivo, por *proceso* se entenderá que se trata de un proceso estocástico continuo unidimensional en $(\Omega, \mathcal{F}, \mathbb{P})$, donde $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in T}, \mathbb{P})$ es su respectivo espacio de probabilidad filtrado. Cuando sea necesario considerar otro tipo de proceso, los cambios se mencionarán de forma explícita.

Dado un proceso (X_t), una filtración de particular interés para (X_t) se conoce como **filtración natural** generada por (X_t), denotada por $(\mathcal{F}_t^X)$, y se define como $\mathcal{F}_t^X = \sigma\{X_s, 0 \leq s \leq t\}$, i.e., la filtración natural está formada por las σ-álgebras más pequeñas con la cual X_s es medible para cada $s \in [0, t]$.

Algunas veces es conveniente escribir $X(\omega, t)$ en lugar de $X_t(\omega)$. Esta notación permite visualizar a los procesos como una función de dos variables

$$(\omega, t) \mapsto X(\omega, t),$$

con dominio $\Omega \times T$ y codominio $\mathbb{R}$.

Nótese que para algún tiempo t fijo, el proceso se convierte en una v.a.. Mientras tanto, fijando $\omega \in \Omega$, éste se convierte en una función que depende del tiempo:

$$t \mapsto X_t(\omega), \quad t \in T.$$

A la función $X(\omega, t) = X_t(\omega)$ se le conoce como ***trayectoria***, *realización o ruta muestral* del proceso (X_t). Debido a que para cada t en particular el proceso toma un valor aleatorio, éste tiene una infinidad de trayectorias, a diferencia de las funciones deterministas que solamente tienen una trayectoria.
Por ejemplo, en la Figura 2.1 se muestra diez trayectorias del proceso (X_t) definido por

$$X_t = \exp(t + 2Z_t), \quad Z_t \sim N(0, 1), \quad t \in [0, 1]. \tag{2.1}$$

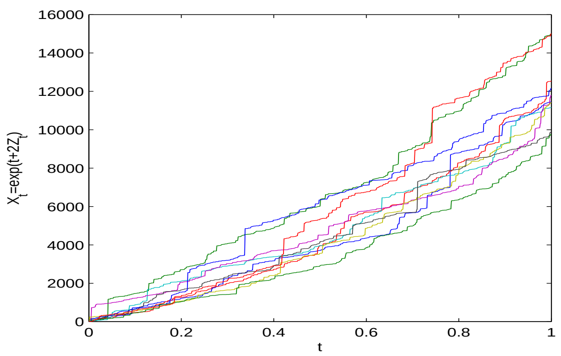

Figura 2.1: Trayectorias de un proceso en particular.

Para describir la distribución de un proceso, y con ello calcular probabilidades sobre un futuro incierto, se necesita conocer las distribuciones conocidas como *distribuciones en dimensión finita*.

Definición 2.1.3. *Sean* (X_t) *un proceso,* $n \in \mathbb{N}$*, y una colección finita* $t_1, \ldots, t_n \in T$ *de valores índice distintos.*

a) *Las distribuciones*

$$\{\mathbb{P}((X_{t_1}, \ldots, X_{t_n}) \in B),\ B \subseteq \mathcal{B}(\mathbb{R}^n), 0 \leq t_1 \leq t_2 \leq \ldots \leq t_n\}$$

son llamadas ***distribuciones en dimensión finita*** *(fidis) de* (X_t).

b) *Se dice que* (X_t) *es* H*-autosimilar si las fidis de* (X_t) *satisfacen la condición*

$$\left(\tau^H X_{t_1}, \ldots, \tau^H X_{t_n}\right) \stackrel{d}{=} \left(X_{\tau t_1}, \ldots, X_{\tau t_n}\right),$$

para cada $\tau > 0$ *y algún* $H > 0$*, i.e.,*

$$\mathbb{P}\Big(\left(\tau^H X_{t_1}, \ldots, \tau^H X_{t_n}\right) \in A\Big) = \mathbb{P}\Big(\left(X_{\tau t_1}, \ldots, X_{\tau t_n}\right) \in A\Big),$$

para todo $A \in \mathcal{B}(\mathbb{R}^n)$.

Los procesos H-autosimilares, son procesos cuyas trayectorias son equivalentes si se reescalan adecuadamente el tiempo y la amplitud.

Como se mencionó anteriormente, un proceso es un objeto matemático, pero el hecho de que tenga una definición no implica que tal objeto exista. El resultado que demuestra la *existencia* de un proceso asociado a una familia de *fidis* se conoce como el *Teorema de Extensión de Kolmogorov*. La demostración de este importante resultado, así como de algunos casos particulares, se pueden encontrar en Billinsgley[1].

En el análisis de la teoría estocástica el concepto de medibilidad permite introducir otros conceptos interesantes.

Definición 2.1.4. *Sea (X_t) un proceso.*

a) *Se dice que (X_t) es **medible** si, para cada $B \in \mathcal{B}(\mathbb{R})$, el conjunto $\{(\omega, t) : X_t(\omega) \in B\}$ pertenece a la σ-álgebra producto $\mathcal{F} \otimes \mathcal{B}(T)$, i.e., si el mapeo*

$$(\omega, t) \mapsto X_t(\omega) : (\Omega \times T, \mathcal{F} \otimes \mathcal{B}(T)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})),$$

es medible.

b) *Si para todo $t \in T$, X_t es una variable aleatoria $\mathcal{F}_t$-medible, entonces se dice que (X_t) es **adaptado a la filtración** $(\mathcal{F}_t)$.*

c) *Se dice que (X_t) es **progresivamente medible** con respecto a la filtración $(\mathcal{F}_t)$ si, para cada $t \geq 0$ y $B \in \mathcal{B}(\mathbb{R})$, el conjunto $\{(\omega, s) : 0 \leq s \leq t, \omega \in \Omega, X_s(\omega) \in B\}$ pertenece a la σ-álgebra producto $\mathcal{F}_t \otimes \mathcal{B}([0, t])$, i.e., si el mapeo*

$$(\omega, s) \mapsto X_s(\omega) : (\Omega \times [0, t], \mathcal{F}_t \otimes \mathcal{B}([0, t])) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})),$$

es medible para cada $t \geq 0$.

Adaptar un proceso (X_t) significa que X_t no contiene más información que $\mathcal{F}_t$. Un proceso (X_t) siempre es adaptado a su filtración natural $(\mathcal{F}_t^X)$. Siempre se cumple que si un proceso es progresivamente medible, entonces el proceso es medible y adaptado; lo inverso también es válido, pero con algunas modificaciones. Para mostrar las respectivas modificaciones, se necesitan introducir otros conceptos.

Existen al menos tres diferentes formas de relacionar el concepto de "igualdad" entre dos procesos.

Definición 2.1.5. *Sean $(X_t), (Y_t)$ dos procesos definidos en el mismo espacio de probabilidad.*

a) *Se dice que (X_t) es una **versión (o modificación)** de (Y_t) si*

$$\mathbb{P}(\{\omega : X_t(\omega) = Y_t(\omega)\}) = 1,$$

o, de forma equivalente,

$$\mathbb{P}(\{\omega : X_t(\omega) \neq Y_t(\omega)\}) = 0,$$

para todo $0 \leq t \leq T$. Esto es, $X_t = Y_t$, $\mathbb{P}$-c.s., para todo $0 \leq t \leq T$.

b) *(X_t) y (Y_t) tienen la misma **distribución en dimensión finita**, y se escribe $(X_t) \stackrel{d}{=} (Y_t)$, si para todo $B \subset \mathcal{B}(\mathbb{R}^n), 0 \leq t_1 \leq \ldots \leq t_n$, se cumple*

$$\mathbb{P}((X_{t_1}, \ldots, X_{t_n}) \in B) = \mathbb{P}((Y_{t_1}, \ldots, Y_{t_n}) \in B).$$

c) *Se dice que (X_t) y (Y_t) son **indistinguibles** si casi todas sus trayectorias coinciden, esto es,*

$$\mathbb{P}(X_t = Y_t, \forall\, 0 \leq t \leq T) = 1.$$

La tercera condición es la más importante; ésta implica la primera, la cual a su vez implica la segunda. Más aún, puede demostrarse que cuando los procesos son continuos (*todas las trayectorias del proceso son continuas*), la noción de indistinguibilidad y modificación coinciden. Por otra parte, puede presentarse el caso en que dos procesos sean versiones de algún otro, y aún así pueden tener trayectorias completamente diferentes.

Ejemplo 2.1. *Considere una distribución continua en el espacio de probabilidad* $([0,\infty), \mathcal{B}([0,\infty)), \mathbb{P})$. *Se definen los procesos* (X_t) *y* (Y_t) *dados por*

$$X_t(\omega) = 0, \quad Y_t(\omega) = \begin{cases} 1 & si\ t = \omega \\ 0 & si\ t \neq \omega, \end{cases}$$

para todo $(\omega, t) \in [0,\infty) \times [0,\infty)$.

(X_t) *es una versión de* (Y_t), *ya que para cada* $t \geq 0$ *se tiene que*

$$\mathbb{P}(\{\omega : X_t(\omega) = Y_t(\omega)\}) = \mathbb{P}(\omega \neq t) = 1 - \mathbb{P}(\omega = t) = 1;$$

además, ambos procesos comparten las mismas fidis.

Por otra parte, las trayectorias $t \mapsto X_t(\omega)$ *son continuas para todo* ω, *mientras que las trayectorias* $t \mapsto Y_t(\omega)$ *son discontinuas para todo* ω, *esto es,* (X_t) *y* (Y_t) *no son indistinguibles.* □

A pesar de que un proceso es una versión de otro, puede ocurrir, como en el ejemplo anterior, que uno tenga trayectorias continuas y el otro no. Condiciones para la existencia de versiones continuas de un proceso son dadas a continuación, siendo éste otro teorema famoso de Kolmogorov.

Teorema 2.1.1 (Teorema Continuidad de Kolmogorov). *Sea* (X_t) *un proceso. Si existen constantes positivas* α, β, D *tales que*

$$E[|X_t - X_s|^\alpha] \leq D|t-s|^{1+\beta}, \quad 0 \leq s, t \leq T,$$

entonces existe una versión continua de (X_t), *tal que sus trayectorias son continuas según Hölder de orden* $0 < h < \frac{\beta}{\alpha}$.

Si es de interés para el lector conocer los detalles de la demostración, véase Karatzas[10], Teorema 2.8, p. 53-55.

Finalmente se enuncia un resultado que relaciona los conceptos de medibilidad.

Teorema 2.1.2. *Sea* (X_t) *un proceso.*

a) *Si* (X_t) *es medible y adaptado a la filtración* $(\mathcal{F}_t)$, *entonces éste tiene una versión medible progresivamente.*

b) *Si* (X_t) *es continua por la derecha (esto es,* $\lim_{u\downarrow t} X_u = X_t$ *en probabilidad) o continua por la izquierda (esto es,* $\lim_{u\uparrow t} X_u = X_t$ *en probabilidad) y adaptado a la filtración* $(\mathcal{F}_t)$*, entonces éste también es medible progresivamente con respecto a* $(\mathcal{F}_t)$*.*

Nótese que todo proceso continuo y adaptado a alguna filtración en particular es progresivamente medible con respecto a la filtración dada.

El concepto de procesos predecibles es de gran relevancia en el cálculo estocástico. Para el caso de procesos a tiempos discretos, se dice que (X_n) es predecible si X_n es $\mathcal{F}_{n-1}$-medible, esto es, (X_n) se conoce con certeza al tiempo n tomando como base la información disponible hasta el tiempo $n-1$. Para procesos a tiempo continuo, este concepto es más difícil de visualizar. La definición exacta de procesos predecibles involucra σ-álgebras generadas en $\Omega \times \mathbb{R}^+$.

Procesos continuos por la izquierda son predecibles, en el sentido de que $X_t = \lim_{s\uparrow t} X_s = X_{s-}$. De esta forma, si los valores del proceso antes de t son conocidos, entonces el valor en t es determinado por el límite. Para nuestros propósitos, es suficiente describir sólo una subclase de procesos predecibles.

Definición 2.1.6. *Un proceso estocástico* (X_t) *es* ***predecible*** *si cumple una de las siguientes características:*

a) *Es un proceso adaptado continuo por la izquierda. En particular, procesos continuos adaptados.*

b) *Es un límite (casi seguramente, en probabilidad) de una sucesión de procesos adaptados continuos por la izquierda.*

c) *Es una función Borel medible de un proceso predecible.*

La teoría de derivados financieros hace gran uso de los valores esperados de los procesos. Generalmente, los valores esperados son tomados como una condicional sobre la historia del proceso. Esta historia formalmente se conoce como *filtración*, término que ya fue definido. En términos matemáticos, estos valores se conocen como *esperanzas condicionales* (véase Apéndice A.4).

Por otra parte, el concepto de *martingala* desempeña un papel esencial en el modelado de los mercados financieros y en la valuación teórica de muchos instrumentos financieros.

Definición 2.1.7. *Un proceso* (X_t) *adaptado a la filtración* $(\mathcal{F}_t)$ *es una* ***martingala*** *con respecto a la filtración si cumple:*

a) $E_{\mathbb{P}}[|X_t|] < \infty, \mathbb{P}$*-c.s., para todo* t*.*

b) *(***Propiedad de martingala***)* $E_{\mathbb{P}}[X_t|\mathcal{F}_s] = X_s, \mathbb{P}$*-c.s., para todo* $s \leq t$*.*

La *propiedad de martingala* expresa que, dada la información disponible hasta el tiempo s, representada por $\mathcal{F}_s$, entonces el mejor pronóstico de X_t es justamente el valor más reciente, X_s, en otras palabras, si se conoce los valores del proceso en el tiempo s, y se cumple la igualdad $X_s = x$, entonces se espera que el valor del proceso en un tiempo futuro sea x.

La martingala juega un rol central en la teoría moderna de procesos estocásticos, y como se verá más adelante, las integrales estocásticas son martingales. Un proceso (X_t) que cumple la definición anterior también se dice que es $\mathcal{F}_t$-martingala o $\mathbb{P}$-martingala.

El siguiente ejemplo muestra el uso, en su forma más simple, de la Definición 2.1.7; además, permite visualizar la necesidad de buscar medidas de probabilidad que no provoquen arbitraje en el modelo, los fundamentos serán mostrados en capítulos posteriores.

Ejemplo 2.2. *Considere un conjunto de tiempos discretos T, en el cual un activo se comercializa. Suponga que en este modelo simple, en cada tiempo de transacción, el precio del activo aumenta un factor u con probabilidad p, y disminuye un factor d con probabilidad $1-p$.*

Sea el proceso (S_n), donde S_n denota el precio del activo al finalizar el tiempo de transacción n. En el tiempo $n+1$, el precio del activo será uS_n con probabilidad p, o dS_n con probabilidad $1-p$.

Luego, se espera que el valor de S_{n+1}, dado que el precio al tiempo n fue S_n, sea $p(uS_n) + (1-p)(dS_n)$, i.e.,

$$E[S_{n+1}|S_n] = p(uS_n) + (1-p)(dS_n).$$

Para ilustrarlo de forma numérica, suponga que $u = 1.25, d = 0.8$, $p = 0.4444$. Si S_n fue registrado con valor 100, entonces S_{n+1} tomará el valor $(1.25)(100) = 125$, si el activo aumenta de valor, o $(0.8)(100) = 80$ si disminuye. De aquí que,

$$E_{\mathbb{P}}[S_{n+1}|S_n] = (0.4444)(125) + (1-0.4444)80 = 100.$$

Para los parámetros considerados, el valor esperado es igual al valor actual conocido, entonces se dice que el proceso de precios (S_n) es una martingala.

Por el contrario, con los mismos valores u, d dados, si se considera otra probabilidad, por ejemplo q, que cumpla $q \neq 0.4444$, entonces $E_{\mathbb{Q}}[S_{n+1}|S_n] \neq S_n$, y así (S_n) no es una martingala.

En efecto, para que (S_n) sea martingala, la probabilidad p debe cumplir

$$p(uS_n) + (1-p)(dS_n) = S_n;$$

$$p(u) + (1-p)(d) = 1,$$

de donde,

$$p = \frac{1-d}{u-d}. \tag{2.2}$$

Sustituyendo los valores de u y d dados anteriormente, se tiene que

$$p = \frac{1-0.8}{1.25-0.8} = 0.4444.$$

Si los valores u, d fueran distintos a los propuestos, utilizando (2.2), la probabilidad para el cual (S_n) es martingala se determina de manera única. $\square$

Otro grupo de procesos de especial interés en la modelación estocástica son los *procesos de Markov.*

Definición 2.1.8. *Un proceso (X_t) adaptado a la filtración $(\mathcal{F}_t)$ es un* ***proceso de Markov*** *con respecto $(\mathcal{F}_t)$ si para cualquiera t y $s > 0$, la distribución condicional de X_{t+s} dado $\mathcal{F}_t$ coincide con la distribución condicional de X_{t+s} dado X_t, i.e.,*

$$\mathbb{P}\Big(X_{t+s} \leq y|\mathcal{F}_t\Big) = \mathbb{P}\Big(X_{t+s} \leq y|X_t\Big), \mathbb{P}\text{-}c.s.. \tag{2.3}$$

A (2.3) se le conoce como la **propiedad de Markov**, y expresa que el comportamiento futuro del proceso es independiente del pasado, sólo depende del estado presente.

Las **probabilidades de transición** (o **funciones de transición de probabilidad**) de un proceso de Markov están definidas por

$$\mathbb{P}(s, x; t, B) = \mathbb{P}(X_t \in B|X_s = x), \quad s < t, B \in \mathcal{B}(\mathbb{R}).$$

Para s, x y t fijos, $\mathbb{P}(s, x; t, \cdot)$ es un medida de probabilidad en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Bajo las hipótesis anteriores, existe una función de densidad $p(s, x; t, \cdot)$, llamada **densidad de transición** (o **función de densidad de transición**), tal que

$$\mathbb{P}(s, x; t, B) = \int_B p(s, x; t, y)dy.$$

Un proceso de Markov se dice que es **homogéneo** si todas sus densidades de transición $p(s, x; t, y)$ dependen solamente de la diferencia de tiempos $t - s$, en lugar de valores específicos de s y t.

Una clase de procesos de Markov que es muy usada debido a las propiedades que la definen son los *procesos de difusión.* El movimiento Browniano, el movimiento Browniano con deriva y el movimiento Browniano geométrico son ejemplos de procesos de difusión. Los procesos de difusión no serán definidos en este trabajo de tesis.

2.2. Movimiento Browniano

En 1827, el médico y botánico R. Brown observó que cuando las partículas de polen se encontraban suspendidas sobre el agua y en otros líquidos, éstas se movían, sin cesar, de forma irregular. Posteriormente, el fenómeno también se asoció a partículas de materia inorgánica como polvo fino de algunos materiales (vidrio, carbón, roca, etc.) Sus investigaciones fueron publicadas en 1828.

En 1905, el físico A. Einstein publica un artículo sobre mecánica estadística que proporciona la formulación matemática del movimiento Browniano (en honor a R. Brown). Él postuló que los movimientos irregulares son causados por las colisiones entre el polen y las moléculas del líquido. Se asume que los golpeteos ocurren muy frecuentemente en cualquier intervalo de tiempo, independientemente de los otros; el efecto de un golpeteo en particular, en un tiempo dado, es pequeño comparado con el efecto total, más aún, es proporcional a dicho tiempo. Así A. Einstein obtiene las ecuaciones para el movimiento Browniano.

Los fundamentos matemáticos para el movimiento Browniano como un proceso estocástico fueron hechos por N. Wiener en 1931. En un sentido estricto, el movimiento Browniano es el fenómeno físico, mientras que su modelo matemático es el *proceso de Wiener*, aunque es común llamar a ambas cosas por el mismo nombre: *movimiento Browniano.*

Aquí no se mostrará la existencia, ni tampoco las diferencias entre la definición del movimiento Browniano y el proceso de Wiener (véase e.g., Oksendal[16] y Venegas[24]). Aquí sólo se limitará a la siguiente definición.

Definición 2.2.1. *Un proceso* (W_t) *es un **movimiento Browniano** (o* **proceso de Wiener***) de parámetro* σ^2 *si cumple las siguientes propiedades:*
a) $W_0 = 0, \mathbb{P}$*-c.s..*
b) *Incrementos independientes: Para* $0 \leq s < t$*,* $W_t - W_s$ *es independiente de* $\mathcal{F}_s^W$*.*
c) *Incrementos normales: Para* $0 \leq s < t$*,* $W_t - W_s \sim N(0, \sigma^2(t-s))$*.*
d) *Trayectorias continuas: Las trayectorias* $t \mapsto W_t$ *son continuas,* $\mathbb{P}$*-c.s..*

Nótese que el parámetro σ^2 es tal que $\sigma > 0$. Como caso particular, si $\sigma^2 = 1$ se dice que (W_t) es un ***movimiento Browniano estándar***. Cualquier movimiento Browniano no estándar puede convertirse en uno estándar a través del cambio de variable $\tau = \sigma^2 t$, o al considerar un nuevo proceso definido por $\frac{W_t}{\sigma}$. Entonces, sin pérdida de generalidad, a partir de ahora por movimiento Browniano se hará referencia a un movimiento Browniano estándar.

Si en la propiedad **c)** se considera $s = 0$, entonces se tiene que $W_t \sim N(0, t)$. Además, ésta propiedad nos asegura que los ***incrementos son estacionarios***, esto es,

$$W_t - W_s \stackrel{d}{=} W_{t+h} - W_{s+h}, \text{ para cualquier } h \geq 0.$$

Las condiciones **b)** y **c)** son equivalentes a solicitar que las *fidis* del proceso cumplan la siguiente propiedad:

b') *Para cualquier sucesión finita de tiempos* $0 < t_1 < \ldots < t_n$ *y conjuntos de Borel* $A_1, \ldots, A_n \subseteq \mathbb{R}$, *se cumple que*

$$\mathbb{P}(W_{t_1} \in A_1, \ldots, W_{t_n} \in A_n) = \int_{A_1} \cdots \int_{A_n} p(0,0;t_1,x_1).p(t_1,x_1;t_2,x_2) \ldots p(t_{n-1},x_{n-1};t_n,x_n)\, dx_n \ldots dx_1,$$

en donde los integrandos de la integral de Lebesgue están definidos por

$$p(s,x;t,y) = \frac{1}{\sqrt{2\pi(t-s)}} \exp\left(-\frac{(y-x)^2}{2(t-s)}\right), \quad \textit{para } x,y \in \mathbb{R}. \tag{2.4}$$

La función definida en (2.4) se conoce como **densidad de transición** del movimiento Browniano.

La propiedad anterior expresa que las *fidis* del movimiento Browniano pueden ser calculadas con la ayuda de la función de probabilidad de transición y usando la hipótesis de incrementos independientes:

$$\begin{aligned}\mathbb{P}(W_{t_1} \leq x_1, W_{t_2} \leq x_2, \ldots, W_{t_n} \leq x_n) &= \int_{-\infty}^{x_1} p(0,0;t_1,y_1)dy_1 \int_{-\infty}^{x_2} p(t_1,y_1;t_2,y_2)dy_2 \\ &\quad \ldots \int_{-\infty}^{x_n} p(t_{n-1},y_{n-1};t_n,y_n)dy_n.\end{aligned}$$

De la expresión anterior se puede derivar que, para cada $\tau > 0$ y cualquier elección de $t_i \geq 0$, $i = 1, \ldots, n$,

$$(\sqrt{\tau}W_{t_1}, \ldots, \sqrt{\tau}W_{t_n}) \stackrel{d}{=} (W_{\tau t_1}, \ldots, W_{\tau t_n}),$$

i.e., el movimiento Browniano es 0.5-autosimilar. Por lo tanto, se cumple que $(\sqrt{\tau}W_t) \stackrel{d}{=} (W_{\tau t})$. Más aún, se cumple la propiedad

$$\left(\sqrt{\tau}W_{\frac{t}{\tau}}\right) \stackrel{d}{=} (W_t), \text{ para cualquier } \tau > 0,$$

la cual es conocida como *propiedad de escalamiento.*

La condición de que el movimiento Browniano inicie en el origen no es absolutamente necesario. Pueden considerarse trayectorias Brownianas que inicien en un punto arbitrario x a través del proceso definido por $W_t + x$.

La condición de *trayectorias continuas* es en el sentido de que dado un proceso Browniano (W_t), siempre existe una versión continua de (W_t). Ésto se puede deducir de la normalidad de los incrementos y apelando al *Teorema de Continuidad de Kolmogorov*, ya que

$$E\Big[|W_t - W_s|^4\Big] = 3(t-s)^2.$$

De aquí en adelante, sin perdida de generalidad, se supondrá que los movimientos Brownianos son versiones continuas.

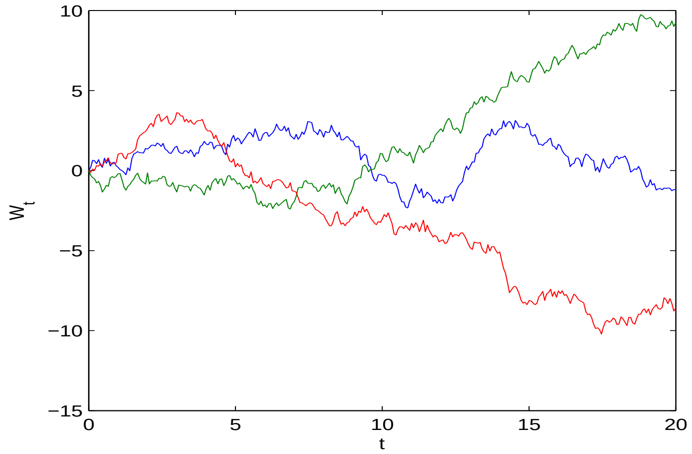

Figura 2.2: Tres trayectorias de (W_t).

En la Figura (2.2) se muestra el comportamiento de tres trayectorias de un movimiento Browniano que inicia en el origen. En el siguiente resultado se resumen algunas de sus propiedades.

Teorema 2.2.1. *Sea (W_t) un movimiento Browniano. Entonces se cumplen las siguientes propiedades:*

a) *Tiene función de covarianza $\gamma(t,s) = Cov(W_t, W_s) = \min(t,s)$.*

b) *La variación cuadrática de las trayectorias sobre $[0,T]$ es T, $\mathbb{P}$-c.s..*

c) *La variación de las trayectorias sobre $[0,T]$ es infinita, $\mathbb{P}$-c.s..*

d) *Las trayectorias no son diferenciables en t, $\mathbb{P}$-c.s., para cualquier $t \geq 0$.*

e) *(W_t) es una martingala con respecto a su filtración natural.*

f) *(W_t) es un proceso de Markov homogéneo con respecto a su filtración natural.*

Demostración.

a) Claramente se cumple que $E[W_t] = 0$. Luego,

$$\gamma(t,s) = E[W_s W_t].$$

Si $0 \leq s < t$, por la independencia de los incrementos, se tiene que

$$\begin{aligned} E[W_s W_t] &= E[W_s(W_t - W_s)] + E[W_s^2] \\ &= E[W_s]E[W_t - W_s] + E[W_s^2] \\ &= s. \end{aligned}$$

Análogamente, si $0 \leq t < s$, entonces $E[W_s W_t] = t$.

Por lo tanto,

$$\gamma(t,s) = \text{mín}(t,s).$$

b) Sea

$$0 = t_0^n < t_1^n < \ldots < t_n^n = T, \qquad \text{donde } t_i^n := \frac{iT}{n}, \tag{2.5}$$

una partición uniforme del intervalo $[0,T]$ en n subintervalos.

Se define los correspondientes incrementos del proceso Browniano por

$$\Delta_i^n W = W_{t_i^n} - W_{t_{i-1}^n}.$$

Puesto que los incrementos $\Delta_i^n W$ son independientes, entonces,

$$E[\Delta_i^n W] = 0, \quad E[|\Delta_i^n W|^2] = \frac{T}{n}, \quad E[|\Delta_i^n W|^4] = \frac{3T^2}{n^2}.$$

Luego,

$$\begin{aligned}
E\left[\left(\sum_{i=1}^n |\Delta_i^n W|^2 - T\right)^2\right] &= E\left[\left(\sum_{i=1}^n \left[|\Delta_i^n W|^2 - \frac{T}{n}\right]\right)^2\right] \\
&= \sum_{i=1}^n E\left[\left(|\Delta_i^n W|^2 - \frac{T}{n}\right)^2\right] \\
&= \sum_{i=1}^n \left(E[|\Delta_i^n W|^4] - 2\frac{T}{n}E[(\Delta_i^n W)^2] + \frac{T^2}{n^2}\right) \\
&= \sum_{i=1}^n \left(\frac{3T^2}{n^2} - \frac{T^2}{n^2}\right) \\
&= \frac{2T^2}{n} \to 0, \text{ cuando } n \to \infty.
\end{aligned}$$

De lo anterior, se concluye que

$$\sum_{i=1}^n |\Delta_i^n W|^2 \xrightarrow{L^2} T, \text{ cuando } n \to \infty,$$

esto es,

$$\lim_{n\to\infty} \sum_{i=1}^n |\Delta_i^n W|^2 = T \text{ en } L^2.$$

En consecuencia, existe una subsucesión $t_i^{n_k}$ tal que

$$\lim_{k\to\infty} \sum_{i=1}^{n_k} |\Delta_i^{n_k} W|^2 = T, \mathbb{P}\text{-c.s.}. \tag{2.6}$$

Por lo tanto, el movimiento Browniano tiene variación cuadrática T sobre $[0,T]$, $\mathbb{P}$-c.s..

c) De forma análoga al inciso anterior, considérese una partición uniforme del intervalo $[0,T]$.

Entonces

$$Q_n = \sum_{i=1}^n |\Delta_i^n W|^2 \leq \Big(\max_{i=1,\ldots,n} |\Delta_i^n W|\Big) \sum_{i=1}^n |\Delta_i^n W|.$$

Puesto que las trayectorias de (W_t) son continuas en $[0, T], \mathbb{P}$-c.s.,

$$\lim_{n\to\infty}\left(\max_{i=0,\ldots,n-1}|\Delta_i^n W|\right) = 0, \ \mathbb{P}\text{-c.s.}$$

Combinando (2.6) y la igualdad anterior, si

$$\lim_{n\to\infty}\sum_{i=1}^{n}|\Delta_i^n W| = k < \infty, \ \mathbb{P}\text{-c.s.},$$

entonces $T = 0(k) = 0, \mathbb{P}$-c.s., lo cual es imposible.

Por lo tanto,

$$\lim_{n\to\infty}\sum_{i=1}^{n}|\Delta_i^n W| = \infty, \ \mathbb{P}\text{-c.s.}.$$

d) Considérese un intervalo de tiempo de longitud $\Delta t = \frac{1}{n}$ que empieza en t arbitrario. La tasa de cambio sobre el intervalo de tiempo $[t, t + \Delta t]$ es

$$V(n) := \frac{W_{t+\Delta t} - W_t}{\Delta t} = n\Big[W_{t+\frac{1}{n}} - W_t\Big].$$

Luego, $V(n)$ tiene distribución normal con parámetros

$$E[V(n)] = nE\Big[W_{t+\frac{1}{n}} - W_t\Big] = n(0) = 0,$$
$$Var(V(n)) = n^2 Var\Big(W_{t+\frac{1}{n}} - W_t\Big) = n,$$

i.e.,

$$V(n) \stackrel{d}{=} \sqrt{n}Z, \quad \text{donde } Z \sim N(0,1).$$

Sea K un número positivo arbitrario, por las propiedades de la probabilidad en procedimientos de límites, se sigue que

$$\begin{aligned}
\mathbb{P}\Big(\lim_{\Delta t\to 0}|V(n)| > K\Big) &= \mathbb{P}\Big(\lim_{n\to\infty}|V(n)| > K\Big) \\
&= \lim_{n\to\infty}\mathbb{P}(|V(n)| > K) \\
&= \lim_{n\to\infty}\mathbb{P}(|\sqrt{n}Z| > K) \\
&= \lim_{n\to\infty}\mathbb{P}\left(|Z| > \frac{K}{\sqrt{n}}\right) \\
&= \mathbb{P}\left(\lim_{n\to\infty}|Z| > \frac{K}{\sqrt{n}}\right) \\
&= \mathbb{P}(|Z| > 0) \\
&= 1.
\end{aligned}$$

Ya que K puede ser elegido arbitrariamente grande, la tasa de cambio en el tiempo t no es finito. Puesto que el tiempo t es arbitrario, entonces las trayectorias Brownianas no son diferenciables en todo $[0, T]$, $\mathbb{P}$-c.s..

e) Por definición, $W_t \sim N(0,t)$, luego entonces $E[|W_t|] < \infty$ para todo t.

Si $s \leq t$, entonces $W_t - W_s$ es independiente de la σ-álgebra $\mathcal{F}_s^W$, por lo tanto

$$\begin{aligned} E\Big[W_t|\mathcal{F}_s^W\Big] &= E\Big[W_s + (W_t - W_s)|\mathcal{F}_s^W\Big] \\ &= E\Big[W_s|\mathcal{F}_s^W\Big] + E\Big[W_t - W_s|\mathcal{F}_s^W\Big] \\ &= W_s, \end{aligned}$$

i.e., $E[W_t|\mathcal{F}_s^W] = W_s$ para todo $s \leq t$.

f) La demostración de este resultado hace uso del Teorema 6 (véase Apéndice A.4). En efecto, para algún $h > 0$ tal que $|u| < h$,

$$\begin{aligned} E\Big[\exp(uW_{t+s})|\mathcal{F}_t^W\Big] &= \exp(uW_t)E\Big[\exp(u(W_{t+s} - W_t))|\mathcal{F}_t^W\Big] \\ &= \exp(uW_t)E\Big[\exp(u(W_{t+s} - W_t))\Big] && (2.7) \\ &= \exp(uW_t)\exp\left(\frac{u^2 s}{2}\right) && (2.8) \\ &= \exp(uW_t)E\Big[\exp(u(W_{t+s} - W_t))|W_t\Big] \\ &= E\Big[\exp(uW_{t+s})|W_t\Big]. \end{aligned}$$

En la igualdad (2.7) se utilizó el hecho de que $\exp(u(W_{t+s} - W_t))$ es independiente de $\mathcal{F}_t^W$; mientras que (2.8) fue consecuencia de que $W_{t+s} - W_t \sim N(0,s)$.

Luego, por el Teorema 6 se concluye que (W_t) cumple la propiedad de Markov.

Por otra parte, nótese que la **densidad de transición** del movimiento Browniano sólo depende de diferencias de tiempo. Por lo tanto, (W_t) es un proceso de Markov homogéneo.

■

A fin de cuentas, en **b)** se desea encontrar $V(T)$ tal que se cumpla

$$\lim_{n\to\infty} E\left[\left(\sum_{i=1}^{n} |\Delta_i^n W|^2 - V(T)\right)^2\right] = 0.$$

Lo siguiente es una de las motivaciones que influye en la elección $V(T) = T$:

$$\begin{aligned} V(T) &= E\left[\sum_{i=1}^{n} |\Delta_i^n W|^2\right] \\ &= \sum_{i=1}^{n} E[(\Delta_i^n W)^2] \\ &= \sum_{i=1}^{n} (t_i - t_{i-1}) \\ &= T. \end{aligned}$$

Además, es independiente de n.

El movimiento Browniano es el proceso que ha causado gran impacto, tanto para la introducción de nuevos objetos matemáticos como en la modelación de fenómenos aleatorios.

Una primera importante generalización del movimiento Browniano es la que a continuación se define.

Definición 2.2.2 (Movimiento Browniano con Deriva). *Sea (W_t) un movimiento Browniano. El proceso (X_t) definido por*

$$X_t = \sigma W_t + \mu t, \quad \mu \in \mathbb{R}, \sigma > 0,$$

se conoce como **movimiento Browniano con deriva** μ. *La constante σ^2 se conoce como* coeficiente de difusión *(o coeficiente de variación).*

Nótese que para cada $t \in T$, $E[X_t] = \mu t$. Esta función determinista influye esencialmente en la forma de las trayectorias del proceso, de aquí que al proceso se le conozca como *movimiento Browniano con deriva.*

Por otra parte, el término "*coeficiente de difusión*" proviene del hecho de que σ^2 coincide con cierto límite que definen a un proceso de difusión (véase e.g., Lefebvre[14], Capítulo 4.).

Aún cuando el movimiento Browniano es una de las bases en la construcción de los modelos financieros, éste no puede, por sí mismo, representar el comportamiento de todas las variables financieras. Los precios de los activos, por ejemplo, no son descritos apropiadamente por el movimiento Browniano, ni por el movimiento Browniano con deriva, ya que los precios no son negativos. Este inconveniente determinó que el modelo de Bachelier no fuera muy convincente. En 1973, Black, Scholes y Merton sugirieron otro proceso como modelo para la especulación de precios, el cual se conoce como *movimiento Browniano geométrico.*

Definición 2.2.3 (Movimiento Browniano Geométrico). *Sea (X_t) un movimiento Browniano con deriva μ y coeficiente de difusión σ^2. Entonces, el proceso (Y_t) definido por*

$$Y_t = \exp(X_t) = \exp(\sigma W_t + \mu t),$$

se llama **movimiento Browniano geométrico**.

El movimiento Browniano geométrico se obtiene por una transformación exponencial del movimiento Browniano. Para un t fijo, la v.a. Y_t tiene una *distribución lognormal* con media μt y varianza $\sigma^2 t$, i.e.,

$$\ln(Y_t) \sim N(\mu t, \sigma^2 t).$$

Se puede generalizar la definición del movimiento Browniano geométrico (Y_t) al definir

$$Y_t = Y_0 \exp(X_t),$$

donde Y_0 puede ser una v.a..

Dado que el movimiento Browniano es independiente del pasado y tiene incrementos estacionarios, entonces se cumple el siguiente resultado.

Teorema 2.2.2. *El movimiento Browniano geométrico es un proceso de Markov con respecto a la filtración* $(\mathcal{F}_t^W)$.

Demostración.

Sea (Y_t) un movimiento Browniano geométrico definido en su forma general por,

$$Y_t = Y_0 \exp(X_t).$$

Por demostrar que (Y_t) cumple con la propiedad de Markov. En efecto,

$$\begin{aligned}
\mathbb{P}\Big(Y_{t+s} \leq y|\mathcal{F}_t^W\Big) &= \mathbb{P}\Big(Y_0 \exp(X_{t+s}) \leq y|\mathcal{F}_t^W\Big) \\
&= \mathbb{P}\Big(Y_t \exp(X_{t+s} - X_t) \leq y|\mathcal{F}_t^W\Big) \\
&= \mathbb{P}\left(\exp(X_{t+s} - X_t) \leq \frac{y}{Y_t}\Big|\mathcal{F}_t^W\right) \\
&= \mathbb{P}\left(\exp(X_{t+s} - X_t) \leq \frac{y}{Y_t}\Big|Y_t\right) \\
&= \mathbb{P}\Big(Y_t \exp(X_{t+s} - X_t) \leq y|Y_t\Big) \\
&= \mathbb{P}(Y_{t+s} \leq y|Y_t).
\end{aligned}$$

Por lo tanto, el movimiento Browniano geométrico es un proceso de Markov. ■

Nótese que si $Z \sim N(0,1)$, entonces

$$\begin{aligned}
E[\exp(\lambda Z)] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(\lambda z) \exp\Big(-\frac{z^2}{2}\Big)\, dz \\
&= \exp\left(\frac{\lambda^2}{2}\right) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\Big(-\frac{(z-\lambda)^2}{2}\Big)\, dz \\
&= \exp\left(\frac{\lambda^2}{2}\right).
\end{aligned}$$

Luego, como el movimiento Browniano es 0.5-autosimilar,

$$\begin{aligned}
E[Y_t] &= Y_0 \exp(\mu t) E\Big[\exp(\sigma W_t)\Big] \\
&= Y_0 \exp(\mu t) E\Big[\exp\Big(\sigma t^{\frac{1}{2}} W_1\Big)\Big] \\
&= Y_0 \exp\left(\Big(\mu + \frac{1}{2}\sigma^2\Big) t\right).
\end{aligned}$$

Esto es, bajo el modelo del movimiento Browniano geométrico, dado un precio inicial Y_0, se espera que el precio del activo aumente en una tasa de $\mu + \frac{\sigma^2}{2}$.

El siguiente resultado muestra algunos ejemplos de procesos martingalas que involucran al movimiento Browniano.

Teorema 2.2.3. *Supóngase que (W_t) es un movimiento Browniano con respecto a la filtración $(\mathcal{F}_t)$. Entonces*

a) *(W_t) es una $\mathcal{F}_t$-martingala,*

b) *$(W_t^2 - t)$ es una $\mathcal{F}_t$-martingala, y*

c) *$\left(\exp\left(\sigma W_t - \frac{\sigma^2}{2}t\right)\right)$ es una $\mathcal{F}_t$-martingala.*

Demostración.
Aquí sólo se demostrará el inciso **c)**.
En efecto, si $s < t$, entonces,

$$E\Big[\exp\Big(\sigma W_t - \frac{\sigma^2}{2}t\Big)\Big|\mathcal{F}_s\Big] = \exp\left(\sigma W_s - \frac{\sigma^2}{2}t\right)E\Big[\exp(\sigma(W_t - W_s))|\mathcal{F}_s\Big].$$

Y utilizando el hecho de que W_s es $\mathcal{F}_s$-medible, se sigue que

$$\begin{aligned} E\Big[\exp\Big(\sigma W_t - \frac{\sigma^2}{2}t\Big)\Big|\mathcal{F}_s\Big] &= \exp\left(\sigma W_s - \frac{\sigma^2}{2}t\right)E\Big[\exp(\sigma(W_t - W_s))\Big] \\ &= \exp\left(\sigma W_s - \frac{\sigma^2}{2}t\right)E\Big[\exp(\sigma W_{t-s})\Big] \\ &= \exp\left(\sigma W_s - \frac{\sigma^2}{2}t\right)E\Big[\exp\Big(\sigma\sqrt{t-s}Z\Big)\Big] \\ &= \exp\left(\sigma W_s - \frac{\sigma^2}{2}t\right)\exp\Big(\frac{1}{2}\sigma^2(t-s)\Big) \\ &= \exp\left(\sigma W_s - \frac{\sigma^2}{2}s\right). \end{aligned}$$

Por lo tanto, $\left(\exp\left(\sigma W_t - \frac{\sigma^2}{2}t\right)\right)$ es una martingala. ∎

El siguiente ejemplo muestra cómo aparece el movimiento Browniano geométrico en la modelación de los precios de un activo financiero.

Ejemplo 2.3. *Sea Δt un incremento pequeño. Supongamos que en cada unidad de tiempo Δt el precio del activo aumenta su valor por un factor u con probabilidad p, o disminuye por un factor d con probabilidad $1-p$, donde*

$$u = \exp\Big(\sigma\sqrt{\Delta t}\Big), \qquad d = \exp\Big(-\sigma\sqrt{\Delta t}\Big),$$

$$p = \frac{1}{2}\left(1 + \frac{\mu}{\sigma}\sqrt{\Delta t}\right).$$

Se define

$$Y_i = \begin{cases} 1 & \textit{si el precio aumenta en el i-ésimo } \Delta t \\ 0 & \textit{si el precio disminuye en el i-ésimo } \Delta t. \end{cases}$$

Ahora, el número de veces en que el activo aumenta de precio en los primeros n *incrementos de tiempo está dado por* $Y = \sum_{i=1}^{n} Y_i$*, y por tanto,* $n - Y$ *es el número de veces que ha disminuido. De aquí que, el precio en el tiempo final puede ser expresado como*

$$S(n\Delta t) = S(0)u^Y d^{n-Y} = S(0)d^n\left(\frac{u}{d}\right)^Y. \tag{2.9}$$

Considérese $n = \frac{t}{\Delta t}$*, luego (2.9) queda expresada como*

$$\frac{S(t)}{S(0)} = d^n\left(\frac{u}{d}\right)^Y.$$

Aplicando logaritmos se tiene

$$\ln\left(\frac{S(t)}{S(0)}\right) = n\ln(d) + Y\ln\left(\frac{u}{d}\right) = -\frac{t\sigma}{\sqrt{\Delta t}} + 2\sigma\sqrt{\Delta t}Y. \tag{2.10}$$

Si $\Delta t \to 0$*, entonces* $n \to \infty$*. Por el* Teorema del Límite Central*,* Y *tiende a tener una distribución normal cada vez más rápido; luego, la ecuación (2.10) implica que* $\ln\left(\frac{S(t)}{S(0)}\right)$ *es una variable aleatoria normal.*

Más aún,

$$\begin{aligned} E\left[\ln\left(\frac{S(t)}{S(0)}\right)\right] &= -\frac{t\sigma}{\sqrt{\Delta t}} + 2\sigma\sqrt{\Delta t}E[Y] \\ &= -\frac{t\sigma}{\sqrt{\Delta t}} + 2\sigma\sqrt{\Delta t}np \qquad (2.11) \\ &= -\frac{t\sigma}{\sqrt{\Delta t}} + \sigma\sqrt{\Delta t}\left(\frac{t}{\Delta t}\right)\left(1 + \frac{\mu}{\sigma}\sqrt{\Delta t}\right) \\ &= \mu t, \end{aligned}$$

$$\begin{aligned} Var\left(\ln\left(\frac{S(t)}{S(0)}\right)\right) &= 4\sigma^2\Delta t Var(Y) \\ &= 4\sigma^2\left(\frac{t}{n}\right)np(1-p) \\ &\approx \sigma^2 t. \qquad (2.12) \end{aligned}$$

La igualdad (2.11) fue consecuencia de que $Y \sim Bin(n, p)$*; mientras que (2.12) se derivó del hecho de que, para* $\Delta << 0$*,* $p \approx 1/2$*.*

Todo lo anterior muestra que cuando $\Delta t << 0$*,* $\ln\left(\frac{S(t)}{S(0)}\right)$ *(de forma análoga para* $\ln\left(\frac{S(t+y)}{S(y)}\right)$*) tiende a comportarse como una v.a. normal con media* μt *y varianza* $\sigma^2 t$*. Además, ya que los cambios de precio en lo sucesivo son independientes y tienen la misma* *probabilidad de aumentar o disminuir, entonces se sigue que* $\frac{S(t+y)}{S(y)}$ *es independiente de los cambios de precios anteriores al tiempo* y. *De aquí que, cuando* $\Delta t \to 0$, *el modelo* $S(t)$ *se aproxima a un movimiento Browniano geométrico.* □

2.3. La integral estocástica

Como se mencionó al principio del capítulo, una de las primeras aplicaciones del movimiento Browniano fue propuesta por Bachelier. En resumen, combinando el modelo propuesto por Bachelier con las terminologías modernas, si X_t representa el precio en el tiempo t, él asumió que los incrementos infinitesimales dX_t son proporcionales a los incrementos dW_t del movimiento Browniano, i.e.,

$$dX_t = \sigma dW_t, \quad \sigma > 0. \tag{2.13}$$

Así, una acción con precio inicial $X_0 = x$, en el tiempo t tendrá el valor

$$X_t = x + \sigma W_t. \tag{2.14}$$

Nótese que el modelo definido por la ecuación (2.13) permite que en algún momento la acción tenga valor negativo. Para remediar esta falla, se observó que el inversionista trabaja en términos de la proporción entre su potencial ganado o perdido $dX(t)$ y la suma invertida $X(t)$. Por lo tanto, es el precio relativo $\frac{dX_t}{X_t}$ de una acción la que reacciona a las fluctuaciones del mercado, i.e., deberá ser proporcional a dW_t :

$$dX_t = \sigma X_t dW_t. \tag{2.15}$$

Formalmente, (2.15) representa una ecuación diferencial, pero esto inmediatamente conduce a dificultades porque las trayectorias de W_t no son diferenciables, y por tanto el cálculo de Leibniz-Newton no aplica.

K. Itô da un significado riguroso a ecuaciones como (2.15), escribiendo éstos como ecuaciones integrales que involucran a un nuevo tipo de integral. En particular, (2.15) queda expresado como

$$X_t = x + \sigma \int_0^t X_t dW_t, \tag{2.16}$$

en donde la integral con respecto a W_t es llamada la ***integral estocástica de Itô***, **integral de Itô**, o simplemente **integral estocástica**.

El concepto de integral, para el caso en que el integrador es un movimiento Browniano y el integrando es un conjunto de funciones aleatorias con determinadas propiedades, fue desarrollado por K. Itô en 1944 en su artículo titulado *Stochastic Integral*, publicado en la

Academia Imperial de Tokio. Aunque, como Itô mismo reconoce, el primero en definir la integral estocástica fue N. Wiener en 1934, pero para integrandos no aleatorios, suaves y de cuadrado integrable. Debido a que fue Itô quien trabajó el caso general, generalmente la teoría del cálculo estocástico se le atribuye a él.

El método que Itô usó para definir la integral estocástica es una combinación de las técnicas de la integral de Riemann-Stieltjes (en lo que se refiere al integrador) y la integral de Lebesgue (en lo que se refiere al integrando).

El cálculo de Itô fue originalmente motivado como un método directo para la construcción de procesos de difusión. La construcción de tales procesos se consigue vía tres pasos fundamentales: la teoría de Hille-Yosida, el teorema de representación de Riesz, y el teorema de extensión de Kolmogorov. Sin embargo, Itô desarrolla la teoría de integración estocástica para construir esos procesos de difusión de un sólo paso como solución de ecuaciones integrales estocásticas.

Durante estas últimas seis décadas, la teoría de la integral estocástica se ha estado estudiando extensivamente, así como aplicando a un amplio rango de campos científicos que involucran el estudio de comportamientos irregulares. Tal vez la más notable aplicación es la teoría de Black-Scholes en finanzas, por la cual R. C. Merton y M. S. Scholes ganaron el Premio Nobel de Economía en 1997.

2.3.1. Construcción y propiedades de la integral estocástica

Existen, al menos, dos grandes diferencias entre la integral de Riemann y la integral de Itô. La primera es el tipo de convergencia. Las aproximaciones de la integral de Riemann convergen en $\mathbb{R}$, mientras que la integral de Itô se aproximará por sucesiones de v.a.'s en L^2.

La otra diferencia es que las sumas de Riemann que aproximan la integral de una función $f : [0, T] \to \mathbb{R}$ son de la forma

$$\sum_{j=1}^{n} f(s_j)(t_j - t_{j-1}),$$

donde $0 = t_0 < t_1 < \ldots < t_n = T$ y $s_j \in [t_{j-1}, t_j]$ es un punto arbitrario para cada j.

El valor de la integral de Riemann *no depende* de la elección de los puntos s_j. Para el caso estocástico, las sumas son de la forma

$$\sum_{j=1}^{n} f(s_j)(W_{t_j} - W_{t_{j-1}}),$$

pero en este caso, el límite de la aproximación *sí depende* de la elección de $s_j \in [t_{j-1}, t_j]$.

Por el Teorema 2.2.1, parece razonable escribir

$$\int_0^T (dW_t)^2 = \lim_{n\to\infty} \sum_{i=1}^{n} (\Delta_i^n W)^2 = T, \quad \text{en } L^2,$$

i.e.,

$$\int_0^T (dW_t)^2 = T. \tag{2.17}$$

En algunos textos es frecuente encontrar otra notación alternativa:

$$\int_0^T d\langle W_t \rangle = T.$$

Con cierto abuso en la notación, (2.17) se puede escribir en forma diferencial como

$$(dW_t)^2 = dt. \tag{2.18}$$

En términos estrictos, los objetos de estudio del cálculo estocástico son integrales y no diferenciales. Cuando se escribe una ecuación diferencial estocástica, realmente se está pensando en una integral estocástica. Así pues, la ecuación (2.18) es una notación simplificada de (2.17).

Como ejemplo ilustrativo, considérese una partición uniforme como en (2.5), y defínase la suma

$$S_n = \sum_{i=1}^{n} W_{t_{i-1}^n} \Delta_i^n W.$$

Luego,

$$\begin{aligned}
\lim_{n\to\infty} S_n &= \lim_{n\to\infty} \sum_{i=1}^{n} W_{t_{i-1}^n}\Big(W_{t_i^n} - W_{t_{i-1}^n}\Big) \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} \left(\frac{1}{2}\Big(W_{t_i^n}^2 - W_{t_{i-1}^n}^2\Big) - \frac{1}{2}(\Delta_i^n W)^2\right) \\
&= \lim_{n\to\infty} \left(\frac{1}{2}W_T^2 - \frac{1}{2}\sum_{i=1}^{n}(\Delta_i^n W)^2\right) \\
&= \frac{1}{2}W_T^2 - \frac{1}{2}T, \quad \text{en } L^2.
\end{aligned}$$

Por lo tanto,

$$\int_0^T W_t dW_t = \frac{1}{2}(W_T^2 - T) = \frac{1}{2}W_T^2 - \frac{1}{2}\int_0^T dt. \tag{2.19}$$

Mediante un análisis similar, es posible comprobar que

$$\int_0^T W_t^2 dW_t = \frac{1}{3}W_T^3 - \int_0^T W_t dt. \tag{2.20}$$

En la definición clásica de la integral de Itô, el integrando es evaluado en el extremo izquierdo de cada subintervalo. Dado que nuestro interés es su aplicación en finanzas, esta elección encaja sin ningún problema, ya que con ello se asegura que el futuro no interviene en las acciones presentes.

El valor de la integral se modifica cuando se elige el extremo derecho, de hecho

$$\int_0^T W_t dW_t = \frac{1}{2}(W_T^2 + T),$$

y cuando se elige el punto medio, se obtiene

$$\int_0^T W_t dW_t = \frac{1}{2}W_T^2.$$

En este último caso, la integral se conoce como la **integral de Stratonovich**.

Dada las motivaciones anteriores, a continuación se define la integral estocástica. Si el lector está interesado en los detalles de la construcción, se recomienda revisar Kloeden and Platen[12] como texto base. En el transcurso de la construcción se notará que las herramientas indispensables son el análisis funcional, teoría de la medida, y más resultados de procesos estocásticos.

Se empezará a estudiar la integral estocástica para una clase de procesos cuyas trayectorias toman un número finito de valores.

Definición 2.3.1. *Sea $0 = t_0 < t_1 < \ldots < t_n = T$ una partición finita del intervalo $[0, T]$. Un proceso estocástico* **simple** *es un proceso (C_t) definido por*

$$C_t = \sum_{i=1}^{n} Z_i \cdot I_{[t_{i-1}, t_i)}(t), \quad C_T = Z_n,$$

en donde $E[Z_i^2] < \infty$ para todo i, y la sucesión de v.a.'s (Z_i) es adaptado a $\left(\mathcal{F}_{t_{i-1}}^W\right)$, i.e., Z_i es una función de W_s, $s \leq t_{i-1}$.

Entonces, un proceso simple es un proceso constante por pedazos, con trayectorias continuas por la derecha, y con límite por la izquierda. Las condiciones solicitadas garantizan que (C_t) sea de cuadrado integrable para cada t y adaptado a la filtración $\mathcal{F}_t$.

La definición intuitiva de integral establece que si el integrando es constante en alguna parte del intervalo de tiempo, entonces la integral debe ser la constante multiplicada por el incremento del movimiento Browniano de dicho intervalo.

Definición 2.3.2. *La* **integral estocástica** *de un proceso simple (C_t) en $[0, T]$ respecto del movimiento Browniano, denotada por $I_T(C)$, está dada por*

$$I_T(C) = \int_0^T C_t dW_t = \sum_{i=1}^{n} C_{t_{i-1}}(W_{t_i} - W_{t_{i-1}}) = \sum_{i=1}^{n} Z_i \Delta_i W.$$

Además, si $t_{k-1} \leq t \leq t_k$, entonces

$$
\begin{aligned}
I_t(C) = \int_0^t C_s dW_s &= \int_0^T C_s I_{[0,t]}(s) dW_s \\
&= \sum_{i=1}^{k-1} Z_i \Delta_i W + Z_k (W_t - W_{t_{k-1}}),
\end{aligned}
$$

donde $\sum_{i=1}^{0} Z_i \Delta_i W = 0$.

Nótese que la integral estocástica es un proceso integrable, pues siendo Z_i y $\Delta_i W$ independientes, cada sumando tiene esperanza cero, y por lo tanto $E[I_t(C)] = 0$. Más aún, $I_t(C)$ es de cuadrado integrable. El siguiente resultado muestra otras propiedades de la integral estocástica de procesos simples.

Teorema 2.3.1. *Sea $I_t(C)$ la integral estocástica de un proceso simple (C_t). Entonces se cumplen las siguientes propiedades:*

a) Isometría de Itô:

$$
E[(I_t(C))^2] = E\left[\int_0^t C_s^2 ds\right] = \int_0^t E[C_s^2] ds, \quad t \in [0, T]. \tag{2.21}
$$

b) *$(I_t(C))$ es una martingala con respecto a la filtración $\left(\mathcal{F}_t^W\right)$.*

c) *$(I_t(C))$ es un proceso con trayectorias continuas.*

d) Linealidad: *Para una constante c y procesos simples $(C_t^{(1)}), (C_t^{(2)})$, se cumple,*

$$
I_t\left(cC^{(1)} + C^{(2)}\right) = cI_t(C^{(1)}) + I_t(C^{(2)}).
$$

e) *Para $0 \leq t_0 \leq t$,*

$$
I_t(C) = I_{t_0}(C) + \int_{t_0}^t C_s dW_s.
$$

Demostración.

Aquí sólo se demostraran las tres primeras propiedades.

a) En efecto, sin perdida de generalidad supóngase que $t = t_k$ para algún k. Luego,

$$
\begin{aligned}
E[(I_t(C))^2] &= E\left[\left(\sum_{i=1}^{k} Z_i \Delta_i W\right)^2\right] \\
&= E\left[\sum_{j=1,i=1}^{k} Z_j Z_i \Delta_j W \Delta_i W\right] \\
&= E\left[\sum_{i=1}^{k} Z_i^2 (\Delta_i W)^2\right] \\
&= \sum_{i=1}^{k} E[Z_i^2] \Delta t_i \\
&= \int_0^t E[Z_s^2] ds.
\end{aligned}
$$

Por lo tanto,

$$E[(I_t(C))^2] = \int_0^t E[C_s^2]ds.$$

b) Por la definición de procesos simples, para t arbitrario, las v.a.'s $Z_1, \ldots, Z_k$ y $\Delta_1 W, \ldots, \Delta_{k-1}W, W_t - W_{t_{k-1}}$ son funciones del movimiento Browniano hasta el tiempo t. Entonces $I_t(C)$ es una v.a. $\mathcal{F}_t^W$-medible, y por tanto $(I_t(C))$ es adaptado a la filtración $(\mathcal{F}_t^W)$.

Por la propiedad de isometría, $E[I_t(C)] < \infty$ para todo $0 \leq t \leq T$.

Sólo resta probar que $E\Big[I_t(C)\Big|\mathcal{F}_s^W\Big] = I_s(C)$, para $s < t$. En efecto, supóngase que $s \in [t_{j-1}, t_j]$ y $t \in [t_{k-1}, t_k]$ para algún $j \leq k$,

$$\begin{aligned}
E\Big[I_t(C)|\mathcal{F}_s^W\Big] &= E\Bigg[I_{t_{j-1}}(C) + Z_j(W_s - W_{t_{j-1}}) + Z_j(W_{t_j} - W_s) + \sum_{i=j+1}^{k-1} Z_i \Delta_i W \\
&\quad + Z_k(W_t - W_{t_{k-1}})\Bigg|\mathcal{F}_s^W\Bigg] \\
&= E\Big[I_s(C)\Big|\mathcal{F}_s^W\Big] + E\Big[Z_j(W_{t_j} - W_s)\Big|\mathcal{F}_s^W\Big] + \sum_{i=j+1}^{k-1} E\Big[Z_i \Delta_i W\Big|\mathcal{F}_s^W\Big] \\
&\quad + E\Big[Z_k(W_t - W_{t_{k-1}})\Big|\mathcal{F}_s^W\Big] \\
&= I_s(C) + Z_j E[W_{t_j} - W_s] + \sum_{i=j+1}^{k-1} E\Big[E\Big[Z_i \Delta_i W\Big|\mathcal{F}_t^W\Big]\Big|\mathcal{F}_s^W\Big] \\
&\quad + E\Big[E\Big[Z_k(W_t - W_{t_{k-1}})\Big|\mathcal{F}_t^W\Big]\Big|\mathcal{F}_s^W\Big] \\
&= I_s(C) + \sum_{i=j+1}^{k-1} E\Big[Z_i E\Big[\Delta_i W\Big|\mathcal{F}_t^W\Big]\Big|\mathcal{F}_s^W\Big] \\
&\quad + E\Big[Z_k E\Big[(W_t - W_{t_{k-1}})\Big|\mathcal{F}_t^W\Big]\Big|\mathcal{F}_s^W\Big] \\
&= I_s(C) + \sum_{i=j+1}^{k-1} E\Big[Z_i 0\Big|\mathcal{F}_s^W\Big] + E\Big[Z_k 0\Big|\mathcal{F}_s^W\Big] \\
&= I_s(C).
\end{aligned}$$

Por lo tanto, $(I_t(C))$ es una martingala con respecto a la filtración $\left(\mathcal{F}_t^W\right)$.

c) Para cada $t \in [0, T]$, existe $i \in \{1, 2, \ldots, n\}$ tal que se cumple

$$I_t(C) = I_{t_{i-1}}(C) + Z_i(W_t - W_{t_{i-1}}), \quad t_{i-1} \leq t \leq t_i.$$

Luego, por la continuidad de (W_t) se sigue que $(I_t(C))$ es continuo.

■

Para la demostración de **d)** se necesita encontrar una partición que funcione para ambos procesos simples, lo cual siempre es posible. El resto de la prueba es utilizar la definición de la integral estocástica. Por otra parte, **e)** se sigue de **d)** al considerar otro proceso (C_s^2) definido por $C_s^2 = C_s I_{(t,T]}$.

Ahora la integral estocástica se extiende a procesos un poco más generales.

Definición 2.3.3. *Sea $\mathcal{V}$ el conjunto de funciones*

$$f : \Omega \times [0,T] \to \mathbb{R},$$

tales que f es medible y adaptado a la filtración $\left(\mathcal{F}_t^W\right)$, y además cumplen

$$E\left[\int_0^T f(\omega,s)^2\,ds\right] < \infty. \tag{2.22}$$

Lo que ahora se pretende es usar la isometría de Itô para extender la definición de integral estocástica de procesos simples a funciones en $\mathcal{V}$. Los procesos simples, las funciones continuas, las funciones deterministas que cumplen (2.22), son ejemplos de funciones que pertenecen a $\mathcal{V}$.

Es posible demostrar que el conjunto de procesos simples es denso en $\mathcal{V}$ bajo la norma definida por

$$||(C_t)|| = \left(E\left[\int_0^T |C_t|^2 dt\right]\right)^{\frac{1}{2}}.$$

Esto significa que para cualquier $f \in \mathcal{V}$ existe una sucesión de procesos simples $(C^{(n)})$ tales que

$$\lim_{n\to\infty} ||f - C^{(n)}|| = 0.$$

Teorema 2.3.2. *Sea $f \in \mathcal{V}$ arbitraria. Entonces existe una sucesión de procesos simples $(C_t^{(n)})$ tal que*

$$\int_0^t E[(f - C_s^{(n)})^2]\,ds \to 0, \quad \textit{cuando } n \to \infty.$$

El procedimiento de aproximación puede llevarse a cabo mediante una sucesión de pasos. En términos generales, primero se aproxima funciones $f \in \mathcal{V}$ acotadas y continuas por procesos simples. Posteriormente, las funciones $f \in \mathcal{V}$ acotadas se aproximan por funciones $f^{(n)} \in \mathcal{V}$ continuas y acotadas. Finalmente, una función $f \in \mathcal{V}$ arbitraria se aproximan por funciones $f^{(n)} \in \mathcal{V}$ acotadas, y así se completa el procedimiento de aproximación.

Definición 2.3.4 (La Integral de Itô). *Sea $f \in \mathcal{V}$. Entonces, para cada $t \in [0,T]$, la integral de Itô de f está definida por*

$$I_t(f) = \int_0^t f(\omega,s)dW_s = \lim_{n\to\infty}\int_0^t C_s^{(n)} dW_s \quad \textit{en } L^2, \tag{2.23}$$

donde $(C_t^{(n)})$ *es una sucesión de procesos simples tal que*

$$E\left[\int_0^t (f - C_s^{(n)})^2\, ds\right] \to 0, \text{ cuando } n \to \infty. \tag{2.24}$$

El límite en (2.23) existe como un elemento de L^2, ya que, por la isometría de Itô, los términos $\int_0^t C_s^{(n)} dW_s$ forman una sucesión una sucesión de Cauchy en L^2. Más aún, el límite en (2.23) no depende de la elección de $(C_t^{(n)})$, siempre y cuando la sucesión cumpla (2.24).

Como era de esperarse, la integral antes definida cumple las propiedades del Teorema 2.3.1. Básicamente, la demostración del siguiente resultado hace uso de (2.23) y del Teorema 2.3.1.

Teorema 2.3.3. *Sean* $f, g \in \mathcal{V}$. *Entonces las integrales estocásticas* $(I_t(f))$ *y* $(I_t(g))$ *cumplen las siguientes propiedades:*

a) *Isometría de Itô:*

$$E[(I_t(f))^2] = E\Big[\int_0^t f^2(\omega, s) ds\Big].$$

b) *Si* $f_n \in \mathcal{V}$, *para* $n \in \mathbb{N}$, *y* $E\Big[\int_0^t (f - f_n)^2\, ds\Big] \to 0$, *cuando* $n \to \infty$, *entonces*

$$\int_0^t f dW_s = \lim_{n\to\infty} \int_0^t f_n dW_s, \quad \text{en } L^2.$$

c) *Media cero:* $E[I_t(f)] = 0$.

d) *Linealidad:* $I_t(cf + g) = cI_t(f) + I_t(g)$, *donde* c *es constante.*

e) $\int_0^t f dW_s = \int_0^u f dW_s + \int_u^t f dW_s, \quad 0 \le u < t$.

El siguiente resultado se obtiene como consecuencia de la isometría de Itô.

Teorema 2.3.4. *Sean* $f(\omega, t), g(\omega, t) \in \mathcal{V}$. *Entonces*

$$E[I_t(f) I_t(g)] = \int_0^t E[fg] ds.$$

Demostración.

El producto de las integrales se puede reescribir como

$$\begin{aligned} I_t(f)I_t(g) &= \frac{1}{2}(I_t(f) + I_t(g))^2 - \frac{I_t(f)^2}{2} - \frac{I_t(g)^2}{2} \\ &= \frac{1}{2}(I_t(f+g))^2 - \frac{I_t(f)^2}{2} - \frac{I_t(g)^2}{2}. \end{aligned}$$

Luego, por la isometría de Itô, se tiene que

$$\begin{aligned} E[I_t(f)I_t(g)] &= \frac{1}{2}E[(I_t(f+g))^2] - \frac{1}{2}E[I_t(f)^2] - \frac{1}{2}E[I_t(g)^2] \\ &= \frac{1}{2}\int_0^t E[(f+g)^2] ds - \frac{1}{2}\int_0^t E[f^2] ds - \frac{1}{2}\int_0^t E[g^2] ds \\ &= \int_0^t E[fg] ds. \end{aligned}$$

Por lo tanto,

$$E[I_t(f)I_t(g)] = \int_0^t E[fg]ds.$$

$\blacksquare$

El procedimiento para verificar si la nueva definición de integral estocástica cumple la propiedad de continuidad y la de martingala es más elaborada. Aquí sólo presentaremos el resultado que relaciona éstas propiedades. En la demostración se aplica la isometría de Itô y el hecho de que la integral estocástica de procesos simples es martingala, así como de otros interesantes resultados tal como la *Desigualdad Martingala de Doob* (véase Steel[22], Teorema 6.2.).

Teorema 2.3.5. *Para $f \in \mathcal{V}$ arbitraria, existe un proceso (X_t) que es continuo, es una martingala con respecto a la filtración $(\mathcal{F}_t^W)$ y es una versión de $(I_t(f))$.*

Sin perdida de generalidad, a partir de ahora se supondrá que la integral estocástica es continua.

La integral estocástica $I_t(f)$ puede ser definida para una clase más amplia de integrandos f. Una primera extensión es cuando se relaja la condición de que la función sea adaptada a la filtración $\left(\mathcal{F}^W\right)$ por la nueva condición:

a') *Existe una filtración $(\mathcal{H}_t)$ tal que:*

- *(W_t) es una martingala con respecto a $(\mathcal{H}_t)$.*
- *f es adaptado a la filtración $(\mathcal{H}_t)$.*

La esencia de la nueva extensión es que a f se le permite depender de más información de la contenida en $\mathcal{F}_t^W$, siempre y cuando W_t sea una martingala con respecto a la "historia" de $f_s, s \leq t$.

Con la nueva condición se puede definir integrales que involucran varias componentes del movimiento Browniano n-dimensional, tales como

$$\int W_2 dW_1 \text{ ó } \int sen(W_1^2 + W_2^2)dW_2,$$

en donde $W_k = W_k(\omega, t)$ es la k-ésima coordenada y $dW_k = dW_k(\omega, t)$. Más aún, es posible definir la **integral estocástica multidimensional.**

Otra segunda generalización de la integral estocástica consiste en debilitar la condición

$$E\left[\int_0^t f(\omega, s)^2 ds\right] < \infty$$

por la nueva condición dada por

b')

$$\mathbb{P}\left(\int_0^t f(\omega,s)^2 ds < \infty\right) = 1,$$

o también denotada como $\int_0^t f(\omega,s)^2 ds < \infty, \mathbb{P}$-c.s..

*A partir de ahora, $\mathcal{H}$ denotará a la clase de funciones medibles que cumplen **a')** y (2.22). Mientras que $\mathcal{W}_{\mathcal{H}}$ denotará a la clase de funciones medibles que cumplen **a')** y **b')**.*

Al procedimiento de extender la integral a ésta nueva clase de integrandos se conoce como *extensión por localización,* y su construcción necesita de nuevos conceptos tal como *tiempos de paro.*

Si $f \in \mathcal{W}_{\mathcal{H}}$, existen resultados que garantizan que para todo t existen funciones simples $f_n \in \mathcal{W}_{\mathcal{H}}$ tal que $\int_0^t |f_n - f|^2 ds \to 0$, $\mathbb{P}$-c.s.. Para tal sucesión se tiene que $\int_0^t f_n(\omega,s)dW_s$ converge, $\mathbb{P}$-c.s., a alguna v.a. y el límite sólo depende de f, no de la sucesión $\{f_n\}$. Luego, se puede definir

$$\int_0^t f(\omega,s)dW_s = \lim_{n\to\infty} \int_0^t f_n(\omega,s)dW_s, \mathbb{P}\text{-c.s., para } f \in \mathcal{W}_{\mathcal{H}}. \qquad (2.25)$$

Nuevamente, es posible demostrar que existe una versión continua de esta integral. Sin embargo, en general no se cumple que sea una martingala. En estos casos se hace uso de la *martingala local.*

A partir de ahora se da por hecho la existencia de la integral estocástica para funciones $f \in \mathcal{V}$, $f \in \mathcal{H}$ o $f \in \mathcal{W}_{\mathcal{H}}$.

2.4. El lema de Itô

En la teoría del cálculo tradicional, en muy raros casos, la integral de Riemman es calculada utilizando su definición de forma directa, en este caso, el teorema fundamental del cálculo y la regla de la cadena permiten cálculos explícitos. Sin embargo, en el contexto estocástico no existe una teoría para la diferenciación, sólo teoría de integración. No obstante, es posible establecer una versión de la regla de la cadena para las integrales estocásticas, conocida como *fórmula de Itô*, *cambio de variable, Lema de Itô* o *regla de la cadena*, siendo ésta la herramienta fundamental en el cálculo estocástico.

En el cálculo clásico, para cualquier función diferenciable $x(t)$ tal que $x(0) = 0$, se cumplen las fórmulas

$$x(t)^2 = 2\int_0^t x(s)dx(s),$$

$$x(t)^3 = 3\int_0^t x(s)^2 dx(s),$$

y en forma general,

$$x(t)^n = n \int_0^t x(s)^{n-1} dx(s).$$

De acuerdo a las ecuaciones (2.19) y (2.20), para el caso estocástico se tiene

$$W_t^2 = \int_0^t ds + 2 \int_0^t W_s dW_s, \tag{2.26}$$

$$W_t^3 = 3 \int_0^t W_s ds + 3 \int_0^t W_s^2 dW_s.$$

Como se verá más adelante, la fórmula para W_t^2 y W_t^3 son casos particulares de la fórmula general de Itô. Nótese que las integrales estocásticas se parecen a las expresiones para funciones continuas diferenciables $x(t)$, sólo que aquí aparece otro término de más. A los términos extras que surgen en la integración estocástica, tales como $\int_0^t ds$, $3\int_0^t W_s ds$, y que no aparecen en el cálculo clásico, son conocidos como la ***corrección de Itô***, y es una característica inherente en la fórmula de Itô.

Existen diferentes versiones de la fórmula de Itô, y estas difieren por la clase de funciones que se consideran como integrador y como integrando, pero en esencia todas expresan lo mismo. Para definir una de las versiones generales de la fórmula de Itô se necesita definir una nueva clase de procesos.

Definición 2.4.1 (Proceso de Itô). *Un proceso de Itô unidimensional es un proceso estocástico* (X_t) *definido por*

$$X_t = X_0 + \int_0^t \mu(\omega, s)\, ds + \int_0^t \sigma(\omega, s)\, dW_s,\ 0 \leq t \leq T, \tag{2.27}$$

donde $\mu(\omega, t)$ *y* $\sigma(\omega, t)$ *son procesos predecibles tales que* $\mu(\omega, s) \in \mathcal{H}$, $\sigma(\omega, s) \in \mathcal{W}_{\mathcal{H}}$, *i.e., satisfacen la condición de integrabilidad*

$$\int_0^t |\mu(\omega, s)|\, ds + \int_0^t |\sigma(\omega, s)|^2\, ds < \infty,\ \mathbb{P}\text{-c.s.,} \qquad \textit{para todo } t \geq 0.$$

Las versiones de la definición anterior difieren dependiendo de las condiciones de integrabilidad. Tales condiciones deben garantizar que existan las integrales que aparecen en la condición (2.27), donde la primera integral es en el sentido de Riemann y la segunda es la integral de Itô.

Generalmente, los procesos μ y σ en (2.27) dependen de X_t o de W_t. Por ejemplo, un caso particular es cuando se pide que la v.a. X_0 sea $\mathcal{F}_0^W$-medible, los procesos μ y σ adaptados a la filtración $(\mathcal{F}_t^W)$, tal que

$$\int_0^T |\mu(\omega, t)| dt < \infty, \quad \int_0^T \sigma(\omega, t)^2 dt < \infty.$$

Además, nótese que nada impide que se considere, por ejemplo, $X_0 = 0, \mu = 0$ y $\sigma = 1$.

La forma diferencial de un proceso de Itô (X_t) definido por la ecuación (2.27) es usualmente escrita en su forma diferencial como,

$$dX_t = \mu\, dt + \sigma\, dW_t, \quad 0 \leq t \leq T. \tag{2.28}$$

Nuevamente se hace énfasis en que la representación (2.28) se debe entender en el sentido de la ecuación (2.27). Así por ejemplo, la fórmula (2.26) se representa en forma diferencial como

$$d\Big(\frac{1}{2}W_t^2\Big) = \frac{1}{2}dt + W_t dW_t.$$

A continuación se enunciará el principal resultado del cálculo estocástico, en el caso unidimensional, así como un bosquejo de su demostración.

Teorema 2.4.1 (Fórmula de Itô). *Sea (X_t) un proceso de Itô dado por*

$$dX_t = \mu\, dt + \sigma\, dW_t, \quad 0 \leq t \leq T, \tag{2.29}$$

y una función $f(x,t) \in C^{2,1}(\mathbb{R} \times \mathbb{R}^+)$. Entonces $Y_t = f(X_t, t)$ es nuevamente un proceso de Itô, y

$$\begin{aligned} dY_t &= f_x(X_t,t)\, dX_t + f_t(X_t,t)\, dt + \frac{1}{2} f_{xx}(X_t,t) \cdot (\, dX_t)^2 \\ &= \Big(f_t(X_t,t) + f_x(X_t,t)\mu + \frac{1}{2} f_{xx}(X_t,t)\sigma^2\Big)\, dt + f_x(X_t,t)\sigma\, dW_t, \end{aligned}$$

donde $(\, dX_t)^2 = (\, dX_t) \cdot (\, dX_t)$ es calculado de acuerdo a las reglas

$$dt \cdot dt = dt \cdot dW_t = dW_t \cdot dt = 0 \quad y \quad dW_t \cdot dW_t = dt.$$

Demostración.

Supóngase que $\mu(\omega, t)$ y $\sigma(\omega, t)$ son funciones simples. Nuevamente considérese una partición $\{t_j^n\}$ de $[0, t]$. Claramente,

$$\begin{aligned} f(X_t, t) - f(X_0, 0) &= \sum_{j=0}^{n-1} (f(X_{t_{j+1}^n}, t_{j+1}^n) - f(X_{t_j^n}, t_j^n)) \\ &= \sum_{j=0}^{n-1} \Delta f(X_{t_j^n}, t_{j^n}) \\ &= \sum_j \Delta f(X_{t_j}, t_j). \end{aligned}$$

Para comodidad en la notación, se omitirá el superíndice n, así por ejemplo, $\Delta f(X_{t_j}, t_j) = f(X_{t_{j+1}}, t_{j+1}) - f(X_{t_j}, t_j)$.

Aplicando la expansión de Taylor de segundo orden a $\Delta f(X_{t_j}, t_j)$ se tiene que

$$\begin{aligned}
\Delta f(X_{t_j}, t_j) &= f_x(X_{t_j}, t_j)\Delta X_j + f_t(X_{t_j}, t_j)\Delta t_j \\
&\quad +\frac{1}{2} f_{xx}(X_{t_j}, t_j)(\Delta X_j)^2 + f_{tx}(X_{t_j}, t_j)(\Delta t_j)(\Delta X_j) \\
&\quad +\frac{1}{2} f_{tt}(X_{t_j}, t_j)(\Delta t_j)^2 + R_j,
\end{aligned}$$

donde $\Delta t_j = t_{j+1} - t_j$, $\Delta X_j = X_{t_{j+1}} - X_{t_j}$, y $R_j = o(|\Delta t_j|^2 + |\Delta X_j|^2)$.
Luego,

$$\begin{aligned}
f(X_t, t) - f(X_0, 0) &= \sum_j f_x(X_{t_j}, t_j)\Delta X_j + \sum_j f_t(X_{t_j}, t_j)\Delta t_j \\
&\quad +\frac{1}{2}\sum_j f_{xx}(X_{t_j}, t_j)(\Delta X_j)^2 + \sum_j f_{tx}(X_{t_j}, t_j)(\Delta t_j)(\Delta X_j) \\
&\quad +\frac{1}{2}\sum_j f_{tt}(X_{t_j}, t_j)(\Delta t_j)^2 + \sum_j R_j.
\end{aligned}$$

Si $\Delta t_j \to 0$, entonces

$$\sum_j f_x(X_{t_j}, t_j)\Delta X_j \xrightarrow{L^2} \int_0^t f_x(X_s, s)\, dX_s,$$

$$\sum_j f_t(X_{t_j}, t_j)\Delta t_j \xrightarrow{L^2} \int_0^t f_t(X_s, s)\, ds.$$

Por otro lado, puesto que μ y σ son simples,

$$\sum_j f_{xx}(\Delta X_j)^2 = \sum_j f_{xx}\mu_j^2(\Delta t_j)^2 + 2\sum_j f_{xx}\mu_j\sigma_j(\Delta t_j)(\Delta W_j) + \sum_j f_{xx}\sigma_j^2(\Delta W_j)^2, \quad (2.30)$$

donde $\mu_j = \mu(\omega, t_j), \sigma_j = \sigma(\omega, t_j)$.

Nótese que cuando $\Delta t_j \to 0$,

$$(\Delta t_j)^2 \to 0, \quad (\Delta t_j)(\Delta W_j) \xrightarrow{L^2} 0,$$

esto es, $dt \cdot dt = 0 = dt \cdot dW_t = dW_t \cdot dt$. Por lo tanto, los dos primeros términos en (2.30) tienden a 0, $\mathbb{P}$-c.s., cuando $\Delta t_j \to 0$. En efecto,

$$\begin{aligned}
E\left[\left(\sum_j f_{xx}\mu_j\sigma_j(\Delta t_j)(\Delta W_j)\right)^2\right] &= \sum_j E\Big[(f_{xx}\mu_j\sigma_j)^2(\Delta t_j)^3\Big] + \sum_j\sum_{i\neq j} 0 \\
&= \sum_j E\Big[(f_{xx}\mu_j\sigma_j)^2\Big](\Delta t_j)^3 \\
&\to 0 \text{ cuando } \Delta t_j \to 0,
\end{aligned}$$

pues si $j < i$, entonces $f_{xx}\mu_j\sigma_j f_{xx}\mu_i\sigma_i(\Delta W_j)$ y ΔW_i son independientes; si $i < j$, entonces $f_{xx}\mu_j\sigma_j f_{xx}\mu_i\sigma_i(\Delta W_i)$ y ΔW_j son independientes.

La convergencia del primer término de (2.30) es similar al caso anterior. Por otro lado, el tercer término de (2.30) converge a

$$\int_0^t f_{xx}\sigma^2\, ds, \quad \text{en } L^2, \text{ cuando } \Delta t_j \to 0.$$

En efecto, sea $a_k := f_{xx}(X_{t_k}, t_k)\sigma^2(\omega, t_k)$. Luego,

$$E\left[\left(\sum_j a_j(\Delta W_j)^2 - \sum_j a_j \Delta t_j\right)^2\right] = \sum_{i,j} E\Big[a_i a_j\Big((\Delta W_i)^2 - \Delta t_i\Big)\Big((\Delta W_j)^2 - \Delta t_j\Big)\Big]. \quad (2.31)$$

Nuevamente, si $i < j$, entonces $a_i a_j\big((\Delta W_i)^2 - \Delta t_i\big)$ y $(\Delta W_j)^2 - \Delta t_j$ son independientes; si $j < i$, entonces $a_i a_j\big((\Delta W_j)^2 - \Delta t_j\big)$ y $(\Delta W_i)^2 - \Delta t_i$ son independientes. Así, los términos en (2.31) se anulan cuando $i \neq j$. Por lo tanto, (2.31) se reduce a lo siguiente:

$$\begin{aligned}
\sum_j E\left[a_j^2\Big((\Delta W_j)^2 - \Delta t_j\Big)^2\right] &= \sum_j E[a_j^2] \cdot E\Big[(\Delta W_j)^4 - 2(\Delta W_j)^2 \Delta t_j + (\Delta t_j)^2\Big] \\
&= \sum_j E[a_j^2] \cdot \Big(3(\Delta t_j)^2 - 2(\Delta t_j)^2 + (\Delta t_j)^2\Big) \\
&= 2\sum_j E[a_j^2] \cdot (\Delta t_j)^2 \\
&\to 0, \text{ cuando } \Delta t_j \to 0.
\end{aligned}$$

Lo anterior prueba que

$$\sum_j a_j(\Delta W_j)^2 \to \int_0^t a(s)\, ds, \quad \text{en } L^2, \text{ cuando } \Delta t_j \to 0.$$

En su forma más simple, lo anterior justifica la fórmula

$$dW_t \cdot dW_t = (\, dW_t)^2 = \, dt.$$

Los argumentos anteriores también prueban que los términos

$$\sum_j f_{tx}(X_{t_j}, t_j)(\Delta t_j)(\Delta X_j), \quad \frac{1}{2}\sum_j f_{tt}(X_{t_j}, t_j)(\Delta t_j)^2, \quad \sum_j R_j,$$

convergen a 0, en L^2, cuando $\Delta t_j \to 0$.

En resumen, como primera instancia se tiene que

$$f(X_t, t) - f(X_0, 0) = \int_0^t f_x(X_s, s)\, dX_s + \int_0^t f_t(X_s, s)\, ds + \frac{1}{2}\int_0^t f_{xx}(X_s, s) \cdot (dX_s)^2.$$

Posteriormente, cuando se sustituye la expresión dada por (2.29), entonces se verifica que

$$f(X_t, t) - f(X_0, 0) = \int_0^t f_x(X_s, s)\, dX_s + \int_0^t f_t(X_s, s)\, ds + \frac{1}{2}\int_0^t f_{xx}(X_s, s)\sigma^2\, ds.$$

O en notación diferencial, primero se tiene que

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

luego, sustituyendo la expresión dX_t dada por (2.29),

$$\begin{aligned} df(X_t,t) &= \Big(f_t(X_t,t) + \frac{1}{2}f_{xx}(X_t,t)\sigma^2\Big)\,dt + f_x(X_t,t)\,dX_t \\ &= \Big(f_t(X_t,t) + \frac{1}{2}f_{xx}(X_t,t)\sigma^2\Big)\,dt + f_x(X_t,t)(\mu dt + \sigma W_t) \\ &= \Big(f_t(X_t,t) + f_x(X_t,t)\mu + \frac{1}{2}f_{xx}(X_t,t)\sigma^2\Big)\,dt + f_x(X_t,t)\sigma\,dW_t. \end{aligned}$$

Así, la prueba de la fórmula de Itô está completa cuando $\mu(\omega,t)$ y $\sigma(\omega,t)$ son funciones simples. El caso general se sigue de (2.25). ■

Ejemplo 2.4. *Ahora se verificarán las fórmulas que se presentaron como motivación en la construcción de la integral estocástica.*

Sean $X_t = W_t$ *y* $f(x,t) = x^n$, *para* $n \geq 2$. *Por la fórmula de Itô,*

$$\begin{aligned} dY_t = d(f(W_t,t)) &= f_x(W_t,t)dX_t + f_t(W_t,t)dt + \frac{1}{2}f_{xx}(dX_t)^2 \\ &= nW_t^{n-1}dW_t + \frac{1}{2}n(n-1)W_t^{n-2}dt. \end{aligned}$$

Luego,

$$f(W_t,t) - f(W_0,0) = \int_0^t nW_s^{n-1}dW_s + \int_0^t \frac{1}{2}n(n-1)W_s^{n-2}ds.$$

Por lo tanto,

$$\int_0^t W_s^{n-1}dW_s = \frac{1}{n}W_t^n - \frac{n-1}{2}\int_0^t W_s^{n-2}ds.$$

Como casos particulares, si $n = 2$,

$$\int_0^t W_s dW_s = \frac{1}{2}W_t^2 - \frac{1}{2}t,$$

y si $n = 3$,

$$\int_0^t W_s^2 dW_s = \frac{1}{3}W_t^3 - \int_0^t W_s ds.$$

□

En el cálculo de variables reales, si t es una variable independiente, el cuadrado de una cantidad infinitesimal dt es una cantidad muy pequeña ($(dt)^2 << 0$). Por tal razón, si $r > 1$, entonces se escribe

$$(dt)^r = 0.$$

Formalmente, en el cálculo estocástico se tiene que

$$\int_0^t (dW_s)^2 = \int_0^t ds = t,$$

y se escribe en su forma diferencial como

$$(dW_t)^2 = dt. \tag{2.32}$$

La regla central del cálculo estocástico, que hace la distinción con el cálculo de variables reales, es que el cuadrado de una cantidad infinitesimal "normal" es significativa. Siendo ésta la principal razón del por qué se afirma que la base del cálculo estocástico es el *movimiento Browniano (o proceso de Wiener)*, ya que es el único proceso que cumple (2.32).

Heurísticamente, obsérvese que

$$(dt)(dW_t) = (dt)(dt)^{\frac{1}{2}} = (dt)^{\frac{3}{2}},$$

el cual sugiere que $(dt)(dW_t)$ sea una cantidad despreciable, i.e.,

$$(dt)(dW_t) = 0.$$

Así, las ***reglas básicas de diferenciación estocástica***, también llamadas *reglas empíricas de diferenciación estocástica*, se resumen en el siguiente cuadro:

$\cdot$	dt	dW_t
dt	0	0
dW_t	0	dt

Si bien es cierto que existen motivaciones para considerar éstas reglas básicas, formalmente no debería interpretarse como en el cálculo clásico, pero la práctica también sugiere que no hay inconvenientes en la interpretación, de aquí que se utilize el término *empírico.*

El siguiente resultado es un caso particular del teorema anterior.

Teorema 2.4.2. *Sea (W_t) un movimiento Browniano y $f(x) \in C^2(\mathbb{R})$. Entonces*

$$f(W_t) = f(0) + \int_0^t f'(W_s)dW_s + \frac{1}{2}\int_0^t f''(W_s)ds, \qquad \textit{para } 0 \leq t \leq T.$$

Una generalización del Teorema 2.4.1 es cuando se consideran dos procesos de Itô.

Teorema 2.4.3. *Sean $f(x,y) \in C^{2,2}(\mathbb{R}\times\mathbb{R})$ y $(X_t),(Y_t)$ dos procesos de Itô dados por*

$$dX_t = \mu_X(\omega,t)\,dt + \sigma_X(\omega,t)\,dW_t, \;\; y \quad dY_t = \mu_Y(\omega,t)\,dt + \sigma_Y(\omega,t)\,dW_t.$$

Entonces

$$\begin{aligned}
df(X_t, Y_t) &= f_x(X_t, Y_t)dX_t + f_y(X_t, Y_t)dY_t \\
&\quad + \frac{1}{2} f_{xx}(X_t, Y_t) \cdot (dX_t)^2 + f_{xy}(X_t, Y_t)dX_t \cdot dY_t + \frac{1}{2} f_{yy}(X_t, Y_t) \cdot (dY_t)^2 \\
&= f_x(X_t, Y_t)dX_t + f_y(X_t, Y_t)dY_t \\
&\quad + \frac{1}{2}\Big(f_{xx}(X_t, Y_t)\sigma_X^2 + 2f_{xy}(X_t, Y_t)\sigma_X\sigma_Y + f_{yy}(X_t, Y_t)\sigma_Y^2\Big)dt.
\end{aligned}$$

Una aplicación directa del teorema anterior es cuando se considera $f(x, y) = xy$, luego

$$d(X_tY_t) = Y_tdX_t + X_tdY_t + dX_t \cdot dY_t,$$

la cual representa a la ***fórmula general de integración por partes***:

$$\int_0^t X_sdY_s = X_tY_t - X_0Y_0 - \int_0^t Y_sdX_s - \int_0^t dX_s \cdot dY_s. \tag{2.33}$$

Por resultados de la integral de Riemann-Stieltjes, cuando g es una función continua y h es de variación acotada, entonces la integral $\int_0^t g(s)dh(s)$ existe. Luego, si en la ecuación (2.33) se considera $X_s = W_s$, $Y_s = g(s)$, entonces la ecuación se reduce a

$$\int_0^t W_sdg(t) = W_tg(t) - \int_0^t g(s)dW_s.$$

Lo anterior demuestra la veracidad del siguiente corolario.

Corolario 2.4.4. *Sea $f(t)$ una función determinista. Si $f(t)$ es continua y de variación acotada en $[0, t]$, entonces*

$$\int_0^t f(s)dW_s = f(t)W_t - \int_0^t W_sdf(s).$$

Sea (W_t) un movimiento Browniano m-dimensional definido por $W_t = \Big(W_1(t), W_2(t), \ldots, W_m(t)\Big)$. Si cada uno de los procesos $\mu_i(\omega, t)$ y $\sigma_{ij}(\omega, t)$, $(1 \leq i \leq n, 1 \leq j \leq m)$ satisfacen la definición de proceso de Itô, entonces es posible formar los siguientes n procesos de Itô dados por

$$\begin{cases}
dX_1 = \mu_1dt + \sigma_{11}dW_1 + \ldots + \sigma_{1m}dW_m, \\
\vdots \qquad \vdots \qquad\qquad\qquad \vdots \\
dX_n = \mu_ndt + \sigma_{n1}dW_1 + \ldots + \sigma_{nm}dW_m,
\end{cases}$$

o en su forma matricial,

$$dX(t) = \mu dt + \sigma dW(t), \tag{2.34}$$

donde

$$X(t) = \begin{pmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{pmatrix}, \ \mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}, \ \sigma = \begin{pmatrix} \sigma_{11} & \dots & \sigma_{1m} \\ \vdots & & \vdots \\ \sigma_{n1} & \dots & \sigma_{nm} \end{pmatrix}, \ dW(t) = \begin{pmatrix} dW_1(t) \\ \vdots \\ dW_n(t) \end{pmatrix}.$$

El proceso $X(t)$ se conoce como **proceso de Itô n-dimensional** (o simplemente proceso de Itô cuando no haya confusión).

A continuación se menciona la fórmula de Itô para procesos de Itô n-dimensional.

Teorema 2.4.5 (Fórmula general de Itô). *Sea*

$$dX(t) = \mu dt + \sigma dW(t)$$

un proceso de Itô n-dimensional definido como en (2.34), y $f(x,t) = \Big(f_1(x,t), \dots, f_p(x,t)\Big) \in C^{2,1}(\mathbb{R}^n \times \mathbb{R}^+)$ un mapeo sobre $\mathbb{R}^p$. Entonces, el proceso

$$Y(t) = f(X(t), t) = f(X_1(t), \dots, X_n(t), t),$$

es nuevamente un proceso de Itô, cuyos componentes Y_k están dados por

$$dY_k = f_{kt} dt + \sum_i f_{kx_i} dX_i + \frac{1}{2} \sum_{i,j} f_{kx_i x_j} dX_i dX_j,$$

en donde la derivadas parciales son evaluadas en $(X(t), t)$, y

$$dW_i dW_j = \delta_{ij} dt, \quad dW_i dt = dt dW_i = 0,$$

donde δ_{ij} denota la función delta de Kronecker.

Capítulo 3

Ecuaciones Diferenciales Estocásticas

El concepto de ecuación diferencial estocástica está íntimamente relacionado con el concepto de integral estocástica. Este nuevo concepto forma una herramienta efectiva para la cosntrucción y el análisis de modelos estocásticos. Ésta teoría también provee un vínculo entre la teoría de la probabilidad y los variados campos de aplicación de las ecuaciones diferenciales ordinarias.

Los modelos estocásticos que se apoyan en resultados clásicos de la física matemática, así como de la matemática pura, ofrecen nuevas interpretaciones intuitivas y en muchos casos más realistas.

El objetivo de este capítulo es encontrar soluciones explícitas de algunas ecuaciones diferenciales estocásticas, así como establecer las condiciones de existencia y unicidad.

3.1. Definición de ecuación diferencial estocástica

En el cálculo de Leibniz-Newton, si $x(t)$ es una función diferenciable definida para $t \geq 0$ y $\mu(x,t)$ es una función que cumple la relación

$$\frac{dx(t)}{dt} = x'(t) = \mu(x(t),t), \quad x(0) = x_0, \tag{3.1}$$

para todo $0 \leq t \leq T$, entonces $x(t)$ es una solución de la ecuación diferencial ordinaria (EDO) $x'(t) - \mu = 0$, con condición inicial x_0.

La ecuación anterior también puede ser expresada en términos diferenciales:

$$dx(t) = \mu(x(t),t)dt, \quad x(0) = x_0,$$

o de forma equivalente, como una ecuación integral:

$$x(t) = x(0) + \int_0^t \mu(x(s),s)ds.$$

Si en la condición inicial de las EDO's se considera una v.a., o si los coeficientes son funciones aleatorias, o cuando se combinan ambos casos, entonces se dice que la ecuación que resulta es una ***ecuación diferencial aleatoria***, y se resuelven trayectorias por trayectorias como una EDO.

Cuando las EDO's son perturbadas por cantidades aleatorias, entonces se obtiene otro tipo de ecuación.

El *proceso ruido blanco* (ξ_t) es formalmente definido como la derivada del movimiento Browniano,

$$\xi_t = \frac{dW_t}{dt} = W'_t.$$

Puesto que W_t es no diferenciable, esto no existe como una función de t en el sentido usual.

Cuando en las EDO's se introducen ruidos aleatorios, tal como el proceso ruido blanco, entonces la nueva ecuación se conoce como ***ecuación diferencial estocástica (EDE)***. Así por ejemplo, cuando la EDO (3.1) es perturbada por el proceso ruido blanco, la ecuación diferencial estocástica que resulta tiene la forma

$$\frac{dX_t}{dt} = \mu(X_t, t) + \sigma(X_t, t)\xi_t, \quad X_0 = x_0. \tag{3.2}$$

La idea innovadora de Itô fue considerar el producto del proceso ruido blanco y la diferencial del tiempo como una diferencial del movimiento Browniano, una cantidad que puede servir como un integrador.

En el cálculo de Itô, la EDE (3.2) se escribe en forma diferencial como

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t, \quad X_0 = x_0, \tag{3.3}$$

o de manera formal como la ecuación integral estocástica

$$X_t = X_0 + \int_0^t \mu(X_t, t)dt + \int_0^t \sigma(X_t, t)dW_t.$$

Ejemplo 3.1. *Supóngase que $x(t)$ denota el valor de una inversión con tasa de interés r compuesto continuamente. Por la definición de interés compuesto, éste satisface la EDO*

$$\frac{dx(t)}{dt} = rx(t).$$

Si la tasa es perturbada por el proceso ruido blanco, entonces se obtiene la EDE

$$\frac{dX_t}{dt} = (r + \sigma\xi_t)X_t,$$

esto es,

$$dX_t = rX_tdt + \sigma X_tdW_t.$$

□

Ejemplo 3.2. *Supóngase que $x(t)$ denota la densidad poblacional, entonces el crecimiento de la población puede ser descrito por la EDO*

$$\frac{dx(t)}{dt} = ax(t)[1 - x(t)],$$

donde a denota la tasa de nacimiento.

Si la tasa de nacimiento es perturbada de forma aleatoria por el ruido blanco, entonces se obtiene la EDE

$$dX_t = aX_t(1 - X_t)dt + \sigma X_t(1 - X_t)dW_t.$$

□

En los ejemplos anteriores, nótese que con $\sigma = 0$ no existe ruido en el modelo, y por tanto la solución de la ecuación resultante, si es que existe, se obtiene por los métodos tradicionales.

Es posible reemplazar el ruido aleatorio por otro que no involucra al movimiento Browniano, pero esto requiere definir la integral estocástica en términos más generales, por ejemplo, con respecto a semimartingalas. Debido al objetivo general que se persigue, aquí sólo se considerará el ruido aleatorio generado por el movimiento Browniano.

Definición 3.1.1. *Sea (W_t) un movimiento Browniano. Una **ecuación diferencial estocástica** (EDE) dirigida por el movimiento Browniano es una ecuación escrita en la forma*

$$dX_t = \mu(X_t, t)\,dt + \sigma(X_t, t)\,dW_t, \quad X_0 = x_0, \tag{3.4}$$

*en donde las funciones $\mu(x,t)$ y $\sigma(x,t)$ se conocen como **coeficiente de deriva** y **coeficiente de difusión**, respectivamente, X_0 es la variable aleatoria inicial y X_t define al proceso desconocido.*

La EDE del proceso desconocido (X_t) frecuentemente es llamada la *dinámica* de (X_t). El coeficiente de difusión $\sigma(X_t, t)$ funciona como una escala de la aleatoriedad generado por el movimiento Browniano. Si éste coeficiente se anula, entonces se tiene la EDO

$$dX_t = \mu(X_t, t)dt,$$

que puede resolverse si se conoce la condición inicial X_0.

Por ejemplo, en el caso de que $\mu(x,t) = a(t)x$, la solución está dada por

$$X_t = X_0 \exp\left(\int_0^t a(s)ds\right).$$

Los modelos planteados como EDE's no son más que la generalización en tiempo continuo de modelos lineales típicos en donde el tiempo es la variable explicativa más una perturbación aleatoria gaussiana; el papel que es jugado en tiempo discreto por la caminata aleatoria, en tiempo continuo es jugado por la diferencial del movimiento Browniano.

3.2. Existencia y unicidad de las soluciones

Recuérdese que se está dando por hecho la existencia de un espacio de probabilidad filtrado $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$. A partir de ahora también se supondrá que $\mu, \sigma : \mathbb{R} \times [0, T] \to \mathbb{R}$ son funciones dadas, X_0 es una v.a. continua $\mathcal{F}_0$-medible y (W_t) es adaptado a la filtración $(\mathcal{F}_t)$.

Usualmente, la primera pregunta que surge cuando se tiene una EDO es que si ésta tiene solución única. En el caso estocástico, el concepto de *solución* de una EDE no es único, de hecho existen dos tipos de solución, una se conoce como *solución fuerte* y la otra como *solución débil.*

Definición 3.2.1. *Una* **solución fuerte** *de una EDE dada como en (3.4) es un proceso (X_t) adaptado a la filtración $(\mathcal{F}_t)$ que satisface:*

a) *Para cada $t \geq 0$, las integrales $\int_0^t \mu(X_s, s)\, ds$ y $\int_0^t \sigma(X_s, s)\, dW_s$ existen, i.e.,*

$$\int_0^t |\mu(X_s, s)|\, ds < \infty, \quad \text{y} \quad \int_0^t |\sigma(X_s, s)|^2\, ds < \infty,\ \mathbb{P}\text{-c.s.}.$$

b) *(X_t) satisface la ecuación (3.4), i.e.,*

$$X_t = X_0 + \int_0^t \mu(X_s, s)\, ds + \int_0^t \sigma(X_s, s)\, dW_s,\ \text{para todo } t \geq 0,\ \mathbb{P}\text{-c.s.}.$$

Una *solución fuerte* está basada en las trayectorias del movimiento Browniano subyacente, i.e., al tiempo t la solución depende de W_s, $s \leq t$. Si se considera otro movimiento Browniano, entonces se podría obtener otra solución fuerte dada por la misma fórmula pero con el nuevo movimiento Browniano en él.

Para el caso de la ***solución débil***, sólo se está interesado en las características de la distribución de (X_t), tales como la esperanza, la varianza y la covarianza. En este caso no es indispensable conocer las trayectorias de (X_t).

Definición 3.2.2. *Si existe un espacio de probabilidad con una filtración, un movimiento Browniano $(\tilde{W}_t)$ y un proceso $(\tilde{X}_t)$ adaptado a la filtración, tal que:*

a) *$\tilde{X}_0$ tiene la distribución de la condición inicial.*

b) *Para todo $t \geq 0$, $(\tilde{X}_t)$ satisface las integrales bien definidas*

$$\tilde{X}_t = \tilde{X}_0 + \int_0^t \mu(\tilde{X}_s, s) ds + \int_0^t \sigma(\tilde{X}_s, s) d\tilde{W}_s.$$

entonces $(\tilde{X}_t)$ es llamada una ***solución débil*** *de la EDE (3.4).*

Claramente, por definición, toda solución fuerte es también una solución débil. Más aún, existe un resultado que garantiza la unicidad de la solución débil cuando existe unicidad de la solución fuerte. Lo inverso no siempre se cumple, existen EDE's que no tienen soluciones fuertes pero sí tienen una única solución débil. En este trabajo se dará más importancia a la solución fuerte.

Análogo al caso de las EDO's, existen muchas EDE's cuyas soluciones no se pueden expresar de forma explícita, en esos casos es necesario el uso de métodos numéricos para su aproximación. Pero antes de aplicar algún método se necesita saber si la ecuación tiene solución, y si es única para un valor inicial dado.

Para poder concluir la unicidad de las soluciones, el uso de la desigualdad de Gronwall es primordial.

Lema 3.2.1 (Desigualdad de Gronwall). *Sean ϕ y f funciones continuas no negativas definidas para todo $0 \leq t \leq T$, y C_0 una constante no negativa. Si*

$$\phi(t) \leq C_0 + \int_0^t f(s)\phi(s)\, ds, \quad \textit{para todo } 0 \leq t \leq T,$$

entonces $\phi(t) \leq C_0 \exp\left(\int_0^t f(s)\, ds\right)$, *para todo* $0 \leq t \leq T$.

Demostración.

Defínase

$$\zeta(t) = C_0 + \int_0^t f(s)\phi(s)ds,$$

entonces $\zeta'(t) = f(t)\phi(t) \leq f(t)\zeta(t)$.

Por otro lado,

$$\left(\zeta(t)\ \exp\left(-\int_0^t f(s)ds\right)\right)' = (\zeta'(t) - f(t)\zeta(t))\exp\left(-\int_0^t f(s)ds\right) \leq 0,$$

de aquí que $\zeta(t)\exp\left(-\int_0^t f(s)ds\right)$ es una función decreciente en $[0,T]$, y en consecuencia

$$\zeta(t)\exp\left(-\int_0^t f(s)ds\right) \leq \zeta(0)\exp\left(-\int_0^0 f(s)ds\right) = C_0.$$

Combinando la hipótesis con la desigualdad anterior, resulta que

$$\phi(t) \leq \zeta(t) \leq C_0 \exp\left(\int_0^t f(s)ds\right).$$

Por lo tanto,

$$\phi(t) \leq C_0 \exp\left(\int_0^t f(s)ds\right).$$

■

Ahora se enunciará un resultado que establece las condiciones para la existencia y unicidad de la solución de las EDE's.

Teorema 3.2.1 (Teorema de existencia y unicidad). *Si para todo $0 \leq t \leq T$, las siguientes condiciones se satisfacen:*

a) *Condición de Lipschitz: Existe una constante K, que no depende de t, tal que*

$$|\mu(x,t) - \mu(y,t)| + |\sigma(x,t) - \sigma(y,t)| < K|x-y|.$$

b) *Crecimiento lineal acotado: Para alguna constante C, que no depende de t,*

$$|\mu(x,t)| + |\sigma(x,t)| \leq C(1+|x|).$$

c) *X_0 es independiente de (W_t), y $E[X_0^2] < \infty$.*

Entonces existe un único proceso continuo (X_t) que es solución fuerte de la EDE (3.4), y además

$$E\left[\int_0^T |X_t|^2 dt\right] < \infty.$$

Demostración.

Aquí sólo se demostrará la *unicidad.*

Supóngase que (X_t) y (Y_t) son soluciones de la ecuación (3.4), entonces para todo $0 \leq t \leq T$,

$$X_t - Y_t = \int_0^t \Big(\mu(X_s,s) - \mu(Y_s,s)\Big)\, ds + \int_0^t \Big(\sigma(X_s,s) - \sigma(Y_s,s)\Big)\, dW_s.$$

Puesto que $(a+b)^2 \leq 2(a^2+b^2)$,

$$\begin{aligned} E\Big[|X_t - Y_t|^2\Big] &\leq 2E\left[\left(\int_0^t \Big(\mu(X_s,s) - \mu(Y_s,s)\Big)\, ds\right)^2\right] \\ &\quad +2E\left[\left(\int_0^t \Big(\sigma(X_s,s) - \sigma(Y_s,s)\Big)\, dW_s\right)^2\right]. \end{aligned}$$

Aplicando la desigualdad de Cauchy,

$$\begin{aligned} \left(\int_0^t \Big(\mu(X_s,s) - \mu(Y_s,s)\Big)\, ds\right)^2 &\leq \left(\int_0^t 1^2 ds\right) \int_0^t \Big(\mu(X_s,s) - \mu(Y_s,s)\Big)^2 ds \\ &= t\int_0^t \Big|\mu(X_s,s) - \mu(Y_s,s)\Big|^2 ds \\ &\leq TK^2 \int_0^t |X_2 - Y_s|^2\, ds. \end{aligned}$$

Por otro lado, por la *isometría de Itô*,

$$\begin{aligned} E\left[\left(\int_0^t \Big(\sigma(X_s,s) - \sigma(Y_s,s)\Big)\, dW_s\right)^2\right] &= 2E\left[\int_0^t \Big|\sigma(X_s,s) - \sigma(Y_s,s)\Big|^2 ds\right] \\ &\leq 2K^2 \int_0^t E\Big[|X_s - Y_s|^2\Big]\, ds. \end{aligned}$$

Por todo lo anterior, para una constante apropiada C,

$$E\Big[|X_t - Y_t|^2\Big] \leq C\int_0^t E\Big[|X_s - Y_s|^2\Big]\, ds, \quad 0 \leq t \leq T. \tag{3.5}$$

Si $\phi(t) = E\Big[|X_t - Y_t|^2\Big]$, entonces la desigualdad anterior se representa por

$$0 \leq \phi(t) \leq C \int_0^t \phi(s)\, ds, \qquad \text{para todo } 0 \leq t \leq T.$$

Por la desigualdad de Gronwall, donde $C_0 = 0$, se concluye que $\phi = 0$. De aquí que $\mathbb{P}\Big(X_t = Y_t\Big) = 1$, para cada $t \in [0,T]$, y entonces

$$\mathbb{P}\Big(X_t = Y_t \text{ para todo } t \in D\Big) = 1,$$

donde D es cualquier subconjunto denso numerable de $[0,T]$.

Puesto que (X_t) y (Y_t) tienen trayectorias continuas, $\mathbb{P}$-c.s., y además coinciden en un subconjunto denso numerable de $[0,T]$, entonces deben coincidir en todo el intervalo $[0,T]$, i.e.,

$$\mathbb{P}\Big(X_t = Y_t \quad \text{para todo } t \in [0,T]\Big) = 1.$$

Por lo tanto, (X_t) y (Y_t) son indistinguibles. ∎

Nótese que la condición de Lipschitz es una condición suficiente para que se cumpla la unicidad de la solución.

La demostración de la existencia de soluciones para EDO's involucra el *método de aproximaciones sucesivas* de Picard-Lindelöf, para el caso de las EDE's se sigue aplicando el mismo método. En este caso, primero se define $X_t^{(0)} = X_0$, y luego de forma inductiva como sigue:

$$X_t^{(n+1)} = X_0 + \int_0^t \mu(X_s^{(n)}, s)ds + \int_0^t \sigma(X_s^{(n)}, s)dW_s,$$

para $n = 0, 1, 2, \ldots$. La continuidad y medibilidad de cada elemento de la sucesión se sigue por inducción, tomando como base que $X_t^{(0)}$ es $\mathcal{F}_t$-medible y continua. La parte laboriosa y que se apoya de resultados interesantes surge cuando se quiere demostrar la convergencia uniforme en $[0,T]$, en el sentido $\mathbb{P}$-c.s., de la sucesión $\{X_t^{(n)}\}$, para luego demostrar que tal límite es una solución de la EDE. Para los detalles de la demostración, veáse Øksendal[16], Teorema 5.2.1., p. 66.

Existen diferentes versiones del teorema anterior. Una versión surge cuando se debilita la condición de crecimiento lineal para el caso del coeficiente de deriva y se reemplaza por

$$x\mu(x,t) \leq C^2(1 + |x|^2). \tag{3.6}$$

En esencia, la nueva condición solicita que el coeficiente no debe crecer más rápido en magnitud para $|x|$ grande.

La ausencia de éstos tipos de condiciones permitirán que las soluciones *exploten* en algún momento, tal y como se muestra en el siguiente ejemplo.

Ejemplo 3.3. *Considérese la EDE*

$$dX_t = X_t^2 dt, \quad X_0 = 1,$$

en donde $\mu(x,t) = x^2$ *y* $\sigma(x,t) = 0$*. Por los métodos tradicionales, se sabe que la única solución está dada por*

$$X_t = \frac{1}{1-t}, \quad 0 \leq t < 1,$$

y que en este caso es imposible encontrar una solución definida para todo t*. Nótese que esta ecuación no cumple las condiciones de crecimiento lineal.* □

El Teorema 3.2.1 es válido para EDE's de varias dimensiones, en donde los valores absolutos son reemplazados por la respectiva norma Euclidiana. El siguiente resultado se aplica sólo al caso unidimensional, y expresa que la condición de Lipschitz sobre el coeficiente de difusión puede ser relajada a la *condición de Hölder de orden* $\alpha \geq \frac{1}{2}$. Su respectiva demostración puede encontrarse en Karatzas and Shreve[10], Proposición 2.13, p. 291.

Teorema 3.2.2 (Yamada-Watanabe). *Supóngase que los coeficientes de una ecuación dada como en (3.4) son unidimensionales, y que para todo* $0 \leq t \leq T$ *satisfacen*

$$|\mu(x,t) - \mu(y,t)| \leq K_1 |x-y|,$$

$$|\sigma(x,t) - \sigma(y,t)| \leq K_2 |x-y|^{\alpha}, \tag{3.7}$$

donde K_1, K_2 *son contantes positivas, que no depende de* t*, y* $\alpha \geq \frac{1}{2}$*.*
Entonces existe una única solución fuerte de la ecuación (3.4).

Ejemplo 3.4. *Para* $\frac{1}{2} \leq r < 1$*, considérese la EDE de Girsanov:*

$$dX_t = |X_t|^r dt, \quad X_0 = 0.$$

Nótese que la función $\mu(x,t) = |x|^r$ *no satisface la condición de Lipschitz, pero sí satisface la* condición de Hölder de orden $\alpha = r \geq \frac{1}{2}$.
Puesto que $X_t = 0$ *es una solución, el teorema de Yamada-Watanabe garantiza la unicidad de ésta solución fuerte.* □

Para el caso de las soluciones débiles, también es posible definir la unicidad de las soluciones.

Definición 3.2.3. *Se dice que una solución débil es única, si siempre que* (X_t) *y* (X_t') *sean dos soluciones (quizas en diferentes espacios de probabilidad), entonces tienen la misma distribución en dimensión finita.*

Como se mencionó en la Sección 2.3, la idea original de Itô fue construir procesos de difusión como soluciones de EDE's. Existe un resultado que garantiza que si los coeficientes μ y

σ son funciones continuas, y se cumplen las hipótesis del Teorema 3.2.1, entonces la solución es un proceso de difusión. De forma similar, para el caso de las soluciones débiles, existe un resultado que garantiza que bajo ciertas hipótesis, la solución es un proceso de difusión. Por esta razón, comúnmente, a la solución de una EDE se le conoce como *proceso de difusión* o simplemente *difusión.*

Como caso particular, si los coeficientes de la EDE son $\mu = 0$ y $\sigma = 1$, entonces se concluye que el movimiento Browniano es un proceso de difusión.

Para los lectores interesados en los detalles de estos temas, se recomienda revisar Kloeden and Platen[12], Capítulo 4.

3.3. Ejemplos y algunos métodos de solución

A continuación se mostrará cómo resolver algunas EDE's clásicas, y a menos de que se especifique lo contrario, por solución se entenderá que es una solución fuerte.

Ejemplo 3.5. *Recuérdese que Bachelier propuso el siguiente modelo:*

$$dX_t = \sigma dW_t,$$

en donde $\sigma > 0$ se llama **volatilidad** *del activo, el cual representa las fluctuaciones del activo.*
Otra forma de modelar tal crecimiento es cuando se asegura que los incrementos en el precio tengan media no cero. Ésto se consigue agregando un pequeño incremento μdt, en donde la constante μ, que se conoce como **deriva***, se interpreta como la tasa de crecimiento esperada de X_t.*
Por ejemplo, en la ausencia de fluctuaciones, el precio del activo está dado por

$$X_t = X_0 + \mu t.$$

Nótese que μ resulta ser la pendiente de los incrementos del precio.

En general, el nuevo modelo satisface la EDE

$$dX_t = \mu dt + \sigma dW_t,$$

con μ y $\sigma > 0$ constantes conocidas.
La única solución del nuevo modelo es el proceso (X_t) definido por

$$X_t = X_0 + \mu t + \sigma W_t, \tag{3.8}$$

y se conoce como **movimiento Browniano aritmético***.*

La Figura 3.1 muestra el comportamiento de dos trayectorias del proceso definido por (3.8), para tres parámetros arbitrarios. □

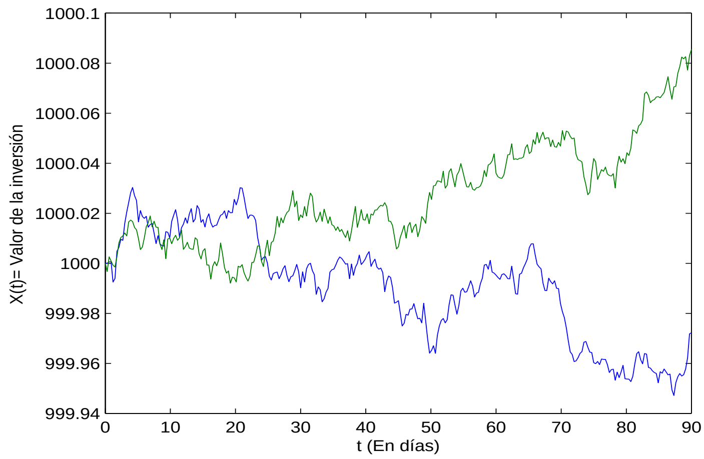

Figura 3.1: Las trayectorias corresponde a los parámetros $\mu = 1.3405 \times 10^{-4}$, $\sigma = 0.0065$, $X_0 = 1000$.

Ejemplo 3.6. *Se sabe que para un activo sin volatilidad, el crecimiento que presenta es exponencial y está expresado por*

$$X_t = X_0 \exp(rt),$$

donde r es la tasa de interés *(compuesta continuamente). Diferenciando se tiene*

$$\frac{dX_t}{X_t} = rdt,$$

el cual representa el incremento relativo en el precio del activo.
Para obtener un modelo más apegado a la realidad, se generaliza la ecuación anterior al adicionar una componente estocástica:

$$\frac{dX_t}{X_t} = rdt + \sigma dW_t.$$

A diferencia del ejemplo anterior, en éste nuevo modelo el inversionista espera que los incrementos relativos *al valor actual sean los mismos. Así, el modelo está representado por la EDE del Ejemplo 3.1:*

$$dX_t = rX_t dt + \sigma X_t dW_t,\ X_0 = x_0, \tag{3.9}$$

donde X_0 representa el capital inicial invertido, las constantes r y σ se conocen como **deriva del activo** *(o rendimiento medio esperado) y* **volatilidad del activo**, *respectivamente.*

Para resolver la EDE (3.9) considérese $f(x) = \ln x$, entonces $f'(x) = \frac{1}{x}$ y $f''(x) = -\frac{1}{x^2}$. Si $Y_t = f(X_t)$, aplicando la fórmula de Itô,

$$\begin{aligned} dY_t = d(\ln X_t) &= \left(\frac{1}{X_t} rX_t + \frac{1}{2}\Big(-\frac{1}{X_t^2}\Big)(\sigma X_t)^2\right)dt + \frac{1}{X_t}\sigma X_t dW_t \\ &= \Big(r - \frac{1}{2}\sigma^2\Big)dt + \sigma dW_t. \end{aligned}$$

Esto es, el nuevo proceso de Itô (Y_t) *está dado por*

$$\begin{aligned} Y_t &= Y_0 + \int_0^t \Big(r - \frac{1}{2}\sigma^2\Big) ds + \int_0^t \sigma dW_s \\ &= Y_0 + \Big(r - \frac{1}{2}\sigma^2\Big) t + \sigma W_t. \end{aligned}$$

Puesto que $Y_t = \ln X_t$, *entonces*

$$\ln\Big(\frac{X_t}{X_0}\Big) = \Big(r - \frac{1}{2}\sigma^2\Big)t + \sigma W_t.$$

Luego,

$$X_t = x_0 \exp\Big(\Big(r - \frac{1}{2}\sigma^2\Big)t + \sigma W_t\Big) \tag{3.10}$$

es la única solución de la ecuación (3.9).

La Figura 3.2 muestra el comportamiento de dos trayectorias del proceso definido por (3.10), utilizando los mismos parámetros que el ejemplo anterior. □

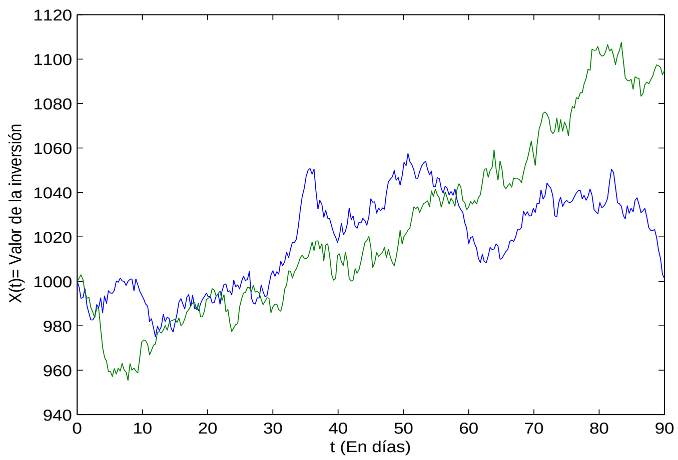

Figura 3.2: Las trayectorias corresponden a los parámetros $\mu = 1.3405 \times 10^{-4}$, $\sigma = 0.0065$, $X_0 = 1000$.

En el ejemplo anterior, nótese que la solución (3.10) representa a un **movimiento Browniano geométrico** (veáse Definición 2.2.3). Este proceso es el modelo tradicional para los precios de las acciones, y éste supone que los precios siguen una distribución lognormal, i.e., el rendimiento tiene distribución normal (véase Apéndice A.1).

Para comprobar si un proceso dado es solución de una EDE, nuevamente es necesario recurrir a la fórmula de Itô. Por ejemplo, para comprobar que (3.10) satisface la EDE (3.9), defínase

$$f(x,t) = x_0 \exp\Big(\Big(r - \frac{1}{2}\sigma^2\Big)t + \sigma x\Big).$$

Si $X_t = f(W_t, t)$, entonces, según la fórmula de Itô,

$$
\begin{aligned}
dX_t = d(f(W_t,t)) &= f_x(W_t,t)dW_t + f_t(W_t,r)dt + \frac{1}{2}f_{xx}(W_t,t)(dW_t)^2 \\
&= \sigma X_t dW_t + \left(r - \frac{1}{2}\sigma^2\right)X_t dt + \frac{1}{2}\sigma^2 X_t dt \\
&= rX_t dt + \sigma X_t dW_t.
\end{aligned}
$$

Ejemplo 3.7. *La EDE*

$$dX_t = -\mu X_t dt + \sigma dW_t, \tag{3.11}$$

con μ y σ constantes positivas, fue postulado por Ornstein y Uhlenbeck en 1930 como una descripción de la aceleración de una partícula de polén sobre un líquido, sujeta a bombardeos por moléculas.
Si X_t representa la velocidad de la partícula en una dimensión, dX_t es el cambio de velocidad por unidad de tiempo, esto es, su aceleración. Ésta aceleración es modelada como si fuera retardada por una fuerza de fricción proporcional a la velocidad, $-\mu X_t$, más una perturbación aleatoria con intensidad σ causada por los bombardeos.

Para resolver la EDE (3.11), considerése $f(x,t) = x\exp(\mu t)$. Por la fórmula de Itô, el proceso $Y_t = X_t \exp(\mu t)$ cumple

$$
\begin{aligned}
dY_t &= \exp(\mu t)dX_t + X_t\mu\exp(\mu t)dt \\
&= -\exp(\mu t)X_t dt + \sigma\exp(\mu t)dW_t + X_t\mu\exp(\mu t)dt \\
&= \sigma\exp(\mu t)dW_t.
\end{aligned}
$$

Luego,

$$Y_t = Y_0 + \int_0^t \sigma\exp(\mu s)dW_s.$$

Por lo tanto, la solución (X_t) de la EDE (3.11) está definida por

$$X_t = \exp(-\mu t)\left(X_0 + \int_0^t \sigma e^{\mu s}dW_s\right),$$

y se conoce como el proceso de Ornstein-Uhlenbeck. □

3.3.1. La exponencial estocástica

Una de las funciones elementales en el cálculo clásico es la función exponencial, cuya propiedad fundamental consiste en que su derivada coincide con ella misma. Ahora se definirá la exponencial en el cálculo de Itô.

Definición 3.3.1. *Si (X_t) y (U_t) son dos procesos tales que la EDE*

$$dU_t = U_t dX_t, \quad U_0 = 1, \tag{3.12}$$

está bien definida, entonces (U_t) *se llama la* **exponencial estocástica de** (X_t)**,** *y se denota por* $\mathcal{E}(X)$.

El siguiente resultado asegura la existencia y unicidad de la exponencial estocástica para procesos de Itô.

Teorema 3.3.1. *Sea* (X_t) *un proceso de Itô. La única solución de (3.12) está definida por*

$$U_t = \mathcal{E}(X)_t := \exp\left(X_t - X_0 - \frac{1}{2}\int_0^t dX_s \cdot dX_s\right).$$

Demostración.

Primero se demostrará que U_t es solución de (3.12).

Defínase $U_t = e^{V_t}$, donde $V_t = X_t - X_0 - \frac{1}{2}\int_0^t dX_s \cdot dX_s$.

Luego,

$$dV_t = dX_t - \frac{1}{2}(dX_t)^2, \quad dV_t \cdot dV_t = (dX_t)^2.$$

Por la fórmula de Itô,

$$\begin{aligned}
dU_t &= d(\exp(V_t)) \\
&= \exp(V_t)dV_t + \frac{1}{2}\exp(V_t)(dV_t)^2 \\
&= \exp(V_t)dX_t - \frac{1}{2}\exp(V_t)(dX_t)^2 + \frac{1}{2}\exp(V_t)(dX_t)^2 \\
&= \exp(V_t)dX_t \\
&= U_t dX_t.
\end{aligned}$$

Así, $U_t = \exp(V_t)$ es una solución de (3.12).

Ahora se demostrará la *unicidad*.

Supóngase que (U_t) y $(\widetilde{U}_t)$ son soluciones de (3.12). Por la fórmula de integración por partes,

$$\begin{aligned}
d\left(\frac{\widetilde{U}_t}{U_t}\right) &= \widetilde{U}_t d\left(\frac{1}{U_t}\right) + \frac{1}{U_t}d\widetilde{U}_t + d\widetilde{U}_t \cdot d\left(\frac{1}{U_t}\right) \\
&= \widetilde{U}_t\left(-\frac{1}{U_t^2}dU_t + \frac{1}{U_t^3}(dU_t)^2\right) + \frac{1}{U_t}\widetilde{U}_t dX_t \\
&\quad + \left(\widetilde{U}_t dX_t\right) \cdot \left(-\frac{1}{U_t^2}U_t dX_t + \frac{1}{U_t^3}U_t^2(dX_t)^2\right) \\
&= \left(-\frac{\widetilde{U}_t}{U_t}dX_t + \frac{\widetilde{U}_t}{U_t}(dX_t)^2\right) + \frac{\widetilde{U}_t}{U_t}dX_t + \left(-\frac{\widetilde{U}_t}{U_t}(dX_t)^2 + 0\right) \\
&= 0.
\end{aligned}$$

Por lo tanto, $U_t = c\tilde{U}_t$, para alguna constante c.

Puesto que $U_0 = 1$, entonces se concluye que $U_t = \tilde{U}_t$, y así el resultado queda demostrado. ■

Si $(U_t) = \mathcal{E}(X)$ es la exponencial de (X_t), entonces el proceso (X_t) se llama el **logaritmo estocástico** de (U_t), y se denota por $\mathcal{L}(U)$.

Por ejemplo, del teorema anterior, la exponencial estocástica del movimiento Browniano está dada por $\exp\Big(W_t - \frac{1}{2}t\Big)$. Por lo tanto, (W_t) es el logaritmo estocástico de $\exp\Big(W_t - \frac{1}{2}t\Big)$.

Formalmente, el logaritmo estocástico de un proceso (U_t) que tiene una diferencial estocástica y que no toma el valor 0, satisface la EDE

$$dX_t = \frac{dU_t}{U_t}, \quad X_0 = 0.$$

Más aún,

$$X_t = \mathcal{L}(U)_t = \ln\left(\frac{U_t}{U_0}\right) + \int_0^t \frac{dU_s \cdot dU_s}{2U_s^2}.$$

3.3.2. Ecuación diferencial estocástica lineal

Las EDE's lineales forman una clase de EDE's que se pueden resolver de forma explícita. En su forma general, una EDE lineal (unidimensional) en el proceso desconocido (X_t) está dada por

$$dX_t = [\mu_1(t) + \mu_2(t)X_t]dt + [\sigma_1(t) + \sigma_2(t)X_t]dW_t, \quad 0 \leq t \leq T, \tag{3.13}$$

donde μ_i y $\sigma_i, i = 1, 2$, son funciones deterministas continuas dadas, y X_0 es una constante. Bajo éstas hipótesis, los coeficientes cumplen las condiciones del Teorema 3.2.1, y en consecuencia, la EDE lineal (3.13) tiene una única solución fuerte.

La ecuación anterior se puede reescribir como

$$\begin{aligned} dX_t &= (\mu_1(t)dt + \sigma_1(t)dW_t) \\ &\quad +(\mu_2(t)X_tdt + \sigma_2(t)X_tdW_t). \end{aligned}$$

Esto sugiere expresar (X_t) como el producto de dos procesos, un movimiento Browniano geométrico (U_t) y un movimiento Browniano aritmético (V_t).

Se propone la solución de la ecuación (3.13) de la forma

$$X_t = U_tV_t,$$

donde

$$dU_t = \mu_2(t)U_tdt + \sigma_2(t)U_tdW_t, \quad U_0 = 1, \tag{3.14}$$

y

$$dV_t = a(t)dt + b(t)dW_t, \quad V_0 = X_0. \tag{3.15}$$

La EDE (3.14) es de la forma

$$dU_t = U_t dY_t, \quad U_0 = 1,$$

donde $dY_t = \mu_2(t)dt + \sigma_2(t)dW_t$, el cual representa la exponencial estocástica del proceso de Itô (Y_t). Luego, su solución está dada por

$$\begin{aligned} U_t &= exp\left(Y_t - Y_0 - \frac{1}{2}\int_0^t dY_s \cdot dY_s\right) \\ &= exp\left(\int_0^t \mu_2(s)ds + \int_0^t \sigma_2(s)dW_s - \frac{1}{2}\int_0^t \sigma_2^2(s)ds\right) \\ &= exp\left(\int_0^t \Big(\mu_2(s) - \frac{1}{2}\sigma_2^2(s)\Big)ds + \int_0^t \sigma_2(s)dW_s\right). \end{aligned} \tag{3.16}$$

Por otra lado, usando la fórmula de integración por partes,

$$\begin{aligned} dX_t = d(U_tV_t) &= U_t dV_t + V_t dU_t + dU_t \cdot dV_t \\ &= U_t a(t)dt + U_t b(t)dW_t + V_t\mu_2(t)U_t dt + V_t\sigma_2(t)U_t dW_t \\ &\quad + b(t)\sigma_2(t)U_t dt \\ &= \Big(a(t)U_t + \mu_2(t)X_t + b(t)\sigma_2(t)U_t\Big)dt + \Big(b(t)U_t + \sigma_2(t)X_t\Big)dW_t. \end{aligned}$$

De donde, para que la relación $X_t = U_tV_t$ se cumpla, los coeficientes $a(t)$ y $b(t)$ deben satisfacer:

$$a(t)U_t + b(t)\sigma_2(t)U_t = \mu_1(t), \quad b(t)U_t = \sigma_1(t),$$

esto es,

$$a(t)U_t + \sigma_1(t)\sigma_2(t) = \mu_1(t), \quad b(t)U_t = \sigma_1(t).$$

Así, los coeficientes $a(t)$ y $b(t)$ quedan expresados en función de coeficientes conocidos. Por lo tanto, la solución de la EDE lineal (3.13) está definida por

$$X_t = U_t\left(X_0 + \int_0^t \frac{\mu_1(t) - \sigma_1(t)\sigma_2(t)}{U_t}ds + \int_0^t \frac{\sigma_1(t)}{U_t}dW_s\right), \tag{3.17}$$

donde U_t está definida por (3.16).

Ejemplo 3.8. *Una generalización de la ecuación (3.11) está dada por la EDE*

$$dX_t = (\beta - \alpha X_t)dt + \sigma dW_t.$$

La solución se obtiene de forma directa al aplicar la fórmula (3.17):

$$X_t = \frac{\beta}{\alpha} + \exp(-\alpha t)\left(X_0 - \frac{\beta}{\alpha} + \int_0^t \sigma e^{\alpha s} dW_s\right). \tag{3.18}$$

□

Ejemplo 3.9. *Uno de los modelos estándar para describir la tasa de interés para préstamos en efectivo, cuando la tasa se asume como una función aleatoria, es el* ***modelo de Vasicek*** *(1977), siendo éste uno de los primeros modelos en términos de estructura.*

El proceso de difusión propuesto por Vasicek es una versión del proceso de Ornstein-Uhlenbeck de tipo **reversión media**, *y está definido como la solución de la EDE*

$$dr_t = b[\mu - r_t]dt + \sigma dW_t, \quad t \in [0, T], \tag{3.19}$$

para r_0 una constante y, en donde b, μ y σ son constantes estrictamente positivas. La constante μ es interpretada como el promedio histórico de la tasa de interés y b como la rapidez con la que los valores r_t que se desvían de μ tienden a regresar a éste; debido a este comportamiento, se dice que (r_t) es de tipo reversión a la media. Por último, el parámetro σ, que se conoce como volatilidad, *mide las fluctuaciones de r_t alrededor de μ.*

Claramente, la EDE (3.19) se puede reescribir como

$$dr_t = [a - br_t]dt + \sigma dW_t, \quad t \in [0, T], \tag{3.20}$$

cuya solución está dada como en (3.18). Por lo tanto, el modelo de Vasicek está dado por

$$r_t = r_0 exp(-bt) + \mu[1 - \exp(-bt)] + \sigma \exp(-bt) \int_0^t \exp(bs)dW_s. \tag{3.21}$$

La Figura 3.3 muestra el comportamiento de dos trayectorias del modelo propuesto por Vasicek. □

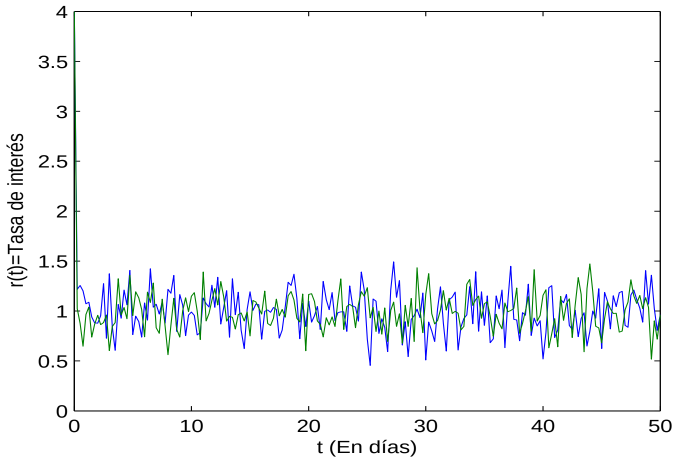

Figura 3.3: Las trayectorias corresponde a los parámetros $\mu = 1$, $b = 4$, $\sigma = 0.4$, $r_0 = 3$.

Para motivar al lector a revisar las aplicaciones de las EDE's a otras áreas de la ciencia, se finaliza este capítulo con el siguiente ejemplo.

Ejemplo 3.10. *En un tiempo t, la carga $Q(t)$ en un punto fijo de un circuito eléctrico satisface la EDO*

$$LQ''(t) + RQ'(t) + \frac{1}{C}Q(t) = F(t), \quad Q(0) = Q_0, \quad Q'(0) = I_0. \tag{3.22}$$

Puede presentarse el caso en el que algún coeficiente, digamos $F(t)$, no se puede controlar totalmente y entonces tiene la forma

$$F(t) = G(t) + \text{“ruido blanco”}.$$

Por lo tanto, la ecuación (3.22) puede ser transformada en la EDE:

$$LQ_t'' + RQ_t' + \frac{1}{C}Q_t = G_t + \alpha\xi_t, \quad Q(0) = Q_0, \quad Q'(0) = I_0. \tag{3.23}$$

Se define el vector

$$X(t) = \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} = \begin{pmatrix} Q_t \\ Q_t' \end{pmatrix},$$

el cual permite representar el siguiente sistema de ecuaciones:

$$\begin{cases} X_1'(t) = X_2(t), \\ LX_2'(t) = -RX_2(t) - \frac{1}{C}X_1(t) + G_t + \alpha\xi_t, \end{cases}$$

cuya notación matricial está dada por

$$dX(t) = \Big(A(t) + BX(t)\Big)dt + CdW_t, \tag{3.24}$$

donde

$$dX(t) = \begin{pmatrix} dX_1(t) \\ dX_2(t) \end{pmatrix}, \; A(t) = \begin{pmatrix} 0 \\ \frac{1}{L}G_t \end{pmatrix}, \; B = \begin{pmatrix} 0 & 1 \\ -\frac{1}{CL} & -\frac{R}{L} \end{pmatrix}, \; C = \begin{pmatrix} 0 \\ \frac{\alpha}{L} \end{pmatrix},$$

y W_t define a un movimiento Browniano unidimensional.

Para resolver la EDE bidimensional (3.24), por la fórmula (3.17) se tiene

$$X(t) = exp(Bt)\left(X(0) + \int_0^t A(t)exp(-Bs)ds + \int_0^t Cexp(-Bt)dW_s\right).$$

Finalmente, por el Corolario 2.4.4,

$$X(t) = exp(Bt)\left(X(0) + Cexp(-Bt)W_t + \int_0^t exp(-Bs)\Big(A(t) + BCW_s\Big)ds\right),$$

en donde

$$exp(D) = \sum_{n=0}^{\infty} \frac{1}{n!}D^n, \;\; \textit{para una matriz } D_{n\times n}.$$

Otra forma para resolver la ecuación anterior es considerar la función $f : \mathbb{R}^2 \times [0, \infty) \to \mathbb{R}^2$ *dada por*

$$f(x_1, x_2, t) = exp(-Bt) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

y luego aplicar la fórmula general de Itô.

Capítulo 4

Introducción a las Finanzas Estocásticas

Una de las herramientas modernas más utilizada por los analistas financieros es el cálculo estocástico, ya que sobre el descansa prácticamente toda la teoría económica y el análisis financiero en tiempo continuo. La martingala es otra herramienta que desempeña un papel esencial en el modelado de los mercados financieros, así como en la valuación teórica de muchos contratos financieros.

El tamaño considerable que han alcanzado los mercados de productos derivados (futuros, warrants, swaps, opciones) se debe, en gran medida, a la flexibilidad que estos instrumentos proporcionan a sus usuarios para entrar o salir rápidamente del mercado. Además, el acelerado desarrollo de las tecnologías de la información ha facilitado su operación y diversificación.

En particular, el crecimiento del mercado de opciones se ve impulsado por su liquidez y por el apalancamiento que éstos presentan, pues la inversión inicial es pequeña comparada con la que requieren otros instrumentos de inversión.

El objetivo principal de este capítulo es hacer notar el formalismo matemático de los derivados financieros, y con ello introducir a los interesados en esta teoría, que en la actualidad está logrando un gran impacto, tanto en la aplicación como en la generación de nuevos resultados teóricos. Además, al final del capítulo se analizará uno de los modelos pioneros en la valuación de opciones: el modelo de Black-Scholes.

4.1. Mercado de opciones

En lo sucesivo, por *mercado financiero* (o simplemente como *mercado*) se entenderá que se trata del lugar, mecanismo o sistema en el cual se compran y venden activos financieros.

En un mercado financiero se pueden apreciar dos grandes grupos. En un primer grupo se comercializan *activos primarios*, tales como acciones, bonos, divisas, activos no financieros (commodities, tales como gas, granos, metales preciosos); mientras que en el segundo grupo se comercializan *contratos financieros*, los cuales prometen algún pago o alguna entrega en un futuro y que éste depende del comportamiento de un determinado activo primario, de aquí que al contrato se le conozca como *derivado financiero* y al activo primario como *activo subyacente.*

A continuación se dará una breve introducción a los contratos financieros conocidos como *opción.* Para detalles de su utilidad y funcionamiento, así como de los términos técnicos, se recomienda revisar De Lara[3] y Hull[8].

Definición 4.1.1. *Un* contrato forward *es un acuerdo para comprar (o vender) un activo en una fecha futura específica, T, por un precio específico, K. Se dice que el comprador mantiene una* posición larga*, y el vendedor una* posición corta.

Nótese que en los contratos forward, tanto el comprador como el vendedor están obligados a cumplir con su parte del contrato.

Un contrato forward también es conocido como un *contrato a plazo.* El "problema de valuación" de un forward es determinar el valor K que debería ser escrito en el contrato.

Un *futuro* es un contrato muy parecido al forward, con la diferencia de que éste se negocia en un mercado organizado, y que tanto el tamaño del contrato como la fecha de vencimiento están estandarizados por la misma; mientras tanto, la transacción de un forward se da directamente entre las partes involucradas, y por tanto, las características son determinadas por los participantes.

En general, en un mercado de forwards y futuros se aprecian dos participantes: los que toman una posición corta, conocidos como *hedgers*, y son quienes buscan cubrirse ante cambios desfavorables en los precios; y los que toman una posición larga, conocidos como *especuladores*, son los que proporcionan liquidez y está constituido por inversionistas que quieren lograr una ganancia anticipándose a los cambios de precios.

Una teoría que está logrando mayor interés es la que se conoce como *teoría de valuación de opciones.* Una **opción** es un contrato, instrumento financiero, título financiero o derivado que otorga a su tenedor el *derecho*, pero no la obligación, de comprar o vender cierto activo subyacente a un precio prefijado en una fecha futura. El término *derecho* es en el sentido de que el tenedor tiene la oportunidad de ejercer el contrato, o de cancelarlo si él lo prefiere.

En el caso de las opciones, existe principalmente dos participantes: el **tenedor** (*holder*) que ha tomado la posición larga (i.e., ha comprado la opción) y el **emisor** (*writer*) que ha tomado la posición corta (i.e., ha vendido o emitido la opción).

Existe básicamente dos tipos de opciones:

- **Opción de compra** (*call option*): otorga al tenedor el derecho de comprar el activo subyacente .

- **Opción de venta** (*put option*): otorga al tenedor el derecho de vender el activo subyacente.

Además de las características referentes al activo subyacente, al momento que se contrata una opción (i.e., en $t = 0$) generalmente se especifican los siguientes datos:

- *Fecha de ejercicio* (*expiration date, exercise date* o *maturity*): fecha de expiración del derecho contenido en la opción.
- *Precio de ejercicio* (*strike price*): precio acordado para la compra o venta del activo subyacente.
- *Prima o precio de la opción*: monto que el tenedor paga al emisor por adquirir el derecho de compra o de venta.
- *Derecho que se adquiere*: o es derecho de compra (call) o derecho de venta (put).

Otra característica esencial que se especifica dentro del contrato de opción es el *tipo de opción*, siendo las Europeas y las Americanas las básicas.

Definición 4.1.2. *Una* ***opción de compra Europea*** *(*European call option*) otorga al tenedor el derecho (pero no la obligación) de comprar en la fecha de ejercicio T un activo a un precio de ejercicio K, el cual fue fijado cuando se firmó el contrato.*
Una ***opción de venta Europea*** *(*European put option*) otorga al tenedor el derecho (pero no la obligación) de vender en la fecha de ejercicio T un activo a un precio de ejercicio K, el cual fue fijado cuando se firmó el contrato.*

Nótese que el término *Europeo* es reservado para las opciones que pueden ser ejercidas sólo en la fecha de ejercicio T.

En términos muy generales, en los siguientes ejemplos se muestra la utilidad de las opciones Europeas.

Ejemplo 4.1. *Supóngase que una empresa importadora mexicana quiere asegurarse contra las alzas en el precio del dólar. Para ello, hoy decide comprar una opción de compra Europea que le da el derecho a comprar 100,000 dólares dentro de tres meses a un precio de ejercicio de \$15.00 por dólar. Para adquirir ese derecho, supóngase que la empresa paga \$0.50 por dólar, i.e., la prima de la opción tiene un costo de*

$$100000 \times \$0.50 = \$50,000.00.$$

Si en la fecha de ejercicio el precio del dólar en el mercado fuese mayor que \$15.00, entonces la empresa ejerce la opción, pues sólo pagaría \$15.00 por dólar. Por el contrario, si en la fecha de ejercicio el precio del dólar en el mercado estuviera por debajo de \$15.00, *la empresa no ejercerá la opción, ya que no tiene sentido pagar \$15 por dólar cuando en el mercado se puede comprar a un precio menor, en este caso la opción vence sin ser ejercida.*

□

Supóngase que S_T denota el valor del activo en el tiempo T. Generalmente en una opción de compra Europea, si $S_T > K$ entonces el tenedor ejerce su derecho y obtiene una ganancia de $S_t - K$, ya sea porque adquiere un activo a menor precio (por ejemplo dólares) o por la entrega en efectivo por parte del emisor. La entrega en efectivo se debe a que los derivados simulan las siguientes transacciones: *el tenedor adquiere activos a un precio K, el cual es menor respecto al precio en el mercado actual, por lo que inmediatamente vende el activo al precio actual S_T, y por lo tanto obtendrá una ganancia de $S_T - K$.*

Pero si se cumple que $S_T \leq K$, entonces el tenedor no ejerce su derecho, *ya que no tiene sentido comprar un activo a un precio mayor que el precio en el mercado actual*, y en consecuencia la ganancia es cero. Debido a este comportamiento, en la fecha de ejercicio, la opción de compra proporciona un pago definido por

$$C^{call} = (S^i - K)^+ = \begin{cases} S^i - K & \text{si } S^i > K \\ 0 & \text{si } S^i \leq K. \end{cases}$$

Ejemplo 4.2. *Ahora considérese el caso de un inversionista que desea tomar una posición que le permita obtener beneficios si aumenta la cotización de las acciones de CEMEX.*

Supóngase que el precio de las acciones de CEMEX es actualmente de \$72.00, y que la correspondiente opción de compra Europea con fecha de ejercicio dentro de tres meses, a un precio de ejercicio de \$80.00, cotiza actualmente a \$3.00.

El inversionista tiene dos posibles alternativas para una inversión de \$7,200.00. La primera alternativa se refiere a una compra directa de 100 acciones, y la segunda a una compra de 2400 opciones de compra sobre acciones de CEMEX (cada opción es escrita sobre 100 acciones, en realidad se compra 24 contratos de opción).

Supóngase que en la fecha de ejercicio el precio de las acciones de CEMEX sube a \$90.00. En la primera alternativa, el inversionista tendrá un beneficio de

$$100 \times (\$90.00 - \$72.00) = \$1,800.00.$$

Respecto a la segunda alternativa, cada opción de compra sobre acciones de CEMEX con precio de ejercicio de \$80.00 proporciona un beneficio de \$10.00, ya que la opción permite comprar por \$80.00 algo que vale \$90.00. Así, el valor total de las opciones compradas es de

$$2400 \times \$10.00 = \$24,000.00.$$

Restando el coste original de las opciones (i.e., la prima), el beneficio neto por la segunda alternativa es de

$$\$24,000.00 - 2400 \times \$3.00 = \$16,800.00.$$

La estrategia de opciones es por lo tanto alrededor de 9 veces más beneficiosa que la estrategia de compra de acciones. Por supuesto las opciones también dan origen a un mayor potencial de pérdida. Supóngase que el precio de las acciones baja a \$68.00 en el tiempo de ejercicio. En la primera alternativa resultaría una pérdida de

$$100 \times (\$72.00 - \$68.00) = \$400.00.$$

Dado que las opciones vencen sin haber sido ejercidas, la estrategia de opciones supondría una pérdida de \$7,200.00. □

En el caso de las opciones *Americanas*, éstas pueden ser ejercidas en cualquier tiempo $t \leq T$. Otros tipos de opciones más complejas son las que se conocen como **opciones exóticas** (u opciones de segunda generación). Éstas se diferencian de las Americanas y Europeas por el grado de dificultad en el cálculo de la prima y porque pueden ser ejercidas de acuerdo a especificaciones particulares del tenedor.

Algunos ejemplos de opciones exóticas son las **opciones Asiáticas**, las **opciones barrera**, y las **opciones lookback**.

Uno de los principales objetivos de la teoría de opción es el cálculo de la prima, la cual principalmente está en función de cinco variables. A continuación se menciona la relación de cada una de ellas cuando se mantiene todas las demás como constantes:

- *Precio del activo subyacente.* Si el precio del activo subyacente tiene una tendencia a aumentar, entonces los derechos a comprar se aprecian. En caso contrario, los derechos se deprecian.
- *Precio de ejercicio.* En una opción de compra, cuanto más se aproxima el precio del activo subyacente al precio de ejercicio, los derechos de compra se deprecian. Y conforme mayor sea la cotización del activo subyacente en el mercado con respecto al precio de ejercicio, los derechos de compra tienden a apreciarse.
- *El tiempo.* Cuando más lejano sea la fecha de ejercicio, mayor será el valor de derecho, sin importar si se trate de una opción de compra o de venta. Y cuando más próximo sea la fecha de ejercicio, éste será menor.
- *El tipo de interés.* Ante variaciones de los tipos de interés, los derechos de compra se mueven en el mismo sentido.
- *Volatilidad del activo subyacente.* La volatilidad indica en qué medida varía el rendimiento del activo subyacente. Cuando el rendimiento es muy volátil, los derechos tanto de compra como de venta se aprecian. Mas no así cuando es poco volátil.

En el análisis anterior, si no se especificó de manera explícita el comportamiento para el caso de las opciones de ventas, se entenderá que es contrario al de las opciones de compra.

En los ejemplos anteriores se puede apreciar las dos principales utilidades de las opciones:

- *Como instrumento de inversión/especulación.* Ofrece a los usuarios un amplio abanico de posibilidades si es capaz de predecir correctamente los potenciales movimientos de los precios o la volatilidad del activo subyacente. Permite lograr unos beneficios muy superiores que si se compra directamente la acción.
- *Como instrumento de cobertura.* Permite cubrirse ante los riesgos generados por el precio del activo subyacente.

En el resto del capítulo se estará interesado en mostrar las herramientas que permiten calcular la prima de una opción.

4.2. Modelos de un periodo

En ésta primera sección se está interesado en la estructura matemática de un modelo simple de un período de un mercado, el cual tiene la característica de valorar los activos al tiempo inicial $t = 0$ y al tiempo final $t = 1$, esto es, $T = \{0, 1\}$. Se partirá del hecho de contar con un número finito de activos, con precio conocido al tiempo $t = 0$, y precio desconocido al tiempo $t = 1$, pero descritos por variables aleatorias en algún espacio de probabilidad.

Se asumirá que comprar y vender activos no crea costos extras, una suposición que puede no ser válida para un inversionista pequeño, pero llega a ser más realista conforme aumenta la inversión. Además, se supondrá que el mercado es completamente líquido, i.e, siempre es posible comprar y/o vender cantidades ilimitadas en el mercado.

4.2.1. Teoría de arbitraje y medidas martingalas

Considérese un mercado con $d+1$ activos y supóngase que el i-ésimo activo está disponible al tiempo $t = 0$ con un precio $\pi^i \geq 0$. La $d + 1$-ada ordenada

$$\overline{\pi} = (\pi^0, \pi^1, \ldots, \pi^d) \in \left(\mathbb{R}^{d+1}\right)^+$$

es llamada un ***sistema de precios***.

Puesto que los precios al tiempo $t = 1$ son usualmente desconocidos, los precios de los activos estarán expresados por las v.a.'s no negativas

$$S^i : (\Omega, \mathcal{F}) \to [0, \infty), \text{ para } i = 0, 1, \ldots, d,$$

en donde el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ es tal que cada $\omega \in \Omega$ corresponde a un *escenario* en particular del mercado evolutivo, y $S^i(\omega)$ es el precio del i-ésimo activo en el tiempo

$t = 1$, si el escenario ω ocurre.

Los activos sin riesgos de inversión se traducen a un pago seguro de cierta cantidad en $t = 1$. En este modelo simple para un periodo, tales oportunidades de inversión serán incluidos en la siguiente hipótesis:

$$\pi^0 = 1, \quad S^0 = 1 + r,$$

donde la constante r representa el *rendimiento* de una unidad monetaria de inversión en un *bono*. Es natural asumir que $r \geq 0$, pero para los propósitos aquí planteados, es suficiente requerir que $S^0 > 0$, o equivalentemente $r > -1$.

Se introducen las siguientes notaciones:

$$\overline{\pi} = (1, \pi),$$

$$\overline{S} = (S^0, S^1, \ldots, S^d) = (S^0, S),$$

donde $S^i, i > 0$, representa el valor del i-ésimo activo con riesgo. Además, de aquí en adelante $a \cdot b$ denotará el producto punto de los vectores a y b.

Al tiempo $t = 0$, un inversionista eligirá un *portafolio*

$$\overline{\xi} = (\xi^0, \xi^1, \ldots, \xi^d) = (\xi^0, \xi) \in \mathbb{R}^{d+1},$$

donde ξ^i representa el número de acciones del i-ésimo activo. Luego, el precio por comprar el portafolio $\overline{\xi}$ está dado por

$$\overline{\pi} \cdot \overline{\xi} = \sum_{i=0}^{d} \pi^i \xi^i. \tag{4.1}$$

Al tiempo $t = 1$, el portafolio tendrá el valor

$$\overline{\xi} \cdot \overline{S}(\omega) = \sum_{i=0}^{d} \xi^i S^i(\omega) = \xi^0(1 + r) + \xi \cdot S(\omega), \tag{4.2}$$

el cual depende del escenario $\omega \in \Omega$.

Nótese que un portafolio tiene un valor de mercado determinista en $t = 0$ y un valor estocástico en $t = 1$. Esta característica permite definir el **proceso de valor** (*value process*) asociado a una estrategia de inversión $\overline{\xi}$ como:

$$V_t^{\overline{\xi}} = \xi^0 S_t^0 + \xi^1 S_t^1 + \ldots + \xi^d S_t^d, \quad t = 0, 1,$$

el cual, es una simplificación de

$$V_0^{\overline{\xi}} = \xi^0 + \xi^1 \pi^1 + \ldots + \xi^d \pi^d,$$

$$V_1^{\overline{\xi}} = \xi^0(1+r) + \xi^1 S^1 + \ldots + \xi^d S^d.$$

El uso de este proceso se vuelve necesario cuando los modelos de mercado son analizados en más de dos pasos. Debido al interés que se persigue, en esta sección se ocupará la notación dada por (4.1) y (4.2).

La definición de portafolio permite que las componentes ξ^i sean negativas. Si $\xi^0 < 0$, esto corresponde a tomar $|\xi^0|$ préstamos con valor $|\xi^0|$ al tiempo $t = 0$ y valor $(1+r)|\xi^0|$ al tiempo $t = 1$. Si $\xi^i < 0$ para $i \geq 1$, esto corresponde a una *venta en corto* [1] del i-ésimo activo. En particular, al inversionista le es permitido tomar una posición corta $\xi^i < 0$, y usar la cantidad $\pi^i|\xi^i|$ para comprar ξ^j activos, $i \neq j$.

Definición 4.2.1. *Una estrategia de inversión $\overline{\xi} \in \mathbb{R}^{d+1}$ es llamada una* ***oportunidad de arbitraje*** *si $\overline{\pi} \cdot \overline{\xi} \leq 0$, pero $\overline{\xi} \cdot \overline{S} \geq 0$, $\mathbb{P}$-c.s, y $\mathbb{P}\Big(\overline{\xi} \cdot \overline{S} > 0\Big) > 0$.*

La definición anterior es equivalente a la siguiente afirmación:
Una oportunidad de arbitraje es un portafolio $\overline{\xi} \in \mathbb{R}^{d+1}$ tal que $\overline{\pi} \cdot \overline{\xi} = 0$, $\overline{\xi} \cdot \overline{S} \geq 0$, y $E\Big[\overline{\xi} \cdot \overline{S}\Big] > 0$.

Intuitivamente, una oportunidad de arbitraje permite al inversionista, con probabilidad positiva, obtener una ganancia positiva sin ser expuesto a riesgos.

La existencia de una oportunidad de arbitraje permite catalogar al mercado como un *mercado ineficiente*, en el sentido de que los precios de ciertos activos no son asignados de forma razonable. En realidad tales mercados son muy difíciles de encontrar, pues el comportamiento está controlado por la oferta y la demanda. En consecuencia, la ausencia de oportunidad de arbitraje será nuestra principal suposición, y ahora se pretende caracterizar a tales modelos.

La ausencia de arbitraje implica que S^i se anula, $\mathbb{P}$-c.s., cuando $\pi^i = 0$. Por lo tanto, sin perdida de generalidad, a partir de ahora se asumirá que

$$\pi^i > 0 \ \text{ para } i = 1, \ldots, d.$$

El siguiente lema muestra que la ausencia de oportunidad de arbitraje es equivalente a la siguiente propiedad del mercado:
Cualquier inversión en activos con riesgo que permita obtener, con probabilidad positiva, un mejor resultado que si la misma cantidad fuese invertida en activos sin riesgos debe estar sujeta a algún riesgo.

[1] El inversionista toma una *posición corta* cuando vende un activo sin aún ser dueño del mismo.

Lema 4.2.1. *Los siguientes enunciados son equivalentes.*

a) *El modelo de mercado admite una oportunidad de arbitraje.*

b) *Existe un vector $\xi \in \mathbb{R}^d$ tal que*

$$\xi \cdot S \geq (1+r)\xi \cdot \pi,\ \mathbb{P}\text{-}c.s.\ , \quad \mathbb{P}(\xi \cdot S > (1+r)\xi \cdot \pi) > 0. \tag{4.3}$$

Demostración.

Primero se verificará que **a)** implica **b)**.

Sea $\overline{\xi}$ una oportunidad de arbitraje. Por definición,

$$0 \geq \overline{\xi} \cdot \overline{\pi} = \xi^0 + \xi \cdot \pi.$$

Luego,

$$\begin{aligned} -\xi \cdot \pi &\geq \xi^0; \\ -(1+r)\xi \cdot \pi &\geq (1+r)\xi^0; \\ \xi \cdot S - (1+r)\xi \cdot \pi &\geq \xi \cdot S + (1+r)\xi^0 = \overline{\xi} \cdot \overline{S}. \end{aligned} \tag{4.4}$$

Puesto que $\overline{\xi} \cdot \overline{S}$ es no negativo, $\mathbb{P}$-c.s., y es estrictamente positivo con probabilidad no cero, entonces, por la ecuación (4.4), se cumple (4.3).

Ahora se verificará que **b)** implica **a)**.

Sea ξ que satisface la condición (4.3). Se afirma que el portafolio (ξ^0, ξ) con $\xi^0 := -\xi \cdot \pi$ es una oportunidad de arbitraje.

En efecto, por construcción,

$$\overline{\xi} \cdot \overline{\pi} = \xi^0 + \xi \cdot \pi = 0.$$

Por otra parte,

$$\overline{\xi} \cdot \overline{S} = -(1+r)\xi \cdot \pi + \xi \cdot S,$$

la cual es no negativa, $\mathbb{P}$-c.s., y estrictamente positiva con probabilidad no cero, pues ξ satisface (4.3). ■

A continuación se caracterizarán aquellos modelos de mercados que no admiten oportunidades de arbitraje. Tales modelos se llamarán *libres de arbitraje*. Una vez alcanzado éste objetivo, se estará dando respuesta a una pregunta clave en las finanzas estocásticas: *¿Qué modelos son correctos para analizar el comportamiento de los mercados financieros?*.

Definición 4.2.2. *Una medida de probabilidad $\mathbb{P}^*$ es llamada una* ***medida de riesgo neutral****, o una* ***medida martingala****, si*

$$\pi^i = E^*\left[\frac{S^i}{1+r}\right],\ i = 0, 1, \dots, d. \tag{4.5}$$

La relación (4.5) establece que el precio del i-ésimo activo es identificado como la esperanza, bajo la probabilidad $\mathbb{P}^*$, del *pago descontado*. Ésta relación se puede reescribir como $E^*[S^i] = (1+r)\pi^i$, la cual expresa que, en términos del valor esperado, al inversionista le es indiferente si la cantidad π^i es invertida en un activo con riesgo o en un activo sin riego, de aquí que a $\mathbb{P}^*$ se le conozca como *medida de riesgo neutral*. Cuando los modelos son analizados en varios periodos, o en el caso continuo, el uso del término *martingala* se vuelve más claro.

Un caso muy particular de un modelo binomial que motiva a los teóricos a considerar la relación (4.5) es cuando $\xi^1 = 1$ y S^1 sólo presenta un aumento u, o una disminución d. Bajo estos supuestos, existe un resultado que garantiza que el modelo es libre de arbitraje si y sólo si $d \leq 1+r \leq u$. Esto es, $1+r$ es una combinación convexa de u y de d :

$$1 + r = q_u \cdot u + q_d \cdot d,$$

donde $q_u, q_d \geq 0$ y $q_u + q_d = 1$. En particular, los pesos q_u y q_d pueden ser interpretados como una nueva probabilidad $\mathbb{Q}$, tal que,

$$E_{\mathbb{Q}}\left[\frac{S^1}{1+r}\right] = \frac{1}{1+r}\left[q_u\pi^1 u + q_d\pi^1 d\right] = \pi^1.$$

Antes de continuar, se recomienda al lector revisar el Apéndice A.2.

En lo sucesivo, por $\mathcal{P}$ se denotará al conjunto convexo definido por

$$\mathcal{P} = \{\mathbb{P}^* | \mathbb{P}^* \text{ es una medida martingala y } \mathbb{P}^* \approx \mathbb{P}\}.$$

Una medida martingala $\mathbb{P}^*$ también recibe el nombre de *medida de valuación* o *medida martingala equivalente.*

Para facilitar la prueba del resultado central en la teoría de las finanzas, se introduce el vector aleatorio $Y = (Y^1, \ldots, Y^d)$, conocido como *vector de flujos descontados* (*discounted net gains*), en donde

$$Y^i := \frac{S^i}{1+r} - \pi^i, \quad i = 1, \ldots, d. \tag{4.6}$$

Por el Lema 4.2.1, la ausencia de arbitraje es equivalente a la siguiente condición:

$$\text{Para } \xi \in \mathbb{R}^d, \text{ si } \ \xi \cdot Y \geq 0, \ \mathbb{P}\text{-c.s., entonces } \ \xi \cdot Y = 0, \ \mathbb{P}\text{-c.s.}. \tag{4.7}$$

Puesto que $-\pi^i$ es cota inferior de Y^i, la esperanza $E^*[Y^i]$ de Y^i bajo cualquier medida $\mathbb{P}^*$ está bien definida. Luego, $\mathbb{P}^*$ es una medida martingala sí y sólo si

$$E^*[Y] = 0.$$

A continuación se demostrará el resultado principal que caracteriza a los modelos de mercado libre de arbitraje en términos de las medidas martingalas. Este teorema fundamental es conocido, de forma abreviada, como *FTAP* (*Fundamental Theorem of Asset Pricing*).

Teorema 4.2.1 (Teorema Fundamental de Valuación de Activos). *Un modelo de mercado es libre de arbitraje sí y sólo si $\mathcal{P} \neq \emptyset$. En este caso, existe $\mathbb{P}^*$ con densidad $\frac{d\mathbb{P}^*}{d\mathbb{P}}$ acotada.*

Demostración.

Primero se demostrará la suficiencia.

Supóngase que existe una medida martingala $\mathbb{P}^* \in \mathcal{P}$. Luego, tómese un portafolio $\overline{\xi} \in \mathbb{R}^{d+1}$ tal que $\overline{\xi} \cdot \overline{S} \geq 0$, $\mathbb{P}$-c.s., y $E[\overline{\xi} \cdot \overline{S}] > 0$.

Puesto que $\mathbb{P}^* \approx \mathbb{P}$, entonces las propiedades previas se siguen cumpliendo si se considera la probabilidad $\mathbb{P}^*$. Así,

$$\overline{\pi} \cdot \overline{\xi} = \sum_{i=0}^{d} \pi^i \xi^i = \sum_{i=0}^{d} E^* \left[\frac{\xi^i S^i}{1+r} \right] = E^* \left[\frac{\overline{\xi} \cdot \overline{S}}{1+r} \right] > 0.$$

Por lo tanto, $\overline{\xi}$ no puede ser una oportunidad de arbitraje.

Ahora se demostrará la necesidad.

Para esta parte, nótese que, por las observaciones previas, el FTAP se reduce en la siguiente afirmación:

Se cumple la condición (4.7) sí y sólo si existe alguna $\mathbb{P}^* \approx \mathbb{P}$ tal que $E^*[Y] = 0$, y en este caso, $\mathbb{P}^*$ puede ser elegido de tal forma que la densidad $\frac{d\mathbb{P}^*}{d\mathbb{P}}$ sea acotada.

En consecuencia, la parte que resta demostrar del FTAP se reduce a demostrar la necesidad de la afirmación anterior.

Como primer caso, supóngase que $E[|Y|] < \infty$.

Sea el conjunto convexo

$$\mathcal{Q} := \left\{ \mathbb{Q} \,|\mathbb{Q} \text{ es medida de probabilidad, } \mathbb{Q} \approx \mathbb{P} \text{ y } \frac{d\mathbb{Q}}{d\mathbb{P}} \text{ es acotada } \right\},$$

y el conjunto

$$\mathcal{C} := \left\{ E_{\mathbb{Q}}[Y] | \mathbb{Q} \in \mathcal{Q} \right\}.$$

Sean $\mathbb{Q}_0, \mathbb{Q}_1 \in \mathcal{Q}$ arbitrarios, y $0 \leq \alpha \leq 1$, entonces

$$\mathbb{Q}_\alpha := \alpha \mathbb{Q}_1 + (1-\alpha)\mathbb{Q}_0 \in \mathcal{Q},$$

luego,

$$\alpha E_{\mathbb{Q}_1}[Y] + (1-\alpha) E_{\mathbb{Q}_0}[Y] = E_{\mathbb{Q}_\alpha}[Y] \in \mathcal{C},$$

por lo tanto, $\mathcal{C}$ es convexo en $\mathbb{R}^d$.

Si se demuestra que $0 \in \mathcal{C}$, entonces la prueba se termina.
Por contradicción, supóngase que $0 \notin \mathcal{C}$. Puesto que $\mathcal{Q} \neq \emptyset$, ya que $\mathbb{P} \in \mathcal{Q}$, entonces $\mathcal{C} \neq \emptyset$. Por la condiciones anteriores, el *teorema de separación del hiperplano en dimensión finita* (véase Apéndice A.4) asegura la existencia de un vector $\xi \in \mathbb{R}^d$ tal que, para todo $x \in \mathcal{C}$,

$$\xi \cdot x \geq 0, \text{ y } \xi \cdot x_0 > 0 \text{ para algún } x_0 \in \mathcal{C}.$$

Esto es, para todo $\mathbb{Q} \in \mathcal{Q}$, ξ satisface

$$E_{\mathbb{Q}}[\xi \cdot Y] \geq 0, \text{ y } E_{\mathbb{Q}_0}[\xi \cdot Y] > 0 \text{ para algún } \mathbb{Q}_0 \in \mathcal{Q}. \tag{4.8}$$

De esta última condición, se concluye que $\mathbb{P}(\xi \cdot Y > 0) > 0$. Si se llegase a demostrar que $\xi \cdot Y \geq 0$, $\mathbb{P}$-c.s., entonces se llegaría a contradecir la condición (4.7); por consiguiente, se estará demostrando que $0 \in \mathcal{C}$.
Para demostrar que $\xi \cdot Y \geq 0$, $\mathbb{P}$-c.s., defínase las funciones φ_n dadas por

$$\varphi_n := \left(1 - \frac{1}{n}\right) I_A + \frac{1}{n} I_{A^c},$$

donde $A := \{\xi \cdot Y < 0\}$. Luego, tómese a φ_n como las densidades para las nuevas probabilidades $\mathbb{Q}_n$:

$$\frac{d\mathbb{Q}_n}{d\mathbb{Q}} := \frac{1}{E[\varphi_n]} \varphi_n, \quad n = 2, 3, \ldots.$$

Puesto que $0 < \varphi_n \leq 1$, entonces $\mathbb{Q}_n \in \mathcal{Q}$. Así, por la condición (4.8),

$$0 \leq \xi \cdot E_{Q_n}[Y] = \frac{1}{E[\varphi_n]} E[\xi \cdot Y \varphi_n]. \tag{4.9}$$

Aplicando el *teorema de convergencia dominada de Lebesgue*,

$$E\Big[\xi \cdot Y I_{\{\xi \cdot Y < 0\}}\Big] = \lim_{n \uparrow \infty} E\Big[\xi \cdot Y \varphi_n\Big] \geq 0.$$

Lo anterior demuestra que $\mathbb{P}(A) = 0$, y entonces se concluye que $\xi \cdot Y \geq 0$, P-c.s.. Por lo tanto, $0 \in \mathcal{C}$.

Para el caso en que Y no es integrable con respecto a la probabilidad $\mathbb{P}$, ésta probabilidad se reemplaza por otra probabilidad apropiada $\widetilde{\mathbb{P}}$, siempre y cuando $\widetilde{\mathbb{P}} \approx \mathbb{P}$, $\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}$ sea acotada, y $E_{\widetilde{\mathbb{P}}}[|Y|] < \infty$. Por ejemplo, defínase $\tilde{\mathbb{P}}$ mediante

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = \frac{c}{1 + |Y|} \text{ para } c := \left(E\left[\frac{1}{1 + |Y|}\right]\right)^{-1}. \tag{4.10}$$

Así, la primera parte de la prueba proporciona una medida martingala $\mathbb{P}^*$ equivalente a $\widetilde{\mathbb{P}}$ cuya densidad $\frac{d\mathbb{P}^*}{d\widetilde{\mathbb{P}}}$ es acotada. Ya que $\mathbb{P}^* \in \mathcal{P}$ y

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = \frac{d\mathbb{P}^*}{d\widetilde{\mathbb{P}}} \cdot \frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}$$

es acotada, entonces $\mathbb{P}^*$ es como el deseado, y así la prueba se termina. ■

De acuerdo al FTAP, ahora ya se está en posición de afirmar que un modelo de mercado viable es aquel en el que es posible encontrar una medida martingala equivalente a la probabilidad original $\mathbb{P}$.

Sea el espacio lineal

$$\mathcal{V} := \left\{ \overline{\xi} \cdot \overline{S} \,|\overline{\xi} \in \mathbb{R}^{d+1} \right\},$$

que denota el conjunto de todos los pagos que pueden ser generados por algún portafolio.

En general, un elemento $V \in \mathcal{V}$ no es generado por un único portafolio. Pero, tal y como el siguiente resultado lo establece, la inversión inicial en cada uno de los portafolios sí coinciden.

Lema 4.2.2. *Sea un modelo de mercado libre de arbitraje y supóngase que $V \in \mathcal{V}$ puede ser escrito como*

$$V = \overline{\xi} \cdot \overline{S} = \overline{\zeta} \cdot \overline{S},\ \mathbb{P}\text{-c.s.},$$

para dos portafolios diferentes $\overline{\xi}$ y $\overline{\zeta}$. Entonces, $\overline{\pi} \cdot \overline{\xi} = \overline{\pi} \cdot \overline{\zeta}$.

Demostración.

Por la hipótesis, se cumple que $(\overline{\xi} - \overline{\zeta}) \cdot \overline{S} = 0$, $\mathbb{P}^*$-c.s., para cualquier $\mathbb{P}^* \in \mathcal{P}$. Luego,

$$\overline{\pi} \cdot \overline{\xi} - \overline{\pi} \cdot \overline{\zeta} = \overline{\pi} \cdot (\overline{\xi} - \overline{\zeta}) = E^* \left[\frac{(\overline{\xi} - \overline{\zeta}) \cdot \overline{S}}{1+r} \right] = 0.$$

■

Por el lema anterior, tiene sentido definir el *precio* de $V \in \mathcal{V}$ como

$$\pi(V) := \overline{\pi} \cdot \overline{\xi} \ \text{ si } V = \overline{\xi} \cdot \overline{S},$$

siempre y cuando el modelo sea libre de arbitraje. Así, para $V \in \mathcal{V}$ y $\mathbb{P}^* \in \mathcal{P}$ arbitrario, se tiene que

$$\pi(V) = E^* \left[\frac{V}{1+r} \right].$$

Definición 4.2.3. *Supóngase que el modelo de mercado es libre de arbitraje y que $V \in \mathcal{V}$ es tal que $\pi(V) \neq 0$. Entonces, el* ***rendimiento*** *de V está definido por*

$$R(V) = \frac{V - \pi(V)}{\pi(V)}.$$

Nótese que un caso especial que ya se había mencionado es el dividendo libre de riesgo:

$$r = \frac{S^0 - \pi^0}{\pi^0} = R(S^0).$$

Teorema 4.2.2. *Supóngase que el modelo de mercado es libre de arbitraje y que $V \in \mathcal{V}$ es tal que $\pi(V) \neq 0$.*

a) *Bajo cualquier medida martingala $\mathbb{P}^*$, la esperanza del rendimiento de V es igual al rendimiento libre de riego r :*

$$E^*[R(V)] = r.$$

b) *Bajo cualquier probabilidad $\mathbb{Q} \approx \mathbb{P}$ tal que $E_{\mathbb{Q}}[|\overline{S}|] < \infty$, la esperanza del rendimiento de V está dado por*

$$E_{\mathbb{Q}}[R(V)] = r - cov_{\mathbb{Q}}\left(\frac{d\mathbb{P}^*}{d\mathbb{Q}}, R(V)\right),$$

donde $\mathbb{P}^*$ es una medida martingala arbitraria en $\mathcal{P}$ y $cov_{\mathbb{Q}}$ denota la covarianza con respecto a $\mathbb{Q}$.

Demostración.

a) La demostración se sigue de la siguiente cadena de igualdades:

$$\begin{aligned} E^*[R(V)] &= E^*\left[\frac{V - \pi(V)}{\pi(V)}\right] \\ &= \frac{E^*[V] - \pi(V)}{\pi(V)} \\ &= \frac{\pi(V)(1+r) - \pi(V)}{\pi(V)} \\ &= r. \end{aligned}$$

b) Sean $\mathbb{P}^* \in \mathcal{P}$ y $\varphi^* := \frac{d\mathbb{P}^*}{d\mathbb{P}}$. Entonces,

$$\begin{aligned} cov_{\mathbb{Q}}(\varphi^*, R(V)) &= E_{\mathbb{Q}}[\varphi^* R(V)] - E_{\mathbb{P}}[\varphi^*]E_{\mathbb{Q}}[R(V)] \\ &= E^*[R(V)] - 1 \cdot E_Q[R(V)] \\ &= r - E_{\mathbb{Q}}[R(V)]. \end{aligned}$$

■

4.2.2. Derivados financieros

Por el momento ya se ha analizado el comportamiento de mercados en donde los activos primarios son los que se comercializan. Generalmente, a tal mercado se le conoce como ***mercado de valores***, y aquí el interés es identificar aquellos modelos "correctos" que permitan analizar el comportamiento de los precios. Sin embargo, como ya se ha mencionado, en los mercados actuales no solo los activos primarios son comerciables, también hay una gran variedad de títulos o instrumentos financieros cuyo valor depende de los activos primarios, y en la mayoría de las veces también de otros factores. A tal mercado se le conoce como ***mercado de derivados financieros***.

Ejemplo 4.3. *En un contrato **forward**, al tiempo $t = 0$, un hedgers conviene en vender el i-ésimo activo en el tiempo $T = 1$ a un especulador, por un* precio de entrega *(*delivery price*) K. El valor de un contrato forward está expresado de forma aleatoria como*

$$C^{fw} = S^i - K.$$

Si $S^i > K$, el tenedor del contrato forward gana la diferencia $S^i - K$, en caso contrario, pierde la cantidad $K - S^i$. □

Ejemplo 4.4. *El tenedor de una* **opción de compra** *sobre el i-ésimo activo tiene el derecho de comprar el i-ésimo activo al tiempo de ejercicio $T = 1$, a un* **precio de ejercicio** *K. Su correspondiente valor está dado por*

$$C^{call} = (S^i - K)^+ = \begin{cases} S^i - K & si\ S^i > K, \\ 0 & si\ S^i \leq K. \end{cases}$$

Por otro lado, una **opción de venta** *da el derecho de vender el i-ésimo activo al tiempo de ejercicio $T = 1$, y a un precio de ejercicio K. En este caso, su correspondiente valor está dado por*

$$C^{put} = (K - S^i)^+ = \begin{cases} K - S^i & si\ S^i < K, \\ 0 & si\ S^i \geq K. \end{cases}$$

Opciones de compra y venta con el mismo precio de ejercicio K y escritos sobre el mismo activo subyacente están relacionados por la formula

$$C^{call} - C^{put} = S^i - K.$$

□

Derivados (*derivative securities, contingent claims*) como el de los ejemplos anteriores son los que se encuentran en el mercado de derivados. Un derivado financiero será interpretado como un contrato que es vendido al tiempo $t = 0$ y el cual paga una cantidad aleatoria $C(\omega) \geq 0$ al tiempo $T = 1$.

Matemáticamente, es conveniente centrarse sólo en derivados cuyo valor sea no negativo. Un derivado cuyo valor terminal pueda llegar a ser también negativo puede usualmente ser reducido de una combinación de derivados no negativos y de posiciones cortas de algunos de los activos primarios $S^0, S^1, \ldots, S^d$.

A continuación se define de manera formal el concepto de derivado financiero.

Definición 4.2.4. *Un **derivado** es una variable aleatoria C en el espacio de probabilidad subyacente $(\Omega, \mathcal{F}, \mathbb{P})$ tal que*

$$0 \leq C < \infty,\ \mathbb{P}\text{-}c.s..$$

Nótese que C depende de la información que describe los precios de los activos primarios. Se dice que C es un derivado de los activos primarios $S^0, \ldots, S^d$ si se cumple que

$$C = f(S^0, \ldots, S^d),$$

donde f es una función medible en $\mathbb{R}^{d+1}$.

Ahora el principal problema es determinar el precio "justo" de un determinado derivado al tiempo $t = 0$. Para caracterizar a este precio, C será considerado como un nuevo activo en el portafolio $\overline{\xi}$. Así,

$$\pi^{d+1} := \pi^C, \quad S^{d+1} := C, \tag{4.11}$$

donde π^C representará el precio inicial del derivado C.

Definición 4.2.5. *Un número real $\pi^C \geq 0$ es llamado* **precio libre de arbitraje** *del derivado C si el modelo de mercado, extendido de acuerdo a (4.11), es libre de arbitraje. El conjunto de todos los precios libres de arbitraje para C es denotado por $\Pi(C)$.*

El siguiente resultado establece que siempre es posible encontrar un precio libre de arbitraje si el modelo inicial es libre de arbitraje.

Teorema 4.2.3. *Supóngase que $\mathcal{P} \neq \emptyset$. Entonces el conjunto de precios libres de arbitraje de un derivado C es no vacío y está dado por*

$$\Pi(C) = \left\{ E^*\left[\frac{C}{1+r}\right] \middle| P^* \in \mathcal{P} \text{ tal que } E^*[C] < \infty \right\}. \tag{4.12}$$

Demostración.

Primero se mostrará que $\Pi(C)$ es no vació. Para ello, se fija una probabilidad $\widetilde{\mathbb{P}} \approx \mathbb{P}$ tal que $E_{\widetilde{\mathbb{P}}}[C] < \infty$ (por ejemplo, la probabilidad puede ser definida mediante (4.10)).

Bajo la medida $\widetilde{\mathbb{P}}$, el modelo de mercado es libre de arbitraje. Luego, por el FTAP, existe $\mathbb{P}^* \in \mathcal{P}$ tal que $\frac{d\mathbb{P}^*}{d\widetilde{P}}$ es acotado.
Así, por construcción $E^*[C] < \infty$, y

$$\pi^C = E^*\left[\frac{C}{1+r}\right] \in \Pi(C),$$

por definición de $\mathcal{P}$. Por lo tanto, $\Pi(C) \neq \emptyset$.

Ahora, por el FTAP, π^C es un precio libre de arbitraje para C sí y sólo si existe una medida martingala $\widehat{\mathbb{P}}$ para el modelo de mercado extendido de acuerdo a (4.11), tal que

$$\pi^i = \hat{E}\Big[\frac{S^i}{1+r}\Big] \text{ para } i = 1, \ldots, d+1.$$

En particular, $\widehat{\mathbb{P}}$ pertenece a $\mathcal{P}$, y entonces se cumple la contención $\subseteq$ en (4.12).

Por otro lado, si $\pi^C = E^*\left[\frac{C}{1+r}\right]$ para alguna $\mathbb{P}^* \in \mathcal{P}$, entonces $\mathbb{P}^*$ es una medida martingala para el modelo de mercado extendido.

Por todo lo anterior, se cumple la igualdad de conjuntos en (4.12). ∎

Por la convexidad de $\mathcal{P}$ se sigue que $\Pi(C)$ también es convexo, esto es, o existe un único precio o una infinidad de ellos, pero no ambos, en otras palabras, $\Pi(C)$ es un intervalo.

En el siguiente resultado se especifica las cotas de los precios de un derivado, en donde Y es el definido en (4.6).

Teorema 4.2.4. *En un modelo libre de arbitraje, las cotas de arbitraje de un derivado C están dadas por*

$$\begin{aligned}
\pi_{\text{ínf}}(C) &= \inf_{\mathbb{P}^* \in \mathcal{P}} E^*\left[\frac{C}{1+r}\right] \\
&= \text{máx}\Big\{m \in [0,\infty) \,|\, \exists\, \xi \in \mathbb{R}^d \textit{ con } m + \xi \cdot Y \leq \frac{C}{1+r},\ \mathbb{P}\text{-}c.s.\Big\},
\end{aligned}$$

y

$$\begin{aligned}
\pi_{\text{sup}}(C) &= \sup_{\mathbb{P}^* \in \mathcal{P}} E^*\left[\frac{C}{1+r}\right] \\
&= \text{mín}\Big\{m \in [0,\infty] \,|\, \exists\, \xi \in \mathbb{R}^d \textit{ con } m + \xi \cdot Y \geq \frac{C}{1+r},\ \mathbb{P}\text{-}c.s.\Big\}.
\end{aligned}$$

La demostración, cuya construcción principalmente hace uso de conceptos definidos anteriormente, se puede encontrar en Föllmer and Schied[5] (véase Teorema 1.31.).

La siguiente definición es esencial para poder asignar un precio inicial a los derivados.

Definición 4.2.6. *Se dice que un derivado C es **replicable** (*replicable, attainable, redundant*) si*

$$C = \overline{\xi} \cdot \overline{S},\ \mathbb{P}\text{-}c.s.,$$

*para algún $\overline{\xi} \in \mathbb{R}^{d+1}$. La estrategia de inversión $\overline{\xi}$ es llamada **portafolio replicante o portafolio de apertura** (*replicating portfolio, hedging portfolio*) para C.*

Si un derivado C es replicable con el portafolio $\overline{\xi}$, entonces, desde un punto de vista financiero, no hay diferencia entre el propietario del derivado y el propietario del portafolio.

Sin importar lo que ocurra en el mercado accionario, el valor del derivado al tiempo $T = 1$ será exactamente igual al valor del portafolio al tiempo $T = 1$.

Apelando al Lema 4.2.2, si se demuestra que un derivado C puede ser replicable por algún portafolio $\overline{\xi}$, entonces el problema de determinar un precio para C tiene una notable solución: *El precio de C es único e igual al costo $\overline{\xi} \cdot \overline{\pi}$ de su replicación.*

El siguiente corolario muestra que sólo los derivados replicables admiten un único precio libre de arbitraje.

Corolario 4.2.5. *Sea un modelo de mercado libre de arbitraje y C un derivado.*

a) *Si C no es replicable, entonces $\pi_{inf}(C) < \pi_{sup}(C)$ y*

$$\Pi(C) = \Big(\pi_{inf}(C), \pi_{sup}(C)\Big).$$

b) *C es replicable sí y sólo si éste admite un único precio libre de arbitraje.*

Demostración.

a) Por resultados previos se sigue que el conjunto $\Pi(C)$ es un intervalo no vacío. Resta mostrar que este intervalo no contiene las cotas $\pi_{\text{ínf}}(C)$ y $\pi_{\text{sup}}(C)$.

Por el Teorema 4.2.4, existe $\xi \in \mathbb{R}^d$ tal que

$$\pi_{\text{ínf}}(C) + \xi \cdot Y \leq \frac{C}{1+r}, \ \mathbb{P}\text{-c.s.}. \tag{4.13}$$

Puesto que C no es replicable, entonces (4.13) no puede ser una identidad, casi seguramente. De aquí que, con $\xi^0 := -(1+r)\pi_{\text{ínf}}(C)$, la estrategia de inversión $(\xi^0, -\xi, 1) \in \mathbb{R}^{d+2}$ es una oportunidad de arbitraje en el modelo de mercado extendido por

$$\pi^{d+1} := \pi_{\text{ínf}}(C) \ \text{ y } \ S^{d+1} := C.$$

En efecto, ya que el mercado es libre de arbitraje, entonces $\xi \cdot \pi > 0$. Luego,

$$\begin{aligned}
\Big(\pi^0, \pi, \pi_{\text{ínf}}(C)\Big) \cdot (\xi^0, -\xi, 1) &= -(1+r)\pi_{\text{ínf}}(C) - \xi \cdot \pi + \pi_{\text{ínf}}(C) \\
&= -r\pi_{\text{ínf}}(C) - \xi \cdot \pi \\
&< 0.
\end{aligned}$$

Por otro lado, de (4.13) se tiene que, $\mathbb{P}$-c.s.,

$$C - (1+r)\pi_{\text{ínf}}(C) - (1+r)\xi \cdot Y \geq 0;$$

$$-(1+r)\pi_{\text{ínf}}(C) - \xi \cdot S + C \geq 0;$$

$$(\xi^0, -\xi, 1) \cdot (S^0, S, S^{d+1}) \geq 0.$$

En consecuencia, $\pi_{\text{ínf}}(C)$ no es un precio libre de arbitraje para C, y entonces $\pi_{\text{ínf}}(C) \notin \Pi(C)$, i.e., es una cota inferior para $\Pi(C)$.
De forma similar se puede concluir que $\pi_{\text{sup}} \notin \Pi(C)$. Por lo tanto, $\Pi(C) = \Big(\pi_{inf}(C), \pi_{sup}(C)\Big)$.

b) Se sigue del inciso anterior y del hecho de que si C es replicable, entonces $|\Pi(C)| = 1$.

■

Ahora se caracterizará aquellas situaciones en las cuáles todos los derivados son replicables.

Definición 4.2.7. *Se dice que un modelo de mercado libre de arbitraje es* ***completo*** *si cada derivado es replicable. En otro caso, se dice que el modelo de mercado es* ***incompleto****.*

El interés de los mercados completos se debe a que permite derivar una teoría de valuación y apertura con derivados. El modelo de Cox-Ross-Rubinstein es un ejemplo simple de un modelo de mercado completo, y éste es una versión a tiempo discreto del modelo de Black-Scholes (véase Lamberton[13], p. 12-16).

El siguiente resultado da una caracterización precisa de los modelos completos.

Teorema 4.2.6. *Un modelo de mercado libre de arbitraje es completo sí y sólo si existe exactamente una medida martingala, i.e., $|\mathcal{P}| = 1$.*
Demostración.
Supóngase que el modelo es completo. Entonces para cada $A \in \mathcal{F}$ se cumple que I_A es un derivado replicable. Luego, por el corolario anterior, I_A admite un único precio libre de arbitraje, lo cual implica que $\mathbb{P}^*(A) = E^*[I_A]$ es independiente de $\mathbb{P}^* \in \mathcal{P}$. Por lo tanto, existe solamente una medida martingala.

Recíprocamente, supóngase que $\mathcal{P} = \{\mathbb{P}^*\}$, y sea C un derivado acotado, esto es, $E^*[C] < \infty$. Entonces C tiene un único precio libre de arbitraje $E^*\left[\frac{C}{1+r}\right]$. Así, por el corolario anterior, C es replicable. ■

Nótese que una consecuencia del teorema anterior es que todos los derivados en los mercados incompletos tienen una infinidad de precios iniciales libres de arbitraje, y esto se debe a que los modelos admiten una infinidad de medidas martingalas.

Ejemplo 4.5. *Para efectos ilustrativos, el modelo que aquí se analizará representa la esencia del modelo binomial de un periodo. Considérese la situación en donde el espacio muestral Ω consiste solamente de dos posibles escenarios: ω^+ y ω^-; y la probabilidad $\mathbb{P}$, es tal que*

$$p := \mathbb{P}(\{\omega^+\}) \in (0, 1).$$

Se asumirá que existe solamente un activo con riesgo con valor π *al tiempo* $t = 0$*, y que al tiempo* $t = 1$ *toma dos posibles valores:* a *con probabilidad* $1 - p$*, y* b *con probabilidad* p*; donde,* a *y* b *son tales que* $0 \leq a < b$*. Esto es,*

$$\mathbb{P}(S(\omega^-) = a) = \mathbb{P}(\{\omega^-\}) = 1 - p,$$

$$\mathbb{P}(S(\omega^+) = b) = \mathbb{P}(\{\omega^+\}) = p.$$

De acuerdo a resultados previos, este modelo es libre de arbitraje si y sólo si

$$\pi(1 + r) \in \left\{E_{\widetilde{\mathbb{P}}}[S] \mid \widetilde{\mathbb{P}} \approx \mathbb{P}\right\} = \left\{\widetilde{p}b + (1 - \widetilde{p})a \mid \widetilde{p} \in (0, 1)\right\} = (a, b).$$

En este caso, el modelo también es completo. En efecto, cualquier medida martingala $\mathbb{P}^*$ *debe satisfacer*

$$\pi(1 + r) = E^*[S] = p^* b + (1 - p^*)a,$$

y esta condición define, de forma única, el parámetro $p^* = \mathbb{P}(\{w^+\})$ *como*

$$p^* = \frac{\pi(1 + r) - a}{b - a} \in (0, 1).$$

De aquí que, $|\mathcal{P}| = 1$*. En consecuencia, por el Teorema 4.2.6, el modelo es completo.*

Ya que el modelo es completo, entonces todo derivado C *es replicable. Luego,*

$$C(\omega) = \xi^0 S^0(\omega) + \xi S(\omega) = \xi^0(1 + r) + \xi S(\omega), \qquad \text{para todo } \omega \in \Omega.$$

Lo anterior determina un sistema de ecuaciones lineales para las variables ξ^0 *y* ξ*, cuya solución está dada por*

$$\xi^0 = \frac{C(\omega^-)b - C(\omega^+)a}{(b - a)(1 + r)}, \quad \xi = \frac{C(\omega^+) - C(\omega^-)}{b - a}.$$

Por lo tanto, el único precio libre de arbitraje para C *es*

$$\begin{aligned}
\pi(C) &= \overline{\pi} \cdot \overline{\xi} \\
&= \frac{C(\omega^-)b - C(\omega^+)a}{(b - a)(1 + r)} + \pi \frac{C(\omega^+) - C(\omega^-)}{b - a} \\
&= \frac{C(\omega^+)}{1 + r} . \frac{\pi(1 + r) - a}{b - a} + \frac{C(\omega^-)}{1 + r} . \frac{b - \pi(1 + r)}{b - a}.
\end{aligned}$$

De esta última igualdad, nótese que $\pi(C) = E^*\left[\frac{C}{1+r}\right]$*, tal y como en su momento ya se había hecho notar. Además, en términos del proceso de valor asociado a la estrategia de inversión* $\overline{\xi} = (\xi^0, \xi)$*,*

$$V_0^{\overline{\xi}} = \overline{\pi} \cdot \overline{\xi} = E^*\left[\frac{C}{1 + r}\right] = E^*\left[\frac{V_1^{\overline{\xi}}}{1 + r}\right].$$

La probabilidad p^* es una "probabilidad artificial" que sólo es usada en la valuación del derivado. Éste no es la probabilidad real que mide el comportamiento de los precios del activo.

Para el caso de una opción de compra $C = (S - k)^+$ con precio de ejercicio $K \in [a, b]$, se sigue que,

$$\begin{aligned} \pi\Big((S-K)^+\Big) &= \frac{b-K}{1+r}\frac{\pi(1+r)-a}{b-a} + \frac{0}{1+r}\frac{b-\pi(1+r)}{b-a} \\ &= \frac{b-K}{b-a}\pi - \frac{(b-K)a}{b-a}\frac{1}{1+r}. \end{aligned} \tag{4.14}$$

Como caso particular, supóngase que el activo con riesgo puede ser comprado al tiempo $t = 0$ por el precio $\pi = 100$. Al tiempo $t = 1$, el precio o es $S(\omega^+) = b = 120$ o $S(\omega^-) = a = 90$. Si se invierte en el activo con riesgo, los correspondientes dividendos están dados por

$$R(S)(\omega^+) = 20\,\% \quad o \quad R(S)(\omega^-) = -10\,\%. \tag{4.15}$$

Ahora supóngase que se invierte en una opción de compra $C := (S - K)^+$ con precio de ejercicio $K = 100$. Eligiendo $r = 0$, por la fórmula (4.14), el precio de la opción es

$$\pi(C) = \frac{20}{3} \approx 6.67.$$

Para la inversión inicial $\pi(C)$, el rendimiento del derivado C está expresado como

$$R(C) = \frac{(S-K)^+ - \pi(C)}{\pi(C)}.$$

De aquí que, al tiempo $t = 1$,

$$R(C)(\omega^+) = \frac{20-\pi(C)}{\pi(C)} = 200\,\% \quad o \quad R(C)(\omega^-) = \frac{0-\pi(C)}{\pi(C)} = -100\,\%. \tag{4.16}$$

Claramente, en las expresiones (4.15) y (4.16) se puede notar el gran beneficio que se obtienen al invertir en un derivado, así como también el gran riesgo al que se está sujeto, lo cual era de esperarse, pues fue la primera afirmación que se resaltó al inicio de ésta sección. Al incremento dramático tanto de la oportunidad de ganancia como del riesgo se le conoce como efecto de apalancamiento *(*leverage effect*) de las opciones.*
Estas características, principalmente, determinan el gran interés que los derivados financieros están adquiriendo en la actualidad □

Finalmente, para motivar el lector interesado en la aplicación de la teoría de análisis convexo, los resultados principales que aquí se plantearon pueden ser tratados mediante caracterizaciones geométricas.

4.3. Teoría de valuación en tiempo continuo

Ya que se ha finalizado el análisis de los modelos de un sólo paso, ahora lo ideal sería analizar los *modelos de mercado de varios pasos (o modelos de mercado de varios periodos)*, i.e., modelos en donde los activos son valuados en los tiempos $t = 0, 1, 2, \ldots, T$. En este caso, el uso de un espacio de probabilidad filtrado y de los procesos estocásticos se vuelven relevantes.

La importancia de estudiar los modelos de mercados en varios periodos es que permite introducir una gran variedad de derivados financieros. El análisis realizado para el caso de un periodo facilita abordar la teoría para el caso de varios periodos, pues se siguen cumpliendo los mismos resultados.

En esta sección se establecerá los principales resultados que sustentan a los modelos de mercado en tiempo continuo, i.e., modelos en un horizonte finito $0 \leq t \leq T$. Análogo a la sección anterior, estos resultados serán aplicados en la valuación de derivados financieros.

4.3.1. Portafolios dinámicos

Considérese un mercado financiero, en donde el tiempo es dividido en periodos de longitud Δt, y en donde la inversión sólo se lleva a cabo en los puntos discretos $n\Delta t, n = 0, 1, \ldots$. Por "periodo t" se estará refiriendo a un periodo fijo $[t, t+\Delta t)$, en donde $t = n\Delta t$ para algún n.

Para comodidad, ahora se introducen las siguientes notaciones:

$$
\begin{aligned}
d &= \text{número de los diferentes tipos de activos.} \\
\xi_t^i &= \text{número de acciones de tipo } i \text{ durante el periodo } [t, t+\Delta t). \\
\xi_t &= \text{portafolio } (\xi_t^1, \ldots, \xi_t^d) \text{ durante el periodo } t. \\
c_t &= \text{cantidad de capital gastado en consumo por unidad de tiempo} \\
& \text{ durante el periodo } [t, t+\Delta t). \\
S_t^i &= \text{precio de una acción de tipo } i \text{ durante el periodo } [t, t+\Delta t). \\
V_t &= \text{valor del portafolio } \xi_t.
\end{aligned}
$$

Las decisiones y la información en el modelo son estructuradas como sigue:

- Al tiempo t, i.e., al inicio del periodo t, se tiene en posesión un portafolio "viejo" $\xi_{t-\Delta t}$ que fue elegido en el periodo previo $t - \Delta t$.
- Al tiempo t se observa el vector de precio $S_t = (S_t^1, \ldots, S_t^d)$.
- Al tiempo t, después de haber observado S_t, se elige un nuevo portafolio ξ_t que será retenido durante el periodo t. Al mismo tiempo se elige la tasa de consumo c_t para el periodo t (puede ser visto como el porcentaje de la fortuna que será retirado de la inversión). Tanto ξ_t como c_t serán considerados constantes en el periodo t.

Obsérvese que, de acuerdo a las suposiciones anteriores, la fortuna del inversionista al inicio del periodo t está dada por

$$V_t = \xi_{t-\Delta t} S_t := \xi_{t-\Delta t} \cdot S_t. \tag{4.17}$$

La igualdad anterior expresa que, al inicio del periodo t, la fortuna es igual a lo que se obtiene si se vende el portafolio, el cual se había formado en el periodo anterior, a un precio actual. Lo que se obtiene de esta venta puede ser usado para dos propósitos:

- Reinvertir en un nuevo portafolio ξ_t.
- Consumir la tasa c_t sobre el periodo t.

El costo del nuevo portafolio ξ_t, el cual ha sido comprado al precio actual, está dado por $\xi_t S_t$, mientras que el costo para la tasa de consumo c_t está dado por $c_t \Delta t$. Así, la ecuación presupuestaria (*budget equation*) para el periodo t está dado como sigue:

$$\xi_{t-\Delta t} S_t = \xi_t S_t + c_t \Delta t,$$

de donde,

$$S_t(\xi_t - \xi_{t-\Delta t}) + c_t \Delta t = 0;$$

$$S_t \Delta \xi_t + c_t \Delta t = 0, \quad \Delta \xi_t := \xi_t - \xi_{t-\Delta t}. \tag{4.18}$$

Puesto que el objetivo es obtener una ecuación presupuestaria en tiempo continuo, nótese que cuando $\Delta t \to 0$ se obtiene la expresión formal

$$S_t d\xi_t + c_t dt = 0.$$

Si se estuviera trabajando con funciones deterministas, el problema ya está resuelto. Pero en este caso, ***ésta expresión es incorrecta***, ya que las diferenciales estocásticas son interpretadas en el sentido de la integral de Itô, y ésta integral está definida como el límite de sumas cuyos términos son de la forma $H_{t_n}(W_{t_{n+1}} - W_{t_n})$, y no con términos de la forma $H_{t_{n+1}}(W_{t_{n+1}} - W_{t_n})$, que es como aparece en (4.18).
Para remediar este falla, a la ecuación (4.18) se le agrega un término 0:

$$S_{t-\Delta t} \Delta \xi_t - S_{t-\Delta t} \Delta \xi_t + S_t \Delta \xi_t + c_t \Delta t = 0.$$

Así, la ecuación anterior se transforma en

$$S_{t-\Delta t} \Delta \xi_t + \Delta S_t \Delta \xi_t + c_t \Delta t = 0,$$

ahora, cuando $\Delta t \to 0$, se tiene

$$S_t d\xi_t + dS_t d\xi_t + c_t dt = 0. \tag{4.19}$$

La ecuación anterior representa la *ecuación presupuestaria* que se cumple para todos los *portafolios autofinanciables* (esto son, como su nombre lo indica, portafolios que no necesitan capital extra para reajustarlos a otro portafolio).

Por otro lado, cuando en (4.17) se toma $\Delta t \to 0$, entonces se tiene

$$V_t = \xi_t S_t.$$

Al aplicar la fórmula de integración por partes, se obtiene la *ecuación general para la dinámica de un portafolio arbitrario*:

$$dV_t = \xi_t dS_t + S_t d\xi_t + dS_t d\xi_t. \tag{4.20}$$

Considerando la ecuación (4.19) y la ecuación (4.20) se obtiene la **dinámica de un portafolio autofinanciable**:

$$dV_t = \xi_t dS_t - c_t dt.$$

En particular, en una situación donde no existe algún consumo, $c_t = 0$, se cumple la dinámica

$$dV_t = \xi_t dS_t. \tag{4.21}$$

Se resalta el hecho de que la interpretación económica de la ecuación (4.21) es, por supuesto, en un modelo sin insumos externos, todos los cambios del valor son consecuencia de los cambios en los precios de los activos.

Ahora se definirá formalmente los conceptos centrales que se abarcaron en el bosquejo anterior.

Definición 4.3.1. *Sea un proceso de precios d-dimensional (S_t) dado.*

a) *Una **estrategia de inversión** (*portafolio estratégico*, o simplemente* portafolio*) es cualquier proceso d-dimensional $\mathcal{F}_t^S$-adaptado*

$$\xi := (\xi_t).$$

b) *El **proceso de valor** (V_t^ξ) **asociado** al portafolio $\xi := (\xi_t)$ está definido por*

$$V_t^\xi = \sum_{i=1}^{d} \xi_t^i S_t^i.$$

c) *Un **proceso de consumo** es cualquier proceso unidimensional $\mathcal{F}_t^S$-adaptado*

$$c := (c_t).$$

d) *Se dice que el par* (ξ, c) *es un* ***portafolio de consumo autofinanciable*** *(portfolio-consumption self-financing) si el proceso de valor* (V_t^ξ) *satisface la condición*

$$dV_t^\xi = \sum_{i=1}^{d} \xi_t^i dS_t^i - c_t dt.$$

Si $c = 0$, *se dice simplemente que el portafolio* ξ *es autofinanciable.*

Nótese que, en general, al portafolio ξ_t se le permite depender de las trayectorias pasadas de los precios $\{S_u, u \leq t\}$.

4.3.2. Arbitraje en modelos a tiempo continuo

Ya que otro de los objetivos principales es el análisis del modelo de Black-Scholes, aquí sólo se analizará un modelo de mercado con dos activos: un activo libre de riesgo con proceso de precios (β_t) y un activo con riesgo cuyo proceso de precios será denotado por (S_t). También se asumirá que $c_t = 0$.

A partir de ahora, por ξ se entenderá de que se trata del portafolio (ξ_t) tal que al tiempo t está definido por

$$\xi_t = (a_t, b_t),$$

en donde a_t denota el número de acciones del activo con riesgo al tiempo t, y b_t denota el número de unidades del activo sin riesgo al tiempo t. Se asumirá que los procesos son tales que las integrales $\int_0^t a_s dS_s$ y $\int_0^t b_s d\beta_s$ existen.

Definición 4.3.2. *Se define el* ***proceso de ganancias*** *para la estrategia de inversión* ξ *como*

$$G_t^\xi = \int_0^t a_s dS_s + \int_0^t b_s d\beta_s.$$

Bajo este concepto, la estrategia de inversión ξ *es* ***autofinanciable*** *si satisface*

$$V_t^\xi = V_0^\xi + G_t^\xi,$$

i.e.,

$$dV_t^\xi = a_t dS_t + b_t d\beta_t.$$

Cuando la estrategia de inversión ξ *es autofinanciable y satisface la restricción de liquidez:*

$$V_t^\xi \geq 0,$$

entonces se dice que la estrategia es ***admisible.***

Puesto que uno de los activos que se está considerando es libre de riesgo, a continuación se define su dinámica.

Definición 4.3.3. *El proceso de precios* (β_t) *es el precio de un activo* ***libre de riesgo*** *si éste tiene la dinámica*

$$d\beta_t = r(t)\beta_t dt,$$

donde $r(t)$ *define a un proceso adaptado* r.

La característica peculiar de un activo libre de riesgo es que ***no es dirigido por los términos del movimiento Browniano:*** dW. De hecho, la dinámica de (β_t) también puede ser escrita como

$$\frac{d\beta_t}{\beta_t} = r(t)dt,$$

de aquí que (β_t) está definida por

$$\beta_t = \beta_0 \exp\left(\int_0^t r(s)ds\right).$$

Un caso importante y especial aparece cuando r es una constante, en este caso el proceso de precios es una función determinista que puede ser interpretada como el precio de un bono.

Por lo anterior, a partir de ahora se asumirá que (β_t) es continuo, es de variación finita, y $\beta_0 = 1$.

En lo sucesivo, por *proceso de precios descontados* se estará refiriendo al proceso $\left(\widetilde{S}_t\right)$, el cual es definido, al tiempo t, mediante la relación

$$\widetilde{S}_t = \frac{S_t}{\beta_t}.$$

Análogamente, por *proceso de valores descontados* se estará refiriendo al proceso $\left(\widetilde{V}_t^\xi\right)$, el cual es definido por

$$\widetilde{V}_t^\xi = \frac{V_t^\xi}{\beta_t}.$$

Teorema 4.3.1. *La estrategia de inversión* ξ *es autofinanciable sí y sólo si*

$$\widetilde{V}_t^\xi = V_0^\xi + \int_0^t a_u d\widetilde{S}_u. \tag{4.22}$$

Demostración.
Primero se demostrará la necesidad.

Supóngase que el portafolio es autofinanciable. Ya que (β_t) es continua y de variación finita,

$$\begin{aligned}
d\widetilde{V}_t^{\xi} &= \frac{1}{\beta_t} dV_t^{\xi} + V_t^{\xi} d\left(\frac{1}{\beta_t}\right) + dV_t^{\xi} d\left(\frac{1}{\beta_t}\right) \\
&= \frac{1}{\beta_t} dV_t^{\xi} + V_t^{\xi} d\left(\frac{1}{\beta_t}\right) \qquad (4.23) \\
&= \frac{1}{\beta_t}\Big(a_t dS_t + b_t d\beta_t\Big) + \Big(a_t S_t + b_t \beta_t\Big) d\left(\frac{1}{\beta_t}\right) \\
&= a_t\left(\frac{dS_t}{\beta_t} + S_t d\left(\frac{1}{\beta_t}\right)\right) + b_t\left(\frac{1}{\beta_t} d\beta_t + \beta_t d\left(\frac{1}{\beta_t}\right)\right) \\
&= a_t\left(\frac{dS_t}{\beta_t} + S_t d\left(\frac{1}{\beta_t}\right)\right) + b_t d\left(\frac{\beta_t}{\beta_t}\right) \qquad (4.24) \\
&= a_t d\left(\frac{S_t}{\beta_t}\right) \\
&= a_t d\widetilde{S}_t,
\end{aligned}$$

de donde, $d\widetilde{V}_t^{\xi} = a_t d\widetilde{S}_t$, i.e., se cumple la condición (4.22).

Ahora se demostrará la suficiencia.

Puesto que

$$V_t^{\xi} = a_t S_t + b_t \beta_t,$$

entonces $b_t = \widetilde{V}_t^{\xi} - a_t \widetilde{S}_t$. Luego, utilizando la condición (4.22) se obtiene

$$b_t = V_0^{\xi} + \int_0^t a_u d\widetilde{S}_u - a_t \widetilde{S}_t,$$

y por otro lado,

$$V_t^{\xi} = V_0^{\xi} \beta_t + \beta_t \int_0^t a_u d\widetilde{S}_u.$$

En consecuencia,

$$\begin{aligned}
dV_t^{\xi} &= V_0^{\xi} d\beta_t + d\beta_t \int_0^t a_u d\widetilde{S}_u + \beta_t a_t d\widetilde{S}_t \\
&= V_0^{\xi} d\beta_t + d\beta_t \Big(b_t - V_0^{\xi} + a_t \widetilde{S}_t\Big) + \beta_t a_t d\widetilde{S}_t \\
&= b_t d\beta_t + a_t \widetilde{S}_t d\beta_t + \beta_t a_t d\widetilde{S}_t \\
&= b_t d\beta_t + a_t d\Big(\widetilde{S}_t \beta_t\Big) \\
&= b_t d\beta_t + a_t dS_t.
\end{aligned}$$

Por lo tanto, el portafolio ξ es autofinanciable.

■

En el resultado anterior se puede apreciar la relevancia de la propiedad de variación finita, ya que permite aplicar el siguiente enunciado:

Si (X_t) y (Y_t) son procesos de Itô y (X_t) es de variación finita, entonces $dY_t dX_t = 0$.
La aplicación de éste enunciado justifica las igualdades dadas en (4.23) y (4.24), conclusiones que ya se esperaban como consecuencia directa de las reglas básicas de diferenciación estocástica (véase Sección 2.4).

Un caso particular del teorema anterior es cuando $\beta_t = \exp(rt)$, para r constante. De hecho, para este caso, la necesidad se prueba como sigue:

$$
\begin{aligned}
d\widetilde{V}_t^{\,\xi} &= d\Big(\exp(-rt)V_t^{\xi}\Big) \\
&= -r\exp(-rt)V_t^{\xi}dt + \exp(-rt)dV_t^{\xi} \\
&= -r\exp(-rt)\Big(a_t S_t + b_t \exp(rt)\Big)dt + \exp(-rt)a_t dS_t + \exp(-rt)b_t d(\exp(rt)) \\
&= a_t\Big(-r\exp(-rt)S_t dt + \exp(-rt)dS_t\Big) \\
&= a_t d\widetilde{S}_t,
\end{aligned}
$$

y así se prueba la condición (4.22). Probar la validez de la suficiencia no es complicado.

Definición 4.3.4. *Una **oportunidad de arbitraje** es una estrategia de inversión admisible ξ tal que $V_0^{\xi} = 0$, pero $E\Big[V_T^{\xi}\Big] > 0$.*

Ya que la estrategia es admisible, entonces $V_T^{\xi} \geq 0$, i.e., $\mathbb{P}\Big(V_T^{\xi} \geq 0\Big) = 1$. Luego, la condición $E\Big[V_T^{\xi}\Big] > 0$ es equivalente a $\mathbb{P}\Big(V_T^{\xi} > 0\Big) > 0$.

Análogo al caso discreto, la posibilidad de arbitraje se traduce en la posibilidad de acumular un capital positivo sin haber tomado algún tipo de riesgo. Y bajo la hipótesis de que el mercado es eficiente, éstas situaciones no deberían de existir. Por tal motivo, se tiene especial interés en buscar modelos que no admitan arbitraje.

Teorema 4.3.2. *Supóngase que existe una probabilidad $\mathbb{P}^*$ tal que el proceso $\Big(\widetilde{S}_t\Big)$ es una $\mathbb{P}^*$-martingala. Entonces para cualquier estrategia de inversión admisible ξ, el proceso $\Big(\widetilde{V}_t^{\xi}\Big)$ es también una $\mathbb{P}^*$-martingala.*

La probabilidad $\mathbb{P}^*$ se llama **medida martingala**, **medida martingala equivalente**, **medida de riesgo neutral o medida de probabilidad de riesgo neutral**.

La prueba se vale del resultado que garantiza que cuando $\Big(\widetilde{S}_t\Big)$ es $\mathbb{P}^*$-martingala, entonces la integral estocástica con integrador $\widetilde{S}_t$ es también una $\mathbb{P}^*$-martingala. Luego, la propiedad de martingala de $\Big(\widetilde{V}_t^{\xi}\Big)$ se verifica como sigue: Para $0 \leq s < t$,

$$
\begin{aligned}
E^*\Big[\widetilde{V}_t^{\xi}\Big|\mathcal{F}_s\Big] &= E^*\left[V_0^{\xi} + \int_0^t a_u d\widetilde{S}_u\middle|\mathcal{F}_s\right] \\
&= V_0^{\xi} + E^*\left[\int_0^t a_u d\widetilde{S}_u\middle|\mathcal{F}_s\right]
\end{aligned}
$$

$$
\begin{aligned}
&= V_0^\xi + \int_0^s a_u d\widetilde{S}_u \\
&= \widetilde{V}_s^\xi.
\end{aligned}
$$

El resultado anterior se aproxima a la condición de no arbitraje en el mercado, ya que garantiza que si se inicia con una inversión de valor cero, entonces el valor positivo no puede crearse si las probabilidades son asignadas por $\mathbb{P}^*$, i.e.,

$$\text{Si } V_0^\xi = 0, \text{ entonces } E^*\left[V_T^\xi\right] = 0 \text{ y } \mathbb{P}^*\left(V_T^\xi = 0\right) = 1.$$

Sin embargo, en la Definición 4.3.4, la esperanza es calculada bajo la probabilidad original $\mathbb{P}$.

El siguiente resultado, conocido como **Primer Teorema Fundamental**, da una condición probabilista que asegura cuándo los modelos son libres de arbitraje.

Teorema 4.3.3. *Un modelo de mercado no admite oportunidades de arbitraje sí y sólo si existe una probabilidad $\mathbb{P}^*$, $\mathbb{P}^* \approx \mathbb{P}$, tal que el proceso $\left(\widetilde{S}_t\right)$ es una $\mathbb{P}^*$-martingala.*

Análogo al caso discreto, cuando el mercado no admite oportunidad de arbitraje se dice que el modelo es **libre de arbitraje**.

Recuérdese que un **derivado financiero** es una variable aleatoria no negativa C en $(\Omega, \mathcal{F}_T)$ ($\mathcal{F}_T$ es simplemente notación, pues en realidad coincide con $\mathcal{F}$). Este objeto matemático es una generalización de contratos sobre activos primarios, el cual paga la cantidad C al tiempo T.

Definición 4.3.5. *Un derivado C es* **replicable** *si $E[C^2] < \infty$ y si existe una estrategia de inversión admisible ξ que replica al derivado, esto es, $V_t^\xi \geq 0$ y $C = V_T^\xi$.*

Tal y como ya se esperaba, para evitar arbitraje en el modelo, los derivados deben ser valuados utilizando los portafolios que lo replican.

Teorema 4.3.4 (**Teorema Fundamental de Valuación de Activos**)**.** *Supóngase que el modelo de mercado es libre de arbitraje y que C es un derivado replicable. Entonces $\Pi(t)$, el precio libre de arbitraje de C en el tiempo $t \leq T$, está dado por V_t^ξ, el valor de un portafolio admisible replicante. Más aún,*

$$\Pi(t) = V_t^\xi = E^*\left[\frac{\beta_t}{\beta_T} C \middle| \mathcal{F}_t\right]. \tag{4.25}$$

Demostración.
Puesto que C es replicable, entonces este es replicado por una estrategia admisible con el valor del portafolio replicante V_t^ξ, para $0 \leq t \leq T$, y se cumple $C = V_t^\xi$.

Para evitar arbitraje, el precio de C en un tiempo arbitrario $t < T$ debe estar dado por el valor del portafolio V_t^ξ, de otra forma es posible obtener una ganancia mediante arbitraje.

Por otro lado, nótese que si el modelo es libre de arbitraje, entonces, por el Teorema 4.3.3, existe una medida martingala $\mathbb{P}^*$ tal que el proceso de precios descontados es una $\mathbb{P}^*$-martingala. Luego, por el Teorema 4.3.2, el proceso de valor descontado $\widetilde{V}_t^\xi$ es una $\mathbb{P}^*$-martingala. De aquí que, por la propiedad martingala,

$$\begin{aligned} \widetilde{V}_t^\xi = \frac{V_t^\xi}{\beta_t} &= E^*\left[\frac{V_T^\xi}{\beta_T}\middle|\mathcal{F}_t\right] \\ &= E^*\left[\frac{1}{\beta_T}C\middle|\mathcal{F}_t\right]. \end{aligned}$$

Por lo tanto,

$$\Pi(t) = V_t^\xi = E^*\left[\frac{\beta_t}{\beta_T}C\middle|\mathcal{F}_t\right].$$

■

De acuerdo al teorema anterior, el valor del derivado C al tiempo $t = 0$ está dado por

$$\Pi(0) = E^*\left[\frac{1}{\beta_T}C\right].$$

El siguiente resultado da una caracterización de los derivados replicables en términos del proceso de precios descontados.

Teorema 4.3.5. *Sea C un derivado integrable y sea $M(t) := E^*\left[\frac{C}{\beta_T}\middle|\mathcal{F}_t\right]$. Entonces C es replicable sí y sólo si $M(t)$ puede ser representado de la forma*

$$M(t) = M(0) + \int_0^t H_u d\widetilde{S}_u$$

para algún proceso predecible (H_t). Además, $\widetilde{V}_t^\xi = M(t)$ es el mismo para cualquier portafolio admisible que replica a C.

Un modelo de mercado libre de arbitraje, en donde todos los derivados son replicables se conoce como **completo.** En otro caso, se dice que el mercado es **incompleto.**

El siguiente resultado, conocido como **Segundo Teorema Fundamental**, caracteriza un modelo completo en términos de la medida martingala.

Teorema 4.3.6. *Las siguientes afirmaciones son equivalentes:*

a) *El modelo de mercado es completo.*

b) *La medida martingala $\mathbb{P}^*$ equivalente a $\mathbb{P}$, que hace a $\widetilde{S}_t$ una $\mathbb{P}^*$-martingala, es única.*

c) *El precio libre de arbitraje dado por (4.25) es único.*

4.4. Modelo de Black-Scholes

El modelo sugerido por Black y Scholes es el modelo de valuación de derivados comúnmente aceptado en la Industria Financiera. Este modelo asume un modelo de mercado formado por dos activos primarios. El primer activo es un bono con proceso de precios (B_t) definido por la EDO

$$\begin{cases} dB_t = rB_t dt, & r \text{ positiva y constante,} \\ B_0 = 1. \end{cases}$$

El segundo activo es un activo con riesgo cuyo proceso de precios (S_t) satisface la EDE

$$\begin{cases} dS_t = \mu S_t dt + \sigma S_t dW_t, \\ S_0 > 0, \end{cases}$$

donde la constante μ denota la media de los dividendos del activo y la constante σ la respectiva volatilidad, $\sigma \neq 0$, y W_t define a un movimiento Browniano (W_t) de valores reales.

Nótese que en el modelo de Black-Scholes se asume que la tasa de interés r y la volatilidad σ son constantes.

Por resultados previos, la dinámica de los activos están dadas de forma explícita por

$$B_t = \exp(rt), \quad S_t = S_0 exp\Bigg(\Big(\mu - \frac{\sigma^2}{2}\Big)t + \sigma W_t\Bigg).$$

Nuevamente, si a_t y b_t denotan el número de acciones y las unidades de bonos que se poseen al tiempo t, respectivamente, entonces la siguiente definición proporciona la condición para que las estrategias de inversión en el modelo de Black-Scholes estén bien definidas.

Definición 4.4.1. *Una estrategia autofinanciable está definida por un par de procesos predecibles (a_t) y (b_t) que satisfacen*

$$\int_0^T |a_t|^2 dt + \int_0^T |b_t| dt < \infty, \mathbb{P}\text{-}c.s.. \tag{4.26}$$

La condición (4.26) garantiza que las integrales que definen al proceso de ganancias está bien definida (véase Definición 4.3.2). Más aún, $\int_0^t a_s dW_s$ es una $\mathbb{P}$-martingala.

A continuación se menciona un resultado clave que permite mostrar la existencia de una medida de probabilidad equivalente a $\mathbb{P}$ bajo la cual el proceso de precios descontados $\big(\widetilde{S}_t\big)$ es una martingala.

Teorema 4.4.1 (Teorema de Transformación de Girsanov). *Sea (θ_t) un proceso adaptado que satisface $\int_0^T \theta_t^2 ds < \infty$, P-c.s. y tal que el proceso (L_t) definido por*

$$L_t = exp\Bigg(- \int_0^t \theta_s dW_s - \frac{1}{2}\int_0^t \theta_s^2 ds\Bigg),$$

es una $\mathbb{P}$-martingala. Entonces, bajo la nueva probabilidad $\mathbb{Q}$ definida en $(\Omega, \mathcal{F}_T)$ por la relación $\mathbb{Q}(A) = E_{\mathbb{P}}[1_A L_T]$, el proceso $\left(\widetilde{W}_t\right)$ definido por $\widetilde{W}_t = W_t + \int_0^t \theta_s ds$, es un movimiento Browniano estándar.

Bajo el contexto del modelo de Black-Scholes, el proceso de precios descontados cumple

$$\begin{aligned}
d\widetilde{S}_t = d\left(\frac{S_t}{\exp(rt)}\right) &= -r\exp(-rt)S_t dt + \exp(-rt)dS_t \\
&= -r\exp(-rt)S_t dt + \exp(-rt)\Big(\mu S_t dt + \sigma S_t dW_t\Big) \\
&= \widetilde{S}_t\Big((\mu - r)dt + \sigma dW_t\Big) \\
&= \sigma\widetilde{S}_t\Big(\frac{\mu - r}{\sigma}dt + dW_t\Big).
\end{aligned}$$

En consecuencia, si se considera $\widetilde{W}_t := W_t + \frac{\mu - r}{\sigma}t$, entonces el proceso de precios descontados cumple la dinámica

$$d\widetilde{S}_t = \sigma\widetilde{S}_t d\widetilde{W}_t. \tag{4.27}$$

Si se define el proceso (θ_t) mediante $\theta_t = \frac{\mu - r}{\sigma}$, entonces se cumple las hipótesis del teorema anterior, y por lo tanto $\widetilde{W}_t$ define a un movimiento Browniano estándar bajo la probabilidad $\mathbb{Q}$.

Por otro lado, de la EDE (4.27) se deduce que

$$\widetilde{S}_t = \widetilde{S}_0 \exp\Big(\sigma\widetilde{W}_t - \frac{\sigma^2}{2}t\Big).$$

Luego, por el Teorema 2.2.3 (véase Sección 2.2) se concluye que el proceso $\left(\widetilde{S}_t\right)$ es una $\mathbb{Q}$-martingala.

Todo lo anterior demuestra la existencia de la medida martingala $\mathbb{P}^*$ tal que el modelo de Black-Scholes es libre de arbitraje (véase Teorema 4.3.3). En este caso, se tiene que

$$\mathbb{P}^*(A) := E[1_A L_T],$$

donde

$$L_T = exp\Bigg(- \frac{\mu - r}{\sigma}W_T - \frac{1}{2}\frac{(\mu - r)^2}{\sigma^2}T\Bigg).$$

Tal y como ya se esperaba, bajo el contexto del modelo de Black-Scholes, el Teorema Fundamental de Valuación de Activos está expresado como sigue.

Teorema 4.4.2. *Sea $\mathbb{P}^*$ la medida martingala en el modelo de Black-Scholes. Supóngase que un derivado al tiempo T está dado por la v.a. no negativa C_T. Si*

$$E^*[C_T^2] < \infty,$$

entonces el derivado es replicable y el valor al tiempo t de un portafolio replicante está dado por

$$V_t = E^*\Big[\exp(-r(T-t))C_T \Big| \mathcal{F}_t\Big].$$

En particular, la prima de la opción es

$$V_0 = E^*[\exp(-rT)C_T].$$

En el caso de las opciones Europeas, tanto el precio de la opción como el portafolio de apertura pueden ser obtenidos explícitamente.

Teorema 4.4.3. *El valor al tiempo t de una opción Europea cuyo pago en el tiempo de ejercicio es $C_T = f(S_T)$ está dado por $V_t = F(S_t, t)$, donde*

$$F(x,t) = \exp(-r(T-t)) \int_{-\infty}^{\infty} f\Bigg(x \exp\Big(\Big(r - \frac{\sigma^2}{2}\Big)(T-t) + \sigma y\sqrt{T-t}\Big)\Bigg) \frac{\exp\Big(-\frac{y^2}{2}\Big)}{\sqrt{2\pi}} dy.$$

Demostración.
Por el Teorema 4.4.2, el valor al tiempo t es

$$V_t = E^*\Big[\exp(-r(T-t))f(S_T)\Big|\mathcal{F}_t\Big], \tag{4.28}$$

en donde la probabilidad $\mathbb{P}^*$ es tal que $\widetilde{W}_t := \frac{\mu - r}{\sigma}t$ es un movimiento Browniano y

$$d\widetilde{S}_t = \sigma \widetilde{S}_t d\widetilde{W}_t.$$

Resolviendo la ecuación anterior,

$$\begin{aligned} \widetilde{S}_T &= \widetilde{S}_0 exp\Big(\sigma\widetilde{W}_T - \frac{\sigma^2}{2}T\Big) \\ &= \widetilde{S}_t \exp\Big(\sigma\Big(\widetilde{W}_T - \widetilde{W}_t\Big) - \frac{\sigma^2}{2}(T-t)\Big). \end{aligned}$$

Luego,

$$S_T = S_t \exp(r(T-t)) exp\Big(\sigma\Big(\widetilde{W}_T - \widetilde{W}_t\Big) - \frac{\sigma^2}{2}(T-t)\Big).$$

Sustituyendo la expresión anterior en (4.28) se obtiene

$$V_t = E^*\Big[\exp(-r(T-t))f\Big(S_t \exp(r(T-t)) \exp\Big(\sigma\Big(\widetilde{W}_T - \widetilde{W}_t\Big) - \frac{\sigma^2}{2}(T-t)\Big)\Big)\Big|\mathcal{F}_t\Big].$$

Ya que la v.a. S_t es $\mathcal{F}_t$ medible y, bajo la probabilidad $\mathbb{P}^*$, $\widetilde{W}_T - \widetilde{W}_t \sim N(0, T-t)$ y es independiente de $\mathcal{F}_t$, por el Teorema 10 del Apéndice A.4 se concluye que

$$\begin{aligned} V(t) &= F(S_t, t) \\ &= \int_{-\infty}^{\infty} \exp(-r(T-t))f\Big(S_t \exp(r(T-t)) \exp\Big(\sigma z - \frac{\sigma^2}{2}(T-t)\Big)\Big) \\ &\quad \times \frac{1}{\sqrt{2\pi(T-t)}} \exp\Bigg(-\frac{z^2}{2(T-t)}\Bigg) dz \end{aligned}$$

$$= \exp(-r(T-t)) \int_{-\infty}^{\infty} f\Big(S_t \exp\Big(\Big(r - \frac{1}{2}\sigma^2\Big)(T-t) + \sigma y\sqrt{T-t}\Big)\Big)$$
$$\times \frac{1}{\sqrt{2\pi}} \exp\Big(-\frac{y^2}{2}\Big) dy.$$

■

La función $F(x,t)$ del teorema anterior puede ser calculada de forma explícita para opciones de compra y de venta Europeas.

Corolario 4.4.4. *Supóngase que se tiene una opción de compra Europea con pago al tiempo de ejercicio por $f(S_T) = (S_T - K)^+$. Entonces, escribiendo $\theta = T - t$, se tiene,*

$$F(x,t) = xN(d_1) - K\exp(-r\theta)N(d_2), \tag{4.29}$$

donde $N(x)$ es la fda de una v.a. normal estándar, i.e.,

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\Big(-\frac{y^2}{2}\Big) dy,$$

donde

$$d_1 = \frac{\ln(\frac{x}{K}) + (r + \frac{\sigma^2}{2})\theta}{\sigma\sqrt{\theta}},$$

y $d_2 = d_1 - \sigma\sqrt{\theta}$.

Demostración.

Por el teorema anterior,

$$F(x,t) = \exp(-r\theta) \int_{-\infty}^{\infty} f\left(x \exp\Big(\Big(r - \frac{\sigma^2}{2}\Big)\theta + \sigma y\sqrt{\theta}\Big)\right) \frac{\exp\Big(-\frac{y^2}{2}\Big)}{\sqrt{2\pi}} dy.$$

Luego,

$$F(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(x \exp\Big(\sigma\sqrt{\theta}y - \frac{\sigma^2}{2}\theta\Big) - K\exp(-r\theta)\right)^+ \exp\Big(-\frac{y^2}{2}\Big) dy. \tag{4.30}$$

El rango de valores de y para los cuales el integrando es no cero debe cumplir la siguiente condición:

$$x \exp\Big(\sigma\sqrt{\theta}y - \frac{\sigma^2}{2}\theta\Big) > K\exp(-r\theta),$$

el cual es equivalente a

$$y > \frac{\ln(\frac{K}{x}) + \frac{\sigma^2}{2}\theta - r\theta}{\sigma\sqrt{\theta}} = -\frac{\ln(\frac{x}{K}) + (r - \frac{\sigma^2}{2})\theta}{\sigma\sqrt{\theta}}.$$

En consecuencia, el integrando en (4.30) es no cero si $y + d_2 \geq 0$. Así,

$$
\begin{aligned}
F(x,t) &= \frac{1}{\sqrt{2\pi}}\int_{-d_2}^{\infty}\left(x\exp\Big(\sigma\sqrt{\theta}y-\frac{\sigma^2}{2}\theta\Big)-K\exp(-r\theta)\right)\exp\Big(-\frac{y^2}{2}\Big)dy \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_2}\left(x\exp\Big(-\sigma\sqrt{\theta}y-\frac{\sigma^2}{2}\theta\Big)-K\exp(-r\theta)\right)\exp\Big(-\frac{y^2}{2}\Big)dy \\
&= x\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_2}\exp\Big(-\sigma\sqrt{\theta}y-\frac{\sigma^2}{2}\theta\Big)\exp\Big(-\frac{y^2}{2}\Big)dy-K\exp(-r\theta)N(d_2) \\
&= x\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_2}\exp\Big(-\frac{(y+\sigma\sqrt{\theta})^2}{2}\Big)dy-K\exp(-r\theta)N(d_2) \\
&= xN(d_1)-K\exp(-r\theta)N(d_2).
\end{aligned}
$$

■

La función (4.29) es conocida como la **fórmula de valuación de Black-Scholes** para una opción de compra Europea. Realizando cálculos similares, el valor de una opción de venta Europea está dado por

$$F(x,t)=K\exp(-r\theta)N(-d_2)-xN(-d_1).$$

Una de las principales características de la fórmula de valuación de Black-Scholes (y una de las razones de su éxito) es que depende solamente del parámetro desconocido σ. El parámetro de tendencia μ desaparece por el cambio de probabilidad.

Ahora sólo resta mostrar cómo construir el portafolio replicante para cubrir una opción.

Teorema 4.4.5. *En la notación del Teorema 4.4.3, el proceso* (a_t,b_t) *que determina el portafolio replicante del Teorema 4.4.2 está definido por*

$$a_t=F_x(S_t,t),\qquad y\qquad b_t=\exp(-rt)(F(S_t,t)-a_tS_t).$$

Demostración.

Combinando el Teorema de Feynman-Kac (véase Apéndice A.4) y las reglas de diferenciación usuales, $F(x,t)$ satisface la siguiente EDP

$$F_t(x,t)+\frac{1}{2}\sigma^2x^2F_{xx}(x,t)-rF(x,t)+rxF_x(x,t)=0,\quad 0\le t\le T,$$

con condición de frontera dado por el valor del derivado al tiempo de ejercicio, $F(x,T)=f(x)$. Ésta ecuación es conocida como la **EDP de Black-Scholes** (o simplemente *ecuación de Black-Scholes*).

Al definir la función

$$\widetilde{F}(x,t)=\exp(-rt)F(x\exp(rt),t),$$

entonces se tiene que

$$\widetilde{V}_t=\widetilde{F}(\widetilde{S}_t,t).$$

Por la fórmula de Itô,

$$\widetilde{F}(\widetilde{S}_t, t) = \widetilde{F}(\widetilde{S}_0, 0) + \int_0^t \widetilde{F}_x(\widetilde{S}_u, u)d\widetilde{S}_u + \int_0^t \widetilde{F}_t(\widetilde{S}_u, u)du + \int_0^t \frac{1}{2}\widetilde{F}_{xx}(\widetilde{S}_u, u)d\widetilde{S}_u \cdot d\widetilde{S}_u. \quad (4.31)$$

Por otro lado, nótese que

$$\begin{aligned}
\widetilde{F}_t(x,t) &= -r\exp(-rt)F(x\exp(rt),t) + \exp(-rt)\Big(rx\exp(rt)F_{x\exp(rt)}(x\exp(rt),t) \\
&\quad +F_t(x\exp(rt),t)\Big) \\
&= \exp(-rt)\Big(- rF(x\exp(rt),t) + F_t(x\exp(rt),t) \\
&\quad +rx\exp(rt)F_{x\exp(rt)}(x\exp(rt),t)\Big) \\
&= -\frac{1}{2}\sigma^2 x^2 \widetilde{F}_{xx}(x,t),
\end{aligned}$$

$$d\widetilde{S}_t \cdot d\widetilde{S}_t = \sigma^2 \widetilde{S}_t^2 dt.$$

Luego, la ecuación (4.31) se reduce a

$$\widetilde{F}(\widetilde{S}_t, t) = \widetilde{F}(\widetilde{S}_0, 0) + \int_0^t \widetilde{F}_x(\widetilde{S}_u, u)d\widetilde{S}_u.$$

Por lo tanto, el candidato natural que define a a_t es entonces

$$a_t = \widetilde{F}_x(\widetilde{S}_t, t) = F_x(S_t, t),$$

y en consecuencia,

$$b_t = \widetilde{F}(\widetilde{S}_t, t) - a_t\widetilde{S}_t = \exp(-rt)(F(S_t,t) - a_t S_t).$$

Así, el portafolio (a_t, b_t) es autofinanciable y su valor descontado al tiempo t está definido por $\widetilde{V}_t = \widetilde{F}(\widetilde{S}_t, t)$. ■

Como aplicación directa del teorema anterior, el portafolio que replica una opción de compra Europea está dado por

$$a_t = N(d_1), \quad b_t = -K\exp(-rT)N(d_2).$$

La EDP de Black-Scholes es un ejemplo de EDP's del tipo *parabólica*. Tales EDP's pueden ser resueltas por varios métodos analíticos o numéricos, pero como ya se mostró anteriormente, existe un método totalmente diferente por el cual se obtiene una solución en la forma de un valor esperado. La conexión entre la valuación de opciones por métodos de EDP's y el método de martingala en tiempo continuo se debe al *teorema, representación o fórmula de Feynman-Kac.*

Existen varios caminos que permiten obtener la EDP de Black-Scholes. Por ejemplo, ahora supóngase que el valor del portafolio V_t puede ser escrito como $V_t = f(S_t, t)$, para una función suave apropiada f. Por la condición de autofinanciabilidad, el modelo debe cumplir

$$\begin{aligned} dV_t = a_t dS_t + b_t dB_t &= a_t(\mu S_t dt + \sigma S_t dW_t) + b_t(rB_t dt) \\ &= (a_t \mu S_t + b_t r B_t)dt + a_t \sigma S_t dW_t. \end{aligned} \tag{4.32}$$

Por otro lado, aplicando la fórmula de Itô a $V_t = f(S_t, t)$ se tiene que

$$dV_t = \Big(f_t(S_t, t) + \frac{1}{2} f_{xx}(S_t, t)\sigma^2 S_t^2 + f_x(S_t, t)\mu S_t\Big)dt + f_x(S_t, t)\sigma S_t dW_t. \tag{4.33}$$

Igualando las ecuaciones (4.32) y (4.33), rápidamente se obtiene el tamaño de la porción del activo que replicará al portafolio:

$$a_t = f_x(S_t, t).$$

Ahora, para determinar el tamaño de la porción de bonos, de las ecuaciones (4.32) y (4.33) solamente se necesita igualar los coeficientes de dt, esto es,

$$\mu f_x(S_t, t)S_t + b_t r B_t = f_t(S_t, t) + \frac{1}{2} f_{xx}(S_t, t)\sigma^2 S_t^2 + f_x(S_t, t)\mu S_t.$$

Luego, resolviendo para b_t :

$$b_t = \frac{1}{rB_t}\left(f_t(S_t, t) + \frac{1}{2} f_{xx}(S_t, t)\sigma^2 S_t^2\right).$$

Debido a que V_t es igual tanto a $f(S_t, t)$ como a $a_t S_t + b_t B_t$, sustituyendo los valores de a_t y b_t dan lugar a una EDP para $f(S_t, t)$:

$$\begin{aligned} f(S_t, t) &= V_t \\ &= a_t S_t + b_t B_t \\ &= f_x(S_t, t)S_t + \frac{1}{rB_t}\left(f_t(S_t, t) + \frac{1}{2} f_{xx}(S_t, t)\sigma^2 S_t^2\right)B_t \\ &= S_t f_x(S_t, t) + \frac{1}{r}\left(f_t(S_t, t) + \frac{1}{2} f_{xx}(S_t, t)\sigma^2 S_t^2\right). \end{aligned}$$

De la expresión anterior, al reemplazar S_t por simplemente S se obtiene justamente la clásica y famosa EDP de Black-Scholes:

$$f_t(S, t) + \frac{1}{2}\sigma^2 S^2 f_{SS}(S, t) + rSf_S(S, t) - rf(S, t) = 0,$$

con la condición de frontera

$$f(S, T) = g(S, K), \qquad \text{para todo } S \in \mathbb{R}.$$

Como caso particular, si el derivado que se está evaluando es una opción de compra Europea con precio de ejercicio K, entonces $f(S,T) = g(S,K) = (S-K)^{+}$.

Capítulo 5

Método de Harper

Desde la propuesta del modelo de Black-Scholes para la valuación teórica de una opción Europea en 1973, la valuación de derivados financieros se ha convertido en el problema central en la teoría financiera. En la actualidad, el principal interés de los inversionistas es encontrar modelos más realistas. Como caso particular, se ha notado que las variables involucradas, tales como la volatilidad y la tasa de interés, no deben permanecer constantes.

En la construcción de modelos más realistas se han implementado diferentes propuestas, y éstas abarcan desde el uso de nuevas distribuciones de probabilidad hasta el uso de nuevos métodos para la solución de EDP's, en donde el teorema de Feynman-Kac funciona como la principal conexión entre estas dos grandes teorías.

Un método para encontrar la solución de la EDP de Black-Scholes con coeficientes dependientes del tiempo, así como otras variaciones, se puede encontrar en Wilmott[25]. Este método consiste en reducir la EDP original a una ecuación de difusión mediante cambios de variables.

En Marianito y Rogermar[20] se puede encontrar otra propuesta para resolver EDP's como la ecuación de Black-Scholes con coeficientes dependientes del tiempo. Este método consiste en transformar la EDP original a una ya conocida. Para obtener soluciones en forma cerrada, este método falla si la solución de la nueva ecuación a resolver no se conoce.

Existen otros métodos para resolver la ecuación de Black-Scholes y algunas de sus versiones, pero la mayoría de ellos hacen uso de herramientas sofisticadas, por ejemplo, la transformada de Laplace, la transformada de Fourier y teoría de grupos de Lie.

En este trabajo de tesis se usará el método propuesto por J. F. Harper en [7] para proponer una solución general de una EDP de tipo parabólica con coeficientes dependientes del tiempo, el cual generaliza la ecuación clásica de Black-Scholes. También se propondrá una segunda solución general de una EDP parabólica menos general que la anterior, pero también de la forma de Black-Scholes.

Cabe resaltar que el contenido de este último capítulo parte solamente de un ejemplo que motiva a considerar una generalización, por lo tanto, todo lo que aquí se presentará es una aportación que el lector no encontrará en la literatura; a diferencia de los capítulos anteriores, en donde se completaron detalles de la demostración de una gran cantidad de resultados, pero que no se mencionaron de forma explícita porque es posible encontrar la idea principal de los mismos en la literatura.

5.1. Solución de la ecuación de Black-Scholes

El método que J. F. Harper propone, permite resolver EDP lineales de la forma

$$C_{x^0x^0} + aC_t + bC + \sum_{i=0}^{i=n} d^i C_{x^i} = 0,$$

donde t es el tiempo, y a, b, d^i son funciones de x^j, t. Éste método consiste en reducir la ecuación original a una ecuación de difusión mediante los siguientes pasos:

1. Ignorar por un momento el término de la segunda derivada.
2. Encontrar la solución general de la ecuación lineal resultante de la forma $C = \Phi f(X^0, \ldots, X^{n-1}, T)$, donde f es una función arbitraria de $n+1$ variables, y Φ, X^i, T son funciones de x^j, t, las cuales son determinadas por el método de las ecuaciones características o **método de Charpit.**
3. Se propone la solución de la ecuación original de la forma $C = \Phi F(X^0, \ldots, X^n, T)$, donde X^n es independiente de x^0 y es elegida de tal manera que transforme la ecuación original a una ecuación de difusión o a una ecuación que se pueda convertir en una ecuación de difusión al aplicar nuevamente el método.

Cabe mencionar que J. F. Harper explica su método durante su misma aplicación en tres ecuaciones particulares, siendo una de ellas la ecuación clásica de Black-Scholes. Para familiarizar al lector con este método, a continuación se mostrará paso a paso cómo resolver dicha ecuación, ya que estos detalles no se encuentran en el artículo Harper[7]. A partir de ahora, por *método de Harper* se entenderá que se trata del método que J. F. Harper propone en [7].

Considérese la EDP de Black-Scholes

$$\frac{1}{2}\sigma^2 S^2 C_{SS}(S,t) + rSC_S(S,t) + C_t(S,t) - rC(S,t) = 0, \tag{5.1}$$

donde σ^2 es la varianza de los dividendos generados por el activo, y r es la tasa de interés en una inversión sin riesgo, ambos términos se consideran constantes.

Para el caso de una opción de compra Europea, las condiciones de frontera están dadas por

$$C(0,t) = 0, \quad C(S,T) = (S-K)^+ = \text{máx}\{S-K, 0\},$$

$$C(S,t) \sim S \text{ cuando } S \to \infty,$$

donde K es el precio de ejercicio, y T es el tiempo de ejercicio. Las condiciones anteriores son equivalentes a

$$C(S,T) = (S-K)H(S-K), \tag{5.2}$$

donde

$$H(z) = \begin{cases} 1 & \text{si } z \geq 0 \\ 0 & \text{si } z < 0, \end{cases}$$

es la función escalonada de Heaviside.

Aplicando el método de Harper a (5.1), ignorando el término C_{SS}, se obtiene

$$rSC_S(S,t) + C_t(S,t) = rC(S,t), \tag{5.3}$$

cuyo sistema de *ecuaciones características asociado* está dado por

$$\frac{dS}{rS} = \frac{dt}{1} = \frac{dC}{rC}. \tag{5.4}$$

Ahora se resuelve el sistema anterior en el intervalo $[t,T]$:

- Sea $\frac{dS}{rS} = \frac{dt}{1}$.
  Entonces,
  $$\int_t^T \frac{dS}{S} = \int_t^T r ds;$$
  $$\ln c_0 - \ln S = r(T-t), \quad c_0 := \text{constante};$$
  $$\ln \frac{S}{c_0} = -r(T-t).$$
  Al proponer $c_0 = K$ resulta la primera integral independiente de (5.4) propuesta en Harper[7]:
  $$X = \ln \frac{S}{K} - r(t-T) = c_1, \tag{5.5}$$
  donde c_1 es una constante arbitraria.

- Sea $\frac{dC}{rC} = \frac{dt}{1}$.
  Luego,
  $$\int_t^T \frac{dC}{C} = \int_t^T r ds;$$
  $$\ln C(S,T) - \ln C(S,t) = r(T-t);$$
  $$\ln \frac{C}{c_2} = -r(T-t);$$
  $$C = c_2 \exp(r(t-T)),$$
  donde c_2 es una constante arbitraria.

  Así se obtiene la segunda integral independiente de (5.4) propuesta en Harper[7]:
  $$f = C \exp(-r(t-T)) = c_2. \tag{5.6}$$

Por lo anterior, la solución general de (5.3) está dada por $c_1 = f(c_2)$, i.e., $C = \exp(r(t-T))f(X)$, donde f es una función arbitraria.

De acuerdo al método de Harper, se propone la solución general de (5.1) de la forma

$$C = \exp(r(t-T))F(X,Y), \tag{5.7}$$

donde $Y := Y(t)$ y X es definida en (5.5).

Al sustituir la expresión (5.7) en la ecuación original (5.1), se tiene

$$\frac{1}{2}\sigma^2 F_{XX} - \frac{1}{2}\sigma^2 F_X + F_Y Y_t = 0.$$

Luego, se propone $Y = \frac{1}{2}\sigma^2(t-T)$, ya que de esta forma la ecuación anterior se reduce a

$$F_{XX} - F_X + F_Y = 0. \tag{5.8}$$

Ahora se necesita resolver la ecuación (5.8), y para ello se repite el método. En este caso, al omitir el término F_{XX}, la nueva ecuación a resolver resulta ser

$$-F_X + F_Y = 0, \tag{5.9}$$

cuyo sistema de *ecuaciones características asociado* está dado por:

$$\frac{dX}{-1} = \frac{dY}{1} = \frac{dF}{0}. \tag{5.10}$$

Al resolver el sistema anterior, para el par $-dX = dY$ se obtiene

$$Z = X + Y = c_3, \tag{5.11}$$

mientras que para el par $dY = \frac{dF}{0}$ resulta

$$F = c_4,$$

en donde c_3 y c_4 son constantes arbitrarias. Luego, la solución general de (5.9) está dada por $c_4 = g(c_3)$, i.e., $F = g(Z)$.

Por el método de Harper, se propone la solución general de (5.8) de la forma

$$F(X,Y) = G(Z,W), \tag{5.12}$$

donde $Z = X + Y$ y $W := W(Y)$.

Al sustituir (5.12) en la ecuación (5.8) se deduce que

$$G_{ZZ} + G_W W_Y = 0.$$

Por lo tanto, al proponer $W = -Y$, la ecuación anterior se reduce a la ecuación de difusión

$$G_{ZZ} = G_W, \quad -\infty < Z < \infty, W > 0, \tag{5.13}$$

donde

$$\begin{aligned} Z &= X + Y = \ln \frac{S}{K} + (r - \frac{1}{2}\sigma^2)(T - t), & (5.14)\\ W &= -Y = \frac{1}{2}\sigma^2(T - t). & (5.15) \end{aligned}$$

En resumen, la solución a la ecuación original (5.1) está dada por

$$C(S,t) = \exp(r(t-T))F(X,Y) = \exp(r(t-T))G(Z,W).$$

En consecuencia, la condición inicial para la ecuación de difusión (5.13) se obtiene como sigue:

$$\begin{aligned} G(Z,0) &= C(K\exp(Z), T) \\ &= (K\exp(Z) - K)H(K\exp(Z) - K) \\ &= K(\exp(Z) - 1)H(Z), & (5.16) \end{aligned}$$

pues cuando $W = 0$ se sigue que $Y = 0$, i.e., $t = T$, y entonces $Z = \ln \frac{S}{K}$.

Ahora se calculará la solución explícita de la ecuación original (5.1).
La ecuación de difusión (5.13) tiene como solución

$$\begin{aligned} G(Z,W) &= \frac{1}{2\sqrt{\pi W}} \int_{-\infty}^{\infty} G(\xi, 0) \exp\left(- \frac{(Z-\xi)^2}{4W} \right) d\xi \\ &= \frac{K}{2\sqrt{\pi W}} \int_{0}^{\infty} (\exp(\xi) - 1) \exp\left(- \frac{(Z-\xi)^2}{4W} \right) d\xi \\ &= K\exp(Z+W)N\left(\frac{2W+Z}{\sqrt{2W}} \right) - KN\left(\frac{Z}{\sqrt{2W}} \right), & (5.17) \end{aligned}$$

donde

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(- \frac{y^2}{2} \right) dy.$$

Finalmente, al sustituir los valores de Z y W se obtiene la solución de (5.1):

$$\begin{aligned} C(S,t) &= \exp(r(t-T))G(Z,W) \\ &= SN\left(\frac{\ln \frac{S}{K} + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \right) \\ &\quad -K\exp(-r(T-t))N\left(\frac{\ln \frac{S}{K} + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \right) \\ &= SN(d_1) - K\exp(-r(T-t))N(d_2), \end{aligned}$$

donde

$$d_1 = \frac{\ln \frac{S}{K} + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},$$

y

$$d_2 = d_1 - \sigma\sqrt{T-t}.$$

Nótese que la solución que se obtiene mediante el método de Harper es precisamente la solución clásica de la ecuación de Black-Scholes.

5.2. Soluciones en forma cerrada

Motivados por el resultado obtenido en la sección anterior, ahora se estará interesado en resolver ecuaciones que tienen la forma de la ecuación de Black-Scholes pero con coeficientes dependientes del tiempo.

5.2.1. Primera solución general

Para $0 \leq t \leq T$, considérese la EDP general

$$\rho(t)S^2C_{SS}(S,t) + (\lambda(t) + \alpha(t)S)C_S(S,t) + C_t(S,t) - \beta(t)C(S,t) + \gamma(t) = 0, \tag{5.18}$$

con la condición terminal

$$C(S,T) = g(S,K). \tag{5.19}$$

De acuerdo al método de Harper, si en la ecuación (5.18) se omite el término C_{SS}, entonces la nueva ecuación a resolver tiene la forma

$$(\lambda(t) + \alpha(t)S)C_S + C_t - \beta(t)C + \gamma(t) = 0,$$

cuyo sistema de *ecuaciones características asociado* está dado por:

$$\frac{dS}{\lambda(t) + \alpha(t)S} = \frac{dt}{1} = \frac{dC}{\beta(t)C - \gamma(t)}. \tag{5.20}$$

Se resuelve el sistema anterior en el intervalo $[t, T]$.

- Sea el par $\frac{dS}{\lambda(t)+\alpha(t)S} = \frac{dt}{1}$.
  La ecuación se puede reescribir como

  $$\frac{dS}{dt} - \alpha(t)S = \lambda(t).$$

  Ya que su factor de integración está dado por

  $$\mu_1(t) = \exp\left(-\int \alpha(t)dt\right),$$

entonces

$$\frac{d}{dt}(\mu_1 S) = \mu_1(t)\lambda(t);$$

$$\mu_1 S - c_0 K = -\int_t^T \mu_1(s)\lambda(s)ds, \tag{5.21}$$

donde $c_0 = \exp\Big(-\int \alpha(T)dT\Big)$. Por lo tanto, se define

$$X = \mu_1 S - c_0 K + \int_t^T \mu_1(s)\lambda(s)ds = c_1,$$

donde c_1 es una constante arbitraria.

- Ahora sea el par $\frac{dt}{1} = \frac{dC}{\beta(t)C-\gamma(t)}$.
  En este caso, la ecuación se puede reescribir como

  $$\frac{dC}{dt} - \beta(t)C = -\gamma(t), \tag{5.22}$$

  cuyo factor integrante está dado por

  $$\mu_2(t) = \exp\left(-\int \beta(t)dt\right).$$

  En consecuencia,

  $$\frac{d}{dt}(\mu_2 C) = -\mu_2(t)\gamma(t);$$

  $$\mu_2(t)C = c_2 + \int_t^T \mu_2(s)\gamma(s)ds.$$

  Por lo tanto, se define

  $$f = \mu_2(t)C - \int_t^T \mu_2(s)\gamma(s)ds = c_2,$$

  para c_2 una constante arbitraria.

De acuerdo al método de Harper, se propone la solución general de (5.18) de la siguiente forma:

$$C = h(t)\left(F(X,Y) + \int_t^T \mu_2(s)\gamma(s)ds\right), \tag{5.23}$$

donde $h(t) = \exp\Big(\int \beta(t)dt\Big), Y := Y(t)$, y $X = \mu_1(t)S - c_0K + \int_t^T \mu_1(s)\lambda(s)ds$.

Al sustituir la expresión (5.23) en la ecuación (5.18), la nueva ecuación a resolver tiene la forma

$$\rho S^2 \mu_1^2 F_{XX} + Y_t F_Y = 0.$$

Así, al sustituir $S^2\mu_1^2$ por la expresión dada en (5.21),

$$\rho(t)\left(c_0 K - \int_t^T \mu_1(s)\lambda(s)ds\right)^2 F_{XX} + Y_t F_Y = 0.$$

De la expresión anterior, si la función $\rho(t)$ es tal que

$$Y(t) = \int_t^T \rho(s)\Big(c_0 K - \int_s^T \exp\Big(-\int \alpha(y)dy\Big)\lambda(y)dy\Big)^2 ds > 0,$$

entonces la función F satisface la ecuación de difusión

$$F_{XX} = F_Y, \quad -\infty < X < \infty, Y > 0, \tag{5.24}$$

donde

$$\begin{aligned} X &= S\exp\left(-\int \alpha(t)dt\right) + \int_t^T \exp\left(-\int \alpha(s)ds\right)\lambda(s)ds - c_0 K, \\ Y &= \int_t^T \rho(s)\left(c_0 K - \int_s^T \exp\Big(-\int \alpha(y)dy\Big)\lambda(y)dy\right)^2 ds, \\ c_0 &= \exp\left(-\int \alpha(T)dT\right), \end{aligned}$$

y con condición inicial

$$\begin{aligned} F(X,0) &= \exp\left(-\int \beta(T)dT\right) g\left((X + c_0 K)\exp\left(\int \alpha(T)dT\right), K\right) \\ &= \exp\left(-\int \beta(T)dT\right) g\left(X\exp\left(\int \alpha(T)dT\right) + K, K\right). \end{aligned}$$

Por lo tanto, **la nueva solución general de la EDP (5.18) que en este trabajo de tesis se propone está dada por**

$$C = \exp\left(\int \beta(t)dt\right)\left(F(X,Y) + \int_t^T \exp\Big(-\int \beta(s)ds\Big)\gamma(s)ds\right), \tag{5.25}$$

en donde la función F satisface la ecuación de difusión (5.24).

5.2.2. Segunda solución general

Ahora, para $0 \le t \le T$, considérese la EDP general

$$\rho(t)S^2 C_{SS}(S,t) + \alpha(t)SC_S(S,t) + C_t(S,t) - \beta(t)C(S,t) + \gamma(t) = 0, \tag{5.26}$$

con condición terminal

$$C(S,T) = g(S,K). \tag{5.27}$$

Nótese que (5.26) corresponde a un caso particular de la EDP (5.18), por lo tanto, **una primera solución de la EDP (5.26) tiene la forma dada en (5.25), pero con $\lambda(t) = 0$.** Para encontrar otra solución, a la ecuación (5.26) nuevamente se le aplicará el método de Harper.

Aplicando el método de Harper, al omitir el término C_{SS} en (5.26), la primera ecuación a resolver es

$$\alpha(t)SC_S + C_t - \beta(t)C + \gamma(t) = 0.$$

Por lo tanto, el sistema de *ecuaciones características asociado* a resolver en el intervalo $[t, T]$ está dado por:

$$\frac{dS}{\alpha(t)S} = \frac{dt}{1} = \frac{dC}{\beta(t)C - \gamma(t)}. \tag{5.28}$$

- Primero considérese $\frac{dS}{\alpha(t)S} = \frac{dt}{1}$.
  A diferencia del caso anterior, ahora la ecuación se puede reescribir como

  $$\frac{dS}{S} = \alpha(t)dt.$$

  Luego,

  $$\int_t^T \frac{dS}{S} = \int_t^T \alpha(s)ds;$$
  $$\ln\frac{S}{K} = -\int_t^T \alpha(s)ds.$$

  Por lo tanto, se define

  $$X = \ln\frac{S}{K} + \int_t^T \alpha(s)ds = c_1,$$

  para c_1 una constante arbitraria.

- Ahora sea $\frac{dt}{1} = \frac{dC}{\beta(t)C-\gamma(t)}$.
  La ecuación a resolver se transforma en

  $$\frac{dC}{dt} - \beta(t)C = -\gamma(t).$$

  Nótese que ésta ecuación es exactamente la misma que en (5.22). Por lo tanto, procediendo como en el caso de la primera EDP general, se define

  $$f = \mu_2(t)C - \int_t^T \mu_2(s)\gamma(s)ds = c_2,$$

  para c_2 una constante arbitraria.

Luego, se propone

$$C = h(t)\left(F(X, Y) + \int_t^T \mu_2(s)\gamma(s)ds\right) \tag{5.29}$$

como solución general de la EDP (5.26), en donde $h(t) = \exp\Big(\int \beta(t)dt\Big)$, $Y := Y(t)$ y $X = \ln\frac{S}{K} + \int_t^T \alpha(s)ds$.

Sustituyendo (5.29) en (5.26), la ecuación para F se reduce a

$$\rho F_{XX} - \rho F_X + Y_t F_Y = 0.$$

Nótese que al proponer $Y(t) = -\int_t^T \rho(s)ds$, entonces se sigue que F debe satisfacer la EDP

$$F_{XX} - F_X + F_Y = 0. \tag{5.30}$$

Para resolver la ecuación anterior nuevamente se aplica el método de Harper, pero en este caso tal ecuación ya fue resuelta en la primera sección del capítulo (véase la ecuación (5.8)). Entonces, bajo la hipótesis de que $\rho(t)$ sea tal que

$$W = \int_t^T \rho(s)ds > 0,$$

la función $F = G$ satisface la ecuación de difusión

$$G_{ZZ} = G_W, \quad -\infty < Z < \infty, W > 0, \tag{5.31}$$

en donde

$$\begin{aligned} Z &= \ln\frac{S}{K} + \int_t^T (\alpha(s) - \rho(s))ds, \\ W &= \int_t^T \rho(s)ds, \end{aligned}$$

y sujeta a la condición inicial

$$G(Z, 0) = \exp\Big(-\int \beta(T)dT\Big)g(K\exp(Z), K).$$

Por lo tanto, **la segunda solución general de la EDP (5.26) que en esta tesis se propone es de la forma**

$$C = \exp\left(\int \beta(t)dt\right)\left(G(Z, W) + \int_t^T \exp\Big(-\int \beta(s)ds\Big)\gamma(s)ds\right), \tag{5.32}$$

en donde ahora la función G satisface la ecuación de difusión (5.31).

5.3. Casos particulares

Para mostrar la utilidad de las soluciones antes obtenidas, en esta sección se resolverán dos ecuaciones particulares, y una de ellas nuevamente será la ecuación clásica de Black-Scholes. De manera similar a la primera sección del capítulo, a partir de ahora $H(x)$ denotará a la función escalonada de Heaviside, $N(x)$ será la fda de una v.a. normal estándar, y la solución de la ecuación de difusión que se ocupará será

$$f(x, t) = \frac{1}{2\sqrt{\pi t}}\int_{-\infty}^{\infty} f(\xi, 0)exp\Big(-\frac{(x-\xi)^2}{4t}\Big)d\xi,$$

donde $f(x, 0)$ es la correspondiente condición inicial.

I) ***Sea la EDP de Black-Scholes***

$$\frac{1}{2}\sigma^2 S^2 C_{SS}(S,t) + rSC_S(S,t) + C_t(S,t) - rC(S,t) = 0, \tag{5.33}$$

sujeta a la condición terminal

$$C(S,T) = (S-E)H(S-E).$$

Debido a que la ecuación (5.33) es un caso particular de (5.26), en donde

$$\rho(t) = \frac{\sigma^2}{2}, \quad \alpha(t) = r, \quad \beta(t) = r, \quad \gamma(t) = 0,$$

entonces (5.33) admite dos soluciones.

a) Primera solución.

Como ya se especificó en su momento, la primera solución hace referencia a la solución general dada en (5.25), en donde $\lambda(t) = 0$.

Las variables involucradas en la solución (5.25) están dadas a continuación:

$$\begin{aligned}
X &= \exp(-rt)S - c_0 K, \\
Y &= c_0^2 K^2 \frac{\sigma^2}{2}(T-t), \\
c_0 &= \exp(-rT), \\
F(X,0) &= \exp(-rT)X\exp(rT)H(X\exp(rT)) \\
&= XH(X).
\end{aligned}$$

Luego, la solución de la respectiva ecuación de difusión (5.24) está dada por

$$\begin{aligned}
F(X,Y) &= \frac{1}{2\sqrt{\pi Y}}\int_{-\infty}^{\infty} F(\xi,0)\exp\left(-\frac{[X-\xi]^2}{4Y}\right)d\xi \\
&= \frac{1}{2\sqrt{\pi Y}}\int_{0}^{\infty} \xi\exp\left(-\frac{[X-\xi]^2}{4Y}\right)d\xi \\
&= XN\left(\frac{X}{\sqrt{2Y}}\right) + \sqrt{\frac{Y}{\pi}}\exp\left(-\frac{X^2}{4Y}\right).
\end{aligned}$$

Sustituyendo los valores de Y y X,

$$\begin{aligned}
C(S,t) &= \exp(rt)\left(XN\left(\frac{X}{\sqrt{2Y}}\right) + \sqrt{\frac{Y}{\pi}}\exp\left(-\frac{X^2}{4Y}\right)\right) \\
&= \Big(S - \exp(-r(T-t))K\Big)N\left(\frac{\exp(r(T-t))S - E}{\sigma K\sqrt{T-t}}\right) \\
&\quad + \frac{\sqrt{T-t}}{\sqrt{2\pi}}\sigma K\exp(-r(T-t)) \cdot \\
&\quad \exp\left(-\frac{(\exp(r(T-t))S - K)^2}{2\sigma^2K^2(T-t)}\right).
\end{aligned}$$

Por lo tanto, utilizando la solución (5.25), ***la nueva solución de la ecuación de Black-Scholes que en esta tesis se obtiene está dada por***

$$C(S,t) = SN(d_0) - K\exp(-r(T-t))\left(N(d_0) + \frac{\sqrt{T-t}}{\sqrt{2\pi}}\sigma\exp\left(-\frac{d_0^2}{2}\right)\right), \quad (5.34)$$

donde

$$d_0 = \frac{\exp(r(T-t))S - K}{\sigma K\sqrt{T-t}}.$$

b) Segunda solución.

Ahora se calcularán los parámetros de la solución general dada en (5.32):

$$\begin{aligned} Z &= \ln\frac{S}{K} + \left(r - \frac{\sigma^2}{2}\right)(T-t), \\ W &= \frac{\sigma^2}{2}(T-t), \\ G(Z,0) &= \exp(-rT)K(\exp(Z)-1)H(Z). \end{aligned}$$

Nótese que ***la expresión para Z y W coincide con las calculadas en la primera sección*** (véase (5.14) y (5.15)), ***pero no así con la condición inicial*** (véase (5.16)), ya que la nueva condición tiene el término constante $\exp(-rT)$.

Así, por la expresión (5.17), la solución de la respectiva ecuación de difusión está dada por

$$G(Z,W) = \exp(-rT)\left(K\exp(Z+W)N\left(\frac{2W+Z}{\sqrt{2W}}\right) - KN\left(\frac{Z}{\sqrt{2W}}\right)\right).$$

Reemplazando los valores Z y W, se tiene que

$$\begin{aligned} C(S,t) &= \exp(rt)G(Z,W) \\ &= \exp(-r(T-t))\left(K\exp(Z+W)N\left(\frac{2W+Z}{\sqrt{2W}}\right)\right. \\ &\quad \left. -KN\left(\frac{Z}{\sqrt{2W}}\right)\right) \qquad (5.35) \\ &= SN(d_1) - K\exp(-r(T-t))N(d_2), \qquad (5.36) \end{aligned}$$

donde

$$d_1 = \frac{\ln\frac{S}{K} + (r+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \quad d_2 = d_1 - \sigma\sqrt{T-t}.$$

Por lo tanto, ***la segunda solución general conduce a la solución clásica de la ecuación de Black-Scholes.***

II) Ahora, ***sea la ecuación de Black-Scholes con parámetros dependientes del tiempo considerada en Marianito y Rogermar[20]:***

$$\frac{1}{2}\sigma(t)^2S^2C_{SS} + (r(t) - D(t))SC_S + C_t - r(t)C = 0, \quad (5.37)$$

con condición terminal

$$C(S,T) = (S-E)H(S-E).$$

Nuevamente, la ecuación (5.37) es un caso particular de (5.26), y en este caso

$$\rho(t) = \frac{\sigma(t)^2}{2}, \quad \alpha(t) = r(t) - D(t), \quad \beta(t) = r(t), \quad \gamma(t) = 0.$$

A continuación se describen las soluciones obtenidas.

a) Primera solución.

Las variables a utilizar en la aplicación de la solución dada en (5.25), en donde $\lambda(t) = 0$, son las siguientes:

$$\begin{aligned}
X &= \exp\Big(-\int (r(t)-D(t))dt\Big)S - c_0K,\\
Y &= c_0^2K^2\frac{1}{2}\int_t^T \sigma(s)^2 ds,\\
c_0 &= \exp\Big(-\int (r(T)-D(T))dT\Big),\\
F(X,0) &= \exp\Big(-\int r(T)dT\Big)\bigg(X\exp\Big(\int (r(T)-D(T))dT\Big)\bigg)\\
&\quad H\bigg(X\exp\Big(\int (r(T)-D(T))dT\Big)\bigg)\\
&= \exp\Big(-\int D(T)dT\Big)XH(X).
\end{aligned}$$

Similar al ejemplo anterior, la solución de la respectiva ecuación de difusión está definida por

$$F(X,Y) = \exp\left(-\int D(T)dT\right)\left(XN\left(\frac{X}{\sqrt{2Y}}\right) + \sqrt{\frac{Y}{\pi}}\exp\left(-\frac{X^2}{4Y}\right)\right).$$

Luego, por (5.25),

$$\begin{aligned}
C(S,t) &= \exp\left(\int r(t)dt\right)F(X,Y)\\
&= \exp\left(\int r(t)dt - \int D(T)dT\right)\left(XN\left(\frac{X}{\sqrt{2Y}}\right)\right.\\
&\quad \left.+\sqrt{\frac{Y}{\pi}}\exp\left(-\frac{X^2}{4Y}\right)\right)\\
&= \left(\exp\Big(-\int_t^T D(s)ds\Big)S - \exp\Big(-\int_t^T r(s)ds\Big)K\right)N(d_0)\\
&\quad +\frac{K}{\sqrt{2\pi}}\sqrt{\int_t^T \sigma(s)^2 ds}\,\exp\left(-\int_t^T r(s)ds - \frac{d_0^2}{2}\right),
\end{aligned}$$

donde

$$\begin{aligned} d_0 &= \frac{X}{\sqrt{2Y}} \\ &= \frac{\exp\Big(\int_t^T (r(s)-D(s))ds\Big)S-K}{K\sqrt{\int_t^T \sigma(s)^2 ds}}. \end{aligned}$$

En resumen, ***la solución general (5.25) permite obtener, de forma explícita, la solución dada como***

$$\begin{aligned} C(S,t) &= \exp\Big(-\int_t^T D(s)ds\Big)SN(d_0) - \exp\Big(-\int_t^T r(s)ds\Big)K\Bigg(N(d_0) \\ &\quad + \frac{1}{\sqrt{2\pi}}\sqrt{\int_t^T \sigma(s)^2 ds}\,\exp\Bigg(-\frac{d_0^2}{2}\Bigg)\Bigg), \end{aligned}$$

donde

$$d_0 = \frac{\exp\Big(\int_t^T (r(s)-D(s))ds\Big)S-K}{K\sqrt{\int_t^T \sigma(s)^2 ds}}.$$

b) Segunda solución.

De acuerdo a la solución dada en (5.32), las variables independientes y la condición terminal para la respectiva ecuación de difusión son

$$\begin{aligned} Z &= \ln\frac{S}{K} + \int_t^T (r(s)-D(s))ds - \int_t^T \frac{\sigma(u)^2}{2}du, \\ W &= \int_t^T \frac{\sigma(s)^2}{2}ds, \\ G(Z,0) &= \exp\Big(-\int r(T)dT\Big)(K\exp(Z)-K)H(K\exp(Z)-K) \\ &= \exp\Big(-\int r(T)dT\Big)K(\exp(Z)-1)H(Z). \end{aligned}$$

Luego,

$$\begin{aligned} G(Z,W) &= \exp\Big(-\int r(T)dT\Big)\Bigg(K\exp(Z+W)N\Bigg(\frac{2W+Z}{\sqrt{2W}}\Bigg) \\ &\quad -KN\Bigg(\frac{Z}{\sqrt{2W}}\Bigg)\Bigg). \end{aligned}$$

Similar al ejemplo anterior, por (5.32) se tiene que

$$\begin{aligned} C(S,t) &= \exp\Big(\int r(t)dt\Big)G(Z,W) \\ &= S\exp\Big(-\int_t^T D(s)ds\Big)N(d_1) - K\exp\Big(-\int_t^T r(s)ds\Big)N(d_2), \end{aligned}$$

donde

$$\begin{aligned} d_1 &= \frac{2W+Z}{\sqrt{2W}} \\ &= \frac{\ln\frac{S}{K}+\int_t^T(r(s)+\frac{1}{2}\sigma(s)^2-D(s))ds}{\sqrt{\int_t^T\sigma(s)^2ds}}, \\ d_2 &= d_1-\sqrt{2W} \\ &= d_1-\sqrt{\int_t^T\sigma(s)^2ds}. \end{aligned}$$

En conclusión, ***la segunda solución que se obtiene al aplicar la solución general (5.32) está dada por***

$$C(S,t) = \exp\left(-\int_t^T D(s)ds\right)SN(d_1)-K\exp\left(-\int_t^T r(s)ds\right)N(d_2),$$

donde

$$\begin{aligned} d_1 &= \frac{\ln\frac{S}{K}+\int_t^T(r(s)+\frac{1}{2}\sigma(s)^2-D(s))ds}{\sqrt{\int_t^T\sigma(s)^2ds}}, \\ d_2 &= d_1-\sqrt{\int_t^T\sigma(s)^2ds}. \end{aligned}$$

Se hace notar que ninguna de las soluciones obtenidas en el segundo ejemplo coinciden con la solución que se propone en Marianito y Rogermar[20].

En general, se resalta el hecho que las soluciones explícitas que se obtuvieron en los ejemplos anteriores, y que son consecuencias de la aplicación de las soluciones generales que en esta tesis se proponen, tienen la forma de la fórmula de valuación de Black-Scholes, lo cual es siempre deseable, pues la interpretación se sigue conservando.

Nótese que para la obtención de las soluciones generales, así como de las soluciones explícitas de los casos particulares, **en ningún momento se recurrió a herramientas sofisticadas** que hacen ver a las EDP's difíciles de abordar, en general, **sólo se aplicó resultados básicos del cálculo elemental**.

En J. Zavaleta[1] p. 108 se realiza una comparación gráfica entre la solución clásica de la ecuación de Black-Scholes y la nueva solución que propone N. Sukhomlin[2] mediante aplicaciones de grupos Lie. Tomando ciertos valores específicos, en tal análisis se puede apreciar

[1]Zavaleta-Sanchez, J. (2008). *Aplicaciones computacionales de las EDP's aasociadas a procesos estocásticos*. Tesis de Licenciatura, Instituto de Física y Matemáticas Aplicadas, Universidad Tecnológica de la Mixteca, Oaxaca, México.

[2]Sukhomlin, N. (2004). "Simetría y nuevas soluciones de la ecuación de Black-Scholes". Asociación Matemática Venezolana, **XI**, pp. 175-189.

que para valores del subyacente menores que el precio de ejercicio la nueva solución toma valores negativos, lo que no ocurre con la solución clásica.

En esta tesis también se realizó comparaciones entre la nueva solución de la ecuación de Back-Scholes que aquí se obtuvo y la solución clásica, y en el análisis se notó que en efecto tales fórmulas proporcionan o valores muy muy cercanos o valores totalmente idénticos (véase Apéndice A.3). La diferencia entre éstos valores depende de los parámetros de la aplicación en particular.

Lo anterior motiva a utilizar la solución general (5.25) cuando el modelo a emplear sea de la forma

$$\rho(t)S^2C_{SS}(S,t) + (\lambda(t) + \alpha(t)S)C_S(S,t) + C_t(S,t) - \beta(t)C(S,t) + \gamma(t) = 0,$$

para $\lambda(t) \neq 0$, que es cuando la solución general (5.32) no aplica.

Conclusiones

En este trabajo se han expuesto los fundamentos matemáticos de los derivados financieros y se ha realizado un análisis detallado del modelo de Black-Scholes. Éste análisis finaliza con la introducción de la ecuación de Black-Scholes, ecuación que posteriormente se generaliza y se resuelve mediante el método de Harper. Una vez cumplido con los objetivos plateados originalmente, se obtuvieron los elementos necesarios para presentar las siguientes conclusiones.

El comportamiento de un sistema dinámico generalmente es modelado por medio de las ecuaciones diferenciales, las cuales pueden ser ordinarias o parciales, pero cuando el sistema evoluciona de forma aleatoria, su estudio requiere de un análisis mediante conceptos de probabilidad. En el caso particular de los modelos financieros, tal análisis permite comprender su formulación, su utilidad y sus limitantes. En esta formulación es necesario introducir una gran variedad de objetos matemáticos; por ejemplo, espacio de probabilidad, filtración, procesos estocásticos, integral estocástica, ecuación diferencial estocástica, sólo por mencionar algunos.

A pesar de que existe suficiente material bibliográfico para el análisis de los derivados financieros, la gran mayoría parte de que el lector ya domina una gran cantidad de conceptos, situación que generalmente influye en que se pierda el interés. Uno de los logros de este trabajo consiste en que se presenta de forma ordenada una gran cantidad de conceptos y resultados, y que posteriormente son aterrizados en la valuación teórica de los derivados.

Ahora que se cuenta con un panorama sobre la estrecha relación entre los problemas que involucran a las distribuciones de probabilidad y la teoría de las EDP's, el proponer nuevos métodos de solución adquiere un mayor interés. Este mismo interés se sigue identificando, principalmente por parte de los inversionistas y los analistas financieros, cuando se proponen soluciones cerradas para ciertos tipos de EDP's, ya que al trabajar con tales soluciones las respectivas simulaciones son mucho más eficientes. Y si el proceso para la obtención de las soluciones sólo se vale de matemáticas básicas, entonces tanto el método como las soluciones adquieren especial atención en la comunidad científica.

Una de las principales aportaciones de este trabajo, y que no se encuentra en la literatura, es que se presentan dos soluciones cerradas a ecuaciones más generales que la ecuación clásica de Black-Scholes, y tales soluciones se obtienen sin hacer uso de herramientas sofisticadas. Se hace énfasis en que las soluciones fueron obtenidas mediante la generalización de una aplicación particular del método de Harper, por lo tanto, para lograrlo primero se tuvo que detallar tal aplicación.

Otra de las aportaciones consiste en que, mediante la aplicación de las soluciones generales que se proponen, se resuelven dos ecuaciones conocidas y por cada ecuación se obtienen dos soluciones de forma explícita. Una de las ecuaciones que se resolvió fue precisamente la ecuación clásica de Black-Scholes, y una de las soluciones obtenidas es la solución clásica, mientras que la otra solución no se encuentra en trabajos previos, por lo tanto, en este trabajo se obtiene una nueva solución de la ecuación de Black-Scholes, siendo lo anterior otro de los logros de este trabajo.

Los temas aquí tratados son la base para abordar una gran variedad de problemas de la teoría financiera, así como otras modernas teorías. Por ejemplo, ahora se está en condiciones de empezar a analizar la valuación de opciones exóticas mediante los procesos de Lévy, siendo lo anterior una de las posibles continuaciones de este trabajo. Otra posible continuación consiste en la aplicación de las soluciones generales que aquí se proponen a problemas prácticos.

Más aún, a pesar de que los conceptos están muy enfocados en su aplicación en las Matemáticas Financieras, también permiten abordar diversos problemas de otras áreas de la ciencia. Por ejemplo, en Física es posible analizar problemas de turbulencia, teoría del campo y física cuántica; en Biología se pueden abordar problemas relacionados con la dinámica de poblaciones y los modelos ecológicos; en Ingeniería se aplican a problemas de filtros, teoría de control, estabilidad y reacciones químicas.

Apéndice

A.1 Movimiento Browniano Geométrico

El concepto de distribución normal ha sido esencial para el desarrollo de la teoría financiera. Históricamente, los resultados teóricos más conocidos se han obtenido bajo la hipótesis de normalidad. Por supuesto, los precios de un activo no pueden ser normales, simplemente porque éstos son positivos. Sin embargo, el rendimiento puede tomar valores tanto positivos como negativos.

Si los rendimientos son de alta frecuencia, casi nunca se encuentra que siguen una distribución normal. De hecho, dichos rendimientos casi nunca se comportan de acuerdo a distribuciones teóricas comúnmente conocidas. Por ejemplo, las distribuciones empíricas de los rendimientos diarios de varios de los títulos que se cotizan en los mercados de capitales presentan, en general, sesgo, exceso de curtosis, colas pesadas, y no siempre es posible ajustar una distribución teórica simple a un conjunto de observaciones de mercado.

Para modelar el comportamiento de los rendimientos de un activo financiero, bajo el supuesto de normalidad, se considera un registro histórico. Como caso particular, considérese el registro de los precios mensuales de un activo, $S_0, S_1, \ldots, S_N$. Los rendimientos mensuales del activo son:

$$R_t = \frac{S_t - S_{t-1}}{S_{t-1}}, \quad t = 1, 2, \ldots, N,$$

o, equivalentemente,

$$R_t\Big(\frac{1}{mes}\Big) = \frac{S_t - S_{t-1}}{S_{t-1}}\Big(\frac{1}{mes}\Big), \quad t = 1, 2, \ldots, N.$$

Es importante enfatizar que la unidad de tiempo en cuestión es un mes. El valor medio mensual de los rendimientos del activo está dado por

$$\mu = \frac{1}{N}\sum_{t=1}^{N} R_t\Big(\frac{1}{mes}\Big),$$

y con una varianza mensual

$$\sigma^2 = \frac{1}{N}\sum_{t=1}^{N}(R_t - \mu)^2\Big(\frac{1}{mes}\Big).$$

Si se estandarizan los rendimientos y su histograma de frecuencias coincide con la función de densidad de una v.a. $Z \sim N(0,1)$, entonces se podría pensar que los rendimientos tienen una distribución normal con media $\mu \times mes$ y varianza $\sigma^2 \times mes$. En este caso, se puede escribir

$$\begin{aligned} \frac{S_t - S_{t-1}}{S_{t-1}} &= \frac{1}{N}\sum_{t=1}^{N} R_t + \sqrt{\frac{1}{N}\sum_{t=1}^{N}(R_t-\mu)^2} Z \\ &= \mu \times (mes) + \sqrt{\sigma^2 \times (mes)} Z. \end{aligned}$$

Si Δt denota un mes, y en general, denota la unidad de tiempo que separa las observaciones S_t, entonces

$$\frac{S_t - S_{t-1}}{S_{t-1}} = \mu\Delta t + \sigma\sqrt{\Delta t} Z,$$

o, equivalentemente,

$$\frac{\Delta S_t}{S_{t-1}} = \mu\Delta t + \sigma\Delta W_t,$$

donde $\Delta W_t \sim N(0, \Delta t)$. Si ahora se supone que Δt se hace cada vez más y más pequeño, en el límite se tiene que

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t, \tag{A.1}$$

donde $dW_t \sim N(0, dt)$.

La v.a. dW_t modela el riesgo de mercado. Se dice, en este caso, que el precio S_t del activo sigue un movimiento Browniano geométrico (o movimiento geométrico Browniano) o, simplemente, que el precio S_t es lognormal.

El proceso definido por (A.1) fue introducido, por primera vez, por Paul Samuelson en 1965.

A.2 Cambio de medida

Supóngase que $\mathbb{P}$ y $\mathbb{Q}$ son dos medidas de probabilidad en el espacio medible $(\Omega, \mathcal{F})$.

Definición 1. *La medida de probabilidad $\mathbb{Q}$ se dice que es* ***absolutamente continua*** *con respecto a $\mathbb{P}$ en la σ-álgebra $\mathcal{F}$, y se denota por $\mathbb{Q} \ll \mathbb{P}$, si para todo $A \in \mathcal{F}$ tal que $\mathbb{P}(A) = 0$, entonces $\mathbb{Q}(A) = 0$.*
Si se cumple $\mathbb{Q} \ll \mathbb{P}$ y $\mathbb{P} \ll \mathbb{Q}$, entonces se dice que $\mathbb{Q}$ y $\mathbb{P}$ son ***equivalentes****, y se denota por $\mathbb{Q} \approx \mathbb{P}$.*

De la definición anterior, las medidas equivalentes tienen los mismos conjuntos nulos. Puesto que las medidas de probabilidad son finitas o σ-finitas, el término *continuidad* encaja en el sentido de que la condición $\mathbb{Q} \ll \mathbb{P}$ implica que $\mathbb{Q}(A)$ se puede hacer arbitrariamente pequeño eligiendo A tal que $\mathbb{P}(A)$ sea suficientemente pequeña.

El siguiente teorema, conocido como el Teorema de Radon-Nikodym, da una caracterización de las medidas de probabilidad absolutamente continuas.

Teorema 1. *$\mathbb{Q} \ll \mathbb{P}$ si y sólo si existe una única v.a. φ, $\mathbb{P}$-c.s., tal que $\varphi \geq 0$, $E_{\mathbb{P}}[\varphi] = 1$, y*

$$\mathbb{Q}(A) = E_{\mathbb{P}}[\varphi I_A] = \int_A \varphi d\,\mathbb{P}, \tag{A.2}$$

para cualquier conjunto medible A.

La v.a. φ se conoce como la *derivada de Radon-Nikodym* o la *densidad* de $\mathbb{Q}$ con respecto a $\mathbb{P}$, y es denotado por $\frac{d\mathbb{Q}}{d\mathbb{P}}$.

Nótese que de la condición (A.2), si $\mathbb{Q} \ll \mathbb{P}$, entonces la esperanza bajo $\mathbb{P}$ y $\mathbb{Q}$ están relacionadas por

$$E_{\mathbb{Q}}[X] = E_{\mathbb{P}}[\varphi X],$$

para cualquier v.a. X integrable con respecto a $\mathbb{Q}$.

Teorema 2. *Si $\mathbb{Q} \ll \mathbb{P}$, entonces*

$$\mathbb{Q} \approx \mathbb{P} \text{ si y sólo si } \tfrac{d\mathbb{Q}}{d\mathbb{P}} > 0,\ \mathbb{P}\text{-c.s.}.$$

En este caso, la densidad de $\mathbb{P}$ con respecto a $\mathbb{Q}$ está dado por

$$\frac{d\mathbb{P}}{d\mathbb{Q}} = \left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)^{-1}.$$

A.3 Comparación gráfica de la nueva solución

A continuación se presenta un análisis gráfico en donde se compara la solución clásica de Black-Scholes contra la nueva solución que se obtuvo en (5.34).

Puesto que la tendencia del subyacente (i.e., μ) es irrelevante en las fórmulas, sólo se especificará los valores particulares del resto de los parámetros.

Por ejemplo, considérese los siguientes parámetros:

- $S_0 = 21$,
- $K = 20$,
- $r = 0.1$ (tasa anual de un 10 %),
- $T = 0.25$ ($\frac{1}{4}$ de año, 3 meses),
- $\sigma = 0.235$ (una volatilidad de 23.5 %).

Bajo los supuestos anteriores, en la Figura 5.1 se muestra una trayectoria de los precios del activo subyacente (etiquetado por Activo). Además, utilizando tales precios del subyacente, en la figura también se incluye los valores obtenidos por la solución clásica (etiquetado

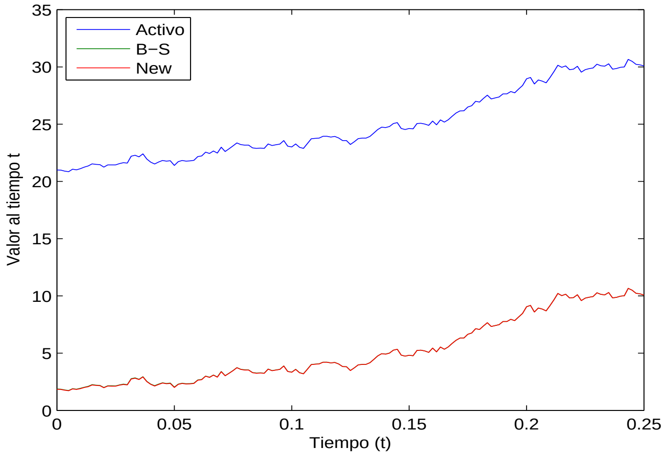

Figura 5.1: Primera comparación.

por B-S), y los valores obtenidos por la solución que aquí se propone (etiquetado por New).

Ahora, en la Figura 5.2 se ven reflejados los siguientes nuevos parámetros:

- $S_0 = 21$,
- $K = 20$,
- $r = 0.3$ (tasa anual de un 30 %),
- $T = 0.5$ ($\frac{1}{2}$ de año, 6 meses),
- $\sigma = 0.75$ (una volatilidad de 75 %).

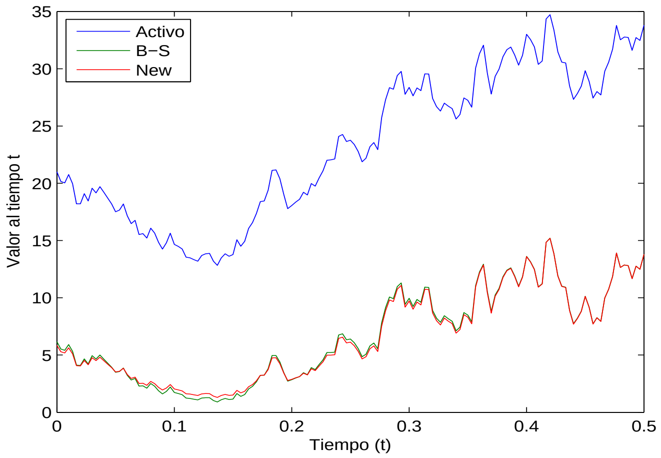

Figura 5.2: Segunda comparación.

Se realizaron varias pruebas con diferentes parámetros y se notó que los valores que se obtienen por ambas soluciones son exactamente los mismos o son muy muy cercanos, tal y como se puede apreciar en las figuras anteriores. También se observó que en situaciones donde el subyacente presenta una tendencia muy pronunciada a la baja los valores no coinciden, comportamiento que se puede apreciar en la Figura 5.2. La diferencia entre éstos valores dependerá de los parámetros de la aplicación en particular.

También se compararon las superficies generadas por cada una de las soluciones, y se notó que es muy difícil diferenciarlas a simple vista, comportamiento que ya se esperaba por el análisis previo. Por esta razón, aquí se omite tal gráfica.

A.4 Resultados auxiliares

Teorema 3. *Las siguientes propiedades son equivalentes:*

a) *Sea* $\{A_\alpha | \alpha \in \mathcal{A}\}$ *una familia no vacía de conjuntos. Si cada* $A_\alpha \neq \emptyset$*, entonces* $\prod_\alpha A_\alpha \neq \emptyset$*, en donde* $\prod_\alpha A_\alpha$*, conocido como producto cartesiano, es el conjunto de todos las funciones* $c : \mathcal{A} \to \bigcup_\alpha A_\alpha$.

b) *El **axioma de elección**: Dada cualquier familia no vacía de conjuntos no vacíos y disjuntos a pares* $\{A_\alpha | \alpha \in \mathcal{A}\}$*, existe un conjunto* S *formado de exactamente un elemento de cada* A_α.

Teorema 4 (Teorema de Convergencia Monótona). *Sea* $(X_n)_{n\geq 1}$ *una sucesión de v.a.'s no negativas tal que* $X_n \uparrow X$*. Entonces* $\lim_{n\to\infty} E[X_n] = E[X]$.

Teorema 5 (Teorema de Convergencia Dominada de Lebesgue). *Supóngase que* $|X_n| \leq Y$*, para todo* $n \in \mathbb{N}$*, donde* $E[Y] < \infty$*, y que* $X_n \to X$*,* $\mathbb{P}$*-c.s., cuando* $n \to \infty$*. Entonces*

$$E[X_n - X] \to 0, \ \mathbb{P}\text{-c.s., cuando } n \to \infty.$$

En particular,

$$E[X_n] \to E[X], \ \mathbb{P}\text{-c.s., cuando } n \to \infty.$$

Definición 2. *Una función* f *satisface la condición de Hölder (o continuidad de Hölder) de orden* α*,* $0 < \alpha \leq 1$*, en* $[a,b] \subseteq \mathbb{R}$ *si existe una constante* $K > 0$*, tal que para todo* $x, y \in [a,b]$

$$|f(x) - f(y)| \leq K|x-y|^\alpha.$$

Una condición de Lipschitz es una condición de Hölder con $\alpha = 1$,

$$|f(x) - f(y)| \leq K|x-y|.$$

Teorema 6. *Sean* X, Y *v.a.'s. Si* $E[e^{tX}] = E[e^{tY}]$ *cuando* $|t| < h$ *para algún* $h > 0$*, entonces* $X \stackrel{d}{=} Y$.

Teorema 7. *Sean X_1, X_2 v.a.'s en un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. Si $E[|X_1|] < \infty$, entonces la esperanza condicional $E[X_1|\mathcal{F}]$ existe y es única. Además se cumplen las siguientes propiedades:*

a) $E[cX_1 + X_2|\mathcal{F}] = cE[X_1|\mathcal{F}] + E[X_2|\mathcal{F}]$, *para c constante.*

b) $E[X_1] = E[E[X_1|\mathcal{F}]]$.

c) *Si X_1 es independiente de $\mathcal{F}$, entonces $E[X_1|\mathcal{F}] = E[X_1]$.*

d) *Si X_1 es $\mathcal{F}$-medible, entonces $E[X_1|\mathcal{F}] = X_1$.*

e) *Si X_1 es $\mathcal{F}$-medible y X_2 es otra v.a. arbitraria, entonces $E[X_1X_2|\mathcal{F}] = X_1E[X_2|\mathcal{F}]$.*

f) *Si $\mathcal{G} \subseteq \mathcal{F}$ es otra σ-álgebra de Ω, entonces*

$$E[X_1|\mathcal{G}] = E[E[X_1|\mathcal{F}]|\mathcal{G}], \quad E[X_1|\mathcal{G}] = E[E[X_1|\mathcal{G}]|\mathcal{F}].$$

Teorema 8 (Teorema de Separación del Hiperplano en Dimensión Finita). *Sea $C \subset \mathbb{R}^n$ un conjunto convexo no vacío con $0 \notin C$. Entonces existe $\eta \in \mathbb{R}^n$ tal que*

$$\eta \cdot x \geq 0 \text{ para todo } x \in C \text{ y } \eta \cdot x_0 > 0 \text{ para al menos un } x_0 \in C.$$

Teorema 9 (Fórmula de Feynman-Kac). *Para funciones acotadas dadas $r(x,t)$ y $g(x)$ sea*

$$C(x,t) = E\left[\exp\Big(-\int_t^T r(X_s,s)ds\Big)g(X_T)|X_t = x\right].$$

Supóngase que existe una solución de

$$\frac{\partial f}{\partial t}(x,t) + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 f}{\partial x^2}(x,t) + \mu(x,t)\frac{\partial f}{\partial x}(x,t) = r(x,t)f(x,t),$$

con $f(x,T) = g(x)$. Entonces la solución es única y $C(x,t)$ es esa solución.

Teorema 10. *Sea X una v.a. $\mathcal{G}$-medible que toma valores en $(E, \mathcal{E})$ y Y una v.a. independiente de $\mathcal{G}$ que toma valores en $(F, \mathcal{F})$. Para cualquier función Borel medible Φ no negativa (o acotada) en $(E \times F, \mathcal{E} \otimes \mathcal{F})$, la función φ definida por*

$$\varphi(x) = E[\Phi(x,Y)], \quad \forall x \in E,$$

es una función Borel medible en $(E, \mathcal{E})$ y cumple con

$$E[\Phi(X,Y)|\mathcal{G}] = \varphi(X).$$

A.5 Notación

I

1) (A_n) denota a la sucesión numerable de conjuntos $(A_n)_{n\in\mathbb{N}}$.

2) $A_n \uparrow A$ significa que (A_n) es una sucesión no decreciente tal que $A = \bigcup_{n\in\mathbb{N}} A_n$.

3) $A_n \downarrow A$ significa que (A_n) es una sucesión no creciente tal que $A = \bigcap_{n\in\mathbb{N}} A_n$.

4) Sea (x_n) una sucesión en $\mathbb{R}$. Entonces, $\liminf_n x_n = \lim_{m\to\infty} y_m$, donde $y_m = \inf\{x_m, x_{m+1}, \ldots\}$.

5) Sea (x_n) una sucesión en $\mathbb{R}$. Entonces, $\limsup_n x_n = \lim_{m\to\infty} z_m$, donde $z_m = \sup\{x_m, x_{m+1}, \ldots\}$.

6) $F(x+)$ y $F(x-)$ significan el límite por la derecha y por la izquierda de la función F en el punto x, respectivamente.

7) $F_X(x)$ denota la función de distribución de la v.a. X.

8) $\int_\Omega X d\mathbb{P} = \int_\Omega X(\omega)\mathbb{P}(d\omega)$.

9) $E_{\mathbb{P}}[f]$ denota la esperanza de f con respecto a la probabilidad $\mathbb{P}$.

10) $X_n \uparrow X$ significa que $X_1 \leq X_2 \leq \ldots$, tal que tal que $\lim_{n\to\infty} X_n(\omega) = X(\omega)$.

11) $X_n \overset{\mathbb{P}\text{-c.s.}}{\to} X$ o $\lim_{n\to\infty} X_n = X, \mathbb{P}$-c.s., denotan la convergencia *casi segura* (o con probabilidad 1) de (X_n) a la v.a. X.

12) $X_n \overset{L^p}{\to} X$ o $\lim_{n\to\infty} X_n = X$, en L^p, denotan la convergencia en *p-ésima media* de (X_n) a la v.a. X.

II

1) (X_t) y $(\mathcal{F}_t)$ denotan a los procesos estocásticos $(X_t)_{t\in T}$ y $(\mathcal{F}_t)_{t\in T}$, respectivamente.

2) $\mathcal{F} \otimes \mathcal{G} := \sigma(A \times B : A \in \mathcal{F}, B \in \mathcal{G})$.

3) $C^{2,1}(A \times B)$ denota a todas las funciones $f(x, y)$ definidas en $A \times B$, con segunda derivada parcial continua con respecto a x, y primera derivada parcial continua con respecto a y.

Bibliografía

[1] Billingsley, P. (1985). *Probability and Measure.* Second Edition. John Wiley & Sons.

[2] Björk, T. (2004). *Arbitrage Theory in Continuous Time.* Second Edition. Oxford University Press.

[3] De Lara, A. (2005). *Productos derivados financieros: Instrumentos, valuación y cobertura de riesgos.* Limusa.

[4] Etheridge, A. (2002). *A Course in Financial Calculus.* Cambridge University Press.

[5] Fölmer, H and A. Schied. (2004). *Stochastic Finance, An Introducction in Discrete Time.* Second Edition and Extended Edition. Walter de Gruyter, Berlin, New York.

[6] Gut, A. (2005). *Probability: A Graduate Course.* Springer Science+Business Media, Inc.

[7] Harper, J. F.(1994). "Reducing parabolic partial differential equations to canonical form". European Journal of Applied Mathematics, **5**, pp. 159–164.

[8] Hull, J. y V. Morales (1996). *Introducción a los mercados de futuros y opciones (Traducción del inlges al español 1995).* Segunda Edición. Prentice Hall.

[9] Jacod, J. and P. Protter (2002). *Probability Essential.* Second Edition. Springer-Verlag.

[10] Karatzas, I. and S. E. Shreve (1998). *Brownian Motion and Stochastic Calculus.* Springer-Verlag.

[11] Keblaner, Fima C. (2005). *Introduction to Stochastic Calculus with Applications.* Second Edition. Imperial College Press.

[12] Kloeden, P. E. and E. Platen (1995). *Numerical Solution of Stochastic Differential Equations.* Springer-Verlag.

[13] Lamberton, D, B. Lepeyre, N. Rabeau, and F. Mantion. (2000). *Introduction to stochastic calculus applied to finance (English translation from the French 1996).* Chapman & Hall/CRC.

[14] Lefebvre, M. (2007) *Applied Stochastic Processes.* Springer Science+Business Media LLC.

[15] Mikosch, T. (1998). *Elementary Stochastic Calculus with Finance in View.* World Scientific.

[16] Øksendal, B. (2000). *Stochastic Differential Equations. An Introduction with Applications.* Fifth Edition, Corrected Printing. Springer-Verlag.

[17] Rao, M. M. and R.J. Swift (2006). *Probability theory with applications.* Second Edition. Springer Science+Business Media, Inc.

[18] Rincón, L. (2007). *Curso Intermedio de Probabilidad.* Departamento de Matematicas, Facultad de ciencias UNAM, Octubre.

[19] Rincón, L. (2008). *Introducción a los Proceso Estocásticos.* Departamento de Matematicas, Facultad de ciencias UNAM, Agosto.

[20] Rodrigo, Marianito R. and Rogermar S. Mamon (2005). "An alternative approach to solving the Black-Scholes equation with time-varying parameters". Applied Mathematics Letters, **19**, pp. 398-402.

[21] Roberts, A. J. (2009). *Elementary Calculus of Financial Mathematics.* Society for Industrial and Applied Mathematics.

[22] Steel, J. M. (2001). *Stochastic calculus and financial applications.* Springer-Verlag.

[23] Taylor, J. C. (1997). *An Introduction to Measure and Probability.* Springer-Verlag.

[24] Venegas-Martínez, F. (2006). *Riesgos financieros y económicos. Productos derivados y decisiones económicas bajo incertidumbre.* Thomson.

[25] Wilmott, P., S. Howison, and J. Dewynne (1996). *The mathematics of financial derivatives.* Cambridge.

[26] Wiersema, Ubbo F. (2008). *Brownian Motion Calculus.* John Wiley & Sons.