



DEPARTAMENTO DE  
ECONOMÍA APLICADA I

UNIVERSIDAD DE SEVILLA

**MODELIZACIÓN DE LOS MERCADOS  
FINANCIEROS MEDIANTE ECUACIONES  
DIFERENCIALES ESTOCÁSTICAS.  
APLICACIÓN DE TÉCNICAS NO  
PARAMÉTRICAS AL CASO DE LOS TIPOS DE  
INTERÉS A CORTO PLAZO EN ESPAÑA.**

Trabajo presentado por **José Manuel Gavilán Ruiz** para la obtención del grado de doctor, dirigido por **Dr. D. José María Alba Riesco** y **Dr. D. Luis González Abril**

Sevilla, Abril de 2007



*A mi hija, Paula.*



## Agradecimientos

Quisiera expresar mi más sincero agradecimiento a mis directores de Tesis, Dr. D. José María Alba Riesco y Dr. D. Luis González Abril, por todo el esfuerzo y tiempo dedicado a la elaboración de este trabajo.

Así como a todos mis compañeros del Departamento de Economía Aplicada I. En particular al Dr. D. Francisco Velasco, por su ayuda desinteresada y su ánimo constante.

Finalmente, un agradecimiento muy especial a mi mujer, Raquel, y a mi hija, Paula.

A todos, muchas gracias.



# Índice de contenido

<b>Introducción</b>	<b>1</b>
<b>1. Procesos estocásticos: Martingalas y movimiento browniano</b>	<b>9</b>
1.1. Introducción . . . . .	9
1.2. Procesos estocásticos. Información . . . . .	10
1.3. Martingalas continuas de cuadrado integrable . . . . .	15
1.4. El movimiento browniano . . . . .	22
<b>2. Introducción al cálculo estocástico</b>	<b>39</b>
2.1. Motivación . . . . .	39
2.2. Integral estocástica: Propiedades básicas . . . . .	42
2.2.1. La regla de Itô . . . . .	52
2.3. Ecuaciones diferenciales estocásticas . . . . .	63
2.3.1. Existencia y unicidad de soluciones . . . . .	69
2.3.2. El modelo de Black-Scholes . . . . .	71
2.3.3. El proceso de Ornstein-Uhlenbeck . . . . .	74
2.4. Tres resultados fundamentales . . . . .	77
<b>3. Principios básicos en modelización de mercados financieros</b>	<b>85</b>
3.1. Introducción . . . . .	85
3.2. El modelo de Black-Scholes generalizado . . . . .	86
3.2.1. El modelo . . . . .	87
3.2.2. Carteras autofinanciadas . . . . .	91
3.2.3. Activos derivados, arbitraje y completitud . . . . .	93
3.2.4. Valoración y cobertura de activos derivados . . . . .	103
3.2.5. Análisis del modelo de Black-Scholes . . . . .	105
3.3. Mercado de bonos y modelos de tipos de interés a corto plazo . . . . .	110
3.3.1. Mercado de bonos y tipos de interés . . . . .	111
3.3.2. Valoración de activos derivados . . . . .	114
3.3.3. Modelos de tipos de interés a corto plazo . . . . .	115
<b>4. Un análisis no paramétrico de los tipos de interés a corto plazo en España</b>	<b>125</b>
4.1. Introducción . . . . .	125

4.2. Los datos: un primer análisis . . . . .	130
4.3. Una modelización no paramétrica . . . . .	145
4.4. Contraste de la no linealidad de la deriva . . . . .	164
4.4.1. El test $\mathbf{J}_n$ . Resultados de su aplicación . . . . .	166
4.5. Aplicación a la valoración de bonos: Comparaciones entre de diferentes modelos . . . . .	172
<b>5. Conclusiones</b>	<b>187</b>
5.1. Conclusiones . . . . .	187
<b>A. Programas Elaborados en MATLAB</b>	<b>193</b>
A.1. El programa bono_ckls . . . . .	194
A.2. El programa bono_tabla . . . . .	196
A.3. El programa BonoCuponCero . . . . .	199
A.4. El programa cambrow . . . . .	204
A.5. El programa ckls_fan_zhang . . . . .	206
A.6. El programa densidadMBB . . . . .	209
A.7. El programa generaCIR . . . . .	212
A.8. El programa histograma . . . . .	213
A.9. El programa LLR . . . . .	215
A.10.El programa LLRCV . . . . .	217
A.11.El programa LLRMBB . . . . .	220
A.12.El programa PRMjiang . . . . .	224
A.13.El programa PRMjiangMBB . . . . .	226
A.14.El programa testJn . . . . .	231
<b>Bibliografía</b>	<b>235</b>



# Índice de tablas

3.1.	<i>Tabla de arbitraje usada para definir los tipos de interés forward.</i>	112
4.1.	<i>Algunos modelos homogéneos unifactoriales paramétricos.</i>	128
4.2.	<i>Algunos modelos no homogéneos unifactoriales paramétricos.</i>	128
4.3.	<i>Estadísticas básicas y test de estacionariedad de la serie de datos semanales de los tipos de interés interbancarios a un mes en España.</i>	134
4.4.	<i>Estadísticas básicas y test de estacionariedad de la serie de datos semanales de los tipos de interés interbancarios a una semana en España.</i>	136
4.5.	<i>Resultados del test <math>J_n</math> con 5000 replicaciones bootstrap para contrastar la linealidad de la deriva en la dinámica de los tipos de interés a corto plazo en España.</i>	170
4.6.	<i>Resultados del test <math>J_n</math> con 1000 replicaciones bootstrap para contrastar un cuadrado de la volatilidad constante y potencial.</i>	172
4.7.	<i>Precios estimados de bonos cupón cero para la Economía española mediante el método Monte Carlo. Se generan, para cada precio, 2000 trayectorias con 100 datos diarios mediante la aproximación de Milstein y reduciendo la variabilidad de las estimaciones mediante el método de las variaciones antitéticas.</i>	181
4.8.	<i>Precios estimados de bonos cupón cero en España imponiendo en la ecuación (4.1) una deriva lineal. Se sigue el método Monte Carlo generando, para cada precio, 2000 trayectorias con 100 datos diarios mediante la aproximación de Milstein y reduciendo la variabilidad de las estimaciones mediante el método de las variaciones antitéticas.</i>	182
4.9.	<i>Estimación del modelo CKLS usando datos semanales de los tipos de interés interbancarios a un mes en España y siguiendo el método propuesto en Fan y Zhang (2003)[41].</i>	184
4.10.	<i>Precios estimados de bonos cupón cero usando el modelo CKLS y con <math>\hat{\lambda}(r) = -0,131863 \cdot r</math>. Se sigue el método Monte Carlo, generando, para cada precio, 2000 trayectorias con 100 datos diarios mediante la aproximación de Milstein y reduciendo la variabilidad de las estimaciones mediante el método de las variaciones antitéticas.</i>	185



# Índice de figuras

1.	<i>Ilustración de los pilares sobre los que descansa la teoría e implementación de la valoración de activos derivados.</i>	2
1.1.	<i>Varias trayectorias del movimiento browniano estándar unidimensional en el intervalo <math>[0, 1]</math>.</i>	24
1.2.	<i>Varias aproximaciones a una trayectoria del movimiento browniano estándar en el intervalo <math>[0, 1]</math>, basadas en el Teorema Central del Límite Funcional y obtenidas a partir de una realización de tamaño 100000 de variables normales estándar.</i>	26
1.3.	<i>Una trayectoria típica del movimiento browniano bidimensional comenzando en el origen.</i>	31
1.4.	<i>Varias trayectorias del proceso puente browniano.</i>	34
1.5.	<i>Ejemplos de trayectorias del movimiento browniano con deriva. Todas se han generado con <math>\mu = 3</math>, las de la gráfica superior con <math>\sigma = 0'5</math> y las de la inferior con <math>\sigma = 1'5</math>.</i>	35
1.6.	<i>Varias trayectorias del movimiento browniano geométrico <math>X(t) = e^{t+W(t)}</math>. La línea discontinua representa la función de medias.</i>	36
2.1.	<i>Una trayectoria de un proceso simple</i>	44
2.2.	<i>Interpretación de la ecuación diferencial estocástica (2.22) como un sistema dinámico.</i>	68
4.1.	<i>Serie de datos semanales de los tipos de interés interbancarios en España a un mes anualizados, desde enero de 1988 hasta mayo de 2006.</i>	132
4.2.	<i>Serie de la primera diferencia de los datos semanales de los tipos de interés interbancarios en España a un mes anualizados, desde enero de 1988 hasta mayo de 2006.</i>	132
4.3.	<i>Autocorrelaciones de los tipos de interés mensuales. Se muestra también una banda dentro de la cual los coeficientes de orden superior a 24 no son significativos.</i>	135
4.4.	<i>Histogramas con origen <math>r_0 = 0</math> y diferentes anchuras de ventana para los datos semanales de los tipos de interés interbancarios a un mes en España.</i>	138
4.5.	<i>Histogramas adaptados con diferentes anchuras de ventana para la serie de datos semanales de los tipos de interés interbancarios a un mes en España.</i>	140

---

4.6.	<i>Algunos núcleos usados en la estimación de densidades de frecuencia: a) Epanechnikov, dado por <math>K(r) = \frac{3}{4\sqrt{5}}(1 - \frac{1}{5}r^2) \cdot I_{(-\sqrt{5},\sqrt{5})}(r)</math>, b) Biweight, dado por <math>K(r) = \frac{15}{16}(1 - r^2)^2 \cdot I_{(-1,1)}(r)</math>, c) Triangular, dado por <math>K(r) = (1 -  r ) \cdot I_{(-1,1)}(r)</math> y d) Rectangular, dado por <math>K(r) = \frac{1}{2} \cdot I_{(-1,1)}(r)</math>.</i> . . . .	141
4.7.	<i>Densidad incondicional estimada de los tipos de interés a corto plazo en España, obtenida a partir de los datos semanales de los tipos de interés interbancarios a un mes usando un núcleo gaussiano y diferentes anchuras de ventana.</i> . . . . .	143
4.8.	<i>Densidad incondicional estimada de los tipos de interés a corto plazo en España y bandas puntuales de confianza al 95 % basadas en el método bootstrap de bloques móviles con 5000 replicaciones.</i> . . . . .	144
4.9.	<i>Ilustración del método bootstrap de bloques móviles, con una serie de tamaño <math>n = 12</math> y bloques de tamaño <math>b = 3</math>.</i> . . . . .	145
4.10.	<i>Nube de puntos de la primera diferencia de los datos semanales de los tipos de interés interbancarios a un mes en España frente a su nivel inicial.</i> . . .	151
4.11.	<i>Determinación de la anchura de la ventana mediante validación cruzada para la regresión no paramétrica conducente a la aproximación de orden uno de la deriva (arriba) y del cuadrado de la volatilidad (abajo) para los tipos de interés a corto plazo en España. Los valores mínimos se alcanzan con <math>h = 0,0068</math> para la deriva y con <math>h = 0,0372</math> para el cuadrado de la volatilidad.</i> . . . . .	154
4.12.	<i>Aproximaciones de primer, segundo y tercer orden a la deriva de la difusión propuesta como modelo de evolución de los tipos de interés a corto plazo en España.</i> . . . . .	155
4.13.	<i>Aproximaciones de primer, segundo y tercer orden a la volatilidad de la difusión propuesta como modelo de evolución de los tipos de interés a corto plazo en España.</i> . . . . .	155
4.14.	<i>Aproximaciones de primer orden y bandas al 95 % confianza de la deriva (arriba) y la volatilidad (abajo) de la difusión propuesta como modelo de evolución de los tipos de interés a corto plazo en España.</i> . . . . .	157
4.15.	<i>Determinación de la anchura de la ventana mediante validación cruzada para la regresión lineal local conducente a la aproximación de orden uno de la deriva (arriba) y del cuadrado de la volatilidad (abajo) de la dinámica de los tipos de interés a corto plazo en España. Los valores mínimos se alcanzan con <math>h = 0,0147</math> para la deriva y con <math>h = 0,0740</math> para el cuadrado de la volatilidad.</i> . . . . .	162
4.16.	<i>Aproximaciones de primer orden a la deriva de los tipos de interés a corto plazo en España mediante el estimador de Nadaraya-Watson y la regresión lineal local.</i> . . . . .	163
4.17.	<i>P-valores del test <math>J_n</math> para contrastar la linealidad de la deriva de los tipos de interés a corto plazo en España en función del valor de la anchura de la ventana. La línea horizontal representa el usual 5 % de significatividad.</i> .	171

---

4.18. *Estimación del precio del riesgo de mercado mediante la ecuación (4.47), usando una aproximación de orden uno y regresiones núcleo con núcleo gaussiano y  $h = 0,10 \cdot n^{-\frac{1}{5}}$  para estimar las correspondientes derivas y volatilidades. Se añade además, banda puntual de confianza al 95 % obtenida por el método bootstrap de bloques móviles. . . . . 175*



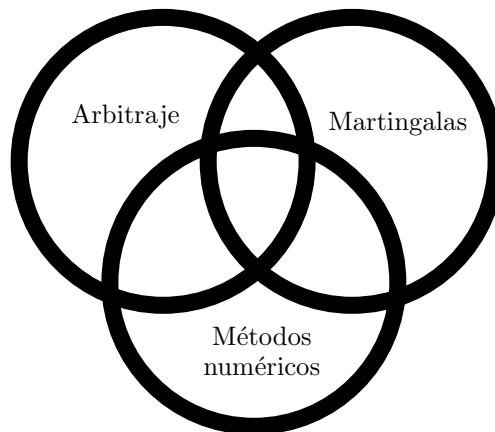
# Introducción

Repasando brevemente la historia, se comprueba que el primer intento de modelizar los precios de activos financieros mediante procesos estocásticos se atribuye a Bachelier (1900)[5]. Sin embargo, este trabajo fue demasiado adelantado para su tiempo y no se le reconoció su importancia en ese momento, de hecho hubo que esperar más de medio siglo para que fuese retomado y generalizado por los trabajos de Samuelson (1965,1973)[91, 92] y Black y Scholes (1973)[11]. En este último trabajo, Black y Scholes, con la ayuda de Merton, desarrollan su célebre fórmula de valoración de opciones, la cual proporciona una solución satisfactoria al importante y práctico problema de encontrar un precio justo para una opción de compra europea (es decir, al derecho de comprar una unidad de un activo a un precio y en un tiempo determinados). Un dato destacable es que ese trabajo fue determinante en la concesión del premio Nobel en Economía en 1997.

En el periodo 1979-1983, Harrison, Kreps y Pliska usan la teoría general de los procesos estocásticos en tiempo continuo para dotar a la fórmula de valoración de opciones de Black-Scholes de una sólida base teórica, mostrando además cómo valorar muchos otros activos derivados<sup>1</sup>. A partir de estos trabajos, y sin perder su aplicación a los aspectos

---

<sup>1</sup>Los activos derivados son aquellos cuyo valor se deduce a partir de otros activos más básicos llamados subyacentes, las opciones de compra europeas antes mencionadas son un claro ejemplo.



**Figura 1:** *Ilustración de los pilares sobre los que descansa la teoría e implementación de la valoración de activos derivados.*

más prácticos, el campo de las finanzas se ha hecho cada vez más matemático, hasta el punto de que algunos problemas surgidos en el campo de la economía financiera inspiran hoy en día investigaciones matemáticas.

La Figura 1, tomada de Pelsser (2000)[86], representa los tres pilares interconectados sobre los que descansa la teoría y la implementación de la valoración de activos derivados: *El arbitraje*, es una oportunidad que ofrece la posibilidad de generar una riqueza en un instante futuro sin inversión inicial ni riesgo, y debe ser excluida de una economía en equilibrio. *Las martingalas* son una formalización matemática de un juego justo. Si los precios de los activos derivados se modelizan como martingalas, ningún participante en el mercado podrá sistemáticamente ganar (o perder) dinero invirtiendo en activos derivados. Por último, *los métodos numéricos* proporcionan las herramientas necesarias para calcular los precios de la gran variedad de activos derivados para los que no existe una fórmula explícita.

En la actualidad, con la creciente matematización de la Economía Financiera, ya no



se pueden usar únicamente herramientas económicas para explicar el precio de los activos, además, se deben resolver ecuaciones en derivadas parciales, ecuaciones diferenciales estocásticas y tener conocimientos básicos de las martingalas. En este sentido, la herramienta principal en la modelización de los mercados financieros en tiempo continuo es el *cálculo estocástico*.

Por otro lado, desde la apertura del primer mercado de opciones en 1973 en Chicago, ha habido un crecimiento explosivo en la negociación de activos derivados, llevando al establecimiento de mercados de derivados, en los que se negocian grandes volúmenes de contratos. Estos mercados, se pueden ver como mercados de seguros para cubrir riesgos financieros. A este respecto, desde que en 1979 la Reserva Federal de Estados Unidos cambiara su política monetaria, las fluctuaciones de los tipos de interés crecieron considerablemente y, como consecuencia de la globalización de los mercados de capitales, esta variabilidad se propagó a lo largo de todo el mundo. Como consecuencia, muchas empresas buscaron un seguro (activo derivado) para protegerse de la creciente incertidumbre en los mercados de tipos de interés. Esto ha propiciado que el mercado de los tipos de interés haya sido uno de los que más ha crecido en las últimas décadas, suscitando un fuerte interés en el área y llevando a una gran cantidad de investigación en la modelización de su comportamiento y en la valoración de activos derivados de ellos.

Centrándonos en la modelización de los tipos de interés, la evidencia empírica (Dybvig (1989)[36], Litterman y Scheinkman (1991)[75]), muestra que es razonable usar una sola variable endógena, el tipo de interés a corto plazo, para la descripción de los movimientos de toda su estructura temporal<sup>2</sup>, lo que justifica el uso de los modelos unifactoriales, los cuales pueden extenderse de forma natural al caso multifactorial. En este contexto, es

---

<sup>2</sup>La estructura temporal de los tipos de interés es la curva que relaciona los tipos de interés proporcionados por activos libres de riesgo con su plazo de vencimiento.

habitual modelizar la dinámica del tipo de interés a corto plazo mediante una ecuación diferencial estocástica.

Los modelos unifactoriales planteados mediante ecuaciones diferenciales estocásticas se pueden clasificar en dos tipos: Aquellos en los que los parámetros del modelo (deriva y volatilidad) son funciones que dependen únicamente del valor actual del tipo de interés y no del tiempo, es decir, la ecuación diferencial estocástica es homogénea (algunos de los más relevantes aparecen recogidos en la Tabla 4.1) y aquellos (no homogéneos) en los que los parámetros del modelo son funciones que dependen tanto del valor actual del tipo de interés como del tiempo (algunos de ellos se muestran en la Tabla 4.2). La principal ventaja de los modelos no homogéneos frente a los homogéneos es que, al permitir que su deriva y/o volatilidad dependan del tiempo, se pueden ajustar perfectamente a la estructura actual observada en los tipos de interés (Björk(1997)[7]). Sin embargo, para mantener este ajuste perfecto, estos modelos deben ser replanteados y reestimados en cada instante de tiempo, lo que conlleva un enorme esfuerzo computacional.

Continuando con los modelos unifactoriales, es importante señalar que se han propuesto numerosos modelos diferentes sin que exista aún un consenso sobre cuál es el más adecuado. La mayoría de estos modelos se han planteado en un contexto paramétrico, atendiendo quizás más a razones de tratabilidad analítica y computacional que a razones de tipo económico. Sin embargo, un problema potencialmente grave de cualquier modelo paramétrico, es que imponer una determinada estructura puede llevar a *errores de especificación* y en consecuencia se pueden tener errores graves al usar el modelo con propósitos de valoración y cobertura de activos de tipos de interés (Canabarro (1995)[18]). En un intento de evitar tales posibles errores de especificación, se han propuesto diferentes métodos no paramétricos, siendo algunos de los más importantes los propuestos por Aït-Sahalia

(1996)[3], Jiang y Knight (1997)[65] y Stanton (1997)[94]. Además, en estos trabajos se proponen también métodos no paramétricos para estimar el precio del riesgo de mercado, parámetro exógeno del modelo que se necesita aproximar para usar el modelo con propósitos de valoración y cobertura de activos derivados de los tipos de interés.

Un hecho destacable del uso de estos modelos no paramétricos para los tipos de interés a corto plazo, es que ha llevado en ocasiones a estimaciones (similares en la forma, con datos de diferentes países) no lineales de la deriva, concluyendo algunos autores en la existencia de evidencia de no linealidad en la deriva del modelo planteado (Stanton (1997)[94], Jiang (1998)[64], Fernández (2001)[45] y Corzo y Gómez (2001)[22]). Sin embargo, Chapman y Pearson (2000)[20] ponen en entredicho esta afirmación tras analizar el comportamiento del estimador empleado. En este trabajo se trata de profundizar en este aspecto, aplicando un método mejorado de estimación y contrastes de hipótesis adecuados, tomando como base para esto último, el trabajo de Kim y Wang (2006)[68].

Por otra parte, al usar un modelo cualquiera para valorar activos derivados, la situación más general es que no exista una fórmula explícita para el cálculo del valor esperado o la solución de la ecuación en derivadas parciales que permiten obtener el precio deseado, debiendo acudir entonces a métodos numéricos (diferencias finitas y método Monte Carlo esencialmente) para obtener aproximaciones.

La estructura del presente trabajo, introduciendo sólo las herramientas matemáticas y financieras necesarias para una exposición coherente y autocontenida<sup>3</sup> de los modelos planteados y su uso en la valoración de activos derivados por ausencia de arbitraje, es la

---

<sup>3</sup>Los conocimientos previos necesarios para abordar la lectura de este trabajo se han situado en la Teoría de la Probabilidad y el Cálculo ordinario. No es necesaria ninguna formación previa en el campo de la finanzas.

siguiente:

En el capítulo 1 se introducen los procesos estocásticos en tiempo continuo y se establecen las propiedades básicas de las martingalas, además, se completa con el estudio de la martingala continua más importante: El movimiento browniano.

El capítulo 2 se dedica al cálculo estocástico (integral estocástica y ecuaciones diferenciales estocásticas), ya mencionado como la herramienta fundamental en la modelización que posteriormente se hace de los mercados financieros.

El capítulo 3 se centra en la modelización de mercados financieros mediante ecuaciones diferenciales estocásticas, teniendo dos partes bien diferenciadas. La primera se dedica a los aspectos básicos y generales de la modelización (modelo generalizado de Black-Scholes) y en la segunda se trata de trasladar estos principios generales al mercado de bonos, considerando modelos unifactoriales para modelizar la dinámica de los tipos de interés.

En el capítulo 4 se hace un estudio empírico de la dinámica de los tipos de interés a corto plazo en España, mediante el planteamiento y estimación (en el contexto no paramétrico) de un modelo homogéneo. Se trata también de profundizar en el problema, antes mencionado, de la linealidad de la deriva del modelo planteado (mediante la estimación lineal local y contrastes de hipótesis). Además, tras la estimación no paramétrica del precio del riesgo de mercado se valoran bonos (por el método Monte Carlo) y se comparan los resultados con los obtenidos usando otros modelos.

Finalmente, el capítulo 5 recoge las conclusiones alcanzadas en este trabajo.

Además, se añade un Apéndice con las líneas de código MATLAB de los programas que ha sido necesario elaborar para llevar a cabo los cálculos surgidos en la aplicación empírica.



# Capítulo 1

## Procesos estocásticos: Martingalas y movimiento browniano

### 1.1. Introducción

En este capítulo se presenta una breve introducción a la teoría general de los procesos estocásticos y se estudia el caso particular de las martingalas, las cuales han llegado a ser una herramienta fundamental en la modelización de los mercados financieros. Se analiza con cierto detalle el ejemplo estándar de martingala continua de cuadrado integrable, que es el movimiento browniano. Este proceso es el integrador básico sobre el que se construyen las integrales estocásticas que se usan para modelizar el comportamiento de los mercados financieros. El objetivo fundamental de este capítulo y el siguiente es llevar a cabo una presentación de los conceptos y resultados necesarios para el desarrollo de este

trabajo dedicado a la modelización de los mercados financieros mediante ecuaciones diferenciales estocásticas. Se dedica una especial atención a las ideas, no a las demostraciones rigurosas, las cuales se pueden consultar en algunos de los textos teóricos dedicados a procesos estocásticos y cálculo estocástico; tales como, Doob (1953)[31], Karatzas y Shreve (1991)[66], Revuz y Yor (1991)[90], Durrett (1996)[35], Oksendal (1998)[84] y Gavilán (1999)[50].

## 1.2. Procesos estocásticos. Información

Un *proceso estocástico* es la herramienta matemática utilizada para la modelización de la dinámica de un fenómeno aleatorio. La aleatoriedad del fenómeno se capta mediante la introducción de un espacio medible  $(\Omega, \mathcal{F})$ , sobre el cual se definen diferentes variables aleatorias. De esta manera, se tiene la siguiente definición:

**Definición 1.1 (Proceso estocástico)** *Un proceso estocástico es una colección parametrizada de variables aleatorias  $X = \{X(t)\}_{t \in T}$  definidas sobre un espacio medible  $(\Omega, \mathcal{F})$  y tomando valores<sup>1</sup> en  $\mathbb{R}^d$ . El conjunto  $T$  en este trabajo se interpreta como el tiempo y por tanto está formado por los valores reales no negativos,  $T = [0, +\infty)$ .*

Por tanto, un proceso estocástico se puede considerar como una función real de dos variables  $(t, \omega)$ , como sigue:

$$\begin{aligned} X : [0, +\infty) \times \Omega &\longrightarrow \mathbb{R}^d \\ (t, \omega) &\longrightarrow X(t, \omega), \end{aligned}$$

---

<sup>1</sup>Las variables aleatorias  $X(t)$  toman valores en el espacio medible  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ , donde, como es habitual,  $\mathcal{B}(U)$  denota la menor  $\sigma$ -álgebra que contiene a todos los conjuntos abiertos del espacio topológico  $U$  ( $\sigma$ -álgebra de Borel).



donde  $X(t, \omega)$  se interpreta como el valor en el instante  $t$  dada la realización  $\omega$ . Nótese que para cada  $t \in [0, \infty)$  fijo, se tiene una variable aleatoria

$$\omega \in \Omega \rightarrow X(t, \omega) \in \mathbb{R}^d,$$

a la que denotaremos por  $X(t)$ . Por otra parte, fijando  $\omega \in \Omega$  se obtiene la función real de variable real

$$t \in [0, \infty) \rightarrow X(t, \omega) \in \mathbb{R}^d,$$

la cual se denomina *camino muestral o trayectoria* del proceso  $X$  asociado a la realización  $\omega$ . Claramente, la clase de procesos estocásticos considerada en este trabajo proporciona un modelo para abordar el estudio de un fenómeno aleatorio cuyo resultado se puede observar continuamente en el tiempo.

Respecto al concepto de igualdad, como funciones de  $t$  y  $\omega$  que son, dos procesos estocásticos  $X$  e  $Y$  definidos sobre el mismo espacio medible son iguales si  $X(t, \omega) = Y(t, \omega)$ ,  $\forall \omega \in \Omega$  y  $\forall t \in [0, \infty)$ . Sin embargo, esta condición es excesivamente fuerte y en presencia de una medida de probabilidad  $\mathbb{P}$  se puede relajar para obtener tres *conceptos de igualdad* entre procesos estocásticos.

**Definición 1.2** *Dados dos procesos estocásticos  $X$  e  $Y$  definidos sobre el mismo espacio medible  $(\Omega, \mathcal{F})$  se dice que:*

1.  *$Y$  es una modificación de  $X$  (o viceversa) si, para cada  $t \geq 0$  se cumple que*

$$\mathbb{P}[X(t) = Y(t)] = 1.$$

2.  *$X$  e  $Y$  tienen las mismas distribuciones finito-dimensionales, si dados cualesquiera*

natural  $n \geq 1$ , reales  $0 \leq t_1 < t_2 < \dots < t_n < \infty$  y  $A \in \mathcal{B}(\mathbb{R}^{n \cdot d})$ , se verifica que

$$\mathbb{P}[(X(t_1), X(t_2), \dots, X(t_n)) \in A] = \mathbb{P}[(Y(t_1), Y(t_2), \dots, Y(t_n)) \in A].$$

3.  $X$  e  $Y$  son indistinguibles si coinciden  $\mathbb{P}$ -en casi todos los caminos muestrales, es decir si

$$\mathbb{P}[X(t) = Y(t), \forall 0 \leq t < \infty] = 1.$$

La tercera definición es la más fuerte de las tres ya que, claramente implica la primera, la cual a su vez es más general que la segunda. Sin embargo, si los procesos son continuos<sup>2</sup> por la derecha o por la izquierda, lo cual incluye a todos los procesos estocásticos con los que se trata en este trabajo, los conceptos 1 y 3 son equivalentes.

Es importante observar que se puede identificar cada  $\omega$  con su correspondiente trayectoria  $t \in T \rightarrow X(t, \omega) \in \mathbb{R}^d$ . Así, se puede considerar que  $\Omega$  es un conjunto de trayectorias, y que es por tanto un subconjunto del espacio  $\tilde{\Omega} = (\mathbb{R}^d)^T$  de todas las funciones de  $T$  en  $\mathbb{R}^d$ , el cual se supone dotado de su  $\sigma$ -álgebra producto  $\mathcal{B}$ . En consecuencia, un proceso estocástico  $X$  se puede considerar también como una medida de probabilidad  $\mathbb{P}$  sobre el espacio medible  $((\mathbb{R}^d)^T, \mathcal{B})$ .

En la teoría de la integración de Lebesgue, las medidas de probabilidad se definen sobre  $\sigma$ -álgebras y las variables aleatorias se suponen medibles respecto a tales  $\sigma$ -álgebras. Por tanto, al decir que el proceso estocástico  $X = \{X(t), 0 \leq t < \infty\}$  es una colección de variables aleatorias se está suponiendo que cada  $X(t)$  es  $\mathcal{F}/\mathcal{B}(\mathbb{R}^d)$  medible. Como  $X$  es realmente una función de dos variables  $(t, \omega)$ , es conveniente tener la misma propiedad

---

<sup>2</sup>Un proceso estocástico se dice continuo si tiene  $\mathbb{P}$ -casi todas sus trayectorias continuas. Y análogamente se tienen procesos continuos por la derecha o por la izquierda.

de medibilidad conjunta respecto a la  $\sigma$ -álgebra producto. Por ello, se da la siguiente definición:

**Definición 1.3** *Un proceso estocástico  $X$  se dice medible si la función  $(t, \omega) \rightarrow X(t, \omega)$ , definida del espacio medible  $([0, \infty) \times \Omega, \mathcal{B}([0, \infty)) \otimes \mathcal{F})$  en el espacio medible  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ , es medible.*

En los procesos estocásticos considerados en este trabajo, es aplicable el teorema de Fubini, y como consecuencia se sigue que las trayectorias de los procesos medibles son funciones Borel-medibles de  $t \in [0, \infty)$  y, supuesto que todas las variables aleatorias  $X(t)$  tengan esperanzas definidas, la función  $m(t) = \mathbb{E}X(t)$  también es medible.

La característica temporal de un proceso estocástico sugiere un flujo de información a lo largo del tiempo y de esta forma, se puede hablar de la información que un observador del proceso conoce hasta un instante determinado. Para formalizar esta idea se dota al espacio medible  $(\Omega, \mathcal{F})$  de una *filtración*, la cual se define como una familia no decreciente  $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$  de  $\sigma$ -álgebras de  $\mathcal{F}$ , es decir,  $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$  para  $0 \leq s < t < \infty$ . En este marco de trabajo,  $\mathcal{F}_t$  representa la información disponible hasta el instante  $t$ , y la filtración  $\mathbb{F}$  representa el flujo de información que evoluciona con el tiempo. Claramente, la propiedad de no decrecimiento de las  $\sigma$ -álgebras  $\mathcal{F}_t$  en la filtración  $\mathbb{F}$  indica cómo la información va aumentando conforme transcurre el tiempo.

Dado un proceso estocástico  $X$  la elección más sencilla de una filtración es la generada por el propio proceso, la cual viene dada por

$$\mathcal{F}_t^X = \sigma(X(s), 0 \leq s \leq t),$$

y es la menor  $\sigma$ -álgebra respecto a la cual la variable aleatoria  $X(s)$  es medible para todo  $s \in [0, t]$ .

Por cuestiones técnicas la filtración asociada a un proceso estocástico se suele ampliar para que verifique las llamadas *condiciones usuales*, que son:

- $\mathcal{F}_0$  contiene todos los sucesos  $\mathbb{P}$ -nulos de  $\mathcal{F}$ ; y
- $\mathbb{F}$  es continua por la derecha, es decir,  $\mathcal{F}_t = \mathcal{F}_{t+} = \bigcap_{s>t} \mathcal{F}_s$ .

A partir de este punto, todos los procesos estocásticos se considerarán, porque siempre es posible esta construcción, definidos sobre espacios de probabilidad filtrados  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$  donde la filtración  $\mathbb{F}$  satisface las condiciones usuales.

Por otro lado, una propiedad deseable para los procesos estocásticos es que el valor  $X(t)$  de tal proceso en el instante  $t$ , por ejemplo el precio de un activo financiero en dicho instante, forme parte de la información disponible hasta dicho instante, lo cual lleva a la siguiente definición.

**Definición 1.4** *Un proceso estocástico  $X$  está adaptado a una filtración  $\mathbb{F}$  si  $X(t)$  es una variable aleatoria  $\mathcal{F}_t$ -medible, para cada  $t \geq 0$ .*

Es importante señalar que en las aplicaciones financieras, la filtración  $\mathbb{F}$  está básicamente generada por los precios de los activos observados. Por ello, un requisito natural y lógico es que la selección de una cartera de valores en el instante  $t$  dependa solo de la información disponible hasta dicho momento. La formalización de esta idea es exigir que la cartera esté adaptada a tal filtración. En este contexto de la modelización financiera, una clase importante de procesos estocásticos son los denominados *martingalas*, ya que

son la herramienta usada para modelizar el ruido en los sistemas dinámicos considerados para los precios de los activos. En consecuencia, vamos a comentar las principales características de estos procesos, en el contexto del presente trabajo.

## 1.3. Martingalas continuas de cuadrado integrable

Ya se ha mencionado, que una clase particularmente importante de procesos estocásticos son los denominados *procesos martingala* o simplemente *martingalas*, ya que constituyen la parte central de todo el análisis estocástico y son, además, procesos muy utilizados en las aplicaciones. Así, por ejemplo, en matemáticas financieras son una herramienta fundamental para deducir el precio de las opciones y construir estrategias de cobertura.

Antes de dar una definición precisa del concepto de martingala, es importante destacar que son modelos de juegos justos. En este contexto, una martingala se describe como sigue: Supongamos que un jugador apuesta en una sucesión de juegos con una estrategia basada en los resultados que va obteniendo. Se podría decir que el juego es justo si la cantidad de dinero que se espera que el jugador tenga en un instante futuro después del  $n$ -ésimo juego, dados los resultados hasta el  $n$ -ésimo juego, es constante. A partir de esta interpretación, se presenta la definición de martingala en un contexto unidimensional y continuo (el cual se mantendrá en toda esta sección).

**Definición 1.5** *Un proceso estocástico  $X = \{X(t), \mathcal{F}_t; 0 \leq t < \infty\}$  definido sobre un espacio de probabilidad filtrado  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$  se llama martingala si para todo*

$0 \leq s < t < \infty$  se cumple que

$$X(s) = \mathbb{E}(X(t) \mid \mathcal{F}_s) \quad \mathbb{P} - \text{casi seguro.}$$

Si el signo de igualdad anterior se reemplaza por  $\leq$  ( $\geq$ ) el proceso se llama submartingala, que corresponde a un juego favorable (supermartingala, que corresponde a un juego desfavorable).

En otras palabras, un proceso estocástico  $X$  es una martingala si la mejor predicción de su estado en el instante  $t$ , condicionada a la información disponible hasta un instante anterior  $s < t$ , es el estado en el instante  $s$ .

Las martingalas representan situaciones donde hay aleatoriedad pero no tendencia o deriva. En un contexto estadístico típico (por ejemplo, en regresión) se desea captar con un modelo cierta regularidad (señal) exhibida por los datos, los cuales están afectados de un inevitable ruido aleatorio, es decir,

$$\text{datos} = \text{señal} + \text{ruido}$$

en este marco, las martingalas se usan para modelizar la componente de ruido.

**Ejemplo 1.6** Un ejemplo típico de una  $\mathbb{F}$ -martingala  $X$  se obtiene considerando una variable aleatoria integrable cualquiera  $Y$  sobre un espacio de probabilidad filtrado  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$  y definiendo  $X(t) = \mathbb{E}[Y \mid \mathcal{F}_t]$ ,  $t \geq 0$ . De esta definición se sigue fácilmente, que sobre un intervalo de tiempo compacto  $[0, T]$  cualquier martingala  $M$  está siempre generada por su valor final mediante la fórmula

$$M(t) = \mathbb{E}[M(T) \mid \mathcal{F}_t], \quad 0 \leq t \leq T.$$

Una cuestión que surge de manera natural es cuándo una martingala  $X$  definida sobre el intervalo infinito  $[0, \infty)$  tiene una representación de esa forma, es decir, en qué condiciones existe una variable aleatoria  $X(\infty)$  tal que  $X(t) = \mathbb{E}[X(\infty) | \mathcal{F}_t]$ . En este punto, el resultado básico es el siguiente.

**Teorema 1.7** *Sea una martingala  $X = \{X(t), \mathcal{F}_t; 0 \leq t < \infty\}$  continua por la derecha, entonces las tres condiciones siguientes son equivalentes:*

1.  $X$  es una familia de variables aleatorias uniformemente integrable, es decir, 
$$\sup_{t \geq 0} \mathbb{E}|X(t)| < \infty.$$
2.  $X$  converge en  $L^1$  cuando  $t \rightarrow \infty$ ; es decir, existe una variable aleatoria  $X(\infty)$  tal que 
$$\lim_{t \rightarrow \infty} \mathbb{E}|X(t) - X(\infty)| = 0.$$
3.  $X$  converge  $\mathbb{P}$ -casi seguro cuando  $t \rightarrow \infty$  a una variable aleatoria integrable  $X(\infty)$ , tal que  $\{X(t), \mathcal{F}_t; 0 \leq t \leq \infty\}$  es una martingala.

Por otro lado, una cuestión interesante, en general, cuando se trabaja con procesos estocásticos es conocer si las propiedades que se obtienen para instantes de tiempo determinísticos  $t$ , también se mantienen para instantes de tiempo aleatorios  $T$ . Para poder profundizar en esta cuestión, es necesario establecer qué se entiende por valor del proceso en un instante aleatorio. En este sentido, dado un proceso estocástico  $X$  y un instante aleatorio  $T$ , se define el valor del proceso en dicho instante  $X(T)$  sobre el suceso  $\{T < \infty\}$  por  $X(T)(\omega) = X(T(\omega), \omega)$ . Si  $X(\infty, \omega)$  está definido para todo  $\omega$ , entonces  $X(T)$  está definido en todo  $\Omega$  tomando  $X(T)(\omega) = X(\infty, \omega)$  sobre  $\{T = \infty\}$ .

Podemos entonces hacernos, por ejemplo, la siguiente pregunta: ¿Cuándo una martingala verifica en instantes aleatorios la propiedad que la caracteriza? Es decir, cuándo una

martingala verifica que

$$X(S) = \mathbb{E}(X(T) \mid \mathcal{F}_S) \quad \mathbb{P} - \text{casi seguro}, \quad (1.1)$$

donde  $S \leq T$  son tiempos aleatorios. Es importante notar que la igualdad anterior tiene una interpretación muy intuitiva en el contexto de los juegos, ya que nos dice que la fortuna  $X(t)$  en un instante  $t$ , de un jugador que juega continuamente a un juego justo, no puede mejorar en valor esperado eligiendo un instante aleatorio en el que abandonar.

Para dar respuesta a la cuestión planteada, en primer lugar, debemos restringirnos a tiempos aleatorios que estén en algún sentido adaptados al flujo de información dado por la filtración. En este sentido se introduce el siguiente concepto.

**Definición 1.8** *Un tiempo de parada con respecto a la filtración  $\mathbb{F}$  es una variable aleatoria no negativa  $T$  tal que  $\{T \leq t\} \in \mathcal{F}_t, \forall t \geq 0$ .*

Esta definición se interpreta de la siguiente manera: En cualquier instante  $t \geq 0$  se conoce, con la información disponible hasta dicho instante, si el suceso  $\{T \leq t\}$  ha ocurrido o no. El ejemplo típico de tiempo de parada son los denominados *tiempos de alcance*, definidos como

$$T = \inf \{t \geq 0; X(t) \in B\}$$

donde<sup>3</sup>  $B$  es un conjunto de Borel<sup>4</sup>.  $T$  así definido es el primer instante en el que  $X$  alcanza al conjunto  $B$ , e intuitivamente es obvio que podemos decidir si el suceso  $\{T \leq t\}$  ha ocurrido en base a la información disponible hasta el instante  $t$ .

---

<sup>3</sup>Por convenio  $\inf \{\emptyset\} = +\infty$ .

<sup>4</sup>Para ver bajo qué condiciones exactamente este tiempo aleatorio es un tiempo de parada se puede consultar Dellacherie (1972)[28].



Para poder dar significado a la igualdad (1.1) se debe definir qué se entiende por la  $\sigma$ -álgebra  $\mathcal{F}_T$  cuando  $T$  es un tiempo de parada. La interpretación es la información disponible hasta el instante aleatorio  $T$ , la cual se consigue a través de la definición siguiente:

**Definición 1.9** *Sea  $T$  un tiempo de parada de la filtración  $\mathbb{F}$ . La  $\sigma$ -álgebra  $\mathcal{F}_T$  viene dada por los sucesos  $\{A \in \mathcal{F} \mid A \cap \{T \leq t\} \in \mathcal{F}_t, \forall t \geq 0\}$ .*

Con esta definición se tiene la propiedad deseable de que si  $X$  es un proceso adaptado con trayectorias continuas por la derecha, la variable aleatoria  $X(T)$ , que da la posición del proceso en el instante de parada  $T$ , es  $\mathcal{F}_T$ -medible, es decir, es conocida con la información disponible hasta el instante aleatorio  $T$ .

A partir de las definiciones y resultados previos, se está ya en condiciones de enunciar el resultado básico para que se verifique la igualdad (1.1) para una martingala  $X$ .

**Teorema 1.10** *Sea  $X = \{X(t), \mathcal{F}_t; 0 \leq t < \infty\}$  una martingala continua por la derecha con último elemento  $X(\infty)$ , y sean  $S < T$  dos tiempos de parada de la filtración  $\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$ . Entonces se cumple que  $\mathbb{E}[X(T)|\mathcal{F}_S] = X(S)$ ,  $\mathbb{P}$ -casi seguro.*

Por otro lado, la herramienta matemática fundamental usada en la modelización en tiempo continuo de los mercados financieros son las ecuaciones diferenciales estocásticas, las cuales se definen en términos de sus correspondientes integrales estocásticas. Los integradores básicos en estas integrales son las martingalas continuas de cuadrado integrable, cuya definición es la siguiente.

**Definición 1.11** Sea  $X = \{X(t), \mathcal{F}_t; 0 \leq t < \infty\}$  una martingala continua por la derecha. Se dice que  $X$  es de cuadrado integrable si  $\mathbb{E}(X(t)^2) < \infty, \forall t \geq 0$ . Si además  $X(0) = 0$   $\mathbb{P}$ -casi seguro, se escribirá  $X \in \mathcal{M}_2$ . Escribiremos  $X \in \mathcal{M}_2^c$  cuando  $X$  sea además continua  $\mathbb{P}$ -casi seguro.

Sin embargo, las integrales estocásticas respecto a martingalas continuas de cuadrado integrable no se pueden definir trayectoria a trayectoria en el sentido de Lebesgue-Stieljes. Esto es debido a que las trayectorias  $t \rightarrow X(t, \omega)$  son extremadamente variables. Concretamente, se tiene que tales trayectorias continuas no son de variación acotada. Antes de establecer el resultado preciso sobre la variabilidad de tales trayectorias, se define el significado de variación de orden  $p$  de un proceso estocástico  $X$ .

**Definición 1.12** Sean  $X = \{X(t), \mathcal{F}_t; 0 \leq t < \infty\}$  un proceso,  $t > 0$  fijo y  $\Pi = \{t_0, t_1, \dots, t_m\}$  con  $0 = t_0 \leq t_1 \leq \dots \leq t_m = t$ , una partición del intervalo  $[0, t]$ . Dado  $p > 0$ , se define la  $p$ -ésima variación de  $X$  sobre la partición  $\Pi$  como

$$V_t^{(p)}(X, \Pi) = \sum_{k=1}^m |X(t_k) - X(t_{k-1})|^p.$$

El diámetro o norma de la partición  $\Pi$  se define como  $\|\Pi\| = \text{Max}_{1 \leq k \leq m} |t_k - t_{k-1}|$ . Si  $V_t^{(p)}(X, \Pi)$  converge en algún sentido cuando  $\|\Pi\| \rightarrow 0$ , el límite se llama variación de orden  $p$  de  $X$  sobre  $[0, t]$ .

En este contexto, un resultado muy importante es que todos los procesos estocásticos continuos tienen una variación natural a estudiar; en el sentido de que todas las variaciones de orden mayor son cero y, excepto en el caso trivial en el que el proceso es constante, todas las variaciones de orden menor son infinito con una probabilidad positiva. Exponemos el resultado preciso.

**Proposición 1.13** Sea  $X = \{X(t), \mathcal{F}_t; 0 \leq t < \infty\}$  un proceso estocástico continuo tal que para todo  $t > 0$  fijo y para algún  $p > 0$ ,  $\lim_{\|\Pi\| \rightarrow 0} V_t^{(p)}(X, \Pi) = L(t)$  (en probabilidad), donde  $L(t)$  es una variable aleatoria no negativa. Se tiene que:

1.  $\forall q > p$ ,  $\lim_{\|\Pi\| \rightarrow 0} V_t^{(q)}(X, \Pi) = 0$  (en probabilidad), y
2.  $\forall 0 < q < p$ ,  $\lim_{\|\Pi\| \rightarrow 0} V_t^{(q)}(X, \Pi) = \infty$  (en probabilidad) sobre el suceso  $\{L(t) > 0\}$ .

De los comentarios anteriores y de este último resultado se deduce que la variación natural a considerar en una martingala continua de cuadrado integrable debe ser  $p > 1$ , y por tanto sus trayectorias son de variación no acotada, y entonces no diferenciables. En concreto, su variación natural a estudiar es la variación cuadrática, como se afirma en el siguiente resultado.

**Teorema 1.14** Sea  $X \in \mathcal{M}_2^c$ . Para  $t > 0$  y particiones  $\Pi$  de  $[0, t]$  se tiene que  $V_t^{(2)}(X, \Pi)$  converge en probabilidad cuando  $\|\Pi\| \rightarrow 0$  a una variable aleatoria no negativa, la cual denotamos por  $\langle X \rangle(t)$ . Además  $X^2 - \langle X \rangle$  es una martingala continua y el proceso variación cuadrática  $\langle X \rangle$ , es el único (salvo indistinguibilidad) proceso adaptado, continuo y creciente que se puede colocar en la expresión anterior para obtener una martingala.

Una consecuencia de enorme importancia de este resultado, es que permite una construcción  $L_2$  (en media cuadrática) de las integrales estocásticas respecto a martingalas continuas de cuadrado integrable.

De manera análoga se tiene la existencia de la variación cruzada entre dos martingalas continuas de cuadrado integrable.

**Proposición 1.15** Sean  $X, Y \in \mathcal{M}_2^c$ . Para  $t > 0$  y particiones  $\Pi$  de  $[0, t]$  se tiene que  $\sum_{k=1}^m (X(t_k) - X(t_{k-1}))(Y(t_k) - Y(t_{k-1}))$  converge en probabilidad cuando  $\|\Pi\| \rightarrow 0$  a una

variable aleatoria, la cual denotamos por  $\langle X, Y \rangle (t)$ . Por tanto,  $\langle X, X \rangle = \langle X \rangle$ . Además  $XY - \langle X, Y \rangle$  es una martingala continua y el proceso variación cruzada  $\langle X, Y \rangle$ , es el único (salvo indistinguibilidad) proceso adaptado, continuo y de variación acotada (diferencia de dos procesos no decrecientes) que se puede colocar en la expresión anterior para obtener una martingala. Además,  $\langle X, Y \rangle = \frac{1}{4} (\langle X + Y \rangle - \langle X - Y \rangle)$ .

Si  $X \in \mathcal{M}_2^c$  con variación acotada, la Proposición 1.13 lleva a que  $\langle X \rangle$  es idénticamente nulo y entonces

$$0 = X(0)^2 = \mathbb{E} [X(t)^2 - \langle X \rangle (t)] = \mathbb{E} [X(t)^2].$$

Por tanto,  $X(t) = 0$  con probabilidad 1,  $\forall t \geq 0$ . Así, por continuidad se tiene que dicha martingala debe ser idénticamente nula. En consecuencia, una martingala continua de cuadrado integrable con variación acotada debe ser constante. Otra propiedad importante que se deduce de las definiciones anteriores es que la variación cruzada entre una martingala continua de cuadrado integrable y un proceso continuo integrable de variación acotada es también idénticamente nula.

## 1.4. El movimiento browniano

El movimiento browniano o proceso de Wiener es, con mucha diferencia, el proceso estocástico más interesante e importante en muchos contextos. El nombre de este tipo de movimiento debe al botánico R. Brown, quien en 1827 observó el desplazamiento irregular de las partículas de polen suspendidas en agua. Sin embargo, el primer trabajo cuantitativo se debe a Bachelier (1900)[5], quién lo usó para estudiar las fluctuaciones de los precios de las acciones en el mercado financiero. Einstein (1905)[39] fue el primero que

explicó el movimiento browniano, para ello usó argumentos probabilísticos y dió la prueba de la existencia de moléculas y átomos. El tratamiento matemático riguroso comienza con Wiener (1923,1924)[100, 101], quién proporcionó la demostración de su existencia. En la actualidad, el campo de aplicación del movimiento browniano es muy amplio, incluyendo la modelización de las perturbaciones aleatorias en una gran variedad de sistemas físicos, biológicos y económicos. Además la integración con respecto al movimiento browniano constituye el núcleo central del cálculo estocástico<sup>5</sup>, el cual, a su vez, es la herramienta fundamental en la modelización de los mercados financieros. A continuación, se expone la definición precisa de esta clase de procesos estocásticos.

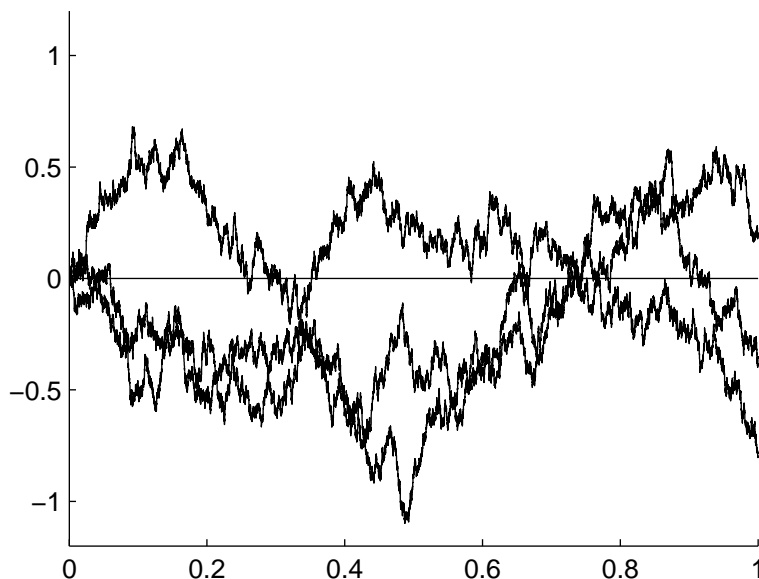
**Definición 1.16** *Un movimiento browniano o proceso de Wiener (estándar unidimensional) es un proceso estocástico continuo y adaptado  $W = \{W(t), \mathcal{F}_t; 0 \leq t < \infty\}$ , definido sobre un espacio de probabilidad filtrado  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ , verificando que  $W(0) = 0$  ( $\mathbb{P}$ -ect) y para  $0 \leq s < t$ , el incremento  $W(t) - W(s)$  es independiente de  $\mathcal{F}_s$  y sigue una distribución normal de media nula y varianza  $t - s$ , es decir,  $W(t) - W(s) \sim \mathcal{N}(0, t - s)$ ,  $\forall 0 \leq s < t$ . De manera similar se define el movimiento browniano  $W = \{W(t), \mathcal{F}_t; 0 \leq t \leq T\}$  sobre un intervalo finito  $[0, T]$  con  $T > 0$ .*

Con objeto de tener una idea del comportamiento de este tipo de procesos estocásticos, en la Figura 1.1 se muestran algunas trayectorias.

Desde un punto de vista matemático, la primera cuestión que se debe abordar al tratar con el movimiento browniano es la de su existencia. Existen varias formas de probarla (Mikosch(2004)[81]), nosotros vamos a comentar aquí su obtención como límite de cierta sucesión de caminos aleatorios, la cual es especialmente útil para realizar aproximaciones y simulaciones con un ordenador.

---

<sup>5</sup>Originado con los trabajos de Itô (1942,1944)[61, 62].



**Figura 1.1:** Varias trayectorias del movimiento browniano estándar unidimensional en el intervalo  $[0, 1]$ .

El resultado que nos muestra por qué la distribución normal juega un papel tan importante en la Teoría de la Probabilidad y la Estadística es el Teorema Central de Límite. En la versión de Lindeberg-Lévy, afirma que si  $\{\xi_i\}_{i \geq 1}$  es una sucesión de variables aleatorias independientes e idénticamente distribuidas de media  $\mathbb{E}(\xi_i) = \mu$  y varianza  $\text{Var}(\xi_i) = \sigma^2$ ,  $\forall i \geq 1$  entonces la sucesión

$$\frac{\sum_{i=1}^n \xi_i - \mathbb{E}\left(\sum_{i=1}^n \xi_i\right)}{\sqrt{\text{Var}\left(\sum_{i=1}^n \xi_i\right)}} = \frac{\sum_{i=1}^n \xi_i - n\mu}{\sigma\sqrt{n}}$$

converge en distribución, cuando  $n$  tiende a infinito, a una variable aleatoria normal estándar  $Z \sim \mathcal{N}(0, 1)$ . Usando este resultado como base se demuestra que una sucesión de caminos aleatorios apropiadamente normalizada converge a un movimiento browniano estándar.

En concreto, sea  $\{\xi_i\}_{i \geq 1}$  una sucesión de variables aleatorias independientes e idénticamente distribuidas de media  $\mu = 0$  y varianza  $\sigma^2 \in (0, \infty)$ . Considerando la sucesión de las sumas parciales  $S_0 = 0$ ,  $S_k = \sum_{i=1}^k \xi_i$  ( $k \geq 1$ ), se puede construir un proceso de tiempo continuo  $Y$  mediante interpolación lineal, como sigue:

$$Y(t) = S_{[t]} + (t - [t]) \xi_{[t]+1}, \quad t \geq 0$$

donde  $[t]$  es la parte entera de  $t$ . Modificando apropiadamente las escalas de tiempo y espacio, se obtiene a partir del proceso  $Y$  una sucesión de procesos estocásticos continuos  $\{X_n\}_{n \geq 1}$  de la forma:

$$X_n(t) = \frac{1}{\sigma\sqrt{n}} Y(nt), \quad t \geq 0.$$

Obsérvese que con  $s = \frac{k}{n}$  y  $t = \frac{k+1}{n}$ , el incremento  $X_n(t) - X_n(s) = \frac{1}{\sigma\sqrt{n}} \xi_{k+1}$  es independiente de la información generada por el proceso hasta el instante  $s$ ,  $\mathcal{F}_s^{X_n} = \sigma(\xi_1, \dots, \xi_k)$  y además tiene media cero y varianza  $t - s$ . Esto sugiere que el proceso  $X_n$  es parecido a un movimiento browniano. De hecho se tiene que la sucesión de procesos  $\{X_n\}_{n \geq 1}$  converge en distribución al movimiento browniano, resultado que se conoce como *Teorema Central del Límite Funcional o Principio de Invarianza de Donsker (1951)*[30]. Esta convergencia en distribución<sup>6</sup> significa dos cosas:

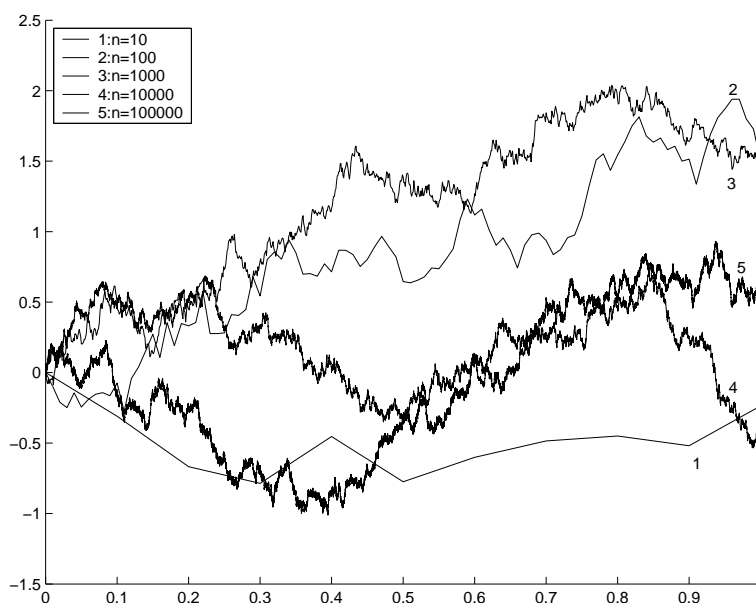
- Las distribuciones finito-dimensionales de  $X_n$  convergen a las correspondientes distribuciones finito-dimensionales del movimiento browniano.
- Como los procesos estocásticos son objetos infinito-dimensionales se necesita además una propiedad de compacidad o estrechez de la sucesión  $X_n$ .

El Teorema Central del Límite Funcional, además de proporcionar una demostración rigurosa de la existencia del movimiento browniano es un método sencillo para simular

---

<sup>6</sup>Una referencia fundamental para la convergencia en distribución de procesos estocásticos es Billingsley (1968)[6].

trayectorias brownianas. Así, del mencionado teorema se sigue que para obtener una aproximación razonable a un camino browniano, basta dibujar las trayectorias del proceso  $X_n$  para  $n$  suficientemente grande. De esta manera, en la Figura 1.2 se muestra trayectorias del proceso  $X_n$  en el intervalo<sup>7</sup>  $[0, 1]$  para diferentes valores de  $n$  y una misma realización de las variables independientes y normales estándar  $\xi_1, \dots, \xi_{100000}$ . Nótese que como el movimiento browniano es un límite en distribución, aunque se mantenga la realización de las variables aleatorias  $\xi_i$  se obtienen gráficas completamente diferentes para los distintos valores de  $n$ .



**Figura 1.2:** Varias aproximaciones a una trayectoria del movimiento browniano estándar en el intervalo  $[0, 1]$ , basadas en el Teorema Central del Límite Funcional y obtenidas a partir de una realización de tamaño 100000 de variables normales estándar.

A continuación, vamos a comentar las propiedades elementales del movimiento browniano.

<sup>7</sup>Para simular trayectorias en un intervalo cualquiera  $[0, T]$ , basta hacerlo en el intervalo  $[0, 1]$  y usar la propiedad de cambio de escala que se menciona en la sección siguiente.



En primer lugar, de la propia definición, se deduce que si  $W$  es un movimiento browniano y  $0 = t_0 < t_1 < \dots < t_n < \infty$ , entonces los incrementos  $W(t_1) - W(t_0)$ ,  $W(t_2) - W(t_1)$ ,  $\dots$ ,  $W(t_n) - W(t_{n-1})$  son independientes y la distribución de  $W(t_j) - W(t_{j-1})$  depende de los instantes  $t_j$  y  $t_{j-1}$  sólo a través de la diferencia  $t_j - t_{j-1}$ ; en concreto sigue una distribución normal de media 0 y varianza  $t_j - t_{j-1}$ . Por esta razón, se dice que el proceso  $W$  tiene *incrementos independientes y estacionarios*. Se sigue fácilmente que  $W$  es una martingala de cuadrado integrable, ya que  $E[W(t)^2] = t$  y dados  $0 \leq s < t < \infty$ :

$$\mathbb{E}[W(t) | \mathcal{F}_s] = \mathbb{E}[W(t) - W(s) | \mathcal{F}_s] + W(s) = \mathbb{E}[W(t) - W(s)] + W(s) = W(s).$$

Obtengamos su variación cuadrática. Nótese que para el movimiento browniano se tiene que

$$\mathbb{E}(W(t)) = 0, \quad \mathbb{E}(W(t)^2) = \text{Var}(W(t)) = t, \quad \text{Var}(W(t)^2) = 2t^2.$$

Así, para  $t > 0$  pequeño se sigue que la varianza de  $W(t)^2$  es despreciable comparada con su valor esperado. Esto sugiere que si se toma una partición  $\Pi$ , dada por  $0 = t_0 < t_1 < \dots < t_n = t$ , del intervalo  $[0, t]$  suficientemente fina ( $n$  grande), entonces, siendo  $\Delta W(t_i) = W(t_i) - W(t_{i-1})$  y  $\Delta t_i = t_i - t_{i-1}$ , se tendrá que la variable aleatoria

$$\sum_{i=1}^n (\Delta W(t_i))^2$$

será muy similar a su media

$$\sum_{i=1}^n \mathbb{E}(\Delta W(t_i))^2 = \sum_{i=1}^n \Delta t_i = \sum_{i=1}^n (t_i - t_{i-1}) = t,$$

lo que sugiere que la variación cuadrática del movimiento browniano viene dada por  $\langle W \rangle(t) = t$ . Para obtener una justificación rigurosa, sea una partición  $\Pi$  del intervalo

$[0, t]$  dada por  $0 = t_0 < t_1 < \dots < t_n = t$ , de aquí se sigue que:

$$\begin{aligned} V^{(2)}(W, \Pi) - t &= \sum_{i=1}^n \left\{ (W(t_i) - W(t_{i-1}))^2 - (t_i - t_{i-1}) \right\} = \\ &= \sum_{i=1}^n \left\{ (\Delta W(t_i))^2 - (\Delta t_i) \right\} = \sum_{i=1}^n Y_i, \end{aligned}$$

donde  $Y_i = (\Delta W(t_i))^2 - (\Delta t_i)$ . Teniendo en cuenta que  $\Delta W(t_i)$  sigue una distribución normal de media cero y varianza  $\Delta t_i$ , entonces las variables  $Y_i$  son de media cero, además, son independientes debido a la independencia de los incrementos de  $W$ . Por tanto,

$$\begin{aligned} \mathbb{E} \left[ (V^{(2)}(W, \Pi) - t)^2 \right] &= \mathbb{E} \left[ \left( \sum_{i=1}^n Y_i \right)^2 \right] = \text{Var} \left( \sum_{i=1}^n Y_i \right) = \sum_{i=1}^n \text{Var} (Y_i) \\ &= \sum_{i=1}^n \text{Var} (\Delta W(t_i)^2) = \sum_{i=1}^n 2 (\Delta t_i)^2 \\ &\leq 2 \|\Pi\| \sum_{i=1}^n (\Delta t_i) = 2 \|\Pi\| t, \end{aligned}$$

que converge a cero cuando  $\|\Pi\| \rightarrow 0$ . En consecuencia  $V^{(2)}(W, \Pi)$  converge en media cuadrática (y por tanto también en probabilidad) a  $t$ , con lo que la variación cuadrática del movimiento browniano sobre el intervalo  $[0, t]$  es  $\langle W \rangle(t) = t$ .

Si se aplica la propiedad anterior sobre los intervalos  $[0, t + dt]$  y  $[0, t]$  y se resta, se puede escribir el resultado formalmente introduciendo el simbolismo

$$(dW(t))^2 = d\langle W \rangle(t) = dt,$$

el cual constituye la esencia del calculo de Itô, que es eje del capítulo siguiente.

Por otro lado, el movimiento browniano es un caso particular de una familia de procesos estocásticos más general, denominados *procesos gaussianos*, que son aquellos procesos  $X = \{X(t), \mathcal{F}_t; 0 \leq t < \infty\}$  cuyas distribuciones finito-dimensionales  $(X(t_1), \dots, X(t_k))$

tienen una distribución conjunta normal,  $\forall k \geq 1$  entero y  $0 \leq t_1 < t_2 \dots < t_k$ . En consecuencia, las distribuciones finito-dimensionales de un proceso gaussiano están determinadas por su vector de medias y su matriz de varianzas-covarianzas. Teniendo en cuenta que el movimiento browniano estándar tiene incrementos normales e independientes, se tiene fácilmente (mediante una transformación lineal) que es un proceso gaussiano. Su función de medias es  $\mu(t) = \mathbb{E}(W(t)) = 0$ ,  $t \geq 0$ , mientras que su función de covarianza para  $0 \leq s < t$  viene dada por

$$\begin{aligned}
 \rho(s, t) &= \text{Cov}(W(s), W(t)) = \mathbb{E}(W(s) \cdot W(t)) \\
 &= \mathbb{E}[(W(t) - W(s) + W(s)) \cdot W(s)] \\
 &= \mathbb{E}[(W(t) - W(s)) \cdot W(s)] + \mathbb{E}[W(s)^2] \\
 &= \mathbb{E}[W(t) - W(s)] \cdot \mathbb{E}[W(s)] + s = s.
 \end{aligned} \tag{1.2}$$

Por tanto,

$$\rho(s, t) = \min\{s, t\}, \quad \forall s, t \in \mathbb{R}. \tag{1.3}$$

De esta forma se tiene caracterizado al movimiento browniano estándar  $W$  como un proceso gaussiano con trayectorias continuas, con función de media nula y función de covarianza dada por (1.3). A partir de esta caracterización, resulta fácil probar que ciertas transformaciones de  $W$  siguen dando como resultado un movimiento browniano estándar. Por ejemplo, sea el *cambio de escala* dado para  $c \geq 0$  por

$$X(t) = \frac{W(ct)}{\sqrt{c}}, \quad t \geq 0.$$

Se comprueba sin dificultad que el proceso  $X$  es un proceso gaussiano con trayectorias continuas, con función de media nula y función de covarianza dada por (1.3), en consecuencia,  $X$  también es un movimiento browniano estándar. Esta propiedad se puede usar

para obtener una trayectoria de un movimiento browniano estándar en un intervalo  $[0, T]$ , con  $T > 0$ , dada una trayectoria en el intervalo unidad  $[0, 1]$ . En concreto, bastaría con tomar  $X(t \cdot T) = W(t)\sqrt{T}$ .

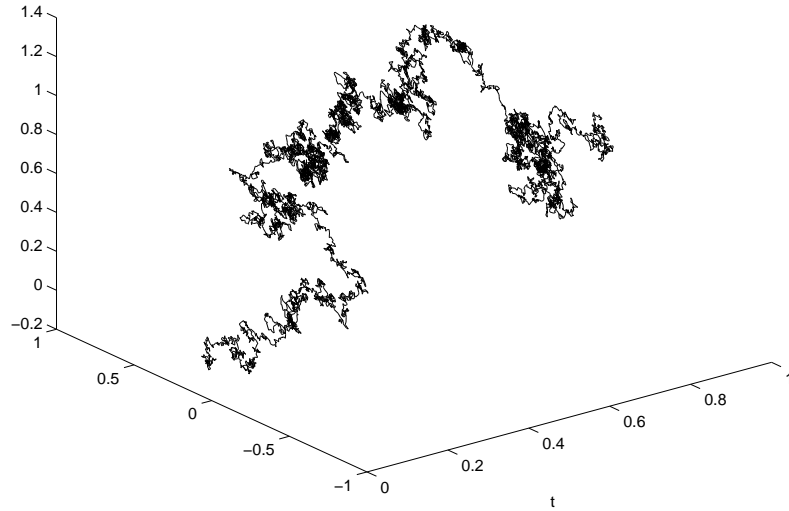
Por otro lado el concepto de movimiento browniano se puede extender sin ninguna dificultad a varias dimensiones.

**Definición 1.17** *Sea  $d$  un entero positivo y  $\mu$  una medida de probabilidad sobre  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ . Un proceso  $W = \{W(t), \mathcal{F}_t; 0 \leq t < \infty\}$  continuo y adaptado, definido sobre algún espacio de probabilidad filtrado  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$  se llama movimiento browniano  $d$ -dimensional con distribución inicial  $\mu$ , si*

1.  $P(W(0) \in \Gamma) = \mu(\Gamma)$ , para todo  $\Gamma \in \mathcal{B}(\mathbb{R}^d)$ .
2. Para  $0 \leq s < t$ , el incremento  $W(t) - W(s)$  es independiente de  $\mathcal{F}_s$  y su distribución es normal de media cero y matriz de covarianzas  $(t - s)I_d$ , donde  $I_d$  es la matriz identidad de orden  $d$ .

*Si  $\mu$  asigna medida uno a algún punto  $\{x\}$ , se dice que  $W$  es un movimiento browniano  $d$ -dimensional que comienza en  $x$ .*

La Figura 1.3 muestra una trayectoria de un movimiento browniano bidimensional comenzando en el origen. Una forma de construir un movimiento browniano  $d$ -dimensional con distribución inicial  $\mu$  es la siguiente: Sea  $X(\omega_0) = \omega_0$  la variable aleatoria identidad sobre  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mu)$  y  $d$  movimientos brownianos estándar  $W_i = \{W_i(t), \mathcal{F}_t^{W_i}; t \geq 0\}$  definidos sobre espacios de probabilidad filtrados  $(\Omega^i, \mathcal{F}^i, \mathbb{P}^i, \mathbb{F}^i)$ ,  $i = 1, \dots, d$ . Sobre el



**Figura 1.3:** Una trayectoria típica del movimiento browniano bidimensional comenzando en el origen.

espacio producto

$$(\mathbb{R}^d \times \Omega^1 \times \dots \times \Omega^d, \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{F}^1 \otimes \dots \otimes \mathcal{F}^d, \mu \times \mathbb{P}^1 \times \dots \times \mathbb{P}^d),$$

sea el proceso

$$W(t, \omega) = X(\omega_0) + (W_1(t, \omega_1), \dots, W_d(t, \omega_d)),$$

y  $\mathcal{F}_t = \mathcal{F}_t^W$ . Entonces, se sigue que  $W = \{W(t), \mathcal{F}_t; 0 \leq t < \infty\}$  es ciertamente un movimiento browniano  $d$ -dimensional con distribución inicial  $\mu$ .

En relación con los procesos martingala, es importante señalar que para un movimiento browniano  $d$ -dimensional se verifica fácilmente que cada  $\{W_i(t), \mathcal{F}_t; 0 \leq t < \infty\}$  ( $i = 1, \dots, d$ ) es una martingala continua de cuadrado integrable con

$$\langle W_i, W_j \rangle(t) = \delta_{ij} \cdot t, \quad \forall i, j = 1, \dots, d$$

o de manera formal

$$dW_i(t) \cdot dW_j(t) = \delta_{ij} \cdot t, \quad \forall i, j = 1, \dots, d.$$

Supongamos que se observa un movimiento browniano  $d$ -dimensional con distribución inicial  $\mu$  justo hasta el instante presente  $s$ , con  $0 \leq s < t$ , siendo  $t$  un instante futuro. En particular, se observa el valor en el instante presente  $W(s)$ , el cual se va a denotar por  $y$ , es decir,  $W(s) = y$ . Condicionando a todas esas observaciones (pasadas y presente), ¿Cuál es la probabilidad de que el proceso en un instante futuro  $W(t)$  esté en un determinado subconjunto  $\Gamma \in \mathcal{B}(\mathbb{R}^d)$ ? Obsérvese que  $W(t) = (W(t) - W(s)) + W(s)$  y que el incremento  $W(t) - W(s)$  es independiente de todas las observaciones hasta el instante presente  $s$  y se distribuye igual que  $W(t-s)$  bajo<sup>8</sup>  $\mathbb{P}^0$ . Por otra parte,  $W(s)$  depende de esas observaciones; en concreto, se conoce que  $W(s) = y$ . Por tanto la suma  $(W(t) - W(s)) + W(s)$  se distribuye como  $W(t-s)$  bajo  $\mathbb{P}^y$ . En consecuencia, el conocimiento de todo el pasado y el presente hasta el instante  $s$ , no aporta más información sobre el futuro  $W(t)$  que conocer solo el valor presente  $W(s)$ ; en otras palabras, para  $0 \leq s < t$  y  $\Gamma \in \mathcal{B}(\mathbb{R}^d)$  se tiene que

$$\mathbb{P}^\mu(W(t) \in \Gamma | \mathcal{F}_s) = \mathbb{P}^\mu(W(t) \in \Gamma | W(s)), \quad \mathbb{P}^\mu - \text{ect.} \quad (1.4)$$

Un proceso adaptado que verifica la igualdad anterior (1.4) se dice que es un *proceso markoviano*. Así, el movimiento browniano es un proceso markoviano. Es más, la igualdad (1.4) se verifica para un movimiento browniano si los instantes presente y futuro son aleatorios, es decir para  $S < T$  tiempos de parada, en lugar de los instantes fijos  $s < t$ . Por este motivo, se dice que el movimiento browniano es un *proceso fuertemente markoviano*.

---

<sup>8</sup>En general  $\mathbb{P}^\mu$  denota la distribución de un movimiento browniano con distribución inicial  $\mu$ . Si  $\mu(\{x\}) = 1$  se escribe  $\mathbb{P}^x$ .

Además, condicionado a  $W(s) = y$ , el hecho que  $W(t)$  se distribuya como  $W(t - s)$  bajo  $\mathbb{P}^y$  significa que el movimiento browniano “comienza de nuevo” en cada instante de tiempo. Por tanto, para  $0 \leq s < t$  y  $\Gamma \in \mathcal{B}(\mathbb{R}^d)$  se tiene que

$$\mathbb{P}^\mu(W(t) \in \Gamma | W(s) = y) = \mathbb{P}^y(W(t - s) \in \Gamma), \quad \mathbb{P}^\mu W(s)^{-1} - \text{ect} \quad y \in \mathbb{R}^d.$$

Como una primera prueba de la utilidad del movimiento browniano, se tiene que algunos procesos gaussianos y no gaussianos de gran relevancia práctica se deducen a partir de éste. Veamos algunos de ellos, partiendo de un movimiento browniano estándar unidimensional,  $W = \{W(t), \mathcal{F}_t; 0 \leq t < \infty\}$ .

**Ejemplo 1.18 (Puente browniano)** *Sea el proceso  $X = \{X(t), \mathcal{F}_t; 0 \leq t \leq 1\}$  definido por*

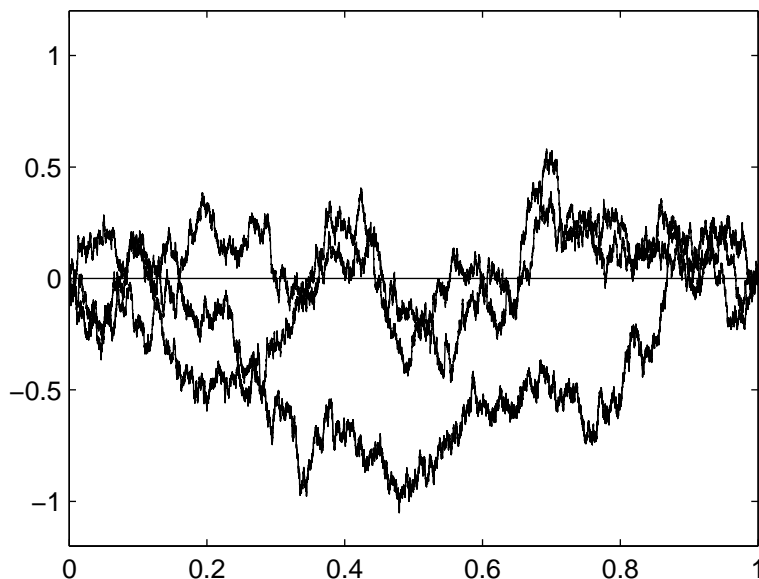
$$X(t) = W(t) - t \cdot W(1).$$

*Claramente*

$$X(0) = W(0) - 0 \cdot W(1) = 0 \quad \text{y} \quad X(1) = W(1) - 1 \cdot W(1) = 0.$$

Por esta razón (siempre empieza y termina en el origen) el proceso  $X$  se llama puente browniano. La Figura 1.4 muestra varias trayectorias de este proceso. Se puede comprobar fácilmente que  $X$  es el proceso gaussiano caracterizado por las siguientes funciones de media y de covarianza:

$$\mathbb{E}(X(t)) = 0 \quad \text{y} \quad \text{Cov}(X(s), X(t)) = \min\{s, t\} - s \cdot t, \quad s, t \in [0, 1].$$



**Figura 1.4:** Varias trayectorias del proceso puente browniano.

**Ejemplo 1.19 (Movimiento browniano con deriva)** Sea el proceso estocástico  $X = \{X(t), \mathcal{F}_t; 0 \leq t < \infty\}$  definido mediante

$$X(t) = \mu t + \sigma W(t), \quad t \geq 0,$$

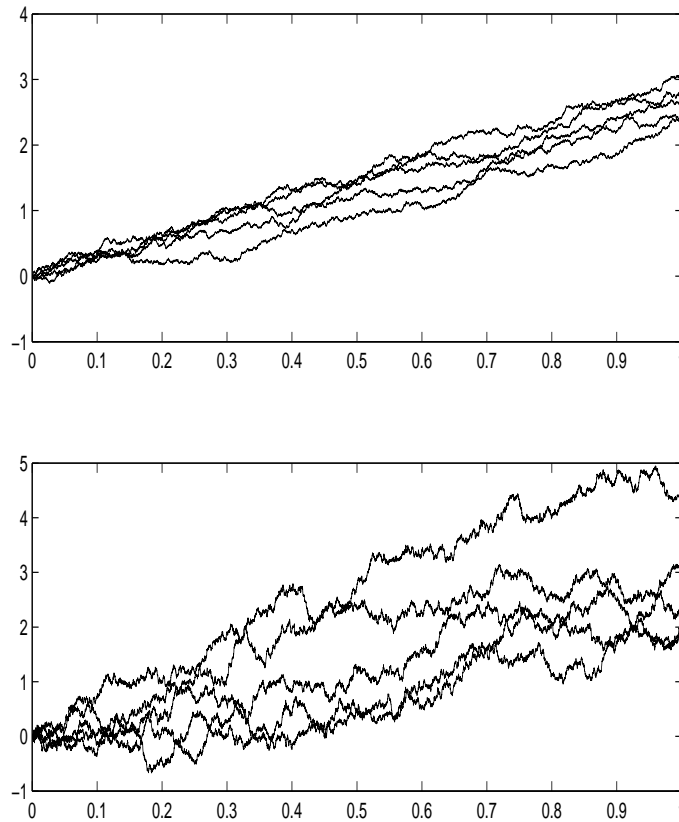
para  $\mu \in \mathbb{R}$  y  $\sigma > 0$  constantes. Se sigue claramente que  $X$  es un proceso gaussiano con funciones de media y covarianza dadas por

$$\mathbb{E}(X(t)) = \mu t \quad \text{y} \quad \text{Cov}(X(s), X(t)) = \sigma^2 \min\{s, t\}.$$

La función esperanza  $\mathbb{E}(X(t)) = \mu t$  es la denominada deriva determinística del proceso y esencialmente determina la tendencia de los caminos muestrales, mientras que  $\sigma$ , llamado coeficiente de volatilidad o de difusión, determina la mayor o menor variabilidad de las trayectorias. Esto puede observarse en la Figura 1.5 donde se han simulado trayectorias



del proceso  $X$  para diferentes valores de la constante  $\sigma > 0$ .

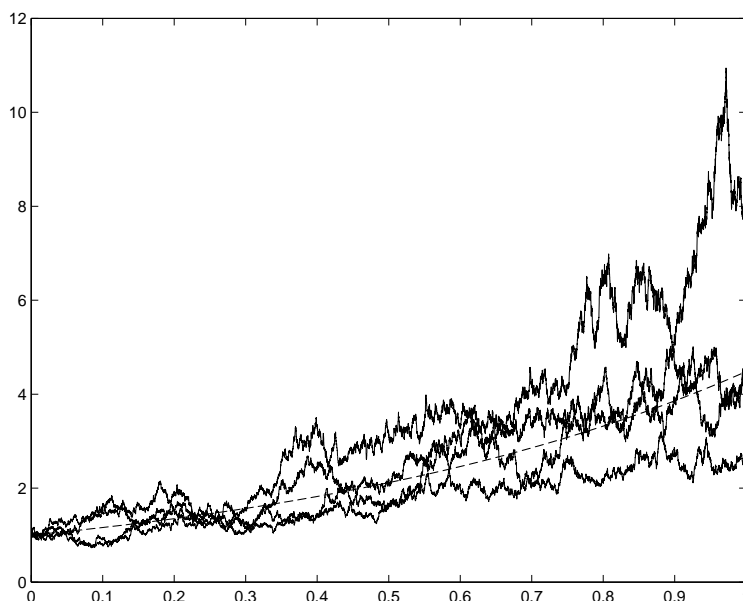


**Figura 1.5:** Ejemplos de trayectorias del movimiento browniano con deriva. Todas se han generado con  $\mu = 3$ , las de la gráfica superior con  $\sigma = 0.5$  y las de la inferior con  $\sigma = 1.5$ .

**Ejemplo 1.20 (Movimiento browniano geométrico)** Con el importante trabajo de Bachelier (1900)[5], anteriormente mencionado, donde se describen los precios de los activos con riesgo (acciones, índices bursátiles, tipos de cambio, etc.) mediante un movimiento browniano nace una nueva área de aplicación de los procesos estocásticos. Sin embargo, un movimiento browniano, al ser un proceso gaussiano, puede tomar valores negativos, lo que no es una propiedad deseable para un precio. En sus conocidos trabajos, Black y Scholes (1973)[11] y Merton (1973)[79] sugieren otro proceso estocástico para modelizar estos precios. Tal proceso  $X = \{X(t), \mathcal{F}_t; 0 \leq t < \infty\}$  viene dado por

$$X(t) = e^{\mu t + \sigma W(t)}, \quad t \geq 0,$$

es decir, la exponencial de un movimiento browniano con deriva y se denomina movimiento browniano geométrico. Claramente, se trata de un proceso no gaussiano, ya que no se tiene la normalidad de  $X(t)$ . En la Figura 1.6 se muestran varias trayectorias del movimiento browniano geométrico. Con el objetivo de conocer un poco mejor este proce-



**Figura 1.6:** Varias trayectorias del movimiento browniano geométrico  $X(t) = e^{t+W(t)}$ . La línea discontinua representa la función de medias.

so, se va a obtener sus funciones de esperanza y covarianza. Para ello se hace uso de la conocida función generatriz de momentos de una variable aleatoria normal estándar  $Z$ , la cual viene dada por

$$M_Z(t) = \mathbb{E} (e^{tZ}) = e^{\frac{t^2}{2}}, \quad t \in \mathbb{R}.$$

La función esperanza del movimiento browniano geométrico es la siguiente, para  $t \in \mathbb{R}$ :

$$\begin{aligned} \mathbb{E}(X(t)) &= \mathbb{E}(e^{\mu t + \sigma W(t)}) = \mathbb{E}(e^{\mu t + \sigma \sqrt{t}Z}) = e^{\mu t} M_Z(\sigma \sqrt{t}) \\ &= e^{\mu t} e^{\frac{t\sigma^2}{2}} = e^{t(\mu + \frac{\sigma^2}{2})}. \end{aligned}$$

En cuanto a la función de covarianza, teniendo en cuenta que dados  $0 \leq s < t$  el incremento  $W(t) - W(s)$  es independiente de  $W(s)$  y sigue una distribución normal de media

cero y varianza  $t - s$ , se tiene que:

$$\begin{aligned}
 \text{Cov}(X(s), X(t)) &= \mathbb{E}(X(s) \cdot X(t)) - \mathbb{E}(X(s)) \cdot \mathbb{E}(X(t)) \\
 &= e^{\mu(s+t)} \mathbb{E}(e^{\sigma(W(s)+W(t))}) - e^{(s+t)(\mu+\frac{1}{2}\sigma^2)} \\
 &= e^{\mu(s+t)} \mathbb{E}(e^{\sigma[(W(t)-W(s))+2W(s)]}) - e^{(s+t)(\mu+\frac{1}{2}\sigma^2)} \\
 &= e^{\mu(s+t)} \mathbb{E}(e^{\sigma(W(t)-W(s))}) \mathbb{E}(e^{2\sigma W(s)}) - e^{(s+t)(\mu+\frac{1}{2}\sigma^2)} \\
 &= e^{\mu(s+t)} M_Z(\sigma\sqrt{t-s}) M_Z(2\sigma\sqrt{s}) - e^{(s+t)(\mu+\frac{1}{2}\sigma^2)} \\
 &= e^{(s+t)(\mu+\frac{1}{2}\sigma^2)} (e^{\sigma^2 s} - 1).
 \end{aligned}$$

En particular, la función varianza de un movimiento browniano geométrico es

$$\text{Var}(X(t)) = e^{(2\mu+\sigma^2)t} (e^{\sigma^2 t} - 1), \quad t \in \mathbb{R}.$$

Para terminar esta sección dedicada al movimiento browniano, se comentan las propiedades más relevantes de sus trayectorias.

A pesar de que, por definición, las trayectorias de un movimiento browniano son continuas, un simple vistazo a las trayectorias simuladas de la Figura 1.1 nos convence de que son funciones extremadamente irregulares. Una prueba de ello, es que son *funciones no diferenciables en todo punto*. La existencia de tales funciones fue descubierta en el siglo *XIX* y fueron consideradas en un principio como curiosidades, lejos de poder ser usadas en aplicaciones prácticas. Sin embargo, hoy día el movimiento browniano es un proceso con aplicación práctica en una gran multitud de disciplinas, como se ha comentado al principio de este capítulo.

Una última indicación de la irregularidad de los caminos de un movimiento browniano, es que, habiendo visto que es una *martingala de cuadrado integrable con variación cuadrática* dada por  $\langle W \rangle(t) = t$ , se tiene que es un *proceso de variación no acotada*, es decir hay una longitud infinita en cualquier trozo de trayectoria entre dos puntos. Este hecho, tiene la consecuencia importante de que la integral estocástica con respecto a un movimiento browniano no se puede definir de la manera clásica como una integral de Lebesgue-Stieljes. Sin embargo, es precisamente el hecho de tener una variación cuadrática finita el que permite construir la integral estocástica mediante convergencia en media cuadrática, tal y como se lleva a cabo en §2.2.

En consecuencia, si se pretende recoger el movimiento del precio de un activo financiero mediante una ecuación diferencial que haga uso del movimiento browniano para modelizar el ruido, se debe desarrollar un cálculo estocástico que será diferente del cálculo clásico, debido a las propiedades antes mencionadas de las trayectorias de este proceso. Éste, es el objetivo fundamental del capítulo siguiente.

# Capítulo 2

## Introducción al cálculo estocástico

### 2.1. Motivación

Teniendo en cuenta que el objetivo fundamental de este trabajo es la modelización de los mercados financieros, supongamos que se pretende representar la evolución en el tiempo del precio  $X$  de un activo financiero, una acción por ejemplo, mediante una ecuación diferencial clásica. Una posibilidad sería usar una ecuación del tipo:

$$\frac{dX(t)}{dt} = a(t) \cdot X(t), \quad X(0) = x_0 \quad (\text{constante}),$$

donde  $a(t)$  es la tasa de apreciación de la acción, midiendo la tasa instantánea de cambio de  $X$  en el instante  $t$ . No parece razonable en este contexto que el parámetro  $a(t)$  se suponga determinista, ya que está sujeto a las decisiones de los participantes en el mercado, a la evolución incierta de la compañía que ha puesto las acciones en el mercado, a la evolución

de los mercados de otros países, etc. Por tanto, en realidad se tiene

$$a(t) = r(t) + \text{“ruido”},$$

donde se supone que  $r(t)$  es no aleatorio y el término de ruido es desconocido. Como primer objetivo, se debe dar una interpretación matemática razonable al “ruido” en la ecuación resultante

$$\frac{dX(t)}{dt} = (r(t) + \text{“ruido”}) \cdot X(t)$$

o de forma más general en ecuaciones de la forma:

$$\frac{dX(t)}{dt} = b(t, X(t)) + \sigma(t, X(t)) \cdot \text{“ruido”},$$

donde  $b$  y  $\sigma$  son funciones dependientes del tiempo y de la variable. Es lógico buscar un proceso estocástico  $U$  que represente al término ruido que añade aleatoriedad a  $\sigma$ , con lo que la ecuación anterior se expresa como:

$$\frac{dX(t)}{dt} = b(t, X(t)) + \sigma(t, X(t)) \cdot U(t). \tag{2.1}$$

Respecto a la elección del término de ruido  $U$ , en base a muchas situaciones, por ejemplo en física, es aconsejable que tenga las siguientes propiedades:

1. Si  $t_1 \neq t_2$  entonces  $U(t_1)$  y  $U(t_2)$  son independientes.
2. Es estacionario, es decir, la distribución conjunta de  $U(t_1 + h), \dots, U(t_k + h)$  no depende de  $h$ , para cualesquiera  $t_1, \dots, t_k \in \mathbb{R}^+$ .
3.  $\mathbb{E}(U(t)) = 0$  para todo  $t \geq 0$ .

Sin embargo, no existe ningún proceso estocástico con trayectorias continuas satisfa-

ciendo las condiciones anteriores (Oksendal(1998)[84]). Para evitar esta complicación, se escribe una versión discreta de la ecuación (2.1). Sea pues  $0 = t_0 < t_1 < \dots < t_m = t$ , se tiene para  $k = 0, \dots, m$

$$X(t_{k+1}) - X(t_k) = b(t_k, X(t_k)) \Delta t_k + \sigma(t_k, X(t_k)) U(t_k) \Delta t_k, \quad (2.2)$$

donde  $\Delta t_k = t_{k+1} - t_k$ . Reemplazando la notación  $U(t_k) \Delta t_k$  por  $\Delta V(t_k) = V(t_{k+1}) - V(t_k)$  donde  $V$  es algún proceso estocástico adecuado, las tres hipótesis anteriores sugieren que este proceso  $V$  debe tener incrementos independientes estacionarios de media cero. En este sentido, resulta que el único proceso con caminos continuos que lo verifica es el movimiento browniano  $W$  (Knight(1981)[69]). Así, sustituyendo  $V(t) = W(t)$  en la ecuación (2.2) se obtiene:

$$X(t_k) = x_0 + \sum_{j=0}^{k-1} b(t_j, X(t_j)) \Delta t_j + \sum_{j=0}^{k-1} \sigma(t_j, X(t_j)) \Delta W(t_j). \quad (2.3)$$

La cuestión que surge de forma inmediata es si existe, en algún sentido, el límite del lado derecho de (2.3); en tal caso, aplicando la notación usual en integración tendremos

$$X(t) = x_0 + \int_0^t b(u, X(u)) du + \int_0^t \sigma(u, X(u)) dW(u).$$

La primera integral se puede ver como una integral de Lebesgue definida trayectoria a trayectoria, pero hay que justificar la existencia de la integral estocástica que aparece en el segundo sumando.

Como las trayectorias del movimiento browniano no son de variación acotada no es posible definir la integral estocástica anterior como una integral de Lebesgue-Stieljes sobre cada trayectoria. Sin embargo, sí es posible construir tales integrales de manera análoga

a la integral de Lebesgue-Stieljes, como se lleva a cabo en la sección siguiente.

## 2.2. Integral estocástica: Propiedades básicas

El concepto de integral estocástica fue introducido inicialmente por Paley, Wiener y Zygmund (1933)[85] para integrandos no aleatorios, y posteriormente por Itô (1942,1944)[61, 62] para el movimiento browniano en la generalidad en la que se presenta en esta sección. El objetivo fundamental en integración estocástica es dar significado a la expresión:

$$\int_0^t X(u)dY(u),$$

para una clase adecuada de procesos  $X$  (integrando) e  $Y$  (integrador). Recordando el ejemplo de la sección anterior nos centramos en el caso donde el integrador es un movimiento browniano:  $Y(t) = W(t)$ . Antes de abordar este caso particular, queremos mencionar que para una mayor generalidad en la construcción de la integral estocástica se puede consultar Karatzas y Shreve (1991)[66] y Revuz y Yor (1991)[90], donde se lleva a cabo el desarrollo para una martingala continua  $Y$ . Un tratamiento en toda su generalidad se desarrolla en Meyer (1976)[80] y Protter (2004)[88].

Sea entonces un movimiento browniano estándar  $W = \{W(t), \mathcal{F}_t; 0 \leq t < \infty\}$  definido sobre un espacio de probabilidad filtrado  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ . Al intentar dar un significado a las integrales estocásticas respecto al movimiento browniano,

$$I_X(t) = \int_0^t X(u)dW(u), \tag{2.4}$$



lo primero que se observa es que al ser  $W$  una martingala continua de cuadrado integrable, se trata de un proceso de variación no acotada sobre cualquier intervalo. En consecuencia, las integrales en la expresión (2.4) no se pueden definir trayectoria a trayectoria, para cada  $\omega \in \Omega$  separadamente, como integrales ordinarias de Lebesgue-Stieljes. Sin embargo, como se probó en §1.4, el movimiento browniano tiene variación cuadrática finita dada por  $\langle W \rangle (t) = t$ , y es precisamente este hecho el que permite construir la integral estocástica con respecto a una clase apropiada de integrandos  $X$ .

La construcción de la integral estocástica se lleva a cabo de forma similar a una integral de Lebesgue-Stieljes, es decir, comenzando con una clase elemental de integrandos  $X$  y después extendiendo la integral para una clase más general de integrandos, aproximando éstos por procesos elementales, que son los denominados procesos simples, cuya definición se introduce a continuación.

**Definición 2.1 (Proceso simple)** *Los procesos simples son aquellos procesos estocásticos  $X$  para los que existe una partición dada por una sucesión de números reales estrictamente creciente  $\{t_n\}_{n=0}^{\infty}$  con  $t_0 = 0$  y  $\lim_{n \rightarrow \infty} t_n = \infty$ , así como una sucesión de variables aleatorias  $\{\xi_n\}_{n=0}^{\infty}$  uniformemente acotada  $\left( \sup_{n \geq 0} |\xi_n(\omega)| \leq c < \infty, \quad \forall \omega \in \Omega \right)$ , tal que  $\xi_n$  es  $\mathcal{F}_{t_n}$ -medible y*

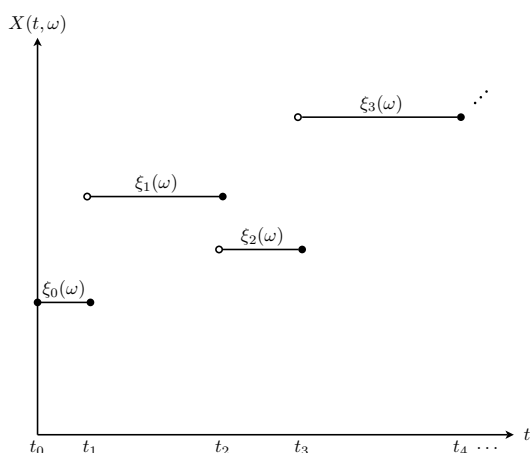
$$X(t, \omega) = \xi_0(\omega) \cdot I_{\{0\}}(t) + \sum_{i=0}^{\infty} \xi_i(\omega) \cdot I_{(t_i, t_{i+1}]}(t), \quad 0 \leq t < \infty, \quad \omega \in \Omega.$$

Un ejemplo de trayectoria típica de un proceso simple se muestra en la Figura 2.1.

Para estos procesos, se tiene la siguiente definición natural de integral estocástica:

**Definición 2.2 (Integral estocástica de procesos simples)** *Dado un proceso simple*

---



**Figura 2.1:** Una trayectoria de un proceso simple

como en la Definición 2.1, se define la integral estocástica  $I_X(t) = \int_0^t X(u)dW(u)$ ,  $0 \leq t < \infty$ , como<sup>1</sup>:

$$\begin{aligned} I_X(t) &= \sum_{i=0}^{\infty} \xi_i (W(t \wedge t_{i+1}) - W(t \wedge t_i)) \\ &= \sum_{i=0}^{n-1} \xi_i (W(t_{i+1}) - W(t_i)) + \xi_n (W(t) - W(t_n)), \end{aligned} \tag{2.5}$$

donde  $n \geq 0$  es el único entero para el que  $t \in [t_n, t_{n+1})$ .

Nótese que en la definición anterior de integral estocástica se han tomado los incrementos hacia delante (*forward*) del movimiento browniano. En concreto, en el término genérico  $\xi_i (W(t_{i+1}) - W(t_i))$ , la variable aleatoria  $\xi_i$  es  $\mathcal{F}_{t_i}$ -medible donde  $t_i$  es el extremo inferior de intervalo  $[t_i, t_{i+1}]$  sobre el que se toma el incremento del proceso  $W$ . Este aspecto es un punto esencial en la teoría (ver Ejemplo 2.7), tanto desde el punto de vista matemático como económico.

Las propiedades básicas de la integral estocástica definida sobre los procesos simples

<sup>1</sup>Se sigue la notación usual  $s \wedge t = \min\{s, t\}$ , para  $s, t \in \mathbb{R}$ .

son las siguientes:

**Proposición 2.3** *La integral estocástica de procesos simples  $X, Y$  respecto a un movimiento browniano estándar  $W$  definida mediante (2.5) verifica las siguientes propiedades, para  $0 \leq s < t < \infty$ :*

1.  $I_X(0) = 0$ ,  $\mathbb{P}$ -ect.
2. (Propiedad martingala)  $\mathbb{E}[I_X(t)|\mathcal{F}_s] = I_X(s)$ ,  $\mathbb{P}$ -ect.
3. (Isometría de Itô)  $\mathbb{E}[I_X(t)^2] = \mathbb{E} \int_0^t X(u)^2 du$ .
4. (Variación cuadrática)

$$\mathbb{E}[(I_X(t) - I_X(s))^2 | \mathcal{F}_s] = \mathbb{E} \left[ \int_s^t X(u)^2 du | \mathcal{F}_s \right], \quad \mathbb{P}\text{-ect.}$$

5. (Variación cruzada)

$$\mathbb{E}[(I_X(t) - I_X(s))(I_Y(t) - I_Y(s)) | \mathcal{F}_s] = \mathbb{E} \left[ \int_s^t X(u)Y(u) du | \mathcal{F}_s \right], \quad \mathbb{P}\text{-ect.}$$

6. (Linealidad)  $I_{\alpha X + \beta Y} = \alpha I_X + \beta I_Y$ ,  $\mathbb{P}$ -ect,  $\forall \alpha, \beta \in \mathbb{R}$ .

En consecuencia, la integral estocástica  $I_X$  es una martingala continua de cuadrado integrable con proceso variación cuadrática dado por

$$\langle I_X \rangle (t) = \int_0^t X(u)^2 du,$$

y la variación cruzada entre las dos integrales estocásticas  $I_X$  e  $I_Y$  es

$$\langle I_X, I_Y \rangle (t) = \int_0^t X(u)Y(u) du.$$

*Demostración.* Las propiedades 1 y 6 son obvias. La 2 se deduce observando que, para cualesquiera  $0 \leq s < t < \infty$  se tiene que

$$\mathbb{E} [\xi_i (W(t \wedge t_{i+1}) - W(t \wedge t_i)) | \mathcal{F}_s] = \xi_i (W(s \wedge t_{i+1}) - W(s \wedge t_i)),$$

lo cual se puede verificar al considerar cada uno de los tres casos  $s \leq t_i$ ,  $t_i < s \leq t_{i+1}$  y  $t_{i+1} < s$  usando la  $\mathcal{F}_{t_i}$ -medibilidad de  $\xi_i$ .

Para establecer 4 en primer lugar se observa que para toda martingala de cuadrado integrable  $M$ , se tiene que el producto de los incrementos sobre intervalos disjuntos tiene media cero. En efecto, dados  $s < t \leq u < v$ , se verifica que

$$\begin{aligned} & \mathbb{E} [(M(v) - M(u))(M(t) - M(s))] \\ &= \mathbb{E} [\mathbb{E} [(M(v) - M(u))(M(t) - M(s)) | \mathcal{F}_u]] \\ &= \mathbb{E} [(M(t) - M(s)) \cdot \mathbb{E} [(M(v) - M(u)) | \mathcal{F}_u]] \\ &= \mathbb{E} [(M(t) - M(s)) \cdot 0] = 0. \end{aligned}$$

Teniendo en cuenta esta propiedad con  $t_{m-1} \leq s < t_m$  y  $t_n \leq t < t_{n+1}$  se obtiene

$$\begin{aligned} & \mathbb{E} [(I_X(t) - I_X(s))^2 | \mathcal{F}_s] \\ &= \mathbb{E} \left[ \left\{ \xi_{m-1} (W(t_m) - W(s)) + \sum_{i=m}^{n-1} \xi_i (W(t_{i+1}) - W(t_i)) \right. \right. \\ & \quad \left. \left. + \xi_n (W(t) - W(t_n)) \right\}^2 | \mathcal{F}_s \right] \\ &= \mathbb{E} \left[ \left\{ \xi_{m-1}^2 (W(t_m) - W(s))^2 + \sum_{i=m}^{n-1} \xi_i^2 (W(t_{i+1}) - W(t_i))^2 \right. \right. \\ & \quad \left. \left. + \xi_n^2 (W(t) - W(t_n))^2 \right\} | \mathcal{F}_s \right] \\ &= \mathbb{E} \left[ \left\{ \xi_{m-1}^2 (\langle W \rangle(t_m) - \langle W \rangle(s)) + \sum_{i=m}^{n-1} \xi_i^2 (\langle W \rangle(t_{i+1}) - \langle W \rangle(t_i)) \right. \right. \\ & \quad \left. \left. + \xi_n^2 (\langle W \rangle(t) - \langle W \rangle(t_n)) \right\} | \mathcal{F}_s \right] \\ &= \mathbb{E} \left[ \int_s^t X(u)^2 d\langle W \rangle(u) | \mathcal{F}_s \right] = \mathbb{E} \left[ \int_s^t X(u)^2 du | \mathcal{F}_s \right]. \end{aligned}$$

Haciendo  $s = 0$  en 4 y tomando esperanzas se obtiene 3. Un cálculo similar lleva a 5. ♣

La isometría de Itô anterior sugiere que la integral estocástica  $\int_0^t X(u)dW(u)$  se debería definir solo para procesos medibles y adaptados con

$$\mathbb{E} \left( \int_0^t X(u)^2 du \right) < \infty, \quad \forall t \geq 0.$$

Entonces se puede transferir la convergencia sobre un espacio  $L^2$  conveniente de procesos estocásticos a un espacio  $L^2$  conveniente de martingalas, lo que proporciona una teoría  $L^2$  de la integral estocástica.

Teniendo en cuenta que en las aplicaciones financieras, existe un intervalo de tiempo fijo  $[0, T]$  en el que se trabaja, el requerimiento anterior se convierte en

$$\mathbb{E} \left( \int_0^T X(u)^2 du \right) < \infty.$$

Por analogía a la construcción de la integral de Lebesgue–Stieljes se define la integral estocástica para una clase apropiada de integrandos que se puedan aproximar por procesos simples. La clase de los procesos en  $L^2$  resulta ser la clase natural.

**Definición 2.4** *Se dice que un proceso  $X$  pertenece a la clase  $L^2[0, t]$  cuando*

1.  $X$  es un proceso  $(\mathcal{B}([0, \infty)) \otimes \mathcal{F})$ -medible,
2.  $X$  está adaptado a la filtración  $\mathbb{F}$  y
3.  $\mathbb{E} \left( \int_0^t X(u)^2 du \right) < \infty$ .

*Se dice que el proceso  $X$  pertenece a la clase  $L^2$  cuando  $X \in L^2[0, t]$  para todo  $t \geq 0$ .*

Dado un proceso  $X \in L^2$  se puede aproximar por una sucesión de procesos simples  $\{X_n\}_{n \geq 1}$  de manera que para  $t \geq 0$

$$\mathbb{E} \int_0^t (X_n(u) - X(u))^2 du \longrightarrow 0, \quad \text{cuando } n \rightarrow \infty.$$

Para cada  $n$ , las integrales de los procesos simples  $\int_0^t X_n(u) dW(u)$  están bien definidas y son variables aleatorias que convergen en  $L^2$  a otra cierta variable aleatoria que se define como la integral estocástica, es decir,

$$I_X(t) = \int_0^t X(u) dW(u) = \lim_{n \rightarrow \infty} \int_0^t X_n(u) dW(u) \quad (\text{límite en } L^2).$$

El proceso integral estocástica  $I_X = \{I_X(t), \mathcal{F}_t; 0 \leq t < \infty\}$  así construido tiene una versión continua, que es la que se toma por definición como la integral. Además, resulta que la integral estocástica así definida no depende de la sucesión de procesos simples aproximantes  $\{X_n\}_{n \geq 1}$  elegida.

Resumiendo, se tiene la siguiente definición de integral estocástica para procesos de  $L^2$  como integrandos y un movimiento browniano como integrador (los detalles teóricos de esta construcción de la integral estocástica y sus propiedades y pueden consultarse en Oksendal[84], Karatzas y Shreve (1991)[66] entre otros):

**Definición 2.5 [La integral de Itô]** *Dado un proceso  $X \in L^2$ , se define la integral estocástica de  $X$  como el proceso continuo definido mediante*

$$I_X(t) = \int_0^t X(u) dW(u) = \lim_{n \rightarrow \infty} \int_0^t X_n(u) dW(u) \quad (\text{límite en } L^2), \quad (2.6)$$

donde  $\{X_n\}_{n \geq 1}$  es cualquier sucesión de procesos simples verificando que

$$\mathbb{E} \int_0^t (X_n(u) - X(u))^2 du \longrightarrow 0, \quad \text{cuando } n \rightarrow \infty.$$

Todas las propiedades de la integral estocástica para procesos simples de la Proposición 2.3 se extienden por paso al límite, en el sentido de (2.6), a todos los procesos de  $L^2$ .

Realmente, como se comenta al principio del capítulo, la construcción anterior de la integral estocástica no es exclusiva para el movimiento browniano como integrador. En concreto, si  $M$  es una martingala continua de cuadrado integrable, para cualquier proceso  $X$  medible y adaptado con  $\mathbb{E} \int_0^t X^2(u) d\langle M \rangle(u) < \infty \quad (\forall t \geq 0)$  se define de manera análoga la integral estocástica de  $X$  respecto a  $M$ ,  $I_X^M(t) = \int_0^t X(u) dM(u) \quad (\forall t \geq 0)$ , obteniendo como resultado una martingala continua de cuadrado integrable con variación cuadrática  $\langle I_X^M \rangle(t) = \int_0^t X^2(u) d\langle M \rangle(u) < \infty \quad (\forall t \geq 0)$ . Como para los propósitos de este trabajo conocer la integral respecto al movimiento browniano es suficiente, no se continúa con esta generalización.

La integral estocástica que se ha definido, se puede generalizar sin dificultad al caso multidimensional de la siguiente manera:

**Definición 2.6 (Integral estocástica multidimensional)** *Sea  $W = (W_1, \dots, W_k)$  un movimiento browniano  $k$ -dimensional estándar definido sobre un espacio de probabilidad filtrado  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$  y sea  $X$  una matriz de procesos donde cada  $X_{ij} \in L^2$ , para todo  $i = 1, \dots, m$  y  $j = 1, \dots, k$ . En esta situación, se define la integral estocástica de  $X$  con*

respecto a  $W$ , denotada por

$$I_X(t) = \int_0^t X(u) dW(u) = \int_0^t \begin{pmatrix} X_{11}(u) & \cdots & X_{1k}(u) \\ \vdots & \ddots & \vdots \\ X_{m1}(u) & \cdots & X_{mk}(u) \end{pmatrix} \begin{pmatrix} dW_1(u) \\ \vdots \\ dW_k(u) \end{pmatrix},$$

como el vector columna  $m$ -dimensional cuya componente  $i$ -ésima es la siguiente suma de integrales de Itô unidimensionales:

$$\sum_{j=1}^k \int_0^t X_{ij}(u) dW_j(u).$$

**Ejemplo 2.7** Para terminar con esta sección dedicada a la definición y propiedades elementales de la integral estocástica se va a calcular la integral de un movimiento browniano con respecto a él mismo, es decir,  $\int_0^t W(u) dW(u)$ .

Sea  $\epsilon$  un número real en  $[0, 1]$  y  $\Pi = \{t_0, t_1, \dots, t_m\}$  una partición del intervalo  $[0, t]$  con  $0 = t_0 < t_1 < \dots < t_m = t$ . Para la integral que se pretende calcular sea la suma aproximante

$$S_\epsilon(\Pi) = \sum_{i=0}^{m-1} [(1 - \epsilon) W(t_i) + \epsilon W(t_{i+1})] [W(t_{i+1}) - W(t_i)].$$

Veamos que si  $\|\Pi\| \rightarrow 0$  entonces  $S_\epsilon(\Pi)$  converge en  $L^2$  a  $\frac{1}{2}W(t)^2 + (\epsilon - \frac{1}{2})t$ .

En efecto, se puede escribir:

$$\begin{aligned} S_\epsilon(\Pi) &= \frac{1}{2} \sum_{i=0}^{m-1} [W(t_{i+1})^2 - W(t_i)^2] + (\epsilon - \frac{1}{2}) \sum_{i=0}^{m-1} [W(t_{i+1}) - W(t_i)]^2 \\ &= \frac{1}{2}W(t)^2 + (\epsilon - \frac{1}{2}) \sum_{i=0}^{m-1} [W(t_{i+1}) - W(t_i)]^2, \end{aligned}$$



usando la independencia y el carácter gaussiano de los incrementos del movimiento browniano, se tiene que:

$$\begin{aligned}
 & \mathbb{E} \left[ S_\varepsilon(\Pi) - \left( \frac{1}{2} W(t)^2 + \left( \varepsilon - \frac{1}{2} \right) t \right) \right]^2 \\
 &= \left( \varepsilon - \frac{1}{2} \right) \mathbb{E} \left[ \sum_{i=0}^{m-1} (W(t_{i+1}) - W(t_i))^2 - t \right]^2 \\
 &= \left( \varepsilon - \frac{1}{2} \right) \mathbb{E} \left\{ \sum_{i=0}^{m-1} [(W(t_{i+1}) - W(t_i))^2 - (t_{i+1} - t_i)] \right\}^2 \\
 &= \left( \varepsilon - \frac{1}{2} \right) \sum_{i=0}^{m-1} \mathbb{E} [(W(t_{i+1}) - W(t_i))^2 - (t_{i+1} - t_i)]^2 \\
 &= \left( \varepsilon - \frac{1}{2} \right) \sum_{i=0}^{m-1} \mathbb{E} (W(t_{i+1}) - W(t_i))^4 - \sum_{i=0}^{m-1} (t_{i+1} - t_i)^2 \\
 &= 2 \left( \varepsilon - \frac{1}{2} \right) \sum_{i=0}^{m-1} (t_{i+1} - t_i)^2 \leq 2 \left( \varepsilon - \frac{1}{2} \right) t \|\Pi\|,
 \end{aligned}$$

lo que concluye la demostración.

Se pueden destacar las situaciones siguientes:

- Cuando  $\varepsilon = 0$ , y solo en ese caso, es decir, cuando  $W$  está evaluado en el extremo izquierdo de cada intervalo  $[t_i, t_{i+1}]$  en la suma aproximante, se tiene que el valor de la integral es  $\frac{1}{2} W(t)^2 - \frac{1}{2} t$ , el cual se prueba sin dificultad que es una martingala. Se tiene así el valor de la integral de Itô, sin más que observar que el proceso  $X^\Pi(s) = \sum_{i=0}^{m-1} W(t_i) I_{(t_i, t_{i+1}]}$ , ( $0 \leq s \leq t$ ), que es constante a trozos, está el  $L^2$ , su integral es  $S_0(\Pi)$  y verifica que

$$\begin{aligned}
 \mathbb{E} \int_0^t |X^\Pi(s) - W(s)|^2 ds &= \sum_{i=0}^{m-1} \int_{t_i}^{t_{i+1}} \mathbb{E} |W(t_i) - W(s)|^2 ds \\
 &= \sum_{i=0}^{m-1} \int_{t_i}^{t_{i+1}} (s - t_i) ds = \frac{1}{2} \sum_{i=0}^{m-1} (t_{i+1} - t_i)^2 \rightarrow 0, \text{ cuando } \|\Pi\| \rightarrow 0.
 \end{aligned}$$

- En el caso  $\varepsilon = \frac{1}{2}$ , es decir, cuando  $W$  está evaluado en el punto medio en cada intervalo  $[t_i, t_{i+1}]$  en la suma aproximante, se obtiene la llamada integral de Fisk-Stratonovich, la cual se denota por  $\int_0^t W(s)_\circ dW(s)$  y obedece a las reglas usuales

del cálculo tales como

$$\int_0^t W(s) \circ dW(s) = \frac{1}{2} W(t)^2.$$

- Finalmente, en el caso  $\epsilon = 1$ , es decir, cuando  $W$  está evaluado en el extremo derecho en cada intervalo  $[t_i, t_{i+1}]$  en la suma aproximante, se obtiene la denominada integral retrospectiva de Itô (McKean (1969)[78], pág. 35).

Llegados a este punto, queda claro que una tarea es desarrollar una teoría de la integración y otra muy diferente es calcular la integral en cualquier caso específico de interés. No se puede esperar repetir de forma satisfactoria el proceso llevado a cabo en el ejemplo anterior en cualquier otra situación. De la misma manera que se dispone de un cálculo integral para la integral de Riemann, el cual proporciona herramientas para efectuar cálculos más o menos mecánicos, se necesita un *cálculo estocástico* para la integral de Itô. Este viene dado por *la regla de Itô*, la cual se presenta a continuación.

### 2.2.1. La regla de Itô

La regla de Itô es la herramienta más importante del cálculo estocástico, ya que proporciona una manera de calcular y manipular integrales estocásticas.

Básicamente, la regla de Itô es el análogo estocástico de la regla de la cadena de la diferenciación clásica. Para un mejor entendimiento de esta regla de la cadena estocástica, y siguiendo a Mikosch (2004)[81], se va a comenzar recordando la regla de la cadena determinística. Se usa la notación habitual  $f'(t), f''(t), \dots$  para las derivadas clásicas de una función  $f$  en un punto  $t$ .

Dadas  $f$  y  $g$  funciones diferenciables, la regla de la cadena clásica afirma que

$$[f(g(u))]' = f'(g(u)) \cdot g'(u),$$

o en forma integral

$$f(g(t)) - f(g(0)) = \int_0^t f'(g(u)) \cdot g'(u) du = \int_0^t f'(g(u)) \cdot dg(u).$$

Si se particulariza la expresión anterior para el caso  $f(u) = u^2$ , con  $g(0) = 0$  se obtiene que

$$\frac{1}{2}g(t)^2 = \int_0^t g(u) \cdot dg(u).$$

Este simple resultado prueba que no se puede reemplazar en la expresión anterior la función determinística  $g(t)$  por un camino muestral  $W(t)$  del movimiento browniano para obtener una versión estocástica de la regla de la cadena, ya que se conoce del Ejemplo 2.7 que la integral estocástica  $\int_0^t W(u)dW(u)$  tiene el valor  $\frac{1}{2}W(t)^2 - \frac{1}{2}t$ , es decir, el valor  $\frac{1}{2}W(t)^2$  de la regla de la cadena clásica, corregido por  $-\frac{1}{2}t$ . Esto sugiere que se debe encontrar un término de corrección a la regla de la cadena clásica. Si se observa detenidamente el Ejemplo 2.7 se tiene que dicho término de corrección resulta del límite en media cuadrática de  $\sum_{i=1}^n (W(t_i) - W(t_{i-1}))^2$ , es decir, de la variación cuadrática del movimiento browniano.

Se ofrece ahora, un argumento heurístico de la validez de la regla de la cadena clásica, el cual motiva la regla de Itô. En concreto, usando el lenguaje de los diferenciales, la regla clásica de la cadena se puede escribir como

$$df(g) = f'dg.$$

Este diferencial se puede interpretar como el término de primer orden en el desarrollo de Taylor:

$$f(g(t) + dg(t)) - f(g(t)) = f'(g(t)) dg(t) + \frac{1}{2} f''(g(t)) (dg(t))^2 + \dots$$

Aquí,  $dg(t) = g(t + dt) - g(t)$  es el incremento de la función  $g$  en el intervalo  $[t, t + dt]$ . Bajo las condiciones adecuadas, el segundo término y los de orden mayor en este desarrollo son despreciables para  $dt$  pequeño. Si ahora se supone que  $f$  es una función de clase  $C^2$  y se reemplaza  $g(t)$  en la expresión anterior por una trayectoria del movimiento browniano  $W(t)$ , se obtiene

$$f(W(t) + dW(t)) = f'(W(t)) dW(t) + \frac{1}{2} f''(W(t)) (dW(t))^2 + \dots \quad (2.7)$$

El cálculo de la variación cuadrática del movimiento browniano llevado a cabo en §1.4 lleva a interpretar el diferencial  $(dW(t))^2$  como  $dt$ . Entonces, a diferencia del caso determinístico, la contribución del término de segundo orden en el desarrollo de Taylor (2.7) no se puede despreciar<sup>2</sup>. Integrando en ambos lados de (2.7) en un sentido formal y despreciando los términos de orden superior a dos, se obtiene para  $0 \leq t < \infty$ :

$$\int_0^t dfW(u) = f(W(t)) - f(W(0)) = \int_0^t f'(W(u)) dW(u) + \frac{1}{2} \int_0^t f''(W(u)) du. \quad (2.8)$$

La primera integral en la expresión (2.8) es la integral estocástica de Itô de  $f'(W(t))$  y la segunda es la integral de Lebesgue de  $f''(W(t))$ . Nótese que además a la integral simbólica  $\int_0^t dfW(u)$  se le ha asignado el valor natural  $f(W(t)) - f(W(0))$ .

**Ejemplo 2.8** Una primera aplicación de esta versión simple de la regla de Itô proporcionada por la expresión (2.8) se obtiene tomando  $f(t) = t^2$ . En ese caso particular  $f'(t) = 2t$

---

<sup>2</sup>Los términos de orden superior al segundo si se pueden despreciar ya que, según la Proposición 1.13, las variaciones de orden mayor a dos son nulas.

y  $f''(t) = 2$  y la regla de Itô lleva a

$$W(t)^2 - W(0)^2 = 2 \int_0^t W(u) dW(u) + \int_0^t du,$$

es decir,

$$\int_0^t W(u) dW(u) = \frac{1}{2} W(t)^2 - \frac{1}{2} t,$$

expresión que ya se conoce del Ejemplo 2.7.

**Ejemplo 2.9 (La función exponencial no es la exponencial de Itô.)** En el cálculo clásico, la función exponencial  $f(t) = \exp(t) = e^t$  tiene la llamativa propiedad de que  $f'(t) = f(t)$ . De forma equivalente,

$$f(t) - f(0) = \int_0^t f(u) du.$$

Se plantea ahora la existencia de un proceso que sea la exponencial de Itô, es decir, que verifique la siguiente ecuación integral estocástica análoga a la anterior:

$$X(t) - X(0) = \int_0^t X(u) dW(u), \quad 0 \leq t < \infty. \quad (2.9)$$

Teniendo en cuenta que para el movimiento browniano las reglas clásicas de la integración no se cumplen, no parece razonable esperar que la exponencial de Itô venga dada por  $f(W(t)) = e^{W(t)}$ . En efecto, haciendo uso del Lema de Itô con  $f(t) = e^t$ , se deduce que

$$e^{W(t)} - e^{W(0)} = \int_0^t e^{W(u)} dW(u) + \frac{1}{2} \int_0^t e^{W(u)} du.$$

Como la segunda integral tiene un valor positivo, no se verifica la relación (2.9). Para

encontrar la exponencial de Itô, se debe extender la regla de Itô a procesos más generales.

Como una primera extensión de la regla de Itô dada por (2.8), se va a considerar procesos estocásticos de la forma  $f(t, W(t))$ , donde la función  $f(t, x)$  tiene derivadas parciales continuas de al menos segundo orden. Modificando apropiadamente el desarrollo de Taylor dado en (2.7) se obtiene la siguiente generalización:

$$f(t + dt, W(t + dt)) - f(t, W(t)) = \frac{\partial}{\partial t} f(t, W(t)) dt + \frac{\partial}{\partial x} f(t, W(t)) dW(t) + \frac{1}{2} \left[ \frac{\partial^2}{\partial t^2} f(t, W(t)) (dt)^2 + \frac{\partial^2}{\partial x^2} f(t, W(t)) (dW(t))^2 + 2 \frac{\partial^2}{\partial t \partial x} f(t, W(t)) dt dW(t) \right] + \dots$$

Por las razones antes comentadas, los términos de orden mayor al segundo en la expresión anterior son despreciables. Además, también lo son los términos con factores  $dt \cdot dW(t) = d\langle t, W(t) \rangle$  (la variación cruzada de una martingala continua de cuadrado integrable y proceso de variación acotada es nula) y  $(dt)^2 = d\langle t \rangle$  (la variación cuadrática de un proceso de variación acotada es nula), y como antes  $(dW(t))^2$  se interpreta como  $dt$ . De esta manera se llega a la siguiente regla de Itô en versión diferencial:

$$df(t, W(t)) = \left[ \frac{\partial}{\partial t} f(t, W(t)) + \frac{1}{2} \frac{\partial^2}{\partial x^2} f(t, W(t)) \right] dt + \frac{\partial}{\partial x} f(t, W(t)) dW(t), \quad (2.10)$$

que es una expresión simbólica de la siguiente igualdad integral:

$$f(t, W(t)) = f(0, W(0)) + \int_0^t \left[ \frac{\partial}{\partial u} f(u, W(u)) + \frac{1}{2} \frac{\partial^2}{\partial x^2} f(u, W(u)) \right] du + \int_0^t \frac{\partial}{\partial x} f(u, W(u)) dW(u). \quad (2.11)$$

**Ejemplo 2.10 (La exponencial de Itô)** Con la extensión de la regla de Itô dada por (2.11) se puede probar que la exponencial de Itô viene dada por el proceso  $f(t, W(t)) = e^{W(t) - \frac{1}{2}t}$ .

En efecto, si se considera la función  $f(t, x) = e^{x - \frac{1}{2}t}$ , la regla de Itô dada por la igualdad (2.11) se transforma en:

$$\begin{aligned} f(t, W(t)) - f(0, W(0)) &= \int_0^t \left[ -\frac{1}{2}f(u, W(u)) + \frac{1}{2}f(u, W(u)) \right] du \\ &+ \int_0^t f(u, W(u)) dW(u) = \int_0^t f(u, W(u)) dW(u). \end{aligned}$$

A continuación, con objeto de introducir otra generalización de la regla de Itô, se necesita el concepto de proceso de Itô, el cual se pasa a definir:

**Definición 2.11 (Proceso de Itô)** *Un proceso estocástico  $X$  se dice que es un proceso de Itô si admite la siguiente representación:*

$$X(t) = x_0 + \int_0^t \mu(u) du + \int_0^t \sigma(u) dW(u), \quad (2.12)$$

donde  $\mu$  y  $\sigma$  son dos procesos adaptados.

Por tanto, un proceso de Itô viene dado por su valor inicial ( $x_0$ ), más un proceso continuo y adaptado que es de variación acotada (la primera integral en la descomposición), más una martingala continua de cuadrado integrable (la segunda de las integrales en la descomposición).

Es habitual, con el objeto de usar una notación menos engorrosa, escribir la igualdad (2.12) en forma diferencial como:

$$dX(t) = \mu(t) dt + \sigma(t) dW(t) \quad (2.13)$$

$$X(0) = x_0. \quad (2.14)$$

En cuyo caso se dice que el proceso  $X$  tiene un *diferencial estocástico* dado por (2.13) con una condición inicial dada por (2.14). Obsérvese que la notación diferencial  $dX(t) = \mu(t) dt + \sigma(t) dW(t)$  es puramente formal y no tiene más significado que una abreviación de la expresión integral (2.12). Sin embargo, desde un punto de vista intuitivo resulta mucho más natural que la correspondiente expresión integral, y esto se debe a que el diferencial estocástico proporciona la “dinámica infinitesimal” del proceso  $X$  en el intervalo  $[t, t + dt]$ , siendo  $\mu(t)$  un término de velocidad local y  $\sigma(t)$  un término que amplifica la perturbación gaussiana dada por el movimiento browniano (ver ecuaciones (2.25) y (2.26)). Por esta razón, se dice que  $\mu$  es la deriva (local) y  $\sigma$  el coeficiente de difusión o volatilidad (local) del proceso  $X$ . Se observa entonces que formalmente el incremento infinitesimal  $dX(t)$  viene dado por un término de deriva local más un término ruido gaussiano.

Sea un proceso estocástico dado por  $f(t, X(t))$ , donde  $f : [0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}$  es una función con derivadas parciales de segundo orden continuas. Un argumento similar con un desarrollo de Taylor, lleva a que, sorprendentemente, la estructura diferencial dada por una deriva más un ruido gaussiano aditivo queda preservada. En concreto, dicho diferencial se expresa como:

$$df(t, X(t)) = \left( \frac{\partial f}{\partial t} + \mu \frac{\partial f}{\partial x} + \frac{1}{2} \sigma^2 \frac{\partial^2 f}{\partial x^2} \right) dt + \sigma \frac{\partial f}{\partial x} dW(t), \quad (2.15)$$

donde por razones de simplicidad se han suprimido muchas variables. Así, por ejemplo, el término  $\mu \frac{\partial f}{\partial x}$ , es una abreviación de  $\mu(t) \cdot \frac{\partial f}{\partial x}(t, X(t))$ , y análogamente para los otros términos.

La expresión (2.15) se puede reescribir de una forma más fácil usando las siguientes



reglas de multiplicación formales:

$$\begin{aligned} (dt)^2 &= 0, \\ dt \cdot dW(t) &= 0, \\ (dW(t))^2 &= dt. \end{aligned}$$

La primera de ellas pone de manifiesto que la variación cuadrática de un proceso de variación acotada es cero; la segunda indica que la variación cruzada entre un proceso de variación acotada y una martingala continua de cuadrado integrable es cero; y la tercera expresa que la variación cuadrática de la martingala continua de cuadrado integrable particular  $W$  (el movimiento browniano) es  $t$ . A partir de estas igualdades, la regla de Itô (2.15) se puede reescribir como:

$$df = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} (dX)^2. \quad (2.16)$$

Expresión que se puede extender fácilmente con las ideas que ya se han usado a procesos más generales, las semimartingalas continuas.

**Definición 2.12** *Una semimartingala continua es un proceso estocástico adaptado  $X = \{X(t), \mathcal{F}_t; 0 \leq t < \infty\}$  que se puede expresar como*

$$X(t) = X_0 + A(t) + M(t), \quad \mathbb{P} - ect,$$

donde  $X_0$  es una variable aleatoria (el valor inicial de la semimartingala),  $A = \{A(t), \mathcal{F}_t; 0 \leq t < \infty\}$  es un proceso continuo y adaptado de variación acotada con  $A(0) = 0$  (la parte de variación acotada de la semimartingala) y  $M = \{M(t), \mathcal{F}_t; 0 \leq t < \infty\}$  es un proceso de  $\mathcal{M}_2^c$  (la parte martingala de la semimartingala).

Para este tipo de procesos las reglas de multiplicación formales anteriores quedan de la siguiente manera:

$$\begin{aligned} (dA(t))^2 &= 0, \\ dA(t) \cdot dM(t) &= 0, \\ (dX(t))^2 &= d\langle X \rangle(t) = (dM(t))^2 = d\langle M \rangle(t), \end{aligned}$$

y la regla de Itô dada por (2.16) sigue siendo válida para esta clase de procesos estocásticos. Una aplicación de la misma, lleva al análogo estocástico de una fórmula bien conocida del cálculo clásico, la fórmula de integración por partes, la cual se introduce en el siguiente ejemplo:

**Ejemplo 2.13 (Fórmula de integración por partes)** *Sean dos semimartingalas continuas  $X(t) = X_0 + A_1(t) + M_1(t)$ ,  $Y(t) = Y_0 + A_2(t) + M_2(t)$  con  $A_1$  y  $A_2$  procesos de variación acotada y  $M_1$  y  $M_2$  martingalas. Entonces, se cumple la siguiente fórmula de integración por partes:*

$$X(t)Y(t) = X(0)Y(0) + \int_0^t X(u)dY(u) + \int_0^t Y(u)dX(u) + \langle M_1, M_2 \rangle(t),$$

o en notación diferencial:

$$dX(t)Y(t) = X(t)dY(t) + Y(t)dX(t) + d\langle M_1, M_2 \rangle(t).$$

*El resultado anterior es una consecuencia directa de la regla de Itô (2.16). En efecto, para deducirla en el caso  $X = Y$  basta aplicarla con  $f(x) = x^2$  y para obtenerla en caso general es suficiente expresar este caso particular en  $X + Y$ , en  $X - Y$  y restar los resultados. Esta fórmula estocástica de la integración por partes sigue poniendo de*

manifiesto cómo el cálculo estocástico difiere del clásico, apareciendo unos términos extra relacionados con las variaciones cuadráticas de los procesos

**Ejemplo 2.14 (La exponencial estocástica)** *Se plantea ahora el problema de encontrar la función exponencial estocástica en un contexto más general que el del Ejemplo 2.10. En concreto, dada una semimartingala continua  $X$  se busca una semimartingala continua  $Z$  que satisfaga la relación*

$$Z(t) - Z(0) = \int_0^t Z(u) dX(u), \quad 0 \leq t < \infty. \quad (2.17)$$

*La existencia y unicidad de soluciones de la ecuación diferencial estocástica (2.17) fue probada por Cathérine Doléans (1970)[29]. Por esta razón, la exponencial estocástica se conoce con el nombre de exponencial de Doléans. Dicha exponencial viene dada por*

$$Z(t) = Z(0) \cdot \mathcal{E}(X)(t),$$

donde

$$\mathcal{E}(X)(t) = \exp \left\{ X(t) - \frac{1}{2} \langle X \rangle (t) \right\}.$$

*La demostración de que ciertamente  $Z$  verifica la ecuación (2.17) es una consecuencia directa de la regla de Itô, ya que se cumple:*

$$d\mathcal{E}(X)(t) = \mathcal{E}(X)(t) dX(t) - \frac{1}{2} \mathcal{E}(X)(t) + \frac{1}{2} \mathcal{E}(X)(t),$$

*con lo cual se ha probado la existencia. Veamos la unicidad.*

Aplicando la regla de Itô al proceso  $Y(t) = \exp\{-X(t) + \frac{1}{2}\langle X \rangle(t)\}$ , se sigue:

$$\begin{aligned} dY(t) &= -Y(t) dX(t) + \frac{1}{2}Y(t) d\langle X \rangle(t) + \frac{1}{2}Y(t) d\langle X \rangle(t) \\ &= -Y(t) dX(t) + Y(t) d\langle X \rangle(t). \end{aligned}$$

Sea ahora  $Z(t)$  cualquier solución a la ecuación (2.17), entonces por la fórmula de integración por partes del Ejemplo (2.13) se tiene que:

$$\begin{aligned} d[Z(t)Y(t)] &= Z(t) dY(t) + Y(t) dZ(t) + d\langle Y, Z \rangle(t) \\ &= Z(t) [-Y(t) dX(t) + Y(t) d\langle X \rangle(t)] + Y(t) [Z(t) dX(t)] \\ &\quad + [-Y(t)Z(t) d\langle X \rangle(t)], \end{aligned}$$

donde la última igualdad se sigue debido a que  $Z(t) = Z(0) + \int_0^t Z(u) dX(u)$  e  $Y(t) = Y(0) - \int_0^t Y(u) dX(u) +$  una integral de Stieljes, y esta integral de Stieljes no contribuye a la variación cruzada  $\langle Y, Z \rangle(t) = \int_0^t Z(u) [-Y(u)] d\langle X \rangle(u) = -\int_0^t Z(u) Y(u) d\langle X \rangle(u)$ . En consecuencia  $d[Z(t)Y(t)] = 0$ , y por tanto  $Z(t)Y(t)$  es una constante (que debe ser  $Z(0)$ , ya que  $Y(0) = 1$ ); entonces  $Z(t) = \frac{Z(0)}{Y(t)} = Z(0) \cdot \mathcal{E}(X)(t)$ .

Para concluir esta sección, se da la versión multidimensional de la regla de Itô, de la que se deducen todos los casos anteriores, y se puede justificar mediante un argumento similar al usado en las unidimensionales. Demostraciones rigurosas se pueden encontrar en la bibliografía citada al comienzo de §2.2.

**Teorema 2.15** Sean  $M = \{M(t) = (M_1(t), \dots, M_d(t)), \mathcal{F}_t; 0 \leq t < \infty\}$  un vector de martingalas de  $\mathcal{M}_2^c$ ,  $A = \{A(t) = (A_1(t), \dots, A_d(t)), \mathcal{F}_t; 0 \leq t < \infty\}$  un vector de procesos de variación acotada con  $A(0) = 0$ ,  $X(t) = X(0) + M(t) + A(t)$  ( $0 \leq t < \infty$ ) donde  $X(0)$  es un vector aleatorio  $\mathcal{F}_0$ -medible en  $\mathbb{R}^d$  y  $f : [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}$  una función

de clase  $C^{1,2}$ , entonces se tiene  $\mathbb{P}$ -ect- y para  $0 \leq t < \infty$ :

$$f(t, X(t)) = f(0, X(0)) + \int_0^t \frac{\partial}{\partial u} f(u, X(u)) du + \sum_{i=1}^d \int_0^t \frac{\partial}{\partial x_i} f(u, X(u)) dX_i(u) + \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \int_0^t \frac{\partial^2}{\partial x_i \partial x_j} f(u, X(u)) d\langle M_i, M_j \rangle(u).$$

**Ejemplo 2.16 (Fórmula de integración por partes)** Como aplicación de la regla de Itô multidimensional al caso particular  $f(x_1, x_2) = x_1 x_2$ , se deduce fácilmente la fórmula de integración por partes vista en el Ejemplo 2.13.

## 2.3. Ecuaciones diferenciales estocásticas

Desde los comienzos del análisis infinitesimal, el estudio de las ecuaciones diferenciales ha constituido uno de los pilares básicos en el desarrollo y las aplicaciones del Análisis Matemático, destacando además por la gran potencia de sus métodos y resultados, a pesar de su aparente sencillez.

El concepto de ecuación diferencial es realmente sencillo, y viene dado por una relación funcional del tipo:

$$b(t, x(t), x'(t), x''(t), \dots) = 0, \quad 0 \leq t \leq T, \quad (2.18)$$

donde aparece una función desconocida del tiempo  $x(t)$  y sus derivadas  $x'(t), x''(t)$ , etc. Bajo ciertas condiciones de regularidad de la función  $b$ , se demuestra que existe una única solución a la ecuación (2.18), dada unas condiciones iniciales.

Las ecuaciones diferenciales más simples son las de primer orden, que son aquellas en las que aparece solo  $t$ ,  $x(t)$  y su primera derivada  $x'(t)$ , las cuales se suelen escribir de la siguiente manera:

$$x'(t) = b(t, x(t)), \quad x(0) = x_0, \quad (2.19)$$

para una función conocida  $b(t, x(t))$ . También es común escribir la ecuación (2.19) en su forma diferencial siguiente:

$$dx(t) = b(t, x(t)) dt, \quad x(0) = x_0. \quad (2.20)$$

Así, si se interpreta  $x(t)$  como la posición en el instante  $t$  de un punto que se mueve en la recta 1-dimensional, la ecuación (2.20) proporciona el cambio de posición del punto en un intervalo de tiempo pequeño  $[t, t + dt]$ . En concreto, la igualdad (2.20) indica que tal cambio de posición  $dx(t) = x(t + dt) - x(t)$  es proporcional al incremento de tiempo con factor  $b(t, x(t))$ . En otras palabras, la ecuación (2.20) establece que la velocidad  $\frac{dx(t)}{dt} = x'(t)$  del punto es una función dada del tiempo  $t$  y de su posición  $x(t)$ .

Al tratar con cualquier ecuación diferencial, estocástica o no, las dos primeras cuestiones que se deben resolver son la existencia y la unicidad de las soluciones. Antes de considerar el caso de las ecuaciones diferenciales estocásticas, se presentan dos ejemplos sencillos de ecuaciones diferenciales ordinarias:

**Ejemplo 2.17** *La ecuación diferencial*

$$x'(t) = 3x(t)^{\frac{2}{3}}, \quad x(0) = 0, \quad (2.21)$$

*admite infinitas soluciones. En concreto, dado cualquier  $a > 0$ , la función  $x(t) = (t - a)^3 \cdot I_{\{t > a\}}$ , resuelve la ecuación como puede comprobarse fácilmente. En la teoría de ecuaciones*

diferenciales ordinarias es habitual imponer una condición local de Lipschitz en el espacio variable  $x$  para obtener unicidad de soluciones de la ecuación (2.20). En este contexto, la función  $b(t, x)$  se dice que es lipschitziana si existe una constante  $K \geq 0$  tal que para cualesquiera  $t, x_1$  y  $x_2$  se tiene que

$$\|b(t, x_1) - b(t, x_2)\| \leq K \|x_1 - x_2\|.$$

Se dice que es localmente lipschitziana en el espacio variable  $x$  si verifica la propiedad anterior en un entorno de cualquier punto  $(t, x)$ . Tal condición no es una restricción muy fuerte para una función. La función  $b(t, x) = 3x^{\frac{2}{3}}$  de este ejemplo no es localmente lipschitziana en el espacio variable  $x$  en ningún punto de la forma  $(t, 0)$  como puede comprobarse.

**Ejemplo 2.18** Sin embargo, incluso para las ecuaciones diferenciales ordinarias, una condición de Lipschitz local no es suficiente para garantizar la existencia global de una solución. A este respecto, sea la siguiente ecuación diferencial:

$$x'(t) = x(t)^2, \quad x(0) = 1.$$

Esta ecuación posee una única (porque  $b(t, x) = x^2$  es localmente lipschitziana) solución, la cual viene dada por

$$x(t) = \frac{1}{1-t}, \quad 0 \leq t < 1.$$

Sin embargo, es importante observar que esta solución “explota” cuando  $t \uparrow 1$ . Por tanto, no es posible encontrar una solución global definida para todo  $0 \leq t < \infty$ .

El resultado más básico para garantizar la existencia y unicidad local de soluciones a la ecuación (2.20) es el Teorema de Picard (ver por ejemplo el Teorema I.5.3 en

Hale(1969)[53]), el cual afirma que si la función  $b(t, x)$  satisface una condición local de Lipschitz en el espacio variable  $x$  y está acotada en los conjuntos compactos, para  $t > 0$  suficientemente pequeño, las denominadas iteraciones de Picard

$$x^{(0)}(t) = x_0, \quad x^{(n+1)}(t) = x_0 + \int_0^t b(u, x^{(n)}(u)) du,$$

convergen a la única solución de la ecuación (2.20), cuando  $n$  tiende a infinito.

Respecto a las ecuaciones diferenciales estocásticas, éstas se obtienen añadiendo en la ecuación clásica (2.20) un término de ruido de la siguiente manera:

$$dX(t) = b(t, X(t)) dt + \sigma(t, X(t)) dW(t), \quad X(0) = \xi, \quad (2.22)$$

donde, como es usual,  $W$  es un movimiento browniano estándar y  $b(t, x)$  y  $\sigma(t, x)$  son funciones determinísticas conocidas. Este tipo de ecuaciones son conocidas como ecuaciones diferenciales estocásticas de Itô.

La ecuación diferencial estocástica (2.22) puede interpretarse como sigue: El cambio en el intervalo de infinitesimal  $[t, t + dt]$  sufrido por el proceso estocástico  $X(t)$ , es decir,  $dX(t) = X(t + dt) - X(t)$ , es causado por el cambio de tiempo  $dt$  con factor  $b(t, X(t))$  en combinación con un cambio  $dW(t) = W(t + dt) - W(t)$  del movimiento browniano con factor  $\sigma(t, X(t))$ .

Es importante tener en cuenta que el significado preciso de la ecuación diferencial



estocástica de Itô (2.22) es el que se sigue de su expresión integral:

$$X(t) = \xi + \int_0^t b(u, X(u)) du + \int_0^t \sigma(u, X(u)) dW(u), \quad (2.23)$$

donde la primera integral se considera en el sentido de Lebesgue y la segunda es una integral estocástica de Itô.

Antes de enunciar el resultado general de existencia y unicidad de soluciones para la ecuación diferencial estocástica (2.22) se debe tener claro qué es una solución de tal ecuación. Aunque existe más de una definición de solución para dicha ecuación, la definición más natural resulta de trasladar al caso estocástico el concepto de solución de la teoría no estocástica, es decir, tomar cada trayectoria muestral como una función y trabajar con ella. Para ser precisos, se da la definición de este tipo de soluciones, llamadas *soluciones fuertes* siguiendo a Karatzas y Shreve(1991)[66].

**Definición 2.19 (Solución fuerte de una ecuación diferencial estocástica)** *Sea la ecuación diferencial estocástica de Itô dada, para  $0 \leq t < \infty$ , por (2.23) o equivalentemente por su versión diferencial (2.22), la cual se puede escribir componente a componente como*

$$dX_i(t) = b_i(t, X(t)) dt + \sum_{j=1}^r \sigma_{ij}(t, X(t)) dW_j(t), \quad X_i(0) = \xi_i, \quad (2.24)$$

donde el vector velocidad  $d$ -dimensional  $b(t, x) = \{b_i(t, x)\}_{1 \leq i \leq d}$  y la matriz de dispersión  $d \times r$ -dimensional  $\sigma(t, x) = \{\sigma_{ij}(t, x)\}_{1 \leq i \leq d, 1 \leq j \leq r}$  tienen componentes que son funciones medibles de  $[0, \infty) \times \mathbb{R}^d$  en  $\mathbb{R}$  dadas,  $\xi$  es un vector aleatorio  $d$ -dimensional de cuadrado integrable independiente de  $W$  y  $W = \{W(t), \mathcal{F}_t; 0 \leq t < \infty\}$  es un movimiento browniano  $d$ -dimensional estándar fijo dado<sup>3</sup>. Una solución fuerte de dicha ecuación es un

---

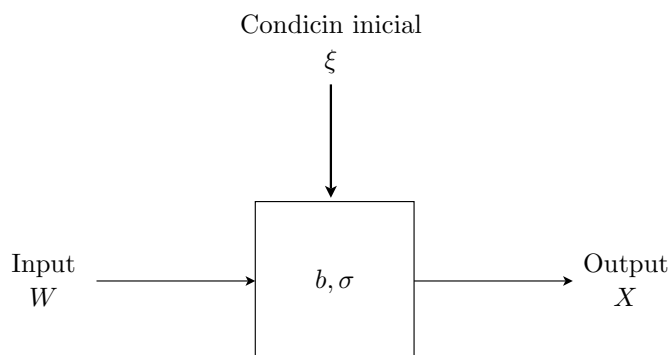
<sup>3</sup>Aquí la filtración  $\mathbb{F}$  es básicamente la generada por el movimiento browniano y por la condición inicial

proceso adaptado  $\{X(t), \mathcal{F}_t; 0 \leq t < \infty\}$  con caminos continuos tal que

1. Satisface la condición inicial, es decir,  $\mathbb{P}(X(0) = \xi) = 1$ .
2. Se tiene garantizada la existencia de las integrales que aparecen en la ecuación, es decir,  $b_i(t, X(t))$  es de variación acotada y  $\sigma_{ij}(t, X(t))$  está en  $L^2$ .
3. Verifica la ecuación, es decir, para  $0 \leq t < \infty$  se tiene  $\mathbb{P}$ -ect que

$$X_i(t) = \xi_i + \int_0^t b_i(u, X(u)) du + \sum_{j=1}^r \int_0^t \sigma_{ij}(u, X(u)) dW_j(u) \quad (1 \leq i \leq d).$$

Una condición crucial de la definición corresponde a la comprensión intuitiva de  $X$  como el “output” de un sistema dinámico que está descrito por los coeficientes  $(b, \sigma)$  y cuyo “input” es el movimiento browniano  $W$  y que además tiene un dato inicial  $\xi$ , (ver Figura 2.2).



**Figura 2.2:** Interpretación de la ecuación diferencial estocástica (2.22) como un sistema dinámico.

El principio de causalidad para un sistema dinámico, requiere que  $X(t)$  (el output en el instante  $t$ ) dependa sólo de la condición inicial  $\xi$  y del input hasta  $t$ ,  $\{W(u), 0 \leq u \leq t\}$ . De aquí que este principio encuentre su expresión matemática exigiendo que la solución  $X$  sea un proceso adaptado a la filtración  $\mathbb{F}$ .

$\xi$ .

La noción de unicidad de soluciones de la ecuación (2.24) con coeficientes dados  $(b, \sigma)$  queda entonces clara. En concreto, se dice que se tiene unicidad si, dados el movimiento browniano y la condición inicial, para dos soluciones cualesquiera  $X$  y  $X^*$  de la ecuación se verifica que  $\mathbb{P}(X(t) = X^*(t); 0 \leq t < \infty) = 1$ , es decir, coinciden  $\mathbb{P}$ -ect sus trayectorias.

### 2.3.1. Existencia y unicidad de soluciones

El resultado fundamental sobre existencia y unicidad de soluciones de la ecuación diferencial estocástica (2.24), de igual manera que su análogo no estocástico, impone condiciones de regularidad sobre los coeficientes y construye la única solución mediante iteraciones de Picard. Así, se presenta, tomado de §5.2 en Karatzas y Shreve (1991)[66], el siguiente resultado:

**Teorema 2.20 (Existencia y unicidad de soluciones fuertes.)** *Sea la ecuación diferencial estocástica (2.22) bajo las condiciones dadas en la Definición 2.19. Si los coeficientes satisfacen las condiciones siguientes de Lipschitz y de crecimiento lineal  $\forall t \geq 0, x, y \in \mathbb{R}^d$  y alguna constante  $K > 0$ :*

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

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

*entonces las iteraciones de Picard dadas por:*

$$X^{(0)}(t) = \xi,$$

$$X^{(n+1)}(t) = \xi + \int_0^t b(u, X^{(n)}(u)) du + \int_0^t \sigma(u, X^{(n)}(u)) dW(u),$$

convergen a la única solución fuerte  $X$  de la ecuación (2.22). Además, este proceso solución es de cuadrado integrable, en concreto satisface la siguiente condición de crecimiento:

$$\exists C = C(K, T) \mid E \|X(t)\|^2 \leq C \cdot (1 + E \|\xi\|^2) \cdot e^{Ct}, \quad \forall t \in [0, T].$$

Como propiedad particularmente destacable de los procesos  $X$  que se pueden representar mediante una ecuación diferencial estocástica de tipo (2.22), se tiene su carácter fuertemente markoviano. Es más, bajo ciertas condiciones<sup>4</sup>, se tiene que  $X$  pertenece a la clase de los denominados *procesos de difusión*, para estos procesos continuos fuertemente markovianos se tiene, dado un instante de tiempo  $t$  y un estado  $x$ , la existencia de los siguientes límites  $\forall x \in \mathbb{R}^d, 1 \leq i, k \leq d$ :

$$\lim_{h \downarrow 0} \frac{1}{h} \mathbb{E} [X_i(t+h) - X_i(t) \mid X(t) = x] = b_i(t, x), \quad (2.25)$$

$$\lim_{h \downarrow 0} \frac{1}{h} \mathbb{E} [(X_i(t+h) - X_i(t))(X_k(t+h) - X_k(t)) \mid X(t) = x] = a_{ik}(t, x), \quad (2.26)$$

de donde se sigue que el vector deriva  $b(t, x)$  mide localmente la velocidad media del movimiento aleatorio modelizado por  $X$  y la matriz  $a(t, x)$  mide localmente la razón de cambio en su matriz de varianzas-covarianzas. De esta manera, siendo  $X$  un proceso estocástico que satisface la ecuación diferencial estocástica (2.24) se tiene que  $a(t, x) = \sigma(t, x) \cdot \sigma'(t, x)$ . Las ecuaciones diferenciales estocásticas proporcionan una herramienta útil para estudiar una amplia clase de procesos de difusión.

---

<sup>4</sup>Para profundizar en este aspecto se puede consultar el Capítulo 5 de Karatzas y Shreve (1991)[66].

En relación a la modelización de los mercados financieros mediante ecuaciones diferenciales estocásticas, donde  $X$  modeliza básicamente precios de activos, es importante mencionar que el carácter fuertemente markoviano es compatible con la denominada *hipótesis de eficiencia de los mercados*, bajo la cual los precios incorporan en cada instante toda la información disponible; y en consecuencia, la evolución futura de los precios depende sólo de su estado presente y no del camino recorrido para llegar hasta él.

A continuación, se analizan algunos ejemplos de ecuaciones diferenciales estocásticas que aparecen en la modelización de los mercados financieros.

### 2.3.2. El modelo de Black-Scholes

En el célebre modelo de Black-Scholes se considera que el precio  $S$  de un activo con riesgo (una acción, por ejemplo) cambia en un intervalo pequeño de tiempo  $[t, t + dt]$  de manera que su rendimiento  $\frac{dS(t)}{S(t)}$  se descompone en una parte sistemática y una parte aleatoria. La parte sistemática se modeliza mediante  $\mu dt$ , donde  $\mu$  es un parámetro que representa el tipo de rendimiento medio del activo. La parte aleatoria se modeliza mediante  $\sigma dW(t)$ , donde  $dW(t)$  representa el término ruido de la dinámica del precio del activo y  $\sigma$ , que es el segundo parámetro del modelo, describe cuánto efecto tiene este ruido, es decir, cuánto fluctúa el precio del activo. Como ya se indicó anteriormente, este parámetro  $\sigma$  indica cómo de volátil es el precio, por eso se llama volatilidad del activo. Básicamente, el papel del término de ruido es recoger el efecto aleatorio de la gran cantidad de factores que entran en juego en el entorno económico donde se forma el precio.

Teniendo en cuenta todo lo anteriormente expuesto, se propone la siguiente ecuación

diferencial estocástica para describir los movimientos de precio en el activo:

$$dS(t) = S(t) \cdot (\mu dt + \sigma dW(t)), \quad S(0) > 0. \quad (2.27)$$

Esta ecuación corrige la propuesta de Bachelier (1990)[5] que no tiene el factor  $S(t)$  en el lado derecho de la ecuación, perdiendo así la interpretación en términos de rendimientos, además de presentar la poco deseable propiedad de poder resultar precios negativos.

En virtud del Teorema 2.20, con  $b(t, x) = \mu x$  y  $\sigma(t, x) = \sigma x$ , se tiene garantizada la existencia y unicidad de solución para la ecuación (2.27). A continuación, siguiendo a Mikosch (2004)[81], se va a obtener la expresión explícita de dicha solución, usando para ello como herramienta fundamental la regla de Itô.

Si se supone que la solución a la ecuación (2.27) viene dada por  $S(t) = f(t, W(t))$  para alguna función diferenciable  $f(t, x)$ , entonces por la regla de Itô se sigue que:

$$dS(t) = \left[ \frac{\partial}{\partial t} f(t, W(t)) + \frac{1}{2} \frac{\partial^2}{\partial x^2} f(t, W(t)) \right] dt + \frac{\partial}{\partial x} f(t, W(t)) dW(t). \quad (2.28)$$

El proceso  $S$  es un proceso de Itô; por tanto, se puede identificar los integrandos en las integrales de Lebesgue y de Itô en las ecuaciones (2.27) y (2.28). Así, se obtiene el siguiente sistema de ecuaciones en derivadas parciales:

$$\frac{\partial}{\partial t} f(t, x) + \frac{1}{2} \frac{\partial^2}{\partial x^2} f(t, x) = \mu f(t, x) \quad (2.29)$$

$$\frac{\partial}{\partial x} f(t, x) = \sigma f(t, x). \quad (2.30)$$

De la ecuación (2.30) se deduce que

$$\frac{\partial^2}{\partial x^2} f(t, x) = \sigma^2 f(t, x),$$

y aplicando esta relación en (2.29) se llega al siguiente sistema de ecuaciones:

$$\begin{cases} (\mu - \frac{1}{2}\sigma^2) f(t, x) = \frac{\partial}{\partial t} f(t, x) \\ \sigma f(t, x) = \frac{\partial}{\partial x} f(t, x). \end{cases} \quad (2.31)$$

Si se supone que  $f(t, x)$  es de variable separable, es decir, que se puede expresar como el producto de funciones  $f(t, x) = g(t) \cdot h(x)$ , el sistema (2.31) se convierte en:

$$\begin{cases} (\mu - \frac{1}{2}\sigma^2) g(t) = g'(t) \\ \sigma h(x) = h'(x), \end{cases}$$

donde cada ecuación se puede resolver sin dificultad por separado, obteniéndose

$$\begin{cases} g(t) = g(0) \cdot e^{(\mu - \frac{1}{2}\sigma^2)t} \\ h(x) = h(0) \cdot e^{\sigma x}. \end{cases}$$

Por tanto,  $f(t, x) = g(0)h(0) \cdot e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma x}$ , y como  $S(0) = f(0, W(0)) = f(0, 0) = g(0)h(0)$ , se tiene que la expresión explícita de la única solución de la ecuación diferencial estocástica (2.27) es

$$S(t) = S(0) \cdot e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)}, \quad t \geq 0.$$

Esta deducción ilustra cómo la solución de una ecuación diferencial estocástica de Itô se puede obtener, en algunas ocasiones, con el auxilio de ecuaciones en derivadas parciales

determinísticas. Sin embargo, no siempre es posible realizar una construcción de este tipo. En realidad, éste es uno de los pocos casos en los que una ecuación diferencial estocástica se puede resolver explícitamente. Usualmente, se establece la existencia y unicidad de soluciones mediante el Teorema 2.20 o algún resultado más específico para el caso de estudio, y se aplican algoritmos numéricos para obtener aproximaciones a tal solución.

Por otro lado, nótese que la solución  $S(t)$  es la exponencial estocástica del movimiento browniano con deriva  $\mu t + \sigma W(t)$  (ver Ejemplos 1.19 y 2.14). En particular

$$\log S(t) = \log S(0) + \left( \mu - \frac{1}{2}\sigma^2 \right) t + \sigma W(t)$$

tiene una distribución normal y  $S(t)$  tiene una distribución lognormal. Este modelo de movimiento browniano geométrico (ver Ejemplo 1.20) y la distribución lognormal que involucra, son la base del modelo de Black-Scholes para la dinámica en tiempo continuo de los precios de los activos con riesgo. En este sentido, la generalización multidimensional y con coeficientes del tipo  $\mu(t, S(t))$ ,  $\sigma(t, S(t))$  constituye la base teórica fundamental del tema siguiente.

### 2.3.3. El proceso de Ornstein-Uhlenbeck

Otro proceso importante, originado en la mecánica estadística y que aparece también en finanzas en el modelo considerado por Vasicek (1977)[97] para modelizar la estructura temporal de los tipos de interés, es el proceso de Ornstein-Uhlenbeck.

En concreto, se trata de la ecuación de Langevin para el movimiento browniano de



una partícula con fricción y viene dada por la siguiente ecuación diferencial estocástica lineal:

$$dX(t) = -\mu X(t)dt + \sigma dW(t), \quad \mu, \sigma > 0. \quad (2.32)$$

Nótese que a diferencia de la ecuación lineal (2.27), el proceso  $X$  no viene ligado al movimiento browniano en la integral de Itô, por eso se dice que este proceso presenta un *ruido aditivo*.

Para resolver la ecuación planteada se aplica la fórmula de integración por partes (ver Ejemplo 2.13) al proceso  $e^{\mu t} \cdot X(t)$  y se obtiene

$$\begin{aligned} d(e^{\mu t} X(t)) &= e^{\mu t} dX(t) + X(t) \mu e^{\mu t} dt + 0 \\ &= e^{\mu t} (dX(t) + X(t) \mu dt) = e^{\mu t} \sigma dW(t). \end{aligned}$$

Por tanto, el proceso

$$X(t) = e^{-\mu t} X(0) + \sigma e^{-\mu t} \int_0^t e^{\mu u} dW(u), \quad (2.33)$$

conocido como proceso de Ornstein-Uhlenbeck (1930)[96], resuelve la ecuación diferencial estocástica (2.32). En concreto, se tiene que  $X(t) = e^{-\mu t} Z(t)$  donde  $Z(t) = X(0) + \sigma \int_0^t e^{\mu u} dW(u)$ , y aplicando la regla de Itô (2.15) con  $f(t, z) = e^{\mu t} z$  se tiene que el proceso  $X$  resuelve la ecuación (2.32). Además, al tratarse de una ecuación lineal se verifican las condiciones del Teorema 2.20 y entonces  $X$  es la única solución fuerte de la ecuación.

Si se considera que la condición inicial  $X(0)$  es independiente del movimiento browniano  $W$  y sigue una distribución normal, se deduce que el proceso de Ornstein-Uhlenbeck

(2.33) es gaussiano. En efecto, según la definición de integral de Itô, la integral estocástica

$$\int_0^t e^{\mu u} dW(u)$$

es el límite en media cuadrática de las sumas aproximantes

$$S_n = \sum_{i=1}^n e^{\mu t_{i-1}} (W(t_i) - W(t_{i-1}))$$

para particiones  $\Pi$  de  $[0, t]$  cuando  $\|\Pi\| \rightarrow 0$ . Esta suma tiene una distribución normal de media cero y varianza

$$\sum_{i=1}^n e^{2\mu t_{i-1}} (t_i - t_{i-1}),$$

la cual es, a su vez, una suma aproximante de la integral de Riemann

$$\int_0^t e^{2\mu t} dt = \frac{1}{2\mu} (e^{2\mu t} - 1).$$

Teniendo en cuenta que la convergencia en media cuadrática implica la convergencia en distribución, se concluye que el límite  $X(t)$  de las variables normales  $S_n$  tiene una distribución normal. En consecuencia,  $X(t)$  sigue una distribución normal con

$$\mathbb{E}X(t) = \mathbb{E}X(0) e^{-\mu t}$$

y

$$\mathbb{V}ar(X(t)) = \frac{\sigma^2}{2\mu} + e^{-2\mu t} \left( \mathbb{V}ar(X(0)) - \frac{\sigma^2}{2\mu} \right).$$

Un argumento similar lleva a que

$$\mathbb{C}ov(X(s), X(t)) = e^{-\mu(s+t)} \left( \mathbb{V}ar(X(0)) + \frac{\sigma^2}{2\mu} (e^{2\mu s} - 1) \right), \quad s < t.$$

En concreto, si  $X(0)$  sigue una distribución normal de media cero y varianza  $\frac{\sigma^2}{2\mu}$ , se tiene que  $X$  es un proceso estacionario, gaussiano de media cero y con función de covarianza

$$\mathbb{Cov}(X(s), X(t)) = \frac{\sigma^2}{2\mu} e^{-\mu|t-s|}, \quad \forall s, t \in \mathbb{R}.$$

## 2.4. Tres resultados fundamentales

Para terminar este capítulo dedicado a los aspectos fundamentales del cálculo estocástico, se exponen tres resultados que son de vital importancia en la modelización de mercados financieros mediante las ecuaciones diferenciales estocásticas presentadas en la sección anterior.

El primero de ellos es el *teorema de Girsanov*, un resultado que permite hacer un cambio de medida en el modelo de Black–Scholes para obtener la medida martingala de riesgo neutral, la cual garantiza un modelo de mercado libre de arbitraje y se usa para valorar activos derivados. Este resultado, se puede encontrar formulado y demostrado con varios grados de generalidad en Karatzas y Shreve (1991)[66], Protter (2004)[88], Revuz y Yor (1991)[90] y Dothan (1990)[33].

Para motivar el teorema de Girsanov, se va a considerar  $Z_1, \dots, Z_n$  variables aleatorias normales estándar independientes sobre un espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$ . Dado un vector  $(\mu_1, \dots, \mu_n) \in \mathbb{R}^n$ , sea la nueva medida de probabilidad  $\mathbb{P}^*$  sobre  $(\Omega, \mathcal{F})$  dada por

$$\mathbb{P}^*(d\omega) = \exp \left\{ \sum_{i=1}^n \mu_i Z_i(\omega) - \frac{1}{2} \sum_{i=1}^n \mu_i^2 \right\} \mathbb{P}(d\omega).$$

Nótese que  $\mathbb{P}^*$  así definida es en efecto una medida de probabilidad sobre  $(\Omega, \mathcal{F})$  sin más que recordar la función generatriz de momentos de una variable normal<sup>5</sup>. Además la medida de probabilidad  $\mathbb{P}^*$  es equivalente a  $\mathbb{P}$  (tienen los mismos conjuntos de medida nula).

Con esta definición, se tiene que

$$\begin{aligned} \mathbb{P}^*(Z_1 \in dz_1, \dots, Z_n \in dz_n) &= \exp \left\{ \sum_i \mu_i Z_i - \frac{1}{2} \sum_i \mu_i^2 \right\} \mathbb{P}(Z_1 \in dz_1, \dots, Z_n \in dz_n) \\ &= (2\pi)^{-\frac{n}{2}} \exp \left\{ \sum_i \mu_i z_i - \frac{1}{2} \sum_i \mu_i^2 - \frac{1}{2} \sum_i z_i^2 \right\} dz_1 \dots dz_n \\ &= (2\pi)^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2} \sum_i (z_i - \mu_i)^2 \right\} dz_1 \dots dz_n \\ &= \prod_i \left[ (2\pi)^{-\frac{1}{2}} \exp \left\{ -\frac{(z_i - \mu_i)^2}{2} \right\} dz_i \right]. \end{aligned}$$

Por tanto, siendo las variables aleatorias  $Z_i$  normales estándar independientes bajo  $\mathbb{P}$ , también lo son las variables aleatorias trasladadas  $Z_i^* = Z_i - \mu_i$  bajo  $\mathbb{P}^*$ . Así, el efecto principal del cambio de la medida original  $\mathbb{P}$  a la medida equivalente  $\mathbb{P}^*$  es cambiar el vector de medias de  $(0, \dots, 0)$  a  $(\mu_1, \dots, \mu_n)$ . Con el teorema de Girsanov, se extiende esta idea de invarianza de distribuciones gaussianas finito-dimensionales bajo traslaciones apropiadas y cambio en la medida de probabilidad subyacente al caso continuo.

En este contexto, sea  $W = (W_1, \dots, W_d)$  un movimiento browniano  $d$ -dimensional estándar definido sobre un espacio de probabilidad filtrado  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ ,  $\gamma = \{\gamma(t), \mathcal{F}_t; 0 \leq t \leq T\}$  un proceso  $d$ -dimensional medible y adaptado con

---

<sup>5</sup>Dada una variable  $X$  normal de media  $\mu \in \mathbb{R}$  y varianza  $\sigma^2 > 0$ , su función generatriz de momentos viene dada por  $M(t) = \mathbb{E}(e^{tX}) = e^{t\mu + \frac{1}{2}t^2\sigma^2}$ ,  $\forall t \in \mathbb{R}$ .

$\mathbb{E} \int_0^T \|\gamma(t)\|^2 dt < \infty$  y  $L = \{L(t), \mathcal{F}_t; 0 \leq t \leq T\}$  el proceso definido mediante

$$L(t) = \exp \left\{ \int_0^t \gamma'(u) dW(u) - \frac{1}{2} \int_0^t \|\gamma(u)\|^2 du \right\}. \quad (2.34)$$

Entonces,  $L$  es un proceso continuo y, bajo ciertas condiciones de integrabilidad sobre  $\gamma$ , es una martingala (nótese que se trata de la exponencial estocástica –ver Ejemplo 2.14– del proceso  $\int_0^t \gamma'(u) dW(u)$ ). Para ello la condición siguiente, conocida como condición de Novikov, es suficiente:

$$\mathbb{E} \left( \exp \left\{ \frac{1}{2} \int_0^T \|\gamma(u)\|^2 du \right\} \right) < \infty. \quad (2.35)$$

Se está ya en condiciones de dar una versión del Teorema de Girsanov.

**Teorema 2.21 (Teorema de Girsanov (1960)[51])** *Sea un proceso  $\gamma$  satisfaciendo las condiciones anteriormente mencionadas incluida la de Novikov (2.35) y sea  $L$  la correspondiente martingala continua dada por (2.34). Se definen los procesos  $W_i^*$  mediante las traslaciones:*

$$W_i^*(t) = W_i(t) - \int_0^t \gamma_i(u) du, \quad 0 \leq t \leq T, \quad i = 1, \dots, d.$$

*Entonces, bajo la medida de probabilidad equivalente  $\mathbb{P}^*$  definida sobre  $(\Omega, \mathcal{F}_T)$  mediante la derivada de Radon–Nikodym*

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = L(T),$$

*el proceso  $W^* = (W_1^*, \dots, W_d^*)$  es un movimiento browniano  $d$ -dimensional estándar.*

Si la filtración  $\mathbb{F}$  es la conocida como filtración browniana (básicamente es la filtración generada por el movimiento browniano  $\mathcal{F}_t^W = \sigma(W(u), 0 \leq u \leq t)$  ligeramente aumentada para satisfacer las condiciones usuales), cualquier par de medidas de probabilidad equivalentes  $\mathbb{P} \sim \mathbb{Q}$  sobre  $\mathcal{F} = \mathcal{F}_T$  es un par de Girsanov, es decir, existe un proceso  $\gamma$  tal que, para la correspondiente martingala  $L(t)$  dada por (2.34) se tiene:

$$\left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_t} = L(t).$$

Nótese que el caso particular de  $\gamma$  constante corresponde a un cambio en la deriva. Además, se tiene que las integrales respecto a  $dW_i^*(t)$  tienen dos posibles interpretaciones: En primer lugar, como es un movimiento browniano bajo  $\mathbb{P}^*$ , se puede ver como una integral de Itô bajo  $\mathbb{P}^*$ ; y en segundo lugar, se puede interpretar bajo  $\mathbb{P}$  reemplazando  $dW_i^*(t)$  por  $dW_i(t) - \gamma_i(t)dt$  como una integral de Itô y una integral de Lebesgue. Afortunadamente, se demuestra que ambas interpretaciones coinciden.

Por otro lado, el segundo resultado fundamental que se va a comentar en este trabajo surge al considerar cuestiones referentes a la alcanzabilidad y la completitud en mercados financieros.

Si se considera  $W = \{W(t), \mathcal{F}_t; 0 \leq t < \infty\}$  un movimiento browniano  $d$ -dimensional estándar definido sobre un espacio de probabilidad filtrado  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$  donde  $\mathbb{F}$  es la filtración browniana, se tiene que, dado un proceso  $d$ -dimensional  $X \in L^2$ , la integral de Itô

$$M(t) = M(0) + \int_0^t X(u) dW(u), \quad t \geq 0, \quad (2.36)$$

es una martingala continua de cuadrado integrable (ver §2.2). El resultado que se presenta

es un recíproco de la propiedad anterior.

**Teorema 2.22 (Representación Martingala)** *Sea  $W = \{W(t), \mathcal{F}_t; 0 \leq t < \infty\}$  un movimiento browniano  $d$ -dimensional estándar definido sobre un espacio de probabilidad filtrado  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$  donde  $\mathbb{F}$  es la filtración browniana. Si  $M$  es una  $\mathbb{F}$ -martingala continua de cuadrado integrable, entonces existe un único proceso  $X \in L^2$  tal que la martingala  $M$  admite la representación (2.36).*

La demostración de este resultado admite diferentes niveles de generalidad y puede encontrarse en Karatzas y Shreve (1991)[66], Revuz y Yor (1991)[90] y Oksendal (1998)[84], entre otros.

El tercer y último resultado de esta sección muestra la rica relación existente entre la teoría de la probabilidad y el análisis matemático, ya que enlaza las ecuaciones diferenciales estocásticas con problemas de contorno de ecuaciones en derivadas parciales.

Para ver esta relación, se considera el siguiente problema de contorno, denominado *problema de Cauchy*: Dadas las funciones escalares  $b(t, x)$ ,  $\sigma(t, x)$  y  $\Phi(x)$ , encontrar una función  $F(t, x)$  que sobre  $[0, T] \times \mathbb{R}$  verifique:

$$\frac{\partial F}{\partial t}(t, x) + b(t, x) \frac{\partial F}{\partial x}(t, x) + \frac{1}{2} \sigma^2(t, x) \frac{\partial^2 F}{\partial x^2}(t, x) = 0 \quad (2.37)$$

$$F(T, x) = \Phi(x) \quad (2.38)$$

En lugar de intentar resolver este problema usando herramientas puramente analíticas, se va encontrar la solución en términos de una ecuación diferencial estocástica obteniendo la llamada *fórmula de representación estocástica*.

Sea  $F$  una solución al problema de Cauchy (2.37-2.38) y sean  $t$  un instante temporal y  $x$  un punto espacial fijos. A partir de aquí, se define el proceso  $X$  sobre el intervalo  $[t, T]$  como una solución a la ecuación diferencial estocástica

$$\begin{aligned} dX(u) &= b(u, X(u)) du + \sigma(u, X(u)) dW(u) \\ X(t) &= x. \end{aligned}$$

Usando el generador infinitesimal<sup>6</sup>  $\mathcal{L}$  de este proceso, dado por,

$$\mathcal{L} = \frac{\partial}{\partial t} + b(t, x) \frac{\partial}{\partial x} + \frac{1}{2} \sigma^2(t, x) \frac{\partial^2}{\partial x^2},$$

el problema de Cauchy (2.37-2.38) se puede escribir como:

$$\mathcal{L}F(t, x) = 0 \tag{2.39}$$

$$F(T, x) = \Phi(x). \tag{2.40}$$

Aplicando la regla de Itô al proceso  $F(u, X(u))$  se obtiene que

$$F(T, X(T)) = F(t, X(t)) + \int_t^T \mathcal{L}F(u, X(u)) du + \int_t^T \sigma(u, X(u)) \cdot \frac{\partial}{\partial x} F(u, X(u)) dW(u).$$

Teniendo en cuenta que  $F$  satisface (2.39), la integral de Lebesgue de la expresión anterior se anula. Además, si se supone que el proceso  $\sigma(u, X(u)) \cdot \frac{\partial}{\partial x} F(u, X(u))$  es de cuadrado integrable y se toma valores esperados, la integral estocástica también se anula. Teniendo en cuenta también el valor inicial  $X(t) = x$  y la condición frontera  $F(T, x) = \Phi(x)$  se

---

<sup>6</sup>La teoría de semigrupos de operadores y del generador infinitesimal, íntimamente relacionada con los procesos de Markov, puede verse en Dynkin (1965)[37], Dynkin y Yushkevich (1969)[38] y Feller (1971)[44].



llega a la siguiente fórmula:

$$F(t, x) = \mathbb{E}[\Phi(X(T)) | X(t) = x] \stackrel{\text{def.}}{=} \mathbb{E}_{t,x}[\Phi(X(T))].$$

En consecuencia, se ha probado la siguiente fórmula de representación estocástica:

**Teorema 2.23 (Fórmula de Feynman-Kac)** *Sea  $F$  una solución al problema de contorno*

$$\begin{aligned} \frac{\partial F}{\partial t}(t, x) + b(t, x) \frac{\partial F}{\partial x}(t, x) + \frac{1}{2} \sigma^2(t, x) \frac{\partial^2 F}{\partial x^2}(t, x) &= 0 \\ F(T, x) &= \Phi(x). \end{aligned}$$

*Si se supone además que el proceso  $\sigma(u, X(u)) \cdot \frac{\partial}{\partial x} F(u, X(u))$  está en  $L^2$ , donde el proceso  $X$  satisface la ecuación diferencial estocástica:*

$$dX(u) = b(u, X(u)) du + \sigma(u, X(u)) dW(u) \tag{2.41}$$

$$X(t) = x, \tag{2.42}$$

*entonces  $F$  admite la representación*

$$F(t, x) = \mathbb{E}_{t,x}[\Phi(X(T))].$$

**Observación 2.24** *Respecto a este último resultado, es importante mencionar que un razonamiento similar lleva, bajo las condiciones adecuadas (Teorema 7.6 en Karatzas y Shreve (1991)[66]) a un resultado un poco más general. En concreto, en las condiciones del teorema anterior y con  $k(t, x)$  una función escalar, se verifica que si  $F$  es una solución*

al problema de contorno:

$$\begin{aligned} \frac{\partial F}{\partial t}(t, x) + b(t, x) \frac{\partial F}{\partial x}(t, x) + \frac{1}{2} \sigma^2(t, x) \frac{\partial^2 F}{\partial x^2}(t, x) &= k(t, x) F(t, x) \\ F(T, x) &= \Phi(x). \end{aligned}$$

entonces  $F$  admite la representación

$$F(t, x) = \mathbb{E}_{t,x} \left[ \exp \left\{ - \int_t^T k(s, X(s)) ds \right\} \cdot \Phi(X(T)) \right],$$

donde el proceso  $X$  satisface la ecuación diferencial estocástica dada por (2.41) y (2.42).

Llegados a este punto, se dispone de las herramientas necesarias para introducir la modelización de mercados financieros mediante ecuaciones diferenciales estocásticas, que es el objetivo fundamental de este trabajo. En consecuencia, en el capítulo siguiente se describen los aspectos básicos de esta modelización, dedicando el último capítulo al desarrollo de una aplicación empírica con datos de la Economía española.

# Capítulo 3

## Principios básicos en modelización de mercados financieros

### 3.1. Introducción

Habiendo introducido en los capítulos anteriores las herramientas necesarias para la modelización de mercados financieros mediante ecuaciones diferenciales estocásticas (esencialmente el cálculo estocástico), este capítulo se dedica al análisis de algunos modelos importantes. Para ello, se divide en dos partes: La primera, §3.2, se dedica a la teoría general de los mercados financieros donde los precios de los activos sujetos a riesgo se modelizan mediante semimartingalas positivas, adaptadas y continuas; lo que lleva esencialmente al modelo de Black-Scholes generalizado. En la segunda parte, §3.3, se analiza el caso particular del mercado de bonos y su modelización mediante la dinámica de los

tipos de interés a corto plazo.

Las referencias más importantes que se han seguido en este capítulo son Karatzas y Shreve (1991,1998)[66, 67] y Björk (1997,2004)[7, 8].

## 3.2. El modelo de Black-Scholes generalizado

El primer intento de modelizar los precios de activos financieros mediante procesos de difusión se atribuye a Bachelier (1900)[5], quien utiliza lo que ahora llamamos movimiento browniano para describir la evolución de los precios de las acciones en el mercado. Sin embargo, hubo que esperar más de medio siglo para que apareciera la primera generalización de este modelo, debida Samuelson (1965,1973)[91, 92] y más conocido hoy día como modelo de Black-Scholes (1973)[11], después del influyente trabajo de estos autores donde se analiza y resuelve el problema de la valoración de un tipo particular de activos financieros: las opciones de tipo europeo.

En el modelo de mercado financiero de Black-Scholes se consideran dos procesos estocásticos que representan el precio por unidad de dos activos financieros: El primero de ellos, denotado por  $B$ , representa un activo "instantáneamente libre de riesgo", típicamente un bono o cuenta bancaria, y que se suele denominar genéricamente *cuenta bancaria*. La dinámica de este proceso viene dada por la ecuación  $dB(t) = rB(t)dt$ , donde  $r$  es una constante, y cuya solución es  $B(t) = B(0)e^{rt}$ , quedando clara la interpretación del parámetro  $r$  como *el tipo de interés compuesto continuo*. Respecto al otro activo, comentado ya en §2.3.2, se denomina genéricamente *acción* (aunque también puede ser por ejemplo una

mercancía o una divisa) y está “sujeto a riesgo sistemático”. La evolución de su precio viene dada por la ecuación diferencial estocástica lineal

$$dS(t) = S(t) (bdt + \sigma dW(t)),$$

donde  $b$  y  $\sigma > 0$  son constantes y  $W$  es un movimiento browniano unidimensional estándar. Se sigue de §2.3.2 que la única solución a esta ecuación viene dada por el movimiento browniano geométrico  $S(t) = S(0) \cdot \exp \left\{ \left( b - \frac{\sigma^2}{2} \right) \cdot t + \sigma W(t) \right\}$ . En cuanto a la interpretación de los parámetros,  $b$  es *la tasa de beneficio instantánea esperada* y  $\sigma$  la desviación estándar local de la tasa de beneficio, conocida también como *volatilidad*. En este sentido, se tiene formalmente que  $\mathbb{E} \left[ \frac{dS(t)}{S(t)} \right] = bdt$  y  $\text{Var} \left[ \frac{dS(t)}{S(t)} \right] = \sigma^2 dt$ . En consecuencia, la tasa de beneficio infinitesimal es  $r$  (no estocástico) para la cuenta bancaria y  $b + \sigma \dot{W}$  (estocástico)<sup>1</sup> para la acción, de ahí la expresión “instantáneamente libre de riesgo” para la cuenta bancaria y “sujeto a riesgo sistemático” para la acción.

El modelo sobre el que se van a desarrollar los conceptos y resultados fundamentales de este capítulo, es la generalización del modelo de Black-Scholes a  $k+1$  activos (la cuenta bancaria y  $k$  acciones) con coeficientes dados por procesos estocásticos.

#### 3.2.1. El modelo

Sea un modelo para un mercado financiero donde se negocia continuamente un número finito  $k+1$  de activos. A los procesos estocásticos que recogen la evolución en el tiempo de los precios por unidad de tales activos se les denota por  $S_0, S_1, \dots, S_k$ . Más concretamente,

---

<sup>1</sup>Se ha usado aquí la notación  $\dot{W}$  para representar la derivada de  $W(t)$  respecto a  $t$ . Notación muy del agrado de los físicos, a quienes esta teoría debe importantes contribuciones.

se supone que los procesos  $S_i$  ( $i = 0, \dots, k$ ) son *semimartingalas continuas* definidas sobre un espacio de probabilidad filtrado  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ .

En el modelo planteado, el activo  $S_0$  juega el papel particular de activo numerario, es decir, el activo con precios positivos respecto al cual se miden los demás, éste se suele elegir como un activo localmente “libre de riesgo” y su  $\mathbb{P}$ -dinámica viene dada por<sup>2</sup>:

$$dS_0(t) = S_0(t)r(t)dt, \quad S_0(0) = 1,$$

donde  $r$  es un proceso estocástico medible y adaptado dado. Este activo se suele interpretar como una cuenta bancaria o un bono sujeto a un tipo de interés compuesto continuo  $r$  estocástico, y como antes se denomina genéricamente cuenta bancaria. En adelante, como es habitual, se denota también a este activo por  $B$ . De esta forma, la ecuación que modeliza la dinámica de este activo se escribe también como:

$$dB(t) = B(t)r(t)dt, \quad B(0) = 1. \tag{3.1}$$

Los precios de los restantes  $k$  activos (genéricamente acciones) están sujetos a un riesgo sistemático y se supone que vienen descritos mediante un sistema de ecuaciones diferenciales estocásticas en el que el ruido se modeliza por un número finito  $d$  de movimientos brownianos independientes. En concreto, el precio del activo  $i$ -ésimo viene dado por la ecuación diferencial estocástica lineal siguiente:

$$\begin{aligned} dS_i(t) &= S_i(t) \left[ b_i(t) dt + \sum_{j=1}^d \sigma_{ij}(t) dW_j(t) \right], \\ S_i(0) &= s_i \in (0, \infty), \end{aligned} \tag{3.2}$$

---

<sup>2</sup>El valor inicial  $S_0(0)$  no es relevante, y por simplicidad se toma unitario.

donde  $b_i$  y  $\sigma_{ij}$  son procesos estocásticos medibles adaptados y  $W_j$  movimientos brownianos estándar independientes ( $1 \leq i \leq k, 1 \leq j \leq d$ ).

De esta forma se ha planteado el modelo dado por las ecuaciones (3.1) y (3.2) conocido como modelo de Black-Scholes generalizado, sobre el cual es importante hacer las siguientes observaciones:

- Aunque en realidad los precios de los activos no varían en tiempo continuo, si lo hacen en intervalos de tiempo muy cortos, y por ello el modelo resulta útil como aproximación a la realidad. Por otro lado, el hecho de considerar procesos con trayectorias continuas se suele interpretar como que no hay “sorpresas” en el modelo.
- Como es usual, en las aplicaciones financieras que se desarrollan en este capítulo se considera un horizonte de tiempo  $T$  finito.
- Las fuentes de riesgo sistemático se modelizan mediante las componentes independientes de un movimiento browniano  $d$ -dimensional estándar  $W(t) = (W_1(t), \dots, W_d(t))$  ( $0 \leq t \leq T$ ). Por tanto, el *coeficiente de volatilidad*  $\sigma_{ij}$  se puede interpretar como la intensidad instantánea con la cual la  $j$ -ésima fuente de incertidumbre influye sobre el precio del activo  $i$ -ésimo.
- Al proceso estocástico *tipo de interés*  $\{r(t), \mathcal{F}_t; 0 \leq t \leq T\}$ , al proceso vectorial de *tasa media de beneficios*  $\{b(t), \mathcal{F}_t; 0 \leq t \leq T\}$  y al proceso matricial de *volatilidades*  $\{\sigma(t), \mathcal{F}_t; 0 \leq t \leq T\}$  se les conoce conjuntamente como *los coeficientes del modelo*, los cuales se suponen adaptados y que verifican la siguiente condición de integrabilidad:

$$\mathbb{E} \left[ \int_0^T (|r(t)| + \|b(t)\| + \|\sigma(t)\|^2) dt \right] < \infty.$$

Bajo estas condiciones, una aplicación directa de la regla de Itô demuestra que la solución al sistema de ecuaciones diferenciales estocásticas lineales propuestas para

el mercado financiero viene dada por:

$$B(t) = \exp \left( \int_0^t r(u) du \right),$$

$$S_i(t) = S_i(0) \cdot \exp \left( \int_0^t \left[ b_i(u) - \frac{1}{2} \sum_{j=1}^d \sigma_{ij}^2(u) \right] du + \sum_{j=1}^d \int_0^t \sigma_{ij}(u) dW_j(u) \right), \quad i = 1, \dots, k.$$

- Respecto a la filtración  $\{\mathcal{F}_t\}$  considerada, ésta es la *filtración browniana*, que es, esencialmente, la generada por el movimiento browniano  $W$ . En consecuencia, al imponer que los coeficientes del modelo sean procesos  $\{\mathcal{F}_t\}$ -adaptados, se tiene que éstos son básicamente funciones del movimiento browniano, lo que lleva a las deseables propiedades de que se excluya la anticipación del futuro (información privilegiada) y se permita que sólo dependan del pasado del movimiento browniano y de los precios de los activos.
- El modelo se puede extender para permitir pagos de dividendos y saltos (discontinuidades) en los precios de las acciones, aunque éste es lo suficientemente general como para proporcionar respuestas concretas a muy diversas cuestiones de interés, algunas de las cuales se analizan en este capítulo.
- Como consecuencia de los teoremas de representación martingala, el modelo está “cerca” de ser el modelo más general que se puede considerar para un proceso  $B$  de variación acotada y semimartingalas  $S_i$ ,  $i = 1, \dots, k$  siendo todos ellos procesos continuos, adaptados y estrictamente positivos (ver Karatzas y Shreve (1998)[67] Apéndice B para mayores detalles).
- En el caso particular en el que los coeficientes se pueden expresar en todo  $t \in [0, T]$  como funciones del tiempo  $t$  y del vector de precios  $S(t) = (S_1(t), \dots, S_k(t))$  (el cual incluye al caso de coeficientes constantes) se tiene que el proceso  $\{S(t), \mathcal{F}_t; 0 \leq t \leq T\}$  no sólo es markoviano, sino que se trata de un proceso de



difusión (ver §2.3.1).

### 3.2.2. Carteras autofinanciadas

Una vez planteado el modelo de mercado financiero dado por las ecuaciones (3.1) y (3.2), se considera sobre el mismo a un inversor que decide en cada instante  $t \in [0, T]$  la cantidad de dinero que quiere invertir en cada uno de los activos del mercado. En este sentido, sea el proceso vectorial  $h(t) = (h_0(t), \dots, h_k(t))$ , donde  $h_i(t)$  es el número de unidades del activo  $i$  que el inversor tiene en el instante  $t$ . Este proceso, llamado *cartera* o *portfolio* debe ser  $\{\mathcal{F}_t\}$ -adaptado, es decir, la anticipación del futuro o información privilegiada queda excluida del modelo. Asociado a este proceso  $h$ , se tiene el *proceso valor* de la cartera, el cual viene dado por

$$V_h(t) = \sum_{i=0}^k h_i(t) S_i(t),$$

y ofrece el valor de mercado de la cartera  $h$  en el instante  $t$ .

En este contexto, se dice que una cartera  $h$  es *autofinanciada* si verifica:

$$dV_h(t) = \sum_{i=0}^k h_i(t) dS_i(t), \tag{3.3}$$

expresión que formaliza la idea de una cartera en la que la variación de su valor se debe únicamente a las variaciones en los precios de los activos. Por tanto, la compra de un nuevo activo se financia con la venta de activos de la propia cartera y, además, no hay inyección externa ni retirada de dinero<sup>3</sup>

---

<sup>3</sup>La expresión (3.3) se deduce de la correspondiente en el caso de tiempo discreto. En esa situación,

Es importante mencionar que todos los precios y valores anteriores se supone que vienen dados en algún activo numerario, típicamente una divisa local como euros (€), dólares americanos (\$) u otros. Sin embargo, se suele expresar todos los precios en función de algún otro activo numerario. Así, eligiendo el activo libre de riesgo  $B$  como activo numerario, se definen:

- El proceso vectorial de los *precios descontados*  $Z = (Z_0(t), \dots, Z_k(t))$  como  $Z(t) = \frac{S(t)}{B(t)}$ , y
- el proceso *valor descontado* asociado a la cartera  $h$  mediante  $V_h^Z(t) = \frac{V_h(t)}{B(t)}$ .

Como es lógico, la propiedad de autofinanciación de una cartera  $h$  no depende de la elección del activo numerario; resultando que la cartera es autofinanciada si y solo verifica la condición:

$$dV_h^Z(t) = \sum_{i=0}^k h_i(t) dZ_i(t),$$

hecho que se puede comprobar directamente usando la regla de Itô.

En ocasiones, generalmente con propósitos computacionales, es conveniente describir una cartera en términos relativos en lugar de en términos absolutos. Así, en lugar de especificar el número de cada tipo de activo en la cartera, se especifica la proporción de valor total de la cartera que se invierte en cada activo. En este sentido, dada una cartera  $h$ , la correspondiente *cartera relativa*  $u(t) = (u_0(t), \dots, u_k(t))$  viene dada por

$$u_i(t) = \frac{h_i(t)S_i(t)}{V_h(t)}, \quad i = 0, 1, \dots, k,$$

---

existen unos instantes concretos de tiempo  $t_j$  en los que el inversor decide cuántas unidades del activo  $i$ ,  $h_i(t_j)$ , tendrá en su cartera durante todo el intervalo de tiempo  $[t_j, t_{j+1})$ , con  $j = 0, \dots, n-1$  y  $0 = t_0 < t_1 < \dots < t_n = T$ . La variación del valor de la cartera autofinanciada en cada uno de esos intervalos de tiempo, teniendo en cuenta la continuidad de los precios, viene dada por:  $\Delta V_h(t_j) = V_h(t_{j+1}^-) - V_h(t_j) = \sum_{i=0}^k h_i(t_j) S_i(t_{j+1}) - \sum_{i=0}^k h_i(t_j) S_i(t_j) = \sum_{i=0}^k h_i(t_j) \Delta S_i(t_j)$ .

---

y la condición de autofinanciación de la cartera, dada en (3.3), se puede reescribir en términos de la cartera relativa como

$$dV_h(t) = V_h(t) \sum_{i=0}^k u_i(t) \frac{dS_i(t)}{S_i(t)}. \quad (3.4)$$

#### 3.2.3. Activos derivados, arbitraje y completitud

Uno de los objetivos fundamentales en este capítulo es valorar, en algún caso particular, los denominados *activos derivados*, que son aquellos que se definen en términos de los activos considerados en el modelo,  $S_0, \dots, S_k$ , también llamados *activos subyacentes*. Antes de dar la definición general de este tipo de activos se va a considerar un par de casos particulares e importantes, el de las opciones europeas y americanas, cuyo concepto se introduce a continuación.

**Definición 3.1 (Opciones europeas y americanas)** *Una opción de compra (o call) europea con precio de ejercicio  $K$  y fecha de vencimiento  $T$  sobre el activo subyacente  $S$ , es un contrato que otorga a su comprador el derecho (pero no la obligación) de comprar una unidad del activo subyacente al vendedor en el instante  $T$  (y únicamente en ese instante). El precio de ejercicio  $K$  y la fecha de vencimiento  $T$  están fijadas en el momento que se establece el contrato, que se denomina instante  $t = 0$ .*

*Una opción de venta (o put) europea es un contrato en los mismos términos que el anterior que da a su comprador el derecho a vender una unidad del activo subyacente al precio de ejercicio predeterminado.*

*En las opciones de tipo americano, el derecho a comprar o vender la unidad del activo subyacente que adquiere el comprador de la opción la puede ejercer en cualquier instante desde el inicial hasta el de vencimiento.*

Está claro que el comprador de la opción debería pagar un precio positivo al vendedor por adquirir este derecho, ya que en el momento de ejercer la opción puede obtener beneficios pero nunca pérdidas. Por ejemplo, si pensamos en una opción de compra europea, en el instante de vencimiento  $T$  el comprador tiene derecho a comprar una unidad del activo subyacente al precio  $K$ , cuando su precio en el mercado es  $S(T)$ . De esta manera, si  $S(T) \leq K$  el comprador simplemente no ejercerá su derecho a comprar un activo a un precio más caro que en el mercado. Sin embargo, si  $S(T) > K$  el tenedor de la opción comprará el activo por debajo de su precio de mercado, obteniendo un beneficio de  $S(T) - K$ . En consecuencia, el beneficio (nunca negativo) que obtiene el comprador de una opción *call* europea en el instante de vencimiento viene dado por:

$$\Phi(S(T)) = \max\{S(T) - K, 0\}.$$

De forma análoga, el beneficio (nunca negativo) que obtiene el comprador de una opción *put* europea en el instante de vencimiento viene dado por  $\Phi(S(T)) = \max\{K - S(T), 0\}$ . Entonces, estas opciones se pueden ver como un contrato que especifica que la cantidad de dinero estocástica  $X = \Phi(S(T))$  tiene que ser pagada en un instante futuro  $T$  al comprador del contrato. Además, como el pago  $X$  es  $\mathcal{F}_T$ -medible, se tiene que es conocido en el instante de vencimiento.

Una vez introducidos estos casos particulares de activos derivados, resulta natural la definición general que se da a continuación.

**Definición 3.2 (Activo derivado)** *Un activo derivado con fecha de vencimiento  $T$  es cualquier variable aleatoria  $X$  que sea  $\mathcal{F}_T$ -medible, la cual representa la cantidad de dinero (conocida en el instante de vencimiento) que el vendedor de la opción debe entregar al comprador. Como caso particular, se dice que el activo derivado es simple si es de la forma  $X = \Phi(S(T))$ , donde la función  $\Phi$  se denomina función contrato.*

Una consecuencia inmediata de la definición anterior es que las opciones *call* y *put* europeas son activos derivados simples, con funciones contrato  $\Phi(x) = \max\{x - K, 0\}$  y  $\Phi(x) = \max\{K - x, 0\}$  respectivamente.

La consideración de activos derivados en un mercado lleva a dos problemas fundamentales:

- Determinar un “precio justo” en algún sentido para el activo derivado considerado, y
- “cubrir” el riesgo financiero que el vendedor asume, ya que se compromete a pagar la cantidad estocástica  $X$  en un instante futuro  $T$ .

Para abordar estos problemas, es conveniente estudiar ciertas carteras autofinanciadas, en particular las de arbitraje y las de cobertura. A continuación se introduce la primera de ellas.

**Definición 3.3 (Cartera de arbitraje)** *Una cartera autofinanciada  $h$  se dice que es una cartera de arbitraje<sup>4</sup> si su proceso valor asociado es no negativo y además verifica que  $V_h(0) = 0$  y  $\mathbb{P}(V_h(T) > 0) > 0$ . Se dice que un modelo está libre de arbitraje si no existen carteras de arbitraje para cualquier  $T > 0$ .*

---

<sup>4</sup>También llamada oportunidad de arbitraje y posibilidad de arbitraje.

En relación a la definición anterior, una cartera de arbitraje se interpreta como una máquina determinista de hacer dinero, ya que es una posibilidad de hacer una cantidad positiva de dinero sin asumir ningún riesgo. Así, una posibilidad de arbitraje es un caso de malformación de precios en el mercado y una hipótesis usual que se hace sobre el mercado es que es *eficiente* en el sentido que no admite posibilidades de arbitraje.

En cuanto a las posibilidades de arbitraje, un ejemplo típico de potencial arbitraje ocurre cuando además del activo libre de riesgo  $B$  antes considerado, existe otro activo libre de riesgo con dinámica dada por

$$dY(t) = k(t)Y(t)dt, \quad (3.5)$$

donde  $k$  es un proceso medible y adaptado. Si el modelo está libre de arbitraje se debe tener en casi todo  $t$ ,  $k(t) = r(t)$ ,  $\mathbb{P}$ -casi seguro. En efecto, ya que si en algún intervalo de tiempo se tiene que  $k > r$ , se puede obtener una oportunidad de arbitraje tomando prestado del activo  $B$  e invirtiendo en  $Y$ , y al contrario si  $k < r$ . Esta es la razón por la que se considera un sólo activo libre de riesgo en los modelos financieros que se plantean.

Para trabajar con mercados eficientes, es muy importante conocer cuándo un modelo dado está libre de arbitraje, y a ese respecto la herramienta principal es el concepto de medida equivalente de martingalas, cuya definición se da a continuación.

**Definición 3.4 (Medida equivalente de martingalas)** *Se dice que una medida de probabilidad  $\mathbb{P}^*$  es una medida equivalente de martingalas si verifica:*

- $\mathbb{P}^* \sim \mathbb{P}$  (son medidas equivalentes) y
- el proceso de los precios descontados  $Z = \frac{S}{B}$  es una  $\mathbb{P}^*$ -martingala.

Un criterio útil para determinar si una medida equivalente dada  $\mathbb{P}^* \sim \mathbb{P}$  es una medida martingala, es que esto ocurre si y solo si los procesos  $S_i$  que modelizan los precios de los activos tienen una dinámica bajo  $\mathbb{P}^*$  de la siguiente forma:

$$dS_i(t) = r(t)S_i(t)dt + dM_i(t),$$

donde  $M_i$  es una  $\mathbb{P}^*$ -martingala ( $\forall i = 1, \dots, k$ ). Igualdad que pone de manifiesto que, bajo  $\mathbb{P}^*$ , el rendimiento medio de todas las acciones es el mismo y además coincide con el del activo libre de riesgo. La demostración de esta equivalencia es directa a partir del teorema de Girsanov (Teorema 2.21) y la regla de Itô.

El resultado básico de la teoría, vincula la ausencia de arbitraje a la existencia de una medida equivalente de martingalas. Sin embargo, para construir una teoría “razonable” se debe reducir la clase de las carteras autofinanciadas imponiendo alguna condición de integrabilidad, ya que si se trabaja con un mercado donde se exige solamente que las carteras sean procesos vectoriales medibles y  $\mathbb{F}$ -adaptados se pueden construir oportunidades de arbitraje en mercados muy generales (Björk (2004)[8], capítulo 10).

En este sentido, es importante observar que como por definición  $Z$  es una  $\mathbb{P}^*$ -martingala y su proceso valor descontado  $V_h^Z$  es la integral estocástica de  $h$  con respecto a  $Z$ , se tiene que para toda cartera autofinanciada suficientemente integrable (por ejemplo, uniformemente acotada) el proceso  $V_h^Z$  es una  $\mathbb{P}^*$ -martingala. Las carteras autofinanciadas  $h$  que llevan a procesos  $V_h^Z$  que son  $\mathbb{P}^*$ -martingalas se denominan *carteras admisibles*.

La necesidad de la condición de admisibilidad de las carteras autofinanciadas consideradas en la relación entre la existencia de medidas equivalentes de martingalas y la

exclusión de arbitraje en el mercado, queda patente en la prueba del siguiente resultado fundamental:

**Teorema 3.5 (Primer teorema fundamental de valoración de activos)** *Si existe una medida equivalente de martingalas  $\mathbb{P}^*$ , entonces el modelo está libre de arbitraje, en el sentido de que no existen carteras de arbitraje  $\mathbb{P}^*$ -admisibles.*

*Demostración.* Sea  $h$  una cartera de arbitraje  $\mathbb{P}^*$ -admisibles, entonces la variable aleatoria  $V_h(T)$  es no negativa, y por tanto  $V_h^Z(T)$  también lo es. En consecuencia, como  $Z$  es una  $\mathbb{P}^*$ -martingala se llega a

$$V_h(0) = V_h^Z(0) = \mathbb{E}^* [V_h^Z(T)] > 0,$$

donde aquí (y en adelante)  $\mathbb{E}^*$  denota la esperanza respecto a la medida  $\mathbb{P}^*$ , lo cual contradice la condición de arbitraje  $V_h(0) = 0$ . ♣

En realidad, el resultado anterior no es exactamente el conocido como primer teorema fundamental, ya que éste proporciona una doble implicación y el Teorema 3.5 nos dice que una condición suficiente para la ausencia de arbitraje es la existencia de una medida martingala. En este contexto, es importante mencionar que en Harrison y Pliska (1981)[55] se prueba que el recíproco del Teorema 3.5 es cierto en modelos de tiempo discreto con  $\Omega$  finito. Sin embargo, en casos más generales la situación es más delicada. Concretamente, Delbaen y Schachermayer (1994)[26] prueban que la existencia de una medida martingala es equivalente a un concepto similar al de ausencia de arbitraje pero más general, al cual denominan *ausencia de comida gratis con riesgo nulo (no free lunch with vanishing risk)*. Una recopilación de éste y otros resultados matemáticos referentes al arbitraje se encuentra en el reciente texto de Delbaen y Schachermayer (2006)[27]. En este trabajo no se considera tal nivel de generalidad, sino el del Teorema 3.5. Así, se consideran mercados



donde la propiedad de ausencia de arbitraje se garantiza mediante la existencia de una medida equivalente de martingalas.

Partiendo de un mercado libre de arbitraje se plantea el problema de cubrir un activo derivado mediante una cartera basada en los activos subyacentes, para lo cual resulta necesario introducir los siguientes conceptos:

**Definición 3.6** *Un activo derivado  $X$  con vencimiento en  $T$ , también llamado  $T$ -activo derivado, se dice que es alcanzable si existe una cartera autofinanciada  $h$  para la que su correspondiente proceso valor verifica que  $V_h(T) = X$ ,  $\mathbb{P}$ -casi seguro. Un mercado se dice completo si todo  $T$ -activo derivado  $X$ , con  $\frac{X}{B(T)}$  una variable aleatoria  $\mathbb{P}^*$ -integrable, es alcanzable.*

*La cartera  $h$  anterior se llama cartera de cobertura o replicante del activo derivado  $y$ , desde un punto de vista financiero, es indistinguible del activo derivado  $X$ .*

En este marco de trabajo, el problema principal es determinar los activos derivados que se pueden cubrir, lo cual resulta más conveniente expresarlo en términos de precios descontados. A este respecto, el siguiente resultado expresa que *la cobertura es equivalente a la existencia de una representación integral estocástica para el activo derivado descontado.*

**Proposición 3.7** *Sea  $\mathbb{P}^*$  una medida equivalente de martingalas y supongamos que el activo derivado descontado  $\frac{X}{B(T)}$  es  $\mathbb{P}^*$ -integrable. Si la  $\mathbb{P}^*$ -martingala  $M$  definida por*

$$M = \mathbb{E}^* \left[ \frac{X}{B(T)} \middle| \mathcal{F}_t \right] \tag{3.6}$$

admite una representación integral de la forma

$$M(t) = x + \sum_{i=1}^k \int_0^t h_i(u) dZ_i(u), \quad (3.7)$$

con  $h = (h_0, \dots, h_k)$  un proceso medible y adaptado, entonces  $X$  es alcanzable. Además, la cartera de cobertura está dada por  $(h_1, \dots, h_k)$  junto con  $h_0(t) = M(t) - \sum_{i=1}^k h_i(t) Z_i(t)$ .

*Demostración.* Teniendo en cuenta que, como se mencionó anteriormente, la condición de cartera autofinanciada no depende de la elección de numerario, se está buscando, en términos de precios descontados, un proceso  $\tilde{h} = (\tilde{h}_0, \dots, \tilde{h}_k)$  que verifique las siguientes condiciones:

$$V_h^Z(T) = \frac{X}{B(T)} \quad \mathbb{P} - \text{ect}, \quad (\text{Alcanzabilidad}) \quad (3.8)$$

$$dV_h^Z(t) = \sum_{i=1}^k \tilde{h}_i(t) dZ_i(t) \quad (\text{Autofinanciación}), \quad (3.9)$$

donde el proceso valor descontado viene dado por

$$V_h^Z(t) = \tilde{h}_0(t) \cdot 1 + \sum_{i=1}^k \tilde{h}_i(t) Z_i(t). \quad (3.10)$$

Comparando las igualdades (3.7) y (3.9) es razonable tomar  $(\tilde{h}_0, \dots, \tilde{h}_k) = (h_0, \dots, h_k)$  y así, con  $\tilde{h}_0(t) = h_0(t) = M(t) - \sum_{i=1}^k h_i(t) Z_i(t)$ , se tiene que  $M(t) = h_0(t) + \sum_{i=1}^k h_i(t) Z_i(t) = V_h^Z(t)$ . De esta manera, la representación (3.7) proporciona la condición de autofinanciación escrita en (3.9) y la igualdad

$$V_h^Z(T) = M(T) = \mathbb{E}^* \left[ \frac{X}{B(T)} \middle| \mathcal{F}_T \right] = \frac{X}{B(T)},$$

aporta la condición de alcanzabilidad (3.8). ♣

En consecuencia, aparte de algunas condiciones de integrabilidad, la completitud es equivalente a una representación martingala para el activo derivado descontado. En este contexto, Harrison y Pliska (1981)[55], usando resultados de la teoría semimartingala que relacionan las propiedades de representación martingala para  $Z$  con el conjunto de puntos de extremos del conjunto  $\mathcal{P} = \{\text{medidas equivalentes de martingalas}\}$ , llegan al siguiente resultado fundamental:

**Teorema 3.8 (Segundo teorema fundamental de valoración de activos)** *Sea un mercado libre de arbitraje. Entonces, éste es completo (en el sentido de que todo activo derivado  $X$  con  $\frac{X}{B(T)}$  una variable aleatoria  $\mathbb{P}^*$ -integrable se puede cubrir) si y solo si la medida martingala es única.*

Una vez establecido que la existencia de medidas equivalentes de martingalas garantiza que el mercado está libre de arbitraje, y que el mercado es completo cuando ésta medida es única, surge la cuestión de determinar si un modelo concreto dado, como el modelo descrito por las ecuaciones (3.1) y (3.2), está libre de arbitraje y es completo.

En este sentido, en Björk (1997)[7] se encuentra el siguiente razonamiento heurístico: Supongamos que en el proceso de los precios de los activos subyacentes aparecen  $R$  “fuentes aleatorias” (en el caso típico del modelo dado por las ecuaciones (3.1) y (3.2) las fuentes aleatorias son las componentes del movimiento browniano, que en este caso son  $d$ ). Al analizar la ausencia de arbitraje y la completitud se observa que estos dos conceptos van en “direcciones opuestas”. Tratemos de entender esta idea pensando en un modelo con  $R$  fuentes aleatorias fijas; en este modelo, cada nuevo activo subyacente añadido (sin modificar  $R$ ) dará más oportunidades para crear arbitraje. En consecuencia, para tener un modelo libre de arbitraje el número  $K$  de activos subyacentes (aparte del numerario) debe ser pequeño en comparación con  $R$ . Por otra parte, cada nuevo activo subyacente

añadido al modelo da nuevas posibilidades de cubrir un activo derivado, así la completitud requiere  $K$  grande en comparación con  $R$ . En consecuencia, la siguiente conjetura resulta de gran utilidad.

**Conjetura 3.9** *Sea  $K$  el número de activos subyacentes en un modelo, excluyendo el activo numerario, y sea  $R$  el número de fuentes aleatorias. En general se tienen las siguientes relaciones:*

- *El modelo está libre de arbitraje cuando  $K \leq R$ , y*
- *El modelo es completo cuando  $K \geq R$ .*

En algunos casos concretos, se puede dar una formulación precisa al resultado anterior por ejemplo, en el modelo de Black-Scholes generalizado se tiene el siguiente resultado (Karatzas y Shreve (1998)[67]):

**Teorema 3.10** *El modelo de mercado financiero dado por las ecuaciones (3.1) y (3.2) está libre de arbitraje y es completo cuando el número de activos sujetos a riesgo  $k$  es igual a la dimensión  $d$  del movimiento browniano del modelo y la matriz de volatilidad  $\sigma$  es no singular (en casi todo  $t \in [0, T]$ ,  $\mathbb{P}$ -casi seguro).*

Una vez introducidos los conceptos de mercado libre de arbitraje y completo y dadas reglas simples para su identificación, se pueden abordar los problemas de la valoración y la cobertura de activos derivados.

### 3.2.4. Valoración y cobertura de activos derivados

Tomando como punto de partida un modelo para el que existe alguna medida equivalente de martingalas, lo cual garantiza que el modelo está libre de arbitraje (Teorema 3.5), se plantea el problema de determinar un precio “justo” (en algún sentido), denotado por  $\Pi_X(t)$ , ( $0 \leq t \leq T$ ) para un activo derivado  $X$  que vence en el instante  $T$ . Para realizar tal valoración, existen fundamentalmente dos enfoques:

- Valorar el activo derivado de forma consistente con los precios de los activos subyacentes, en el sentido de que el mercado extendido  $(\Pi_X, S_0, \dots, S_k)$  esté libre de arbitraje.
- Si el activo derivado es alcanzable con cartera de cobertura  $h$ , el único precio razonable es el que viene dado por la igualdad  $\Pi_X(t) = V_h(t)$ .

En relación al primero de los enfoques, se exige que exista una medida equivalente de martingalas para el mercado extendido  $(\Pi_X, S_0, \dots, S_k)$ . Si  $\mathbb{P}^*$  es una de esas medidas martingalas, se deduce, suponiendo las condiciones de integrabilidad adecuadas, que:

$$\frac{\Pi_X(t)}{S_0(t)} = \mathbb{E}^* \left[ \frac{\Pi_X(T)}{S_0(T)} \middle| \mathcal{F}_t \right] = \mathbb{E}^* \left[ \frac{X}{S_0(T)} \middle| \mathcal{F}_t \right],$$

de donde se obtiene la siguiente *fórmula de valoración*:

$$\Pi_X(t) = S_0(t) \mathbb{E}^* \left[ \frac{X}{S_0(T)} \middle| \mathcal{F}_t \right], \quad \forall 0 \leq t \leq T, \quad (3.11)$$

donde  $\mathbb{P}^*$  es una medida martingala para el mercado dado  $(S_0, \dots, S_k)$  y considerando  $S_0$  como activo numerario. Es importante observar que diferentes elecciones de medidas martingalas  $\mathbb{P}^*$  llevan en general a diferentes procesos de precios, situación típica de los

mercados incompletos (ver Björk (2004) [8], capítulo 15, para profundizar en esta materia).

En referencia al segundo enfoque, en esa situación tener el activo derivado  $X$ , que se puede cubrir con la cartera  $h$ , y tener la cartera  $h$  es equivalente desde un punto de vista financiero. Por lo tanto, el precio del activo derivado se debe dar, en ausencia de arbitraje, mediante la igualdad  $\Pi_X(t) = V_h(t)$ . Sin embargo, una vez definido así el precio del activo derivado, se llega a la misma fórmula (3.11) que en el enfoque anterior, puesto que, siendo  $h$  una cartera admisible,  $V_h^Z(t) = \frac{\Pi_X(t)}{S_0(t)}$  es una  $\mathbb{P}^*$ -martingala. Así, para un activo derivado alcanzable se verifica la fórmula

$$V_h(t) = S_0(t) \mathbb{E}^* \left[ \frac{X}{S_0(T)} \middle| \mathcal{F}_t \right],$$

para cualquier cartera de cobertura  $h$  y cualquier medida equivalente de martingalas  $\mathbb{P}^*$  (si no fuera así se podrían construir fácilmente oportunidades de arbitraje). En consecuencia, los dos enfoques planteados coinciden sobre el conjunto de los activos derivados alcanzables.

A modo de resumen de todo lo anterior, se da el siguiente resultado:

- Proposición 3.11**
1. Si se supone la existencia de una medida martingala para el mercado ampliado  $(\Pi_X, S_0, \dots, S_k)$ , entonces el valor del activo derivado  $X$  con vencimiento en  $T$  viene dado por la fórmula (3.11), donde el valor esperado se calcula bajo una medida equivalente de martingalas  $\mathbb{P}^*$  para  $(S_0, \dots, S_k)$ .
  2. Diferentes elecciones de medidas equivalentes de martingalas  $\mathbb{P}^*$  llevan en general a diferentes procesos de precios, pero si  $X$  es alcanzable, todas las elecciones producen el mismo proceso de precios.
  3. Para un activo derivado alcanzable, el proceso de precios también viene dado por

$\Pi_X(t) = V_h(t)$ , donde  $h$  es la cartera de cobertura. Además, diferentes elecciones de carteras de cobertura producen el mismo proceso de precios.

En consecuencia, en un mercado completo el precio de cualquier activo derivado alcanzable queda unívocamente determinado por la hipótesis de ausencia de arbitraje. Así, el precio es único porque el activo derivado es en cierto sentido superfluo (puede ser reemplazado por su cartera de cobertura). Es importante observar también que el precio no depende de las preferencias de riesgo de los agentes del mercado, lo cual se suele reflejar diciendo que se trata de una *valoración neutral al riesgo*.

#### 3.2.5. Análisis del modelo de Black-Scholes

Una vez presentado el modelo generalizado de Black-Scholes (dado por las ecuaciones (3.1) y (3.2)), se plantea aplicar los resultados previos con objeto de llegar a conclusiones más específicas en el caso particular en el que solo hay una acción ( $k = 1$ ) y los coeficientes son constantes (modelo estándar de Black-Scholes). Por tanto, en este modelo, mencionado al principio de esta sección y en §2.3.2, las dinámicas del activo libre de riesgo  $B$  y del único activo sujeto a riesgo  $S$  vienen dadas por

$$\begin{aligned} dB(t) &= rB(t) dt, & B(0) &= 1 \\ dS(t) &= bS(t) dt + \sigma S(t) dW(t), & S(0) &= s_0 \in (0, \infty), \end{aligned} \tag{3.12}$$

donde los parámetros del modelo,  $r$  (el tipo de interés compuesto continuo),  $b$  (rendimiento medio local) y  $\sigma$  (volatilidad local), son constantes determinísticas.

Como se ha mencionado con anterioridad, la única solución a las ecuaciones del modelo (3.12) planteado viene dada, para  $t \geq 0$ , por:

$$\begin{aligned} B(t) &= B(0) e^{rt}, \\ S(t) &= S(0) \exp \left\{ \left( b - \frac{\sigma^2}{2} \right) t + \sigma W(t) \right\}, \end{aligned}$$

de donde se sigue que  $B$  es un activo libre de riesgo con tipo de interés compuesto continuo  $r$  y  $S$  evoluciona según un movimiento browniano geométrico.

Para conocer si el modelo está libre de arbitraje, el primer teorema fundamental (Teorema 3.5) afirma que esta situación está garantizada si existe una medida equivalente de martingalas en el modelo. En este sentido, se puede determinar el conjunto de medidas martingalas de este modelo, ya que usando el lema de Itô se llega a que la dinámica del precio descontado  $Z(t) = \frac{S(t)}{B(t)}$  viene dada por:

$$dZ(t) = (b - r) Z(t) dt + \sigma Z(t) dW(t).$$

Ahora, según el Teorema de Girsanov (Teorema 2.21), dada cualquier medida martingala  $\mathbb{P}^*$  existe un proceso  $\gamma$  tal que  $dW^*(t) = dW(t) - \gamma(t)dt$ , siendo  $W^*$  un  $\mathbb{P}^*$ -movimiento browniano. Entonces, la  $\mathbb{P}^*$ -dinámica del precio descontado  $Z$  viene dada por:

$$dZ(t) = (b - r + \sigma\gamma(t)) Z(t) dt + \sigma Z(t) dW^*(t).$$

En consecuencia, *para el modelo de Black-Scholes hay una única medida equivalente de martingalas*  $\mathbb{P}^*$ , que es la que obtiene considerando el proceso constante  $\gamma = -\frac{b-r}{\sigma}$ . Como consecuencia de los teoremas fundamentales de valoración de activos (Teoremas 3.5 y 3.8) se tiene que el modelo de Black-Scholes está libre de arbitraje y es completo.



Con propósitos ilustrativos, se va a probar que todo activo derivado  $X$  con vencimiento en  $T$  de cuadrado  $\mathbb{P}^*$ -integrable es alcanzable. Para ello, considerando la martingala

$$M(t) = \mathbb{E}^* [e^{-rT} X | \mathcal{F}_t],$$

se tiene que el teorema de representación martingala (Teorema 2.22) garantiza la existencia de un proceso  $g$  continuo,  $\mathbb{P}^*$ -adaptado, de cuadrado  $\mathbb{P}^*$ -integrable tal que

$$M(t) = M(0) + \int_0^t g(u) dW^*(u).$$

Teniendo además en cuenta que la  $\mathbb{P}^*$ -dinámica de  $Z$  viene dada por  $dZ(t) = \sigma Z(t) dW^*(t)$ , esta última igualdad se puede escribir como:

$$M(t) = M(0) + \int_0^t h_1(u) dZ(u),$$

donde  $h_1(t) = \frac{g(t)}{\sigma Z(t)}$ . Entonces, de la Proposición 3.7 se deduce que el activo derivado  $X$  es alcanzable usando la cartera dada por:

$$\begin{aligned} h_1(t) &= \frac{g(t)}{\sigma Z(t)}, \\ h_0(t) &= M(t) - h_1(t) Z(t) = M(t) - \frac{g(t)}{\sigma}. \end{aligned}$$

En el caso particular en el que el activo derivado es simple, es decir,  $X = \Phi(S(T))$  con  $\Phi$  una función real suficientemente integrable, se pueden obtener resultados más específicos. En concreto, en ese caso se verifica que

$$M(t) = \mathbb{E}^* [e^{-rT} \Phi(S(T)) | \mathcal{F}_t],$$

y usando la fórmula de Feynman-Kac (Teorema 2.23) se deduce que  $M(t) = F(t, S(t))$ ,

donde  $F(t, x)$  resuelve el siguiente problema de contorno:

$$\begin{aligned}\frac{\partial F}{\partial t}(t, x) + rx\frac{\partial F}{\partial x}(t, x) + \frac{1}{2}x^2\sigma^2\frac{\partial^2 F}{\partial x^2}(t, x) &= 0, \\ F(T, x) &= e^{-rT}\Phi(x).\end{aligned}\tag{3.13}$$

En cuanto a la cartera de cobertura, la regla de Itô lleva a que:

$$dM(t) = \sigma S(t) \frac{\partial F}{\partial x}(t, S(t)) dW^*(t),$$

con lo que, en términos de la notación anterior, se tiene que  $g(t) = \sigma S(t) \frac{\partial F}{\partial x}(t, S(t))$ , lo cual proporciona la siguiente cartera de cobertura:

$$\begin{aligned}h_0(t) &= F(t, S(t)) - S(t) \frac{\partial F}{\partial x}(t, S(t)), \\ h_1(t) &= S(t) \frac{\partial F}{\partial x}(t, S(t)),\end{aligned}$$

siendo  $F(t, S(t)) = V_h^Z(t)$  el valor descontado del activo derivado.

Por otra parte, el precio del activo derivado  $X$  con vencimiento en  $T$  se puede obtener aplicando la formula de valoración (3.11), la cual se escribe como:

$$\Pi_X(t) = e^{-r(T-t)}\mathbb{E}^*[X | \mathcal{F}_t],\tag{3.14}$$

donde la  $\mathbb{P}^*$ -dinámica de  $S$  está dada por  $dS(t) = rS(t)dt + \sigma S(t)dW^*(t)$ . En el caso particular de un activo derivado simple  $X = \Phi(S(T))$ , el proceso de los precios viene dado por  $\Pi_X(t) = e^{rt}F(t, S(t))$ , donde  $F$  resuelve el problema de contorno (3.13). En el caso particular de una opción de compra europea con precio de ejercicio  $K$  se tiene que  $\Phi(x) = \max\{0, x - K\}$ . En este caso, resulta más fácil calcular el valor esperado (3.14) que resolver el problema de contorno (3.13). En concreto, esta esperanza condicionada

viene dada por:

$$\begin{aligned}\Pi_X(t) &= e^{-r\tau} \mathbb{E}^* [\Phi(S(T)) | \mathcal{F}_t] \\ &= e^{-r\tau} \mathbb{E}^* \left[ \Phi \left( S(t) \cdot \exp \left\{ \left( r - \frac{\sigma^2}{2} \right) \tau + \sigma (W^*(T) - W^*(t)) \right\} \right) \middle| \mathcal{F}_t \right],\end{aligned}$$

donde  $\tau = T - t$  es el tiempo desde el instante actual  $t$  hasta el vencimiento  $T$  del activo. Teniendo en cuenta que  $S(t)$  es  $\mathcal{F}_t$ -medible y que el incremento browniano  $W^*(T) - W^*(t)$  es independiente de  $\mathcal{F}_t$ , se tiene que  $\Pi_X(t) = u(\tau, S(t))$  donde

$$u(\tau, x) = e^{-r\tau} \int_{\mathbb{R}} \Phi \left( x \cdot \exp \left\{ \left( r - \frac{\sigma^2}{2} \right) \tau + \sigma \sqrt{\tau} z \right\} \right) \varphi(z) dz,$$

siendo  $\varphi$  la función de densidad de una variable aleatoria normal estándar. Esta última integral se puede calcular usando que  $\Phi(x) = \max\{0, x - K\}$  y estableciendo la notación

$$Z_1 = \frac{1}{z\sqrt{\tau}} \left[ \log \left( \frac{x}{K} \right) + \left( r + \frac{\sigma^2}{2} \right) \tau \right] \quad \text{y} \quad Z_2 = Z_1 - \sigma\sqrt{\tau},$$

llegando a que el valor de una opción de compra europea está dado por

$$u(\tau, x) = \begin{cases} xN(Z_1) - Ke^{-r\tau}N(Z_2), & 0 < \tau \leq T, \quad s \in (0, \infty) \\ \max\{0, x - K\}, & \tau = 0, \quad s \in (0, \infty) \end{cases}$$

donde  $N$  es la función de distribución de una variable aleatoria normal estándar. Ésta es la conocida fórmula de Black-Scholes (1973)[11], en la que es importante observar que no depende del parámetro  $b$  (rendimiento medio del activo).

Un razonamiento análogo conduce a que el precio de una opción de venta europea,

ahora la función de contrato es  $\Phi(x) = \max\{0, K - x\}$ , viene dado por:

$$u(\tau, x) = \begin{cases} Ke^{-r\tau}N(-Z_2) - xN(-Z_1), & 0 < \tau \leq T, s \in (0, \infty) \\ \max\{0, K - x\}, & \tau = 0, s \in (0, \infty). \end{cases}$$

Una vez estudiados los aspectos básicos de la modelización de los mercados financieros mediante el modelo generalizado de Black-Scholes, se realiza en la sección siguiente un estudio teórico del caso particular del mercado de bonos, poniendo especial atención en su descripción mediante modelos de tipos de interés a corto plazo.

### 3.3. Mercado de bonos y modelos de tipos de interés a corto plazo

El objetivo que se persigue en esta sección, es aplicar las técnicas generales desarrolladas en la sección anterior al caso particular de los mercados de renta fija. En este sentido, queremos recordar que los *activos de renta fija* son títulos que representan deudas de sus emisores con sus tenedores, retribuidos mediante el pago de intereses y con vencimientos determinados, ejemplos típicos son las Letras del Tesoro, los Bonos y Obligaciones del Estado, emitidos por el Tesoro Público, y los depósitos interbancarios. Por contra, los *activos de renta variable* representan participaciones en el capital de los emisores y tienen vencimientos indefinidos o indeterminados; el ejemplo típico son las conocidas acciones.

Con objeto de establecer una teoría de precios para el mercado de renta fija basada en

la hipótesis de no arbitraje, se necesitan los precios de algunos activos subyacentes dados. Así, el elemento básico en la construcción de la teoría que se va a presentar son los bonos cupón cero, los cuales son activos que garantizan a su comprador el pago de una cantidad predeterminada (por convención se va a tomar 1€) en un instante de vencimiento fijo. Entonces, discretizando el espectro de vencimientos, se puede pensar en usar el modelo de Black-Scholes generalizado para los rendimientos de los diferentes bonos cupón cero. En este contexto, es importante señalar que investigaciones empíricas realizadas en Dybvig (1989)[36] y Litterman y Scheinkman (1991)[75], muestran que el grado de correlación entre los rendimientos con diferentes vencimientos es en general muy alto (cercano a 1). Concretamente, mediante el uso de análisis de componentes principales se llega a que la primera componente principal con frecuencia explica entre el 80% y el 90% de la varianza total de los rendimientos. En consecuencia, es razonable usar un tipo de interés específico, normalmente el tipo a corto plazo  $r(t)$ , como variable *proxy* que mejor describe los movimientos de toda la curva de rendimientos, éstos son los llamados *modelos unifactoriales o modelos a corto plazo*, los cuales se presentan en §3.3.3.

#### 3.3.1. Mercado de bonos y tipos de interés

Como se ha mencionado anteriormente, el objetivo fundamental en esta sección es establecer una teoría para los precios de los mercados de renta fija. Los objetos básicos sobre los que se construye dicha teoría, son los bonos cupón cero, cuya definición precisa en el marco estándar de un espacio de probabilidad filtrado  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$  y con un horizonte de tiempo finito  $T^*$  se da a continuación.

**Definición 3.12** *Un bono cupón cero con fecha de vencimiento  $T < T^*$ , también llamado  $T$ -bono, es un contrato que garantiza a su tenedor que le será pagado 1€ en el instante*

$T$ . El precio de un  $T$ -bono en un instante  $t \in [0, T]$  se denota por  $p(t, T)$ .

Está claro que se debe verificar que  $p(t, t) = 1, \forall t \in [0, T]$ , para evitar oportunidades de arbitraje. Además, para garantizar la existencia de un mercado de bonos suficientemente rico y regular, se supone que en cada instante se negocian bonos con todos los vencimientos y que el proceso de los precios  $p(t, T)$  es un proceso adaptado tal que para cada  $t \in [0, T]$  fijo,  $p(t, T)$  es una función diferenciable con respecto a la fecha de vencimiento  $T$ .

Es importante observar que el mercado de bonos es, en principio, bastante diferente del mercado financiero planteado en la sección anterior, ya que contiene un número infinito de activos (un tipo de bono para cada vencimiento).

Dado el mercado de bonos antes descrito, se pueden definir, en base a consideraciones de arbitraje, varios tipos de interés. Para ello, sean  $t < T_1 < T_2$  tres fechas, la cuestión básica es la siguiente: ¿Cuál es el tipo de interés determinista vigente en el instante de contrato  $t$ , para una inversión de 1€ realizada en el instante  $T_1$  sobre el intervalo  $[T_1, T_2]$ ? Para responder a esta pregunta se considera la Tabla 3.1 de varias negociaciones llevadas a cabo por un inversor.

Tiempo	$t$	$T_1$	$T_2$
Inversiones	Vender un $T_1$ -bono Comprar $\frac{p(t, T_1)}{p(t, T_2)}$ $T_2$ -bonos	Pagar 1€	Recibir $\frac{p(t, T_1)}{p(t, T_2)}$ €
Inversión neta (€)	0	-1	$+\frac{p(t, T_1)}{p(t, T_2)}$

**Tabla 3.1:** Tabla de arbitraje usada para definir los tipos de interés forward.

De esta tabla se deduce que para excluir oportunidades de arbitraje en el mercado, el tipo de interés constante (continuo compuesto)  $R$  sobre este periodo (el inversor paga 1€

en  $T_1$  y recibe  $\frac{p(t, T_1)}{p(t, T_2)} \text{€}$  en  $T_2$ ) tiene que verificar la igualdad:

$$1 \cdot e^{R(T_2 - T_1)} = \frac{p(t, T_1)}{p(t, T_2)},$$

la cual motiva las siguientes definiciones.

**Definición 3.13** 1. El tipo de interés a plazo (*forward interest rate*) contratado en el instante  $t$  para el periodo  $[T_1, T_2]$ , denotado por  $R(t; T_1, T_2)$ , se define como:

$$R(t; T_1, T_2) = -\frac{\log p(t, T_2) - \log p(t, T_1)}{T_2 - T_1},$$

y representa el tipo de interés (compuesto continuo) vigente en el instante actual  $t$  para una inversión en el intervalo de tiempo futuro  $[T_1, T_2]$ .

2. El tipo de interés de contado (*spot interest rate*) para el periodo  $[t, T]$ , denotado por  $R(t, T)$ , se define como:

$$R(t, T) = R(t; t, T) = -\frac{\log p(t, T)}{T - t}, \quad (3.15)$$

y representa el tipo de interés (compuesto continuo) vigente en el instante actual  $t$  para una inversión en el intervalo de tiempo  $[t, T]$ . Esta cantidad es conocida también como el rendimiento (compuesto continuo) del  $T$ -bono cupón cero en el intervalo  $[t, T]$ .

3. El tipo de interés a plazo instantáneo contratado en  $t$  y con vencimiento en  $T$ , denotado por  $f(t, T)$ , se define como:

$$f(t, T) = \lim_{T_2 \downarrow T} R(t; T, T_2) = -\frac{\partial \log p(t, T)}{\partial T},$$

y representa el tipo de interés (compuesto continuo) vigente en el instante actual  $t$  para una inversión en el intervalo infinitesimal de tiempo futuro  $[T, T + dt]$ .

4. El tipo de interés de contado instantáneo contratado en  $t$ , denotado por  $r(t)$ , se define como:

$$r(t) = f(t, t),$$

y representa el tipo de interés (compuesto continuo) vigente en el instante actual  $t$  para una inversión en el intervalo infinitesimal de tiempo  $[t, t + dt]$ .

De la misma manera que en la sección anterior se va a usar como activo numerario en el proceso de valoración de activos derivados de los tipos de interés, el proceso *cuenta bancaria*, el cual obedece a la dinámica dada por la ecuación 3.1, es decir  $B(t) = \exp\left(\int_0^t r(u) du\right)$  y describe la situación clásica de un banco que ofrece un tipo de interés estocástico  $r$ .

### 3.3.2. Valoración de activos derivados

Una vez planteado el mercado de bonos, la siguiente tarea es disponer de condiciones que garanticen que esté libre de arbitraje. En este sentido, el Teorema 3.5 asegura la ausencia de arbitraje bajo la existencia de una medida equivalente de martingalas  $\mathbb{P}^*$ , es decir, una medida  $\mathbb{P}^* \sim \mathbb{P}$  con respecto a la cual los procesos de los precios descontados  $Z(t, T) = \frac{p(t, T)}{B(t)}$  ( $0 \leq t \leq T$ ), con  $T < T^*$  fijo, son martingalas.

Entonces, si se supone la existencia de al menos una medida equivalente de martingalas  $\mathbb{P}^*$ , considerando activos derivados como variables aleatorias  $\mathcal{F}_T$ -medibles con  $\frac{X}{B(T)}$  una variable aleatoria  $\mathbb{P}^*$ -integrable, el principio de valoración de riesgo neutral recogido en



(3.11) lleva a que el proceso de los precios del activo derivado viene dado por

$$\Pi_X(t) = B(t) \mathbb{E}^* \left[ \frac{X}{B(T)} \middle| \mathcal{F}_t \right] = \mathbb{E}^* \left[ X \cdot e^{-\int_t^T r(u) du} \middle| \mathcal{F}_t \right]. \quad (3.16)$$

En particular, el proceso de los precios de un  $T$ -bono cupón cero es el siguiente:

$$p(t, T) = \mathbb{E}^* \left[ e^{-\int_t^T r(u) du} \middle| \mathcal{F}_t \right].$$

Como consecuencia importante de las fórmulas anteriores, se observa que la dinámica de los precios bajo la medida de probabilidad objetiva  $\mathbb{P}$  es, hasta cierto punto, irrelevante, ya que lo importante es conocer la dinámica de los precios bajo una medida martingala  $\mathbb{P}^*$ . Esto lleva a que en la construcción de modelos se prefiera en ocasiones modelizar todas las dinámicas directamente bajo una medida equivalente de martingalas  $\mathbb{P}^*$ . Este enfoque es conocido como *modelización martingala*. La desventaja de este enfoque viene dada por los problemas estadísticos asociados con la estimación de parámetros, ya que la evolución de los procesos se observa bajo la medida objetiva  $\mathbb{P}$ .

#### 3.3.3. Modelos de tipos de interés a corto plazo

Teniendo en cuenta las consideraciones realizadas al principio de esta sección justificando el uso del tipo de interés a corto plazo  $r$  como única variable exógena para describir los movimientos de toda la curva de rendimientos, se pasa a desarrollar este enfoque.

En este marco de trabajo, se modeliza el tipo a corto plazo, bajo la medida de proba-

bilidad objetiva  $\mathbb{P}$ , como la solución a la siguiente ecuación diferencial estocástica:

$$dr(t) = \mu(t, r(t)) dt + \sigma(t, r(t)) dW(t), \quad (3.17)$$

donde  $\mu(t, r)$  y  $\sigma(t, r)$  son funciones reales dadas satisfaciendo condiciones de regularidad suficientes para que la ecuación (3.17) tenga una única solución (por ejemplo verificando las hipótesis del Teorema 2.3.1). Como es usual,  $W$  es un  $\mathbb{P}$ -movimiento browniano estándar y la filtración considerada  $\mathbb{F}$  es la filtración browniana.

Es muy importante tener claro cuáles son los objetos que se dan a priori en la modelización (exógenos) y cuáles se deducen endógenamente a partir del modelo planteado. En el caso que se está desarrollando el único objeto dado a priori, modelizado por la dinámica en (3.17), es el tipo de interés a corto plazo  $r$ , y en consecuencia el único activo exógenamente dado es el activo libre de riesgo o cuenta bancaria  $B$ . Entonces, éste es el único activo subyacente dado, mientras que todos los bonos se consideran activos derivados<sup>5</sup> (del tipo de interés a corto plazo), en el mismo sentido que las opciones son activos derivados de las acciones en el modelo de Black-Scholes analizado en §3.2.5. Por tanto, es importante observar una diferencia fundamental con el modelo estándar de Black-Scholes: En éste último, se tienen dos activos (la cuenta bancaria y la acción) dados a priori (exógenos) mediante las dinámicas recogidas en (3.12), resultando un modelo con una única medida equivalente de martingalas (y por tanto un mercado libre de arbitraje y completo). Sin embargo, en la modelización que se está proponiendo para el mercado de bonos, existe solo un activo dado a priori (la cuenta bancaria), lo que resulta, como se expone más adelante, en un modelo con multitud de medidas equivalentes de martingalas (y por tanto en un mercado libre de arbitraje pero no completo).

---

<sup>5</sup>Un bono cupón cero con vencimiento en  $T$ , es un contrato que especifica que  $1\text{€}$  le debe ser pagado a su comprador en el instante de vencimiento. Por tanto, se trata del caso particular de  $T$ -activo derivado (ver Definición 3.2) con  $X = 1$ .

Una vez se tiene el modelo planteado, una cuestión muy importante es analizar si la especificación de la  $\mathbb{P}$ -dinámica de  $r$  en (3.17) junto con la hipótesis de un mercado de bonos libre de arbitraje determina unívocamente los precios de los bonos. La respuesta a esta cuestión es negativa, y la razón se debe a que el modelo planteado no es completo. A este respecto, nótese que como el único activo subyacente exógenamente dado es la cuenta bancaria, sólo se pueden formar carteras poniendo todo el dinero en el banco y esperar pasivamente, y por tanto, no hay posibilidad de cubrir los activos derivados en los que se está interesado. Esta falta de completitud se debe a la ausencia de unicidad de medidas equivalentes de martingalas en el modelo dado. En concreto, como el único proceso de precios descontados  $Z_0(t) = \frac{B(t)}{B(t)} = 1$ , es constante, cualquier medida de probabilidad  $\mathbb{P}^* \sim \mathbb{P}$  es una medida equivalente de martingalas y genera un mercado de bonos libre de arbitraje mediante la igualdad

$$p(t, T) = \mathbb{E}^* \left[ e^{-\int_t^T r(u)du} \middle| \mathcal{F}_t \right].$$

Esta situación se puede entender también a partir de las ideas informales de la Conjetura 3.9, ya que en este caso se tiene que  $K = 0$  (aparte del activo numerario no hay ningún otro activo negociado, puesto que el tipo a corto plazo  $r$  no es el precio de ningún activo negociado) y  $R = 1$  (la dimensión del movimiento browniano subyacente), lo que lleva a un mercado libre de arbitraje pero incompleto. Sin embargo, parece razonable que:

- Los precios de los bonos con diferentes vencimientos satisfagan ciertas relaciones de consistencia para evitar posibilidades de arbitraje en el mercado de bonos, y
- si se añade un activo subyacente más al mercado sin introducir un nuevo proceso de Wiener, es decir, si se toma el precio de un bono “referencia” como dado (exógeno), entonces los precios de todos los otros bonos deben estar unívocamente determinados en términos del precio de este bono de referencia (y la dinámica de  $r$ ).

Estas ideas “informales” están en completo acuerdo con la Conjetura 3.9, según el cual con el bono de referencia más el activo numerario se tiene  $K = R = 1$  y por tanto un mercado libre de arbitraje y completo.

Con objeto de justificar las ideas informales anteriores, se va a considerar que el proceso de los precios de un  $T$ -bono viene dado por  $p(t, T) = F(t, r(t); T)$ , donde  $F$  es una función de tres variables suficientemente diferenciable, la cual se suele denotar también como  $F_T(t, r)$ . Se buscan las relaciones que deben satisfacer las funciones de precio para diferentes vencimientos, y para ello se desea conocer cómo debe ser  $F$  en un mercado libre de arbitraje. En primer lugar, se tiene la siguiente condición frontera  $F(T, r; T) = 1$  ( $\forall r$ ), expresando que el valor de un bono en su instante de vencimiento es 1€. Además, si fijamos dos instantes de vencimiento  $T_1$  y  $T_2$ , la aplicación de la regla de Itô lleva a las siguientes dinámicas para los precios de los bonos:

$$dF_{T_i}(t, r(t)) = F_{T_i}(t, r(t)) \cdot \alpha_{T_i}(t, r(t)) dt + F_{T_i}(t, r(t)) \cdot \sigma_{T_i}(t, r(t)) dW(t), \quad (3.18)$$

donde, omitiendo por comodidad de aquí en adelante los argumentos de las funciones, se tiene que:

$$\begin{cases} \alpha_{T_i} &= \frac{1}{F_{T_i}} \left[ \frac{\partial F_{T_i}}{\partial t} + \mu \frac{\partial F_{T_i}}{\partial r} + \frac{1}{2} \sigma^2 \frac{\partial^2 F_{T_i}}{\partial r^2} \right] \\ \sigma_{T_i} &= \frac{1}{F_{T_i}} \sigma \frac{\partial F_{T_i}}{\partial r}, \end{cases} \quad (3.19)$$

para  $i = 1, 2$ . De esta forma, si se considera una cartera autofinanciada compuesta únicamente por  $T_1$ -bonos y  $T_2$ -bonos, se tiene, en términos de la cartera relativa  $(u_{T_1}, u_{T_2})$  (ver la ecuación (3.4)) la siguiente dinámica del proceso valor asociado a la cartera

$$dV = V \cdot \left( u_{T_1} \frac{dF_{T_1}}{F_{T_1}} + u_{T_2} \frac{dF_{T_2}}{F_{T_2}} \right),$$

y usando los diferenciales en (3.18), esta ecuación se reescribe como:

$$dV = V \cdot (u_{T_1} \alpha_{T_1} + u_{T_2} \alpha_{T_2}) dt + V \cdot (u_{T_1} \sigma_{T_1} + u_{T_2} \sigma_{T_2}) dW. \quad (3.20)$$

Se va a considerar la elección particular de cartera cuyas ponderaciones hacen que el término  $dW$  en (3.20) desaparezca y su dinámica se reduzca a  $dV = V \cdot (u_{T_1} \alpha_{T_1} + u_{T_2} \alpha_{T_2}) dt$ . En concreto, para conseguirlo se debe elegir las ponderaciones (en términos relativos) que verifican:

$$\begin{cases} u_{T_1} + u_{T_2} = 1 \\ u_{T_1} \sigma_{T_1} + u_{T_2} \sigma_{T_2} = 0, \end{cases} \quad (3.21)$$

lo que lleva a tomar,  $u_{T_1} = -\frac{\sigma_{T_2}}{\sigma_{T_1} - \sigma_{T_2}}$  y  $u_{T_2} = -\frac{\sigma_{T_1}}{\sigma_{T_1} - \sigma_{T_2}}$ . De este modo, la dinámica (3.20) se puede reescribir como

$$dV = V \cdot \left( \frac{\alpha_{T_2} \sigma_{T_1} - \alpha_{T_1} \sigma_{T_2}}{\sigma_{T_1} - \sigma_{T_2}} \right) dt.$$

Por otro lado, como el mercado de bonos está libre de arbitraje, esta cartera debe tener un rendimiento medio igual al tipo de interés a corto plazo (ver ecuación (3.5) y discusión posterior), y de aquí, se llega a la condición siguiente:

$$\frac{\alpha_{T_2} \sigma_{T_1} - \alpha_{T_1} \sigma_{T_2}}{\sigma_{T_1} - \sigma_{T_2}} = r(t), \quad \text{para todo } t, \text{ con probabilidad 1,}$$

o escrito de otra forma:

$$\frac{\alpha_{T_1}(t) - r(t)}{\sigma_{T_1}(t)} = \frac{\alpha_{T_2}(t) - r(t)}{\sigma_{T_2}(t)}.$$

Es importante observar que el lado izquierdo de la ecuación anterior no depende de  $T_2$  y el lado derecho no depende de  $T_1$ . En consecuencia, se tiene que ese cociente común no depende de la fecha de vencimiento, llegando al siguiente resultado fundamental:

**Proposición 3.14** *Si el mercado de bonos está libre de arbitraje, entonces existe un proceso  $\lambda_0$  tal que la relación*

$$\frac{\alpha_T(t) - r(t)}{\sigma_T(t)} = \lambda_0(t), \quad (3.22)$$

*se tiene para todo  $0 \leq t \leq T$  sea cual sea la elección del vencimiento  $T$ .*

La ecuación (3.22) tiene la siguiente interpretación: En el numerador aparece  $\alpha_T(t) - r(t)$  que es la diferencia entre la tasa de rendimiento local del  $T$ -bono y la tasa de rendimiento del activo libre de riesgo. Esta cantidad se conoce *prima de riesgo* del  $T$ -bono y se interpreta como el exceso de rendimiento del  $T$ -bono respecto al activo libre de riesgo necesario en el mercado para evitar oportunidades de arbitraje. En el denominador aparece la volatilidad del  $T$ -bono, con lo que  $\lambda_0$  se interpreta como la prima de riesgo por unidad de volatilidad y es conocido como *precio del riesgo de mercado*<sup>6</sup>.

A partir de la igualdad (3.22) se puede obtener una de las ecuaciones más importantes en la teoría de los tipos de interés. En concreto, sustituyendo en ellas las expresiones en (3.19) se llega a la conocida como *ecuación de la estructura temporal (term structure equation)* del  $T$ -bono, la cual tiene la siguiente expresión:

$$\begin{cases} \frac{\partial F_T}{\partial t} + (\mu - \lambda) \frac{\partial F_T}{\partial r} + \frac{1}{2} \sigma^2 \frac{\partial^2 F_T}{\partial r^2} - r F_T = 0, \\ F_T(T, r) = 1, \end{cases} \quad (3.23)$$

donde  $\lambda = \lambda_0 \cdot \sigma$ . Nótese que esta ecuación es similar a la ecuación (3.13) obtenida en el modelo de Black-Scholes; sin embargo, esta última es más complicada debido a la aparición de la prima de riesgo  $\lambda$ , la cual no está especificada de manera endógena en el modelo

---

<sup>6</sup>A menudo ocurre que  $\lambda_0$  negativo. Esto sucede porque el precio  $F_T$  es usualmente una función decreciente de  $r$  y  $\sigma$  se toma positiva, con lo que resulta que  $\sigma_T = \frac{1}{F_T} \sigma \frac{\partial F_T}{\partial r}$  es negativa. De esta manera,  $\lambda_0$  debe ser negativo para que el rendimiento medio del  $T$ -bono,  $\alpha_T$ , sea superior al tipo de interés libre de riesgo  $r$ . Esta situación se da en la aplicación empírica desarrollada en el capítulo siguiente.

sino que debe ser especificada y estimada (de la misma manera que  $\mu$  y  $\sigma$ ) para resolver la ecuación.

Teniendo en cuenta la Observación 2.24, se obtiene la representación de Feynman-Kac del precio  $F_T$ . En concreto, fijados  $t$  y  $r$ , los precios del  $T$ -bono  $p(t, T) = F_T(t, r(t))$  admiten la fórmula de representación estocástica siguiente:

$$F_T(t, r) = \mathbb{E}_{t,r}^* \left[ e^{-\int_t^T r(u) du} \right], \quad (3.24)$$

donde el valor esperado se toma con respecto a la medida equivalente de martingalas  $\mathbb{P}^*$ , bajo la cual el proceso de los tipos de interés a corto plazo tienen la siguiente dinámica:

$$\begin{cases} dr(u) &= (\mu - \lambda) du + \sigma dW^*(u), \\ r(t) &= r, \end{cases}$$

siendo  $W^*$  un  $\mathbb{P}^*$ -movimiento browniano estándar.

Es importante observar que la medida martingala depende de  $\lambda$ , así diferentes elecciones de este parámetro llevan a diferentes medidas martingalas. Esta situación es diferente de la encontrada en el modelo de Black-Scholes donde, debido a la completitud del modelo, la medida martingala es única. Cuando un mercado particular determina, mediante las relaciones entre oferta y demanda, la dinámica del precio de un solo bono, indirectamente especifica  $\lambda$  mediante (3.22) y los precios de todos los demás bonos se pueden obtener resolviendo la ecuación de la estructura temporal (3.23).

Los bonos cupón cero son un caso particular de activos derivados, la ecuación de la estructura temporal (3.23) y la representación estocástica (3.24) se pueden generalizar sin

dificultad a  $T$ -activos derivados de la forma  $X = \Phi(r(T))$  usando argumentos similares a los aplicados para los bonos (o también usando la formula general de valoración (3.16) y la Observación 2.24). Así, a continuación se expone el resultado general al que se llega.

**Proposición 3.15** *Sea  $X$  un  $T$ -activo derivado de la forma  $X = \Phi(r(T))$ . En un modelo libre de arbitraje su precio viene dado por  $\Pi_X(t) = F(t, r(t))$ , donde la función  $F$  resuelve la ecuación general de la estructura temporal, dada por:*

$$\begin{cases} \frac{\partial F_T}{\partial t} + (\mu - \lambda) \frac{\partial F_T}{\partial r} + \frac{1}{2} \sigma^2 \frac{\partial^2 F_T}{\partial r^2} - r F_T = 0, \\ F_T(T, r) = \Phi(r). \end{cases} \quad (3.25)$$

Además,  $F$  admite la siguiente representación estocástica:

$$F_T(t, r) = \mathbb{E}_{t,r}^* \left[ e^{-\int_t^T r(u) du} \cdot \Phi(r(T)) \right], \quad (3.26)$$

donde el valor esperado se toma, fijando  $t$  y  $r$ , con respecto a una medida equivalente de martingalas  $\mathbb{P}^*$ , bajo la que se tiene la siguiente dinámica de los tipos de interés a corto plazo:

$$\begin{cases} dr(u) = (\mu - \lambda) du + \sigma dW^*(u), \\ r(t) = r. \end{cases} \quad (3.27)$$

A primera vista, puede parecer que este resultado contradice las consideraciones realizadas al principio de esta sección sobre la incompletitud del mercado y la imposibilidad establecer un único precio para un activo derivado, todo ello debido a la existencia de múltiples medidas equivalentes de martingalas. Por otro lado, este resultado proporciona una ecuación en derivadas parciales y una fórmula de valoración que proporcionan la única función de precios para un activo derivado simple. La explicación de esta aparente contradicción es que para resolver el problema de contorno (3.25) o calcular el valor es-



perado (3.26) bajo la dinámica (3.27) se debe conocer  $\lambda$  y, debido a la correspondencia biunívoca entre medidas equivalentes de martingalas y precios de riesgo de mercado, elegir  $\lambda$  es equivalente a elegir una  $\mathbb{P}^*$ -particular.

Una vez establecidos los resultados básicos generales de la modelización de mercados financieros mediante el uso de ciertas ecuaciones diferenciales estocásticas y analizado los casos particulares de Black-Scholes y del mercado de bonos, se desarrolla en el capítulo siguiente una aplicación empírica en un mercado particular. En concreto se modeliza el tipo de interés a corto plazo para la economía española y se valoran bonos.



# Capítulo 4

## Un análisis no paramétrico de los tipos de interés a corto plazo en España

### 4.1. Introducción

Es innegable la importancia que la modelización de los tipos de interés tiene en el ámbito económico, y una prueba de ello es la enorme atención que ha recibido por parte de los investigadores en las últimas décadas, lo que ha generado un gran volumen de trabajos y publicaciones. Los motivos más importantes de la relevancia de estas variables económicas son variados, entre ellos destacan los siguientes: Aportan mucha información sobre la situación de la Economía (inflación, política monetaria,...) de una región; en

Teoría Económica se usan, entre otros, en el estudio de la formación de expectativas, las relaciones entre los tipos a corto y a largo plazo, la política monetaria, etc; en Economía Financiera, permite la valoración correcta de activos financieros dependientes de los tipos de interés y de sus activos derivados, lo que lleva a un conveniente diseño de estrategias de valoración y cobertura; etc.

Respecto a la teoría existente, cuyos aspectos básicos se han presentado en el capítulo anterior, nos dice cómo se debe actuar para valorar activos y obtener estrategias de cobertura, una vez que se dispone de un modelo concreto. Sin embargo, en la literatura sobre el tema no existe consenso respecto al modelo que se debe elegir para representar la dinámica de tales variables.

Así, en el presente capítulo, con objeto de realizar un estudio empírico de los tipos de interés libres de riesgo a corto plazo en España, se ha optado, siguiendo la tendencia actual en economía financiera, por representar la evolución temporal de la variable subyacente (el tipo de interés continuo compuesto instantáneo libre de riesgo), a cuyo valor en el instante  $t$  denotamos por  $r(t)$ , mediante un proceso estocástico que evoluciona en tiempo continuo y que verifica una determinada ecuación diferencial estocástica unidimensional.

Siguiendo los comentarios al respecto del capítulo anterior, en éste, nos centramos en los llamados modelos unifactoriales, los cuales describen los movimientos de toda la curva o estructura temporal de los tipos de interés (la que relaciona los tipos de interés proporcionados por activos libres de riesgo con su plazo de vencimiento) a partir de la especificación de una sola variable, el tipo de interés a corto plazo. En consecuencia, estos modelos producen una estructura temporal de los tipos de interés endógena, no ajustando en general a la perfección la curva de tipos de interés inicialmente observada en el mercado.

Los modelos unifactoriales, los cuales se pueden generalizar de varias formas diferentes (ver por ejemplo, Rebonato (1998)[89], Pelsser (2000)[86], James y Webber (2000)[63], Cairns[17], Brigo y Mercurio (2006)[16]), han sido criticados por éste y otros motivos; sin embargo, proporcionan respuestas razonables a una gran variedad de problemas planteados en el marco de la Economía Financiera. Además, son modelos que destacan por su sencillez para explicar la estructura de los tipos de interés y para ser implementados.

Por otro lado, en el contexto de los modelos unifactoriales se han propuesto una gran variedad de modelos diferentes para los tipos de interés a corto plazo, la mayoría de ellos se consideran dentro de un contexto paramétrico, atendiendo quizás más a razones de tratabilidad analítica y computacional que a razones de tipo económico. En concreto, un enfoque muy usado en la práctica consiste en modelizar la dinámica de los tipos de interés a corto plazo y libres de riesgo como la solución de la siguiente ecuación diferencial estocástica homogénea:

$$dr(t) = \mu(r(t)) dt + \sigma(r(t)) dW(t), \quad (4.1)$$

donde  $\{W(t)\}$  es un movimiento browniano estándar, y la deriva  $\mu(\cdot)$  y la volatilidad  $\sigma(\cdot)$  son funciones que dependen únicamente del valor actual del tipo de interés (algunos casos particulares aparecen recogidos en la Tabla 4.1).

Es importante indicar que en este contexto también se han planteado modelos no homogéneos en el tiempo que se apartan de la clase de los dados por la ecuación (4.1); algunos de ellos se recogen en la Tabla 4.2.

Una breve descripción de estos modelos homogéneos y no homogéneos y de su com-

Referencia principal	Ecuación del modelo
Merton(1973)[79]	$dr(t) = \alpha_0 dt + \beta_0 dW(t)$
Black y Scholes(1973)[11]	$dr(t) = \alpha_1 r(t) dt + \beta_1 r(t) dW(t)$
Cox(1975)[23]	$dr(t) = \alpha_1 r(t) dt + \beta_1 r(t)^\tau dW(t)$
Vasicek(1977)[97]	$dr(t) = (\alpha_0 + \alpha_1 r(t)) dt + \beta dW(t)$
Dothan(1978)[32]	$dr(t) = \beta_1 r(t) dW(t)$
Brennan y Schwartz(1980)[15]	$dr(t) = (\alpha_0 + \alpha_1 r(t)) dt + \beta_1 r(t) dW(t)$
Cox, Ingersoll y Ross(1985) [24, 25]	$dr(t) = (\alpha_0 + \alpha_1 r(t)) dt + \beta_1 \sqrt{r(t)} dW(t)$ (Modelo CIR)
Constantinides(1992)[21]	$dr(t) = \left( \alpha_0 + \alpha_1 r(t) + \alpha_2 \sqrt{r(t)} - \alpha_3 \right) dt +$ $+ \beta_0 (r(t) - \beta_1) dW(t)$
Chan <i>et al.</i> (1992)[19]	$dr(t) = (\alpha_0 + \alpha_1 r(t)) dt + \beta_2 r(t)^\tau dW(t)$ (Modelo CKLS)
Duffie y Kan(1996)[34]	$dr(t) = (\alpha_0 + \alpha_1 r(t)) dt + \sqrt{\beta_0 + \beta_1 r(t)} dW(t)$
Ait-Sahalia(1996)[4]	$dr(t) = (\alpha_0 + \alpha_1 r(t) + \alpha_2 r(t)^2 + \alpha_3 r(t)^{-1}) dt +$ $+ (\beta_0 + \beta_1 r(t) + \beta_2 r(t)^\tau) dW(t)$

**Tabla 4.1:** Algunos modelos homogéneos unifactoriales paramétricos.

portamiento empírico se puede encontrar en Moreno (2000)[82].

La principal ventaja de los anteriores modelos no homogéneos en el tiempo frente a los homogéneos recogidos en (4.1) es que, al permitir que su deriva y/o volatilidad dependan del tiempo, se pueden ajustar perfectamente a la estructura actual observada en los tipos de interés (algunos ejemplos aparecen en Björk (1997)[7]). Sin embargo, para mantener este ajuste perfecto, estos modelos deben ser replanteados y reestimados en cada instante de tiempo, lo que lleva a un enorme esfuerzo computacional. Esta “desventaja”, nos lleva

Referencia principal	Ecuación del modelo
Versión continua de Ho y Lee (1986)[57]	$dr(t) = \alpha(t) dt + \beta dW(t)$
Black, Derman y Toy (1990)[9]	$d \ln r(t) = \left[ \alpha(t) - \frac{\beta'(t)}{\beta(t)} \right] dt + \beta(t) dW(t)$
Black y Karasinski (1991)[10]	$d \ln r(t) = \alpha_2(t) [\ln \alpha_1(t) - \ln r(t)] dt +$ $+ \beta(t) dW(t)$
Hull y White (1990)[60]	$dr(t) = \alpha_2(t) [\alpha_1(t) - r(t)] dt$ $+ \beta(t) \sqrt{r(t)} dW(t)$

**Tabla 4.2:** Algunos modelos no homogéneos unifactoriales paramétricos.

a considerar en este trabajo, los modelos homogéneos dados por la ecuación (4.1).

Por otra parte, es también muy conocido que un problema potencialmente grave de cualquier modelo paramétrico, es que imponer una determinada estructura puede llevar a *errores de especificación* y en consecuencia se pueden tener errores graves al usar el modelo con propósitos de valoración y cobertura de activos de tipos de interés (ver Canabarro (1995)[18]). Por ello, en un intento por evitar tales posibles errores de especificación, se han propuesto métodos no paramétricos para estimar la ecuación (4.1). Citamos algunos de los más importantes:

- i) En Aït-Sahalia (1996)[3] se propone un procedimiento semiparamétrico para estimar la volatilidad, bajo una especificación paramétrica de la deriva.
- ii) Jiang y Knight (1997)[65], desarrollan un estimador no paramétrico mediante regresión núcleo para la función de difusión, y después deducen, a partir de éste, un estimador no paramétrico consistente de la deriva.
- iii) Stanton (1997)[94], construye fórmulas de aproximación de varios órdenes para la volatilidad y la deriva. Además, bajo este enfoque se estiman la volatilidad y la deriva separadamente, es decir, para estimar cualquiera de ellos no es necesario disponer de una estimación del otro.

En cuanto al estudio empírico realizado en este capítulo, después de la elección y un primer análisis de datos referentes a los tipos de interés a corto plazo en el marco de la Economía española, se desarrollan y aplican fundamentalmente las técnicas no paramétricas propuestas en Stanton (1997)[94] para estimar el modelo planteado por la ecuación(4.1). De esta forma, las técnicas desarrolladas permiten estimar la deriva y la volatilidad del modelo continuo planteado cuando se dispone de una muestra del proceso en una cantidad finita de instantes de tiempo. Sin embargo, nuestro estudio no finaliza

aquí, ya que una vez estimados los parámetros del modelo, se profundiza en la cuestión de si la no linealidad exhibida en la deriva es una hipótesis estadísticamente aceptable. Para ello, hacemos uso del test  $J_n$  desarrollado por Fan y Li (1997)[43], Li (1999)[73] y Kim y Wang (2006)[68]. Además, se usa también dicho test para contrastar algunas formas funcionales paramétricas propuestas para el cuadrado de la volatilidad. Finalmente, y siempre dentro del contexto no paramétrico, se estima el precio del riesgo de mercado y se usa el modelo estimado para valorar bonos, además de efectuar algunas comparaciones de esta valoración con la que se obtiene imponiendo ciertas condiciones paramétricas sobre el modelo.

## 4.2. Los datos: un primer análisis

La variable subyacente en los modelos que se van a desarrollar para modelizar el tipo de interés a corto plazo en España, es el tipo de interés instantáneo libre de riesgo. Este tipo de interés es una variable no observable en los mercados financieros, lo que hace necesario que se deba usar una variable *proxy* para llevar a cabo el estudio empírico.

Para seleccionar la variable *proxy*, se debe elegir un plazo para el tipo de interés suficientemente corto para que se acerque al concepto de tipo de interés instantáneo. Sin embargo, si se elige un plazo demasiado corto se pueden recoger movimientos no deseables debidos a la microestructura de los mercados. A este respecto, encontramos en la literatura financiera que lo usual es elegir tipos de interés interbancarios o tipos de interés de la deuda pública (para garantizar que son libres de riesgo) con un plazo entre una semana y tres meses. Por ejemplo, para la economía española, García (1998)[46] usa



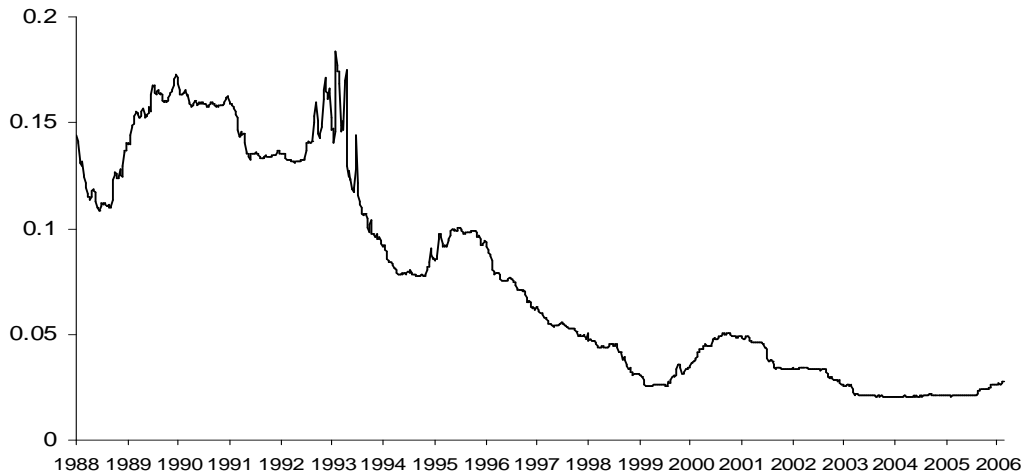
datos mensuales de los tipos de interés interbancarios a un mes para estudiar el modelo paramétrico propuesto en Chan *et al.* (1992)[19]. En esta misma línea, Nowman (1977)[83] usa los tipos de interés interbancarios a un mes en Gran Bretaña y Chan *et al.*(1992)[19], Longstaff y Schwartz (1992)[77] y Vetzal (1997)[98] usan los tipos de interés de las letras del tesoro americanas a un mes. Otros ejemplos son: Andersen y Lund (1997)[1] y Stanton (1997)[94], que utilizan esos mismos títulos pero a tres meses; Aït-Sahalia (1996)[4], que utiliza el tipo Eurodolar a siete días; Arapis y Gao (2006)[2], que usan también estos dos últimos tipos y Fernández (2001)[45], que usa los tipos interbancarios a siete días de la economía chilena.

Nuestra elección, atendiendo a los comentarios anteriores, es emplear la serie de datos semanales del tipo de interés interbancario a un mes anualizado en España, a la que denotamos  $r(t)$ . La serie, está formada por los tipos de interés desde enero de 1988 hasta mayo de 2006, lo que supone un total de 961 observaciones (elaboradas a partir de los datos disponibles en la página web del Banco de España, más concretamente en <http://www.bde.es/banota/series.htm>). La Figura 4.1 muestra la evolución de dicha serie temporal y, por otro lado, en la Figura 4.2 se representa su primera diferencia  $\Delta r(t) = r(t) - r(t - 1)$ , en la cual se observa una volatilidad de los tipos de interés cambiante con el tiempo.

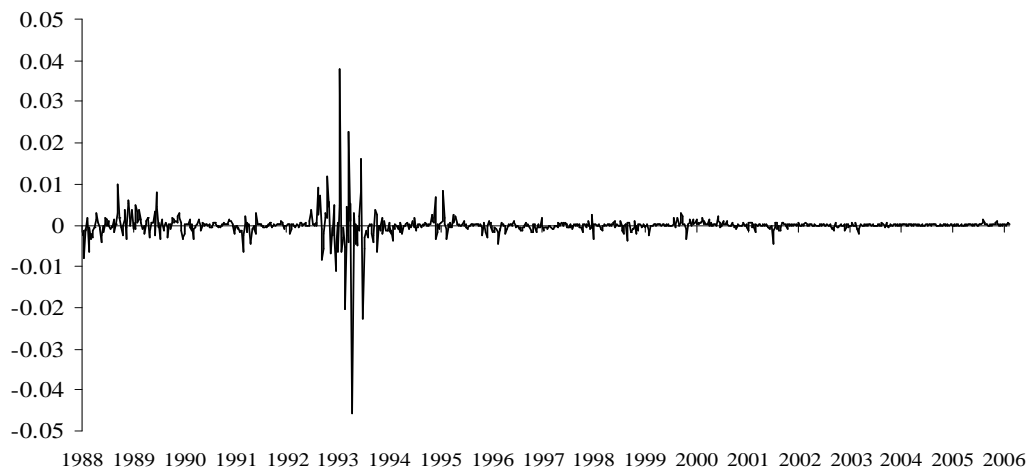
En la Tabla 4.3 se presentan algunas estadísticas básicas de la serie temporal  $r(t)$  de los tipos interbancarios en cuestión y de su primera diferencia. Es destacable como las autocorrelaciones de los tipos de interés decrecen muy lentamente, persistencia que es típica en este tipo de series temporales. Para contrastar la estacionariedad de la serie se lleva a cabo el test aumentado de Dickey–Fuller<sup>1</sup>, cuya hipótesis nula es la de no

---

<sup>1</sup>Ver Hamilton (1994)[54] para una descripción detallada del test.



**Figura 4.1:** *Serie de datos semanales de los tipos de interés interbancarios en España a un mes anualizados, desde enero de 1988 hasta mayo de 2006.*



**Figura 4.2:** *Serie de la primera diferencia de los datos semanales de los tipos de interés interbancarios en España a un mes anualizados, desde enero de 1988 hasta mayo de 2006.*

estacionariedad del proceso (existencia de raíz unitaria). El estadístico del contraste viene dado por

$$\frac{\hat{\gamma} - 1}{ES(\hat{\gamma})},$$

donde  $\hat{\gamma}$  es el estimador por mínimos cuadrados ordinarios y  $ES(\hat{\gamma})$  su error estándar en el modelo

$$r(t) = \alpha + \beta t + \gamma r(t-1) + \sum_{j=1}^p \delta_j \Delta r(t-j) + \varepsilon(t). \quad (4.2)$$

El resultado del test conduce, al nivel usual de significatividad del 5%, a rechazar la hipótesis nula de existencia de raíz unitaria y por tanto a aceptar la estacionariedad del proceso ( $p$ -valor 0,01651). Es importante tener en cuenta que es conocido que este test tiene poca potencia (probabilidad de rechazar la hipótesis nula cuando no es cierta). Así, con un test tan conservador incluso un ligero rechazo significa que la estacionariedad de la serie es muy posible.

El número de retardos incluidos en el modelo (4.2) es  $p = 24$ , y esto se debe a que se observa en el correlograma de la serie  $\Delta r(t)$  que las autocorrelaciones con retardo superior a 24 no son significativas al 5%, y esto mismo no ocurre con ningún retardo inferior a 24 (ver Figura 4.3)<sup>2</sup> (para mayores detalles sobre el contraste aumentado de Dickey–Fuller y la significatividad de las autocorrelaciones se puede consultar Hamilton (1994)[54] y Box, Jenkins y Reinsel (1994)[12]).

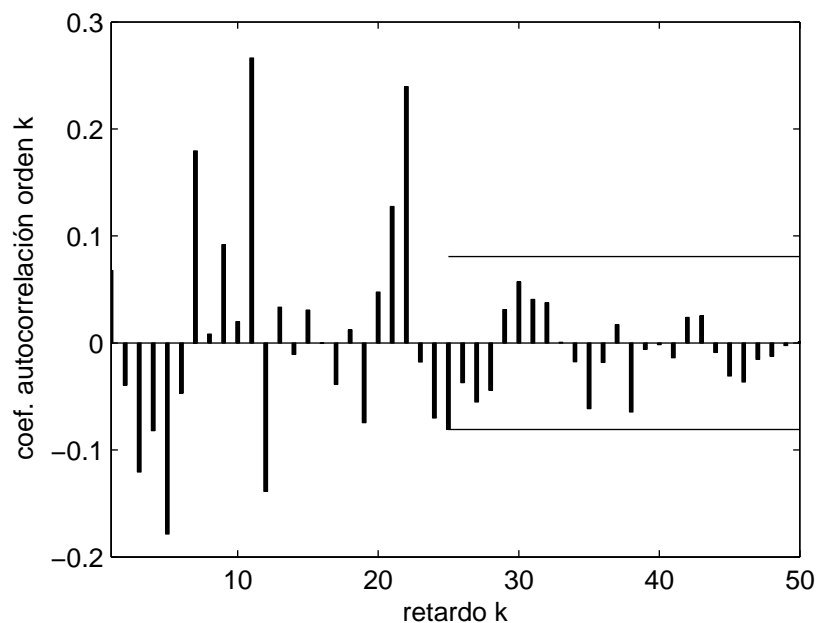
Según la Proposición 3.15, para valorar activos derivados de los tipos de interés por ausencia de arbitraje, es necesario estimar el precio del riesgo de mercado  $\lambda = \lambda_0 \cdot \sigma$ . Como dicho precio está relacionado con el exceso de retorno de los activos dependientes

---

<sup>2</sup>La banda de confianza está calculada a partir de la aproximación de Barlett de la varianza de las autocorrelaciones, resultando en este caso que las autocorrelaciones con retardo superior a 24 dentro del intervalo  $[-0'08077, 0'08077]$  no son significativas al 5%.

	$r(t)$	$\Delta r(t)$
media	0,07779	-0,00012
desv. típica	0,05013	0,00286
mínimo	0,02029	-0,04577
máximo	0,18394	0,03813
$\rho_1$	0,99695	0,06756
$\rho_2$	0,99373	-0,03933
$\rho_3$	0,99066	-0,12055
$\rho_4$	0,98819	-0,08187
$\rho_5$	0,98600	-0,17857
$\rho_6$	0,98443	-0,04688
$\rho_7$	0,98295	0,17946
$\rho_8$	0,98106	0,00828
$\rho_9$	0,97916	0,09170
$\rho_{10}$	0,97704	0,01989
$\rho_{11}$	0,97491	0,26617
$\rho_{12}$	0,97200	-0,13855
$\rho_{13}$	0,96956	0,03330
$\rho_{14}$	0,96702	-0,01061
$\rho_{15}$	0,96451	0,03068
$\rho_{16}$	0,96185	-0,00028
$\rho_{17}$	0,95910	-0,03874
$\rho_{18}$	0,95645	0,01227
$\rho_{19}$	0,95376	-0,07434
$\rho_{20}$	0,95135	0,04753
$\rho_{21}$	0,94889	0,12745
$\rho_{22}$	0,94609	0,23936
$\rho_{23}$	0,94251	-0,01765
$\rho_{24}$	0,93904	-0,07004
$\rho_{25}$	0,93577	-0,08063
$\rho_{26}$	0,93270	-0,03713
$\rho_{27}$	0,92971	-0,05515
$\rho_{28}$	0,92689	-0,04402
$\rho_{29}$	0,92416	0,03103
$\rho_{30}$	0,92133	0,05715
Test aumentado de Dickey-Fuller		
Estadístico	-3,81205	
$p$ -valor	0,01651	

**Tabla 4.3:** Estadísticas básicas y test de estacionariedad de la serie de datos semanales de los tipos de interés interbancarios a un mes en España.



**Figura 4.3:** Autocorrelaciones de los tipos de interés mensuales. Se muestra también una banda dentro de la cual los coeficientes de orden superior a 24 no son significativos.

del tipo de interés, en §4.5 se obtiene una aproximación considerando un segundo activo; en concreto, los tipos de interés interbancarios a una semana. Se consideran, como antes y procedentes de la misma fuente, datos semanales de tales tipos anualizados para el periodo enero de 1988 hasta mayo de 2006. La Tabla 4.4, análoga a la obtenida para los tipos a un mes, muestra un resumen de las estadísticas básicas de esta variable. Nótese como al nivel usual de significatividad del 5% para esta serie también se rechaza la existencia de raíz unitaria; por tanto, se puede considerar que el proceso es estacionario<sup>3</sup>.

Es importante destacar, que el coeficiente de correlación lineal entre las dos series consideradas (las de los tipos interbancarios anualizados a una semana y a un mes) asciende

<sup>3</sup> Siguiendo el mismo procedimiento que con la serie de los tipos a un mes, aquí también se obtiene que para la serie de las primeras diferencias las autocorrelaciones con retardo superior a 24 no son significativas al 5%.

	Tipos a una semana	1 <sup>a</sup> diferencia
media	0,07743	-0,00012
desv típica	0,04981	0,00306
mínimo	0,01932	-0,04340
máximo	0,18481	0,03559
$\rho_1$	0,99659	0,03186
$\rho_2$	0,99308	-0,11134
$\rho_3$	0,99002	-0,10001
$\rho_4$	0,98761	-0,08938
$\rho_5$	0,98555	-0,09703
$\rho_6$	0,98389	-0,01862
$\rho_7$	0,98227	0,12080
$\rho_8$	0,98038	0,04514
$\rho_9$	0,97842	0,02304
$\rho_{10}$	0,97635	0,03829
$\rho_{11}$	0,97414	0,23018
$\rho_{12}$	0,97118	-0,08407
$\rho_{13}$	0,96859	0,01573
$\rho_{14}$	0,96588	-0,03830
$\rho_{15}$	0,96330	0,00260
$\rho_{16}$	0,96071	0,00552
$\rho_{17}$	0,95809	-0,00782
$\rho_{18}$	0,95548	0,04086
$\rho_{19}$	0,95272	-0,05115
$\rho_{20}$	0,95016	-0,00971
$\rho_{21}$	0,94772	0,17485
$\rho_{22}$	0,94468	0,17011
$\rho_{23}$	0,94101	-0,03810
$\rho_{24}$	0,93747	-0,08665
$\rho_{25}$	0,93425	-0,05672
$\rho_{26}$	0,93123	-0,00941
$\rho_{27}$	0,92818	-0,05473
$\rho_{28}$	0,92533	-0,01672
$\rho_{29}$	0,92253	0,06393
$\rho_{30}$	0,91949	0,01283
Test aumentado de Dickey-Fuller		
Estadístico	-3,91769	
$p$ -valor	0,01203	

**Tabla 4.4:** Estadísticas básicas y test de estacionariedad de la serie de datos semanales de los tipos de interés interbancarios a una semana en España.

a 0,99952, lo que muestra evidencia en favor<sup>4</sup> de que las dos series están positivamente y perfectamente correladas como se supone que ocurre en el marco teórico de los modelos unifactoriales.

Continuamos ahora trabajando con la serie principal  $r(t)$  de datos semanales de los tipos de interés interbancarios a un mes, ya que es la que se va a usar para la modelización de los tipos de interés instantáneos libres de riesgo. En concreto, una primera idea de la densidad incondicional del proceso estacionario  $r(t)$  se obtiene trazando su histograma de frecuencias. Esta herramienta estadística básica, se puede ver como una estimación no paramétrica de la densidad estacionaria de la serie temporal. De forma genérica, dada una anchura  $h > 0$  y un origen  $r_0$ , los intervalos sobre los que se construye el histograma vienen dados por  $I_j = [r_0 + jh, r_0 + (j + 1)h]$ , para  $j \in \mathbb{Z}$ , y el propio histograma viene entonces dado por la función

$$\hat{f}(r) = \frac{\sum_{t=1}^n I_{[r_0+jh, r_0+(j+1)h]}(r(t))}{nh} = \frac{f_j}{h}, \quad r \in I_j, \quad (4.3)$$

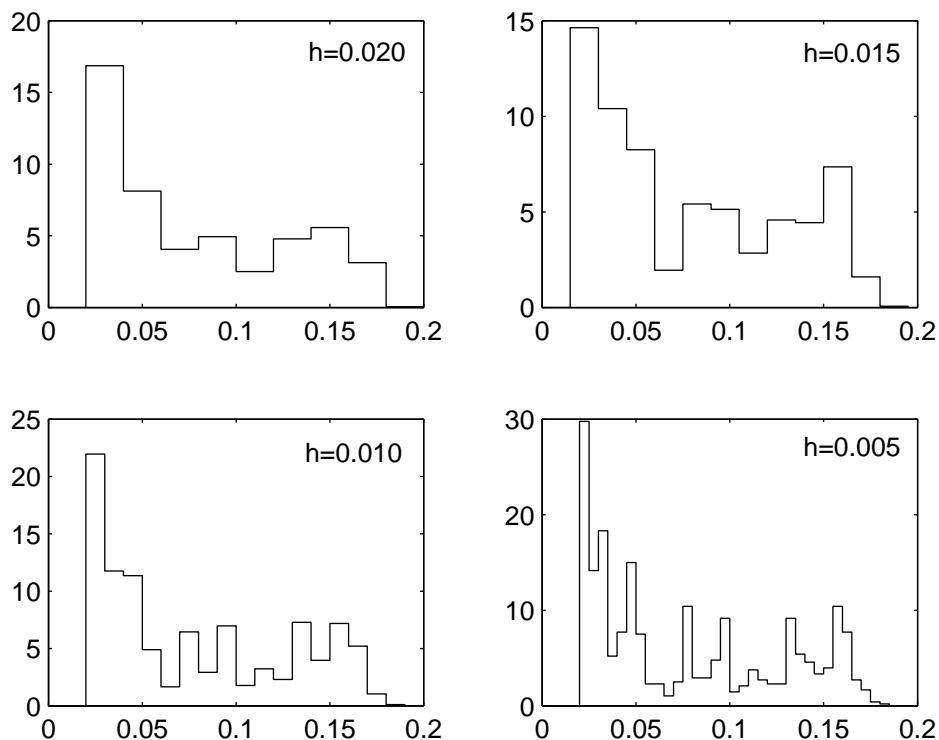
donde  $f_j$  es la frecuencia relativa del intervalo  $I_j$  e  $I_{[r_0+jh, r_0+(j+1)h]}(r)$  es la función indicador del intervalo  $[r_0 + jh, r_0 + (j + 1)h]$ . Esta expresión se deduce directamente aproximando la probabilidad del intervalo  $I_j$  por su frecuencia relativa y suponiendo que en este intervalo la densidad es constante (los datos se reparten uniformemente en cada intervalo).

En la Figura 4.4 aparecen algunos histogramas de  $r(t)$  con  $r_0 = 0$  y diferentes elecciones de anchura de ventana. Se observa cómo la elección del parámetro  $h$  permite controlar la suavidad que se obtiene en la densidad estimada resultante<sup>5</sup>.

---

<sup>4</sup>Más, teniendo en cuenta que la correlación se mantiene en este orden, por encima de 0,90, para los otros plazos de los tipos de interés interbancarios en España.

<sup>5</sup>A lo largo de todo este trabajo se usa una anchura  $h$  constante. Queremos mencionar a este respecto, que una forma de construir histogramas con  $h$  variable y su posterior suavización se propone en Gavilán



**Figura 4.4:** Histogramas con origen  $r_0 = 0$  y diferentes anchuras de ventana para los datos semanales de los tipos de interés interbancarios a un mes en España.

Un importante inconveniente de esta estimación es que todas las observaciones que caen fuera de  $I_j$  no tienen ninguna influencia en  $\hat{f}(r)$  con  $r \in I_j$ , lo cual es particularmente negativo en los extremos de cada intervalo  $I_j$ . Una manera de superar este inconveniente es la construcción de un histograma adaptado, en el cual, para estimar la densidad en un punto  $r$  se tienen en cuenta todas las observaciones en un intervalo centrado en el punto  $r$  y de amplitud  $h$ . Así, su expresión analítica viene dada por la siguiente función:

$$\hat{f}(r) = \frac{\sum_{t=1}^n I_{[r-\frac{h}{2}, r+\frac{h}{2}]}(r(t))}{nh}, \quad (4.4)$$

y González (1999)[48], la cual se aplica a unos datos concretos de la Economía española en Gavilán y González (2000)[49].



la cual se puede poner también bajo la siguiente forma:

$$\hat{f}(r) = \frac{1}{nh} \sum_{t=1}^n K_0 \left( \frac{r - r(t)}{h} \right), \quad r \in \mathbb{R}, \quad (4.5)$$

siendo  $K_0(u) = I_{[-\frac{1}{2}, \frac{1}{2}]}(u)$  el denominado núcleo intuitivo<sup>6</sup>. Además, a diferencia del histograma dado en la ecuación (4.3), el histograma adaptado no depende de la elección de un origen, ya que se toman intervalos centrados en cada observación. En la Figura 4.5 aparecen histogramas adaptados de la serie  $r(t)$  para diferentes elecciones de anchura de ventana  $h$ , donde puede observarse que para los mismos valores de  $h$  anteriores de la Figura 4.4, éstos presentan ahora un contorno mucho más suave.

Para obtener estimaciones más suaves, se suelen usar otros núcleos.<sup>7</sup> Dado un núcleo  $K(\cdot)$ , el estimador de la densidad viene dado por

$$\hat{f}(r) = \frac{1}{nh} \sum_{t=1}^n K \left( \frac{r - r(t)}{h} \right), \quad r \in \mathbb{R}, \quad (4.6)$$

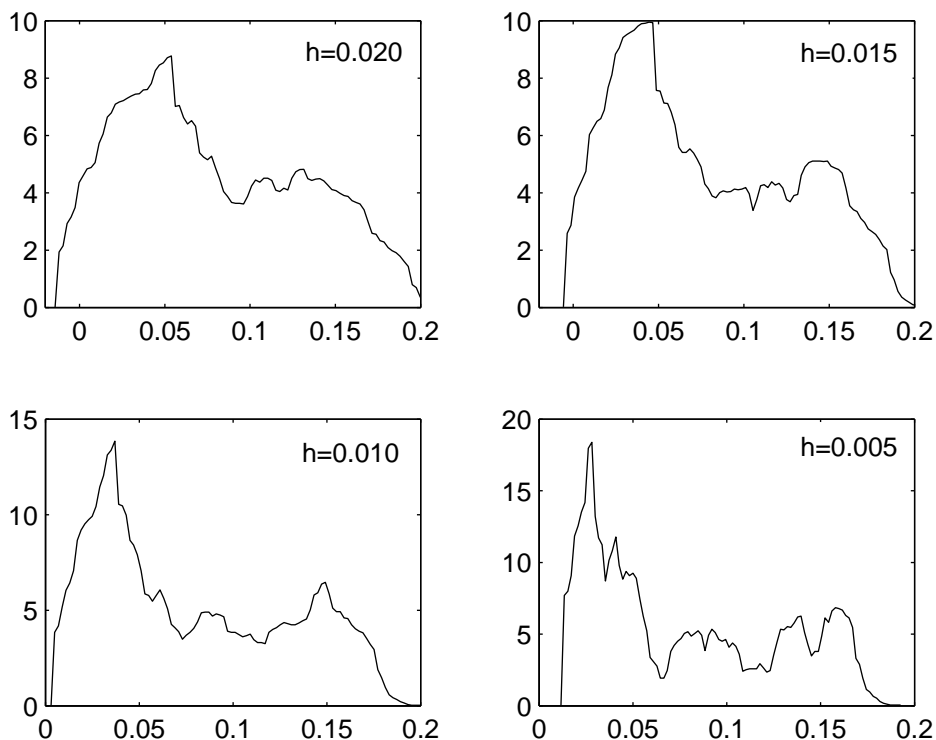
donde el parámetro  $h$  es la anchura de la ventana. En esta situación,  $\hat{f}(r)$  es una mixtura de  $n$  funciones de densidad y por tanto es también una función de densidad.

Luego, para realizar con este método una estimación de la densidad estacionaria de la serie  $r(t)$ , se debe elegir un núcleo  $K(\cdot)$  y una anchura de ventana  $h > 0$ . Algunos resultados teóricos (Silverman (1986)[93]), prueban que la elección del núcleo no es importante desde el punto de vista del error asintótico de estimación y se suele elegir atendiendo a otras consideraciones como diferenciabilidad o esfuerzo computacional asociado a su

---

<sup>6</sup>Una referencia fundamental en lo referente a estimación no paramétrica de densidades es Silverman (1986)[93].

<sup>7</sup>Un núcleo o kernel es una función de densidad simétrica acotada tal que  $|u|K(u) \rightarrow 0$  cuando  $|u| \rightarrow 0$  y con  $\int u^2 K(u) du < \infty$ .



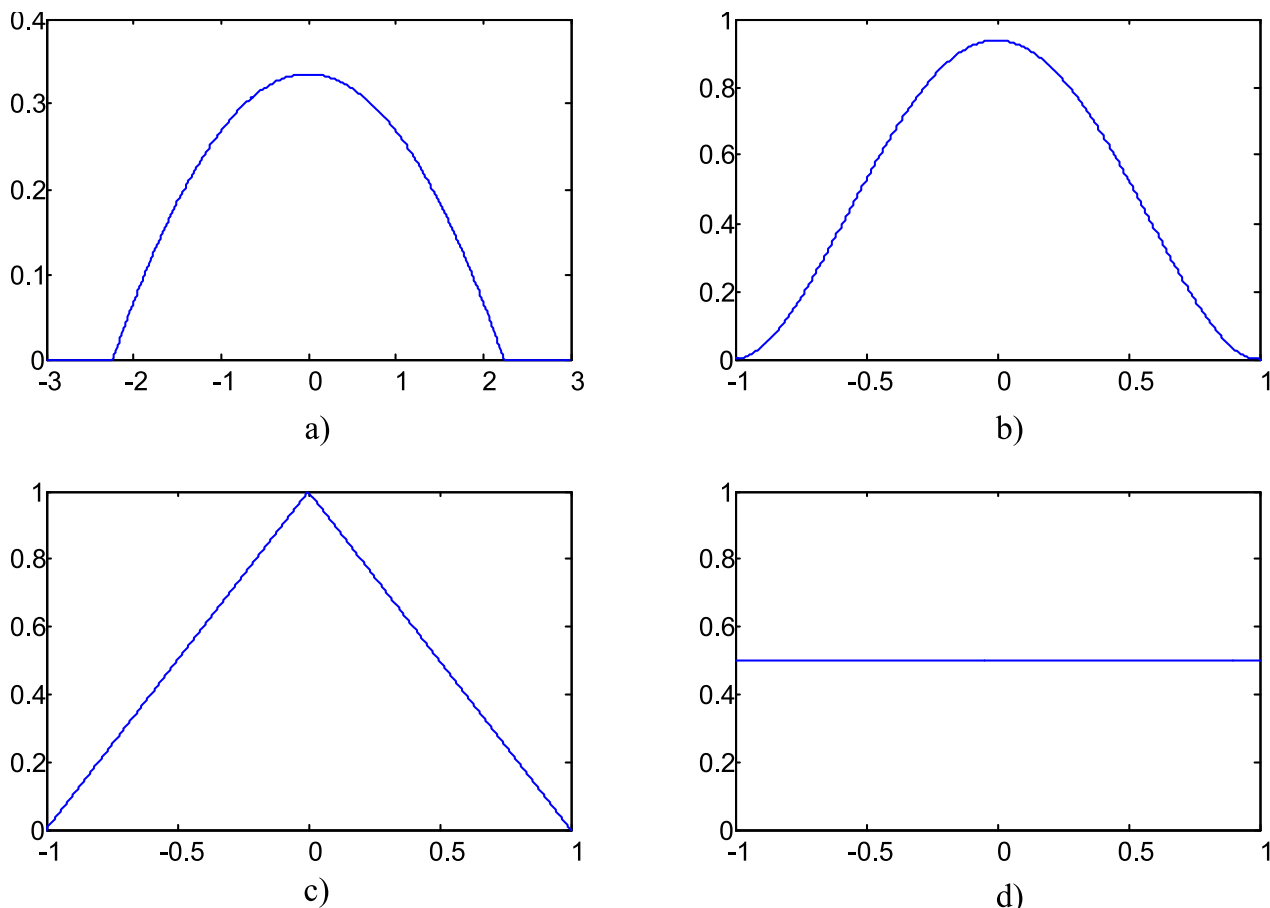
**Figura 4.5:** Histogramas adaptados con diferentes anchuras de ventana para la serie de datos semanales de los tipos de interés interbancarios a un mes en España.

cálculo. Un núcleo muy usado en la práctica, y que utilizamos siempre en este trabajo, es el gaussiano, cuyo núcleo  $K(\cdot)$  es la densidad de la distribución normal estándar

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}, \quad u \in \mathbb{R}. \quad (4.7)$$

En este caso,  $\hat{f}(r)$  es la media aritmética de  $n$  densidades normales centradas en las observaciones  $r(t)$  y de varianza  $h$ . Teniendo en cuenta la forma acampanada de la densidad gaussiana se tiene que cuanto más alejado está un punto muestral del punto de estimación menos contribuirá a la densidad estimada, así la densidad estimada será mayor cerca de las grandes concentraciones de datos muestrales y menor donde las observaciones sean escasas, efecto que no es exclusivo de la densidad gaussiana, sino que es común a la mayoría

de los núcleos habitualmente usados (ver Figura 4.6). Por otra parte, al ser en este caso el núcleo infinitamente diferenciable la densidad estimada también lo es.



**Figura 4.6:** Algunos núcleos usados en la estimación de densidades de frecuencia: a) Epanechnikov, dado por  $K(r) = \frac{3}{4\sqrt{5}} \left(1 - \frac{1}{5}r^2\right) \cdot I_{(-\sqrt{5}, \sqrt{5})}(r)$ , b) Biweight, dado por  $K(r) = \frac{15}{16} (1 - r^2)^2 \cdot I_{(-1,1)}(r)$ , c) Triangular, dado por  $K(r) = (1 - |r|) \cdot I_{(-1,1)}(r)$  y d) Rectangular, dado por  $K(r) = \frac{1}{2} \cdot I_{(-1,1)}(r)$ .

Respecto a la anchura de la ventana  $h$ , su elección es de gran importancia para la precisión del estimador, ya que  $\hat{f}(r)$  es muy sensible a la elección de  $h$ , como ejemplo de ello basta observar la forma de los histogramas adaptados de la Figura 4.5. Es conocido (Silverman (1986)[93]), que en el caso en que la verdadera densidad es gaussiana y se usa

un núcleo gaussiano, la elección óptima, desde el punto de vista de obtener un menor error cuadrático integrado medio, viene dada por

$$h = 1,06 \cdot S \cdot n^{-\frac{1}{5}}, \quad (4.8)$$

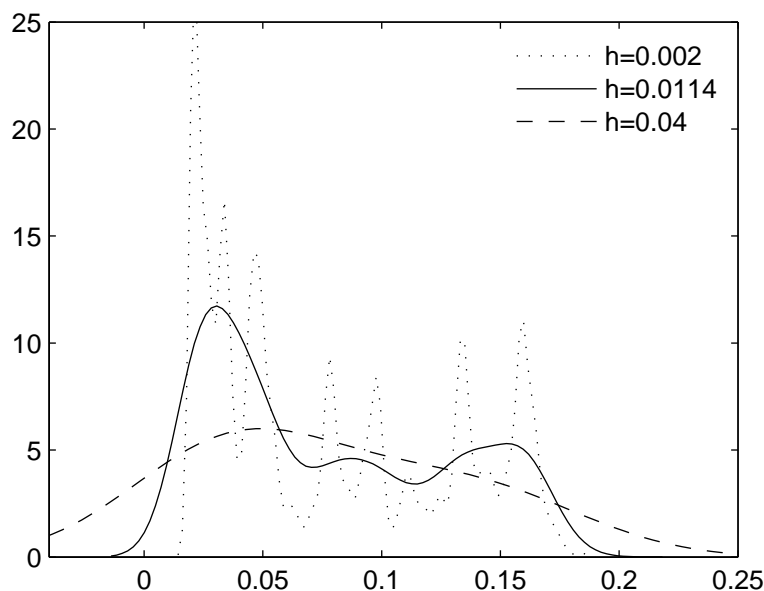
siendo  $S$  una estimación de la desviación estándar de la distribución y  $n$  el número de observaciones. Sin embargo, tal elección lleva a una densidad estimada demasiado suave cuando las verdaderas densidades son multimodales y asimétricas, el que parece ser el caso de la distribución que pretendemos estimar a la vista de los histogramas anteriores de las Figuras 4.4 y 4.5, donde se observan varias modas y una clara asimetría. Según Silverman (1986)[93], la siguiente elección produce mejores resultados para una gran variedad distribuciones:

$$h = 0,9 \cdot \min \left\{ S, \frac{R_Q}{1,34} \right\} n^{-\frac{1}{5}}, \quad (4.9)$$

siendo  $R_Q$  el recorrido intercuartílico de los datos. Para la serie de datos que nos ocupa, esta elección de la anchura de la ventana lleva a tomar  $h = 0,0114$ .

Por otro lado, para ilustrar la gran influencia de  $h$  sobre la densidad estimada se presenta en la Figura 4.7 estimaciones correspondientes a diferentes valores de  $h$ . En concreto, aparte de la mencionada elección  $h = 0,0114$ , se ha tomado  $h = 0,002$ , un valor sensiblemente menor, y  $h = 0,04$ , un valor sensiblemente mayor. Se observa como, en efecto, la densidad estimada es muy sensible al valor de  $h$ , pasando de una estimación con  $h = 0,002$  poco suave y mostrando excesivos detalles, a la obtenida con  $h = 0,04$  que es excesivamente suave y oculta detalles importantes. Parece que la situación intermedia con  $h = 0,0114$  es la más adecuada, ya que se puede considerar como un equilibrio entre suavidad y detalle. Existen métodos más refinados para encontrar una anchura de ventana deseable, sin embargo, hemos observado que en nuestro caso de estudio no

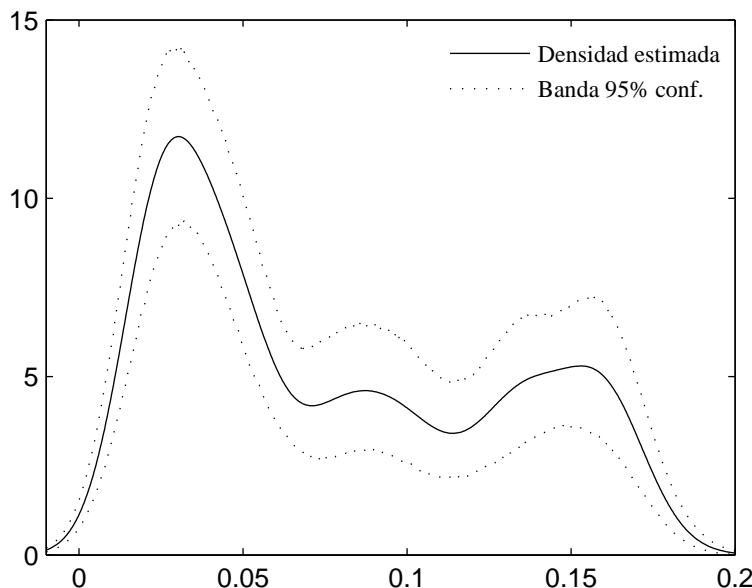
producen mejores resultados respecto a la elección dada por (4.9).



**Figura 4.7:** *Densidad incondicional estimada de los tipos de interés a corto plazo en España, obtenida a partir de los datos semanales de los tipos de interés interbancarios a un mes usando un núcleo gaussiano y diferentes anchuras de ventana.*

Centrándonos en la elección de un núcleo gaussiano y una anchura de ventana  $h = 0,0114$ , se presenta la Figura 4.8. En ésta, además de la representación de la densidad estimada se añaden bandas puntuales de confianza al 95 %, obtenidas utilizando el método bootstrap de bloques móviles desarrollado en Künsch (1989)[70] y Liu y Singh (1992)[76]. Este método, al captar la dependencia débil entre las observaciones, logra mejores resultados que los conseguidos con la clásica normalidad asintótica del estimador.

Comentemos brevemente el funcionamiento de este método de remuestreo: En primer lugar, con este método se obtiene la distribución en el muestreo del estimador  $\hat{f}(r)$  de  $f(r)$  remuestreando de los bloques móviles  $B_1, B_2, \dots, B_{n-b+1}$ , donde  $b$  es el tamaño de cada bloque y  $B_j = \{r(j), r(j+1), \dots, r(j+b-1)\}$  con  $j = 1, 2, \dots, n-b+1$ . Para

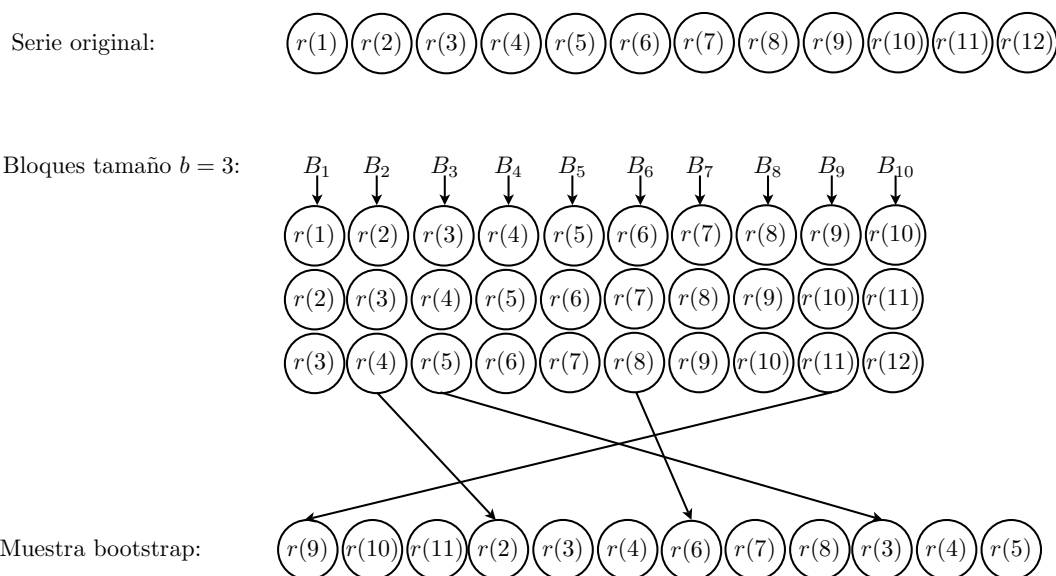


**Figura 4.8:** Densidad incondicional estimada de los tipos de interés a corto plazo en España y bandas puntuales de confianza al 95 % basadas en el método bootstrap de bloques móviles con 5000 replicaciones.

ello, se obtiene una muestra de  $k$  bloques<sup>8</sup> de manera independiente e idénticamente distribuida que se denota por  $\{\xi_1, \xi_2, \dots, \xi_k\}$  (muestra bootstrap) con la cual se calcula el estimador bootstrap  $\hat{f}^*(r)$ . Repitiendo el procedimiento para una gran cantidad de muestras bootstrap, se obtiene una aproximación a la distribución muestral de  $\hat{f}(r)$  y, a partir de ella, un intervalo de confianza para  $f(r)$ . La construcción de una muestra bootstrap a partir de unos datos originales siguiendo este procedimiento se ilustra en la Figura 4.9.

Como conclusión más destacable, se tiene que la densidad estimada multimodal obtenida para los tipos de interés a corto plazo en la Figura 4.8 no invita a pensar que la

<sup>8</sup>Aquí  $k$  se debe tomar para que  $kb$  sea del mismo orden que  $n$ . Por ejemplo, se puede elegir  $k$  como el entero más cercano a  $\frac{n}{b}$ .



**Figura 4.9:** Ilustración del método bootstrap de bloques móviles, con una serie de tamaño  $n = 12$  y bloques de tamaño  $b = 3$ .

dinámica de tales tipos de interés obedezca a modelos simples, como los modelos gaussianos.

### 4.3. Una modelización no paramétrica

Uno de los enfoques más generales para estimar la ecuación diferencial homogénea (4.1) propuesta para la dinámica de los tipos de interés a corto plazo, es el desarrollado en Stanton (1997)[94]. En esta sección, se hace una breve presentación de las aproximaciones desarrolladas dentro del anterior enfoque, se realizan las estimaciones de los parámetros relativos a los datos de la economía española, y además se obtienen algunas conclusiones a partir de los resultados.

La idea fundamental en el trabajo de Stanton (1997)[94] se basa en desarrollar la esperanza condicionada  $\mathbb{E}_t [f(r(t + \Delta), t + \Delta)]$  de forma similar a un desarrollo en serie de Taylor, donde  $f(\cdot, \cdot)$  es una función arbitraria y  $\Delta$  es el paso de tiempo entre observaciones sucesivas. Bajo las condiciones apropiadas (Hille y Phillips (1957)[56]) sobre  $\mu(\cdot)$ ,  $\sigma(\cdot)$  y  $f(\cdot, \cdot)$ , se tiene que

$$\begin{aligned} \mathbb{E}_t [f(r(t + \Delta), t + \Delta)] &= f(r(t), t) + \mathcal{L}f(r(t), t) \cdot \Delta + \frac{1}{2}\mathcal{L}^2 f(r(t), t) \cdot \Delta^2 + \dots + \\ &+ \frac{1}{n!}\mathcal{L}^n f(r(t), t) \cdot \Delta^n + O(\Delta^{n+1}), \end{aligned} \quad (4.10)$$

cuando el paso de tiempo  $\Delta \downarrow 0$  y donde  $\mathcal{L}$  es el generador infinitesimal<sup>9</sup> del proceso  $\{r(t)\}$ , ya citado en §2.4, dado por:

$$\begin{aligned} \mathcal{L}f(r(t), t) &= \lim_{\tau \downarrow t} \frac{\mathbb{E}[f(r(\tau), \tau)/r(t)=r] - f(r, t)}{\tau - t} \\ &= \frac{\partial f(r, t)}{\partial t} + \frac{\partial f(r, t)}{\partial r} \mu(r) + \frac{1}{2} \frac{\partial^2 f(r, t)}{\partial r^2} \sigma^2(r). \end{aligned} \quad (4.11)$$

Nótese que cuando  $\sigma(\cdot) \equiv 0$ , este desarrollo coincide con el desarrollo en serie de Taylor estándar no estocástico. En este marco, si se eligen convenientemente funciones  $f(\cdot, \cdot)$  y se construyen aproximaciones, en base a una cantidad suficiente de datos, de la esperanza condicionada dada en la ecuación (4.10) se obtienen aproximaciones de  $\mu(\cdot)$  y  $\sigma(\cdot)$ .

En concreto, si se considera  $n = 1$  de la expresión (4.10), se deduce que

$$\mathcal{L}f(r(t), t) = \frac{1}{\Delta} \mathbb{E}_t [f(r(t + \Delta), t + \Delta) - f(r(t), t)] + O(\Delta), \quad (4.12)$$

expresión que se usa para obtener aproximaciones de primer orden para las funciones  $\mu(\cdot)$  y  $\sigma(\cdot)$ .

---

<sup>9</sup>Ver §7.3 en Oksendal (1998)[84] para mayores detalles.



Si se quiere obtener aproximaciones de segundo orden, entonces en primer lugar se considera  $n = 2$  en la expresión (4.10) y se toman pasos de tiempo entre observaciones de  $\Delta$  y  $2\Delta$ , obteniendo respectivamente las siguientes igualdades:

$$\mathcal{L}f(r(t), t) = \frac{1}{\Delta} \mathbb{E}_t [f(r(t + \Delta), t + \Delta) - f(r(t), t)] - \frac{1}{2} \mathcal{L}^2 f(r(t), t) \Delta + O(\Delta^2), \quad (4.13)$$

$$\mathcal{L}f(r(t), t) = \frac{1}{2\Delta} \mathbb{E}_t [f(r(t + 2\Delta), t + 2\Delta) - f(r(t), t)] - \frac{1}{2} \mathcal{L}^2 f(r(t), t) 2\Delta + O(\Delta^2). \quad (4.14)$$

Con objeto de eliminar el término  $\mathcal{L}^2 f(r(t), t)$ , que involucra a las funciones desconocidas  $\mu(\cdot)$  y  $\sigma(\cdot)$  y a sus derivadas, se encuentra la combinación lineal de las dos ecuaciones anteriores que lo anula, obteniendo de esta forma la expresión que lleva a aproximaciones de segundo orden de  $\mu(\cdot)$  y  $\sigma(\cdot)$ . En concreto, una vez realizado este proceso, la expresión que se obtiene es la siguiente:

$$\begin{aligned} \mathcal{L}f(r(t), t) &= \frac{1}{2\Delta} \{4\mathbb{E}_t [f(r(t + \Delta), t + \Delta) - f(r(t), t)] \\ &\quad - \mathbb{E}_t [f(r(t + 2\Delta), t + 2\Delta) - f(r(t), t)]\} + O(\Delta^2). \end{aligned} \quad (4.15)$$

De manera análoga, para obtener aproximaciones de tercer orden se considera  $n = 3$  en la expresión (4.10) y se escribe con pasos de tiempo  $\Delta$ ,  $2\Delta$  y  $3\Delta$ , obteniendo:

$$\begin{aligned} \mathcal{L}f(r(t), t) &= \frac{1}{\Delta} \mathbb{E}_t [f(r(t + \Delta), t + \Delta) - f(r(t), t)] - \frac{1}{2} \mathcal{L}^2 f(r(t), t) \Delta - \frac{1}{6} \mathcal{L}^3 f(r(t), t) \Delta^2 + O(\Delta^3). \\ \mathcal{L}f(r(t), t) &= \frac{1}{2\Delta} \mathbb{E}_t [f(r(t + 2\Delta), t + 2\Delta) - f(r(t), t)] - \frac{1}{2} \mathcal{L}^2 f(r(t), t) (2\Delta) \\ &\quad - \frac{1}{6} \mathcal{L}^3 f(r(t), t) (2\Delta)^2 + O(\Delta^3). \\ \mathcal{L}f(r(t), t) &= \frac{1}{3\Delta} \mathbb{E}_t [f(r(t + 3\Delta), t + 3\Delta) - f(r(t), t)] - \frac{1}{2} \mathcal{L}^2 f(r(t), t) (3\Delta) \\ &\quad - \frac{1}{6} \mathcal{L}^3 f(r(t), t) (3\Delta)^2 + O(\Delta^3) \end{aligned}$$

Finalmente, se busca la combinación lineal de las tres ecuaciones que elimina las cantidades  $\mathcal{L}^2 f(r(t), t)$  y  $\mathcal{L}^3 f(r(t), t)$  y se obtiene:

$$\begin{aligned} \mathcal{L}f(r(t), t) &= \frac{1}{6\Delta} \{18\mathbb{E}_t[f(r(t+\Delta), t+\Delta) - f(r(t), t)] - 9\mathbb{E}_t[f(r(t+2\Delta), t+2\Delta) \\ &\quad - f(r(t), t)] + 2\mathbb{E}_t[f(r(t+3\Delta), t+3\Delta) - f(r(t), t)]\} + O(\Delta^3). \end{aligned} \quad (4.16)$$

En consecuencia, para obtener aproximaciones de primer, segundo y tercer orden a la deriva propuesta para los movimientos del tipo de interés a corto plazo libre de riesgo, basta elegir  $f(r, t) = r$ , y se sigue de (4.11) que  $\mathcal{L}f(r(t), t) = \mu(r(t))$ . Las aproximaciones (4.12), (4.15) y (4.16) particularizadas para esta función son las siguientes:

$$\mu(r(t)) = \frac{1}{\Delta} \mathbb{E}_t[r(t+\Delta) - r(t)] + O(\Delta), \quad (4.17)$$

$$\mu(r(t)) = \frac{1}{2\Delta} \{4\mathbb{E}_t[r(t+\Delta) - r(t)] - \mathbb{E}_t[r(t+2\Delta) - r(t)]\} + O(\Delta^2), \quad (4.18)$$

$$\begin{aligned} \mu(r(t)) &= \frac{1}{6\Delta} \{18\mathbb{E}_t[r(t+\Delta) - r(t)] - 9\mathbb{E}_t[r(t+2\Delta) - r(t)] \\ &\quad + 2\mathbb{E}_t[r(t+3\Delta) - r(t)]\} + O(\Delta^3). \end{aligned} \quad (4.19)$$

Para obtener aproximaciones a  $\sigma^2(r(t))$  basta tomar  $f(r, t) = (r - r(t))^2$ , y de (4.11) se sigue que  $\mathcal{L}f(r(t), t) = \sigma^2(r(t))$ . Ahora, las aproximaciones (4.12), (4.15) y (4.16) quedan de la siguiente forma:

$$\sigma^2(r(t)) = \frac{1}{\Delta} \mathbb{E}_t[(r(t+\Delta) - r(t))^2] + O(\Delta), \quad (4.20)$$

$$\sigma^2(r(t)) = \frac{1}{2\Delta} \left\{ 4\mathbb{E}_t[(r(t+\Delta) - r(t))^2] - \mathbb{E}_t[(r(t+2\Delta) - r(t))^2] \right\} + O(\Delta^2), \quad (4.21)$$

$$\begin{aligned} \sigma^2(r(t)) &= \frac{1}{6\Delta} \left\{ 18\mathbb{E}_t[(r(t+\Delta) - r(t))^2] - 9\mathbb{E}_t[(r(t+2\Delta) - r(t))^2] \right. \\ &\quad \left. + 2\mathbb{E}_t[(r(t+3\Delta) - r(t))^2] \right\} + O(\Delta^3). \end{aligned} \quad (4.22)$$

Tomando raíces cuadradas a ambos lados de cada una de las tres ecuaciones anteriores

se obtienen las siguientes aproximaciones<sup>10</sup> a la volatilidad  $\sigma(r(t))$ :

$$\sigma(r(t)) = \sqrt{\frac{1}{\Delta} \mathbb{E}_t [(r(t+\Delta) - r(t))^2]} + O(\Delta), \quad (4.23)$$

$$\sigma(r(t)) = \sqrt{\frac{1}{2\Delta} \left\{ 4\mathbb{E}_t [(r(t+\Delta) - r(t))^2] - \mathbb{E}_t [(r(t+2\Delta) - r(t))^2] \right\}} + O(\Delta^2), \quad (4.24)$$

$$\sigma(r(t)) = \sqrt{\frac{1}{6\Delta} \left\{ 18\mathbb{E}_t [(r(t+\Delta) - r(t))^2] - 9\mathbb{E}_t [(r(t+2\Delta) - r(t))^2] + 2\mathbb{E}_t [(r(t+3\Delta) - r(t))^2] \right\}} + O(\Delta^3). \quad (4.25)$$

De manera análoga se puede continuar para formar aproximaciones de cualquier orden para la deriva y la volatilidad (o su cuadrado). A este respecto, Fan y Zhan (2003)[41] encuentran una expresión general para las aproximaciones anteriores a los parámetros de la difusión y analizan sus propiedades. En concreto, para construir una aproximación de cualquier orden basta usar un método apropiado para estimar la esperanza condicionada

$$\mathbb{E}[Y|X = x] = m(x). \quad (4.26)$$

Así, para obtener aproximaciones de orden  $k$  a la deriva,  $m(x) = \mu(x)$ , se debe tomar:

$$X(t) = r(t), \quad Y(t) = \Delta^{-1} \sum_{j=1}^k a_{kj} \{r(t+j\Delta) - r(t)\}, \quad t = 1, \dots, n-k. \quad (4.27)$$

Mientras que para obtener aproximaciones de orden  $k$  al cuadrado de la volatilidad,  $m(x) = \sigma^2(x)$ , se debe tomar:

$$X(t) = r(t), \quad Y(t) = \Delta^{-1} \sum_{j=1}^k a_{kj} \{r(t+j\Delta) - r(t)\}^2, \quad t = 1, \dots, n-k. \quad (4.28)$$

---

<sup>10</sup>Nótese que si se dispone de una aproximación al cuadrado de la volatilidad de la forma  $\sigma^2(r(t)) = A + O(\Delta^k)$ . Entonces, la aproximación binomial,  $\sigma(r(t)) = \sqrt{A} [1 + O(\Delta^k)]^{1/2} = \sqrt{A} [1 + \frac{1}{2}O(\Delta^k) - \frac{1}{8}O(\Delta^{2k}) + \dots] = \sqrt{A} + O(\Delta^k)$ , prueba que tomando raíz cuadrada se obtiene una aproximación del mismo orden para la volatilidad.

En ambos casos, los coeficientes  $a_{kj}$  responden a la siguiente formula general:

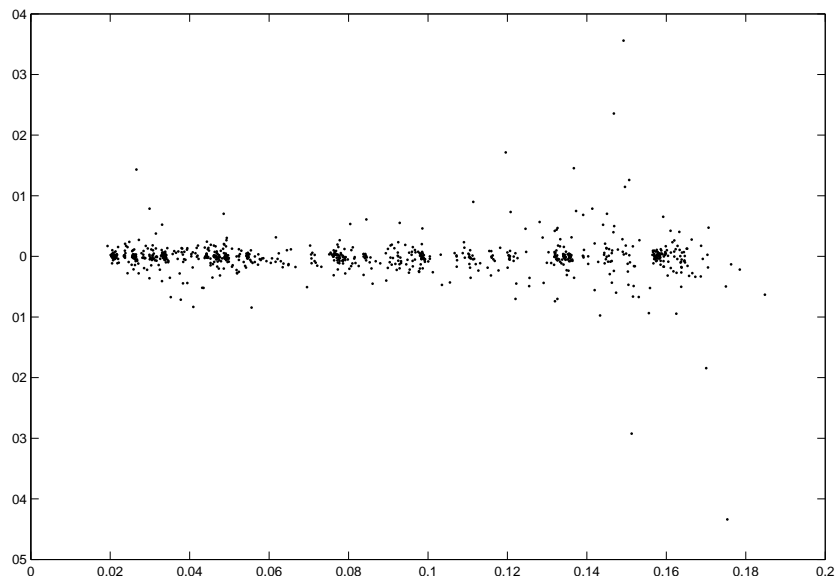
$$a_{kj} = \frac{(-1)^{j+1}}{j} \binom{k}{j}, \quad j = 1, \dots, k.$$

Por ejemplo, para obtener la aproximación de orden 1 a la deriva de los tipos de interés a corto plazo en España a partir de las observaciones semanales de los tipos de interés interbancarios a un mes  $\{r(t)\}$ , se debe estimar la esperanza condicionada  $\mathbb{E}[r(t + \Delta) - r(t) | r(t) = r]$  con el valor usual  $\Delta = \frac{1}{50}$ . Esto corresponde a determinar la línea de regresión de la nube de puntos representada en la Figura 4.10, en la cual se muestran los cambios semanales del tipo de interés  $r(t + 1) - r(t)$  en función del tipo de interés de la primera semana  $r(t)$  para  $t = 1, \dots, n - 1$ . Se puede observar, a partir de la Figura 4.10, cierta evidencia de heteroscedasticidad, ya que las variaciones son mayores cuando los tipos de interés son más altos.

En el contexto general de la regresión no lineal de una variable  $Y$  sobre una variable  $X$ , la estimación análoga al histograma es el regresograma (ver Härdle (1990)[58]), el cual viene dado por

$$\hat{m}(x) = \frac{\sum_{i=1}^n y_i I_{I_j}(x_i)}{\sum_{i=1}^n I_{I_j}(x_i)} = \bar{Y} / \{X \in I_{I_j}\}, \quad (4.29)$$

donde los intervalos  $I_j = [r_0 + jh, r_0 + (j + 1)h]$ ,  $j \in \mathbb{Z}$  son los utilizados en §4.2 para la construcción del histograma. Esta estimación no paramétrica de la línea de regresión adolece de los mismos inconvenientes ya comentados para el histograma y se le da la misma solución. Así, un estimador núcleo propuesto para  $m(x)$ , llamado estimador de



**Figura 4.10:** Nube de puntos de la primera diferencia de los datos semanales de los tipos de interés interbancarios a un mes en España frente a su nivel inicial.

Nadaraya-Watson, viene dado por:

$$\hat{m}(x) = \frac{\sum_{i=1}^n y_i K\left(\frac{x_i - x}{h}\right)}{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right)} = \sum_{i=1}^n y_i w_i(x). \quad (4.30)$$

donde  $w_i(x) = \frac{K\left(\frac{x_i - x}{h}\right)}{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right)}$ . De esta manera, el estimador de  $m(x)$  es una media aritmética ponderada de las respuestas observadas  $y_i$ , donde cada ponderación  $w_i(x)$  depende de todo el vector de observaciones  $X$  y es mayor cuanto más cerca está  $x_i$  de  $x$  (en los núcleos unimodales).

Nuevamente, para poder llevar a cabo esta estimación por núcleo es necesario especificar un núcleo  $K(\cdot)$  y una anchura de la ventana  $h$ , la cual, como se mostró en la estimación de la densidad estacionaria en §4.2, controla el grado de suavidad de la función estimada

$\hat{m}(x)$ . Al igual que en el caso de la estimación de la densidad, la elección del núcleo no es una cuestión crítica para la precisión del estimador (ver Härdle (1990)[58]) y es por ello por lo que continuamos usando el núcleo gaussiano.

Por otro lado, la elección de una anchura de ventana adecuada si es de gran importancia para la precisión del estimador. Una elección común es buscarla minimizando el siguiente error cuadrático medio<sup>11</sup>:

$$\text{ECM}(h) = \frac{\sum_{i=1}^n [m(x_i) - \hat{m}(x_i)]^2}{n}. \quad (4.31)$$

Se observa en esta expresión un grave problema ya que involucra a la función desconocida  $m(\cdot)$  que se quiere estimar. Debido a ello se busca una aproximación a (4.31). Así, un estimador intuitivo de este error cuadrático medio se obtiene reemplazando el valor desconocido  $m(x_i)$  por la observación  $y_i$ , lo que conduce a la media de los cuadrados de los residuos de la regresión, dada por

$$\text{MCR}(h) = \frac{\sum_{i=1}^n [y_i - \hat{m}(x_i)]^2}{n}. \quad (4.32)$$

Sin embargo, éste estimador sesgado del error cuadrático medio lleva siempre a que la mejor elección es la menor anchura de ventana. La razón intuitiva de este hecho, es que para predecir  $y_i$  se usa la propia observación  $y_i$  en  $\hat{m}(x_i)$  y la expresión (4.32) se puede hacer arbitrariamente pequeña tomando  $h \rightarrow 0$  (en el caso extremo  $\hat{m}(\cdot)$  es una interpolación de las observaciones  $y_i$ ). Por tanto, una opción para evitar este problema

---

<sup>11</sup>En esta cantidad, y en consecuencia las correspondientes expresiones en (4.32)-(4.33)-(4.34), suele aparecer en la literatura al respecto un factor  $\alpha(x_i)$  adicional en el sumatorio. Esta función peso se suele usar para excluir del correspondiente cálculo un determinado porcentaje de observaciones extremas en la distribución. En este trabajo se ha optado por incluir todas las observaciones en los cálculos ya que, tras diversas elecciones de esta función, hemos comprobado que con ninguna otra se obtienen mejores resultados.

es excluir la observación  $(x_i, y_i)$  al estimar  $m(x_i)$ . Llamando a este estimador  $\hat{m}_{-i}(x_i)$ , se elige la anchura de la ventana mediante el denominado *método de validación cruzada* minimizando la expresión siguiente:

$$CV(h) = \sum_{i=1}^n [y_i - \hat{m}_{-i}(x_i)]^2. \quad (4.33)$$

Una sencilla manipulación de la expresión anterior prueba que se puede reescribir de una forma más conveniente desde el punto de vista computacional, concretamente:

$$CV(h) = \sum_{i=1}^n [y_i - \hat{m}(x_i)]^2 \cdot \Xi(w_i(x_i)), \quad (4.34)$$

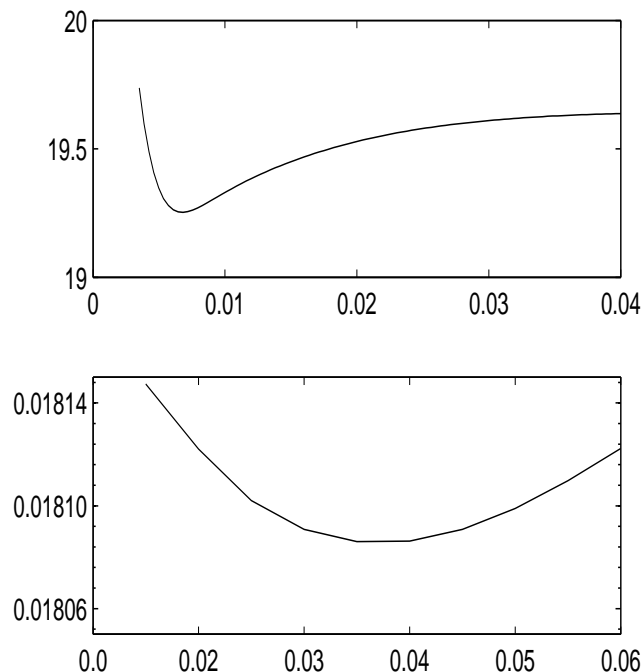
donde<sup>12</sup>  $\Xi(u) = (1 - u)^{-2}$  y  $w_i(x_i) = \frac{K(0)}{\sum_{j=1}^n K\left(\frac{x_i - x_j}{h}\right)}$ .

En consecuencia, el método de validación cruzada, el cual lleva a una anchura de ventana asintóticamente óptima, se puede interpretar también como la minimización de una suma de cuadrados de residuos ponderada, donde los valores pequeños de  $h$  se penalizan mediante la inclusión de la función  $\Xi(\cdot)$ . Para más detalles sobre este punto en particular y sobre la regresión no paramétrica en general se puede consultar Härdle (1997)[58].

En la Figura 4.11 aparecen representadas las funciones  $CV(h)$  correspondientes a las aproximaciones de orden uno a la deriva y al cuadrado de la volatilidad para los datos considerados de la Economía española dada por las ecuaciones (4.17) y (4.20) y usando un núcleo gaussiano. Los valores óptimos obtenidos con este criterio son  $h = 0,0068$  para la aproximación a la deriva y  $h = 0,0372$  para la del cuadrado de la volatilidad.

---

<sup>12</sup>Como es habitual, se usa la letra griega *xi mayúscula*,  $\Xi$ , para designar a la función que multiplica al cuadrado del residuo en la función  $CV$ .

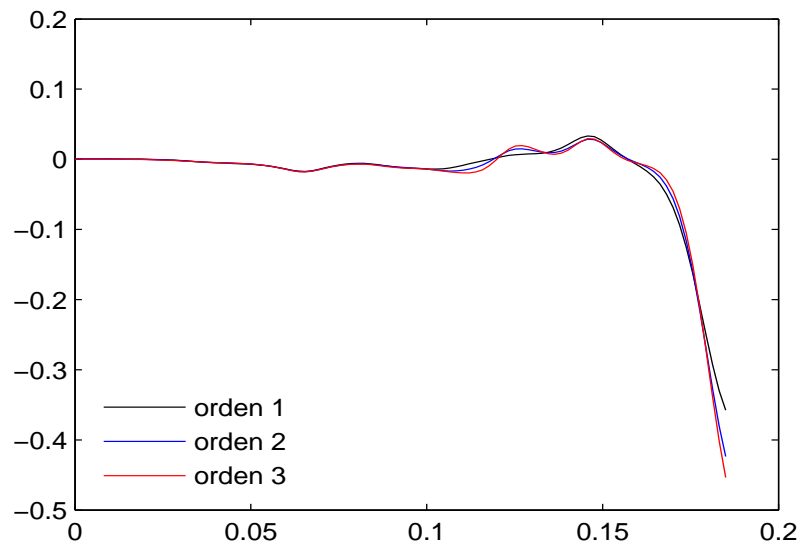


**Figura 4.11:** *Determinación de la anchura de la ventana mediante validación cruzada para la regresión no paramétrica conducente a la aproximación de orden uno de la deriva (arriba) y del cuadrado de la volatilidad (abajo) para los tipos de interés a corto plazo en España. Los valores mínimos se alcanzan con  $h = 0,0068$  para la deriva y con  $h = 0,0372$  para el cuadrado de la volatilidad.*

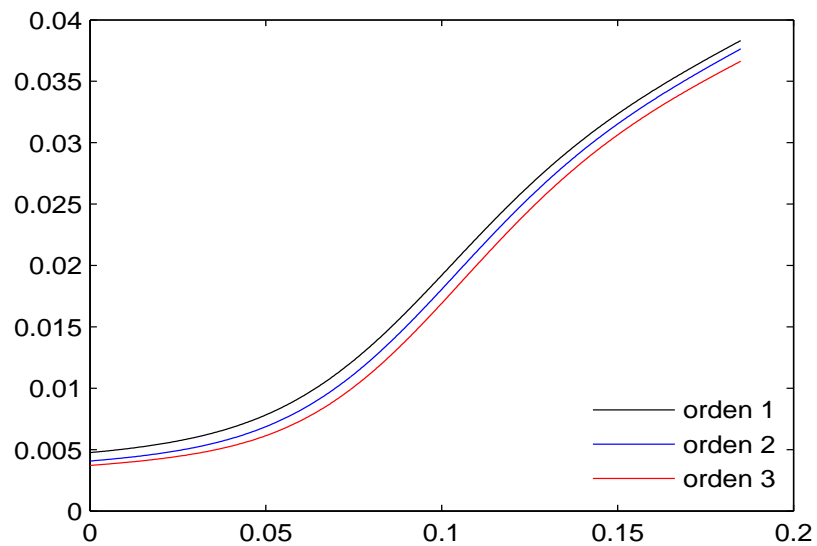
Llegados a este punto, estamos en condiciones de obtener las aproximaciones de orden uno, dos y tres a la deriva y a la volatilidad dadas por las expresiones (4.17)-(4.18)-(4.19) y (4.23)-(4.24)-(4.25). Para ello, usando las anchuras de ventana obtenidas por el método de validación cruzada en (4.34), se estima la esperanza condicionada en (4.26) con los datos apropiados ((4.27) para la deriva y (4.28) para el cuadrado de la volatilidad) y usando el estimador de Nadaraya-Watson en (4.30). En las Figuras 4.12 y 4.13 se recogen las representaciones gráficas de tales estimaciones.

De la observación de dichas gráficas, lo primero a destacar es que las tres estimaciones,





**Figura 4.12:** Aproximaciones de primer, segundo y tercer orden a la deriva de la difusión propuesta como modelo de evolución de los tipos de interés a corto plazo en España.



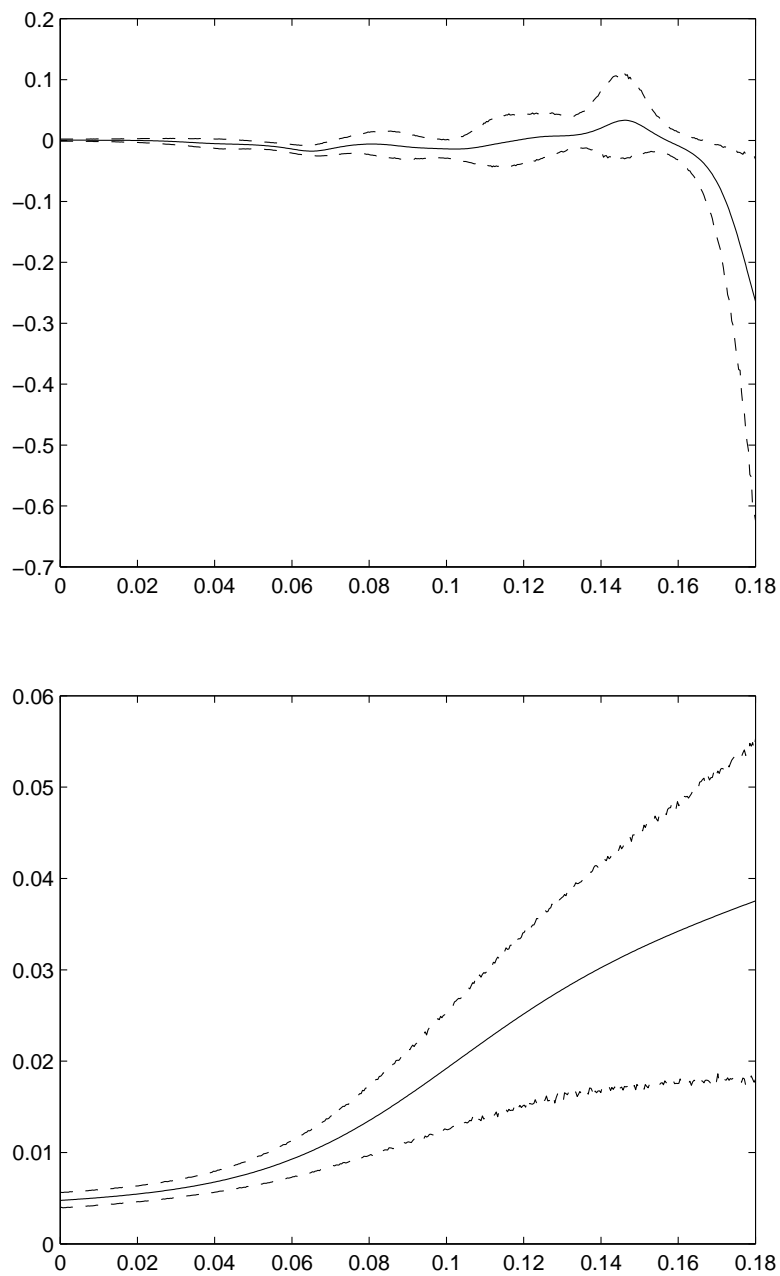
**Figura 4.13:** Aproximaciones de primer, segundo y tercer orden a la volatilidad de la difusión propuesta como modelo de evolución de los tipos de interés a corto plazo en España.

tanto para la deriva como para la volatilidad, son muy similares, que es precisamente lo que cabe esperar que ocurra cuando los errores de aproximación son pequeños. Respecto a las propiedades teóricas de las aproximaciones, es importante tener en cuenta que Fan y Zhan (2003)[41] en una revisión del trabajo de Stanton (1997)[94], encuentran que aunque el uso de las aproximaciones de orden superior reduce el sesgo asintótico de los estimadores de la deriva y de la volatilidad de la difusión, tiene el efecto de que la varianza asintótica crece casi exponencialmente con el orden de la aproximación. Este resultado teórico ensombrece el uso de aproximaciones de orden superior en la estimación de difusiones, y por ello, teniendo en cuenta además las estimaciones obtenidas en particular con los datos de la economía española, decidimos trabajar con las aproximaciones de orden uno.

De esta forma, en la Figura 4.14 se muestran las estimaciones de orden uno de la deriva y de la volatilidad de la difusión homogénea (4.1) que describe el tipo de interés a corto plazo en España; se proporcionan además, bandas puntuales al 95% de confianza construidas mediante el método bootstrap de bloques móviles<sup>13</sup> descrito en §4.2. Las anchuras de ventana del correspondiente estimador de Nadaraya-Watson se han determinado mediante validación cruzada. Es importante señalar, que también se han considerado otras opciones planteadas en la literatura. En concreto, algunos autores toman una anchura de ventana *ad-hoc* de orden  $O\left(n^{-\frac{1}{5}}\right)$ , que es la elección óptima (es la que ofrece un mejor balance entre el sesgo asintótico y la varianza asintótica del estimador). Teniendo en cuenta las elecciones en Stanton (1997)[94], Chapman y Pearson (2000)[20] y Arapis y Gao (2006)[2], se ha probado con las anchuras de ventana  $0'1 \cdot n^{-\frac{1}{5}}$ ,  $0'25 \cdot n^{-\frac{1}{5}}$ ,  $h_0$  y  $4 \cdot h_0$ , donde  $h_0 = S \cdot n^{\frac{1}{5}}$ . Hemos comprobado que ninguna de ellas mejora los resultados obtenidos a partir del método de validación cruzada.

---

<sup>13</sup>El tamaño de los bloques que se ha elegido en ésta y en las estimaciones posteriores basadas en este método es 50, lo que corresponde a un año de observaciones semanales ( $\Delta = \frac{1}{50}$ ).



**Figura 4.14:** Aproximaciones de primer orden y bandas al 95 % confianza de la deriva (arriba) y la volatilidad (abajo) de la difusión propuesta como modelo de evolución de los tipos de interés a corto plazo en España.

A la vista del gráfico de la parte superior de la Figura 4.14, se tiene que la estimación de la deriva es no lineal, mostrando una ligera reversión a la media para tipos de interés medios y bajos y cayendo fuertemente cuando los tipos de interés se sitúan por encima del 15%. Este comportamiento aparentemente no lineal en la deriva es el mismo que aparece en otros trabajos previos, como por ejemplo: Stanton (1997)[94], Jiang (1998)[64], Fernández (2001)[45] y Corzo y Gómez (2001)[22]. Estos autores concluyen, quizás algo precipitadamente, a partir de las estimaciones que realizan que existe evidencia suficiente de no linealidad en la deriva. Sin embargo, Chapman y Pearson (2000)[20] analizan el comportamiento muestral del estimador de la deriva y prueban, usando como referencia el modelo de CIR (ver Tabla 4.1), que aunque la verdadera deriva sea lineal, las estimaciones que se obtienen muestran una no linealidad precisamente del mismo tipo que la exhibida en la Figura 4.12. En este trabajo, tratamos de profundizar en este aspecto, considerando más adelante la estimación lineal local y en la sección siguiente contrastando la hipótesis de una deriva lineal.

Con respecto a la estimación de la volatilidad de la difusión, se observa en la Figura 4.14, que en general aumenta conforme los tipos de interés son mayores. Por tanto, no parece estar en consonancia con la volatilidad constante propuesta por ejemplo en Vasicek (1977)[97], sino más bien con la volatilidad de tipo potencial propuesta en Chan *et al.* (1992)[19]. En el apartado siguiente, se validan estas afirmaciones contrastando, a partir de los datos, estas formas paramétricas para la volatilidad de los tipos de interés a corto plazo en España.

Continuando con los gráficos de la Figura 4.14, es importante observar que las bandas de confianza obtenidas tanto para la deriva como para la volatilidad son bastante estrechas para tipos de interés bajos y medios, donde se dispone de muchas observaciones. Sin

embargo, para tipos de interés altos, donde las observaciones son escasas, las bandas de confianza se hacen mucho más anchas reflejando una menor confianza en las estimaciones.

Todas estas conclusiones sobre la deriva y la volatilidad de los tipos de interés a corto plazo en España estimada a partir de datos semanales de tipos de interés interbancarios en España coinciden con las obtenidas en Gavilán y Alba (2003)[47] a partir de la misma serie pero con una frecuencia de observación diaria.

Respecto a la forma particular que presenta la deriva estimada, ésta es estudiada en Chapman y Pearson (2000)[20]. La conclusión principal a la que llegan estos autores es que este tipo particular de no linealidad mostrado en la deriva estimada, puede ser espúrea y deberse al comportamiento del estimador no paramétrico en la frontera, donde las observaciones son más escasas y hay un mayor sesgo. En este sentido, es conocido que la generalización del estimador de Nadaraya-Watson (4.30) obtenida mediante una estimación local polinomial tiene mejores propiedades teóricas; en concreto, presenta un menor sesgo asintótico y un mejor comportamiento en la frontera. Para más detalles sobre este tema se puede consultar el trabajo de Fan y Gibels (1995)[42], en el cual se recomienda el uso de un ajuste local lineal. En este sentido, nos proponemos usar este tipo de estimador para comprobar si hay una diferencia sustancial con la estimación de la deriva que hemos obtenido mediante el estimador de Nadaraya-Watson.

Antes de pasar a la estimación de los parámetros del modelo propuesto en (4.1) a partir de la regresión lineal local con los datos que disponemos, vamos a describir brevemente en qué consiste el ajuste local polinomial a la esperanza condicionada (4.26). En primer lugar, resulta interesante observar que el estimador de Nadaraya-Watson (4.30), al cual denotamos por  $\hat{m}_{NW}$ , se puede ver como el estimador que minimiza la siguiente suma de

cuadrados ponderada:

$$\min_{a \in \mathbb{R}} \sum_{i=1}^n (y_i - a)^2 K \left( \frac{x_i - x}{h} \right). \quad (4.35)$$

En efecto, a partir del cálculo elemental se tiene que

$$a = \hat{m}_{NW}(x) = \frac{\sum_{i=1}^n y_i K \left( \frac{x_i - x}{h} \right)}{\sum_{i=1}^n K \left( \frac{x_i - x}{h} \right)},$$

que es el estimador de Nadaraya-Watson (4.30). De esta manera, este estimador se puede contemplar como un estimador localmente constante. Esto sugiere que se podría mejorar el estimador usando un polinomio local de grado  $p$  en la expresión (4.35) en lugar de la constante local  $a$ . Por tanto, sea  $x$  un valor fijo en el cual queremos estimar  $m(x)$ . Para valores  $u$  en un entorno de  $x$ , sea el polinomio

$$P_x(u; a) = a_0 + a_1(u - x) + \frac{a_2}{2!}(u - x)^2 + \dots + \frac{a_p}{p!}(u - x)^p.$$

Se puede aproximar una función de regresión diferenciable  $m(u)$  en un entorno del punto objetivo  $x$  mediante el polinomio

$$m(u) \approx P_x(u; a).$$

El vector de parámetros  $a = (a_0, \dots, a_p)'$  se estima eligiendo el valor  $\hat{a} = (\hat{a}_0, \dots, \hat{a}_p)'$  que minimiza la suma de cuadrados localmente ponderada (generalización de (4.35)) siguiente:

$$\sum_{i=1}^n (y_i - P_x(x_i; a))^2 K \left( \frac{x_i - x}{h} \right). \quad (4.36)$$

Es claro que el estimador  $\hat{a}$  depende del valor objetivo  $x$ , con lo que se suele escribir  $\hat{a} = (\hat{a}_0(x), \dots, \hat{a}_p(x))'$  cuando se quiere expresar explícitamente esta dependencia. El

estimador local de  $m$  viene dado por

$$\hat{m}(u) \approx P_x(u; \hat{a}).$$

En particular, en el valor objetivo  $u = x$  se tiene

$$\hat{m}(x) \approx P_x(x; \hat{a}) = \hat{a}_0(x).$$

Tomando  $p = 0$  se obtiene claramente el estimador de Nadaraya-Watson anterior, y el caso particular  $p = 1$  se denomina regresión lineal local, que es la elección que se recomienda. Para obtener la expresión explícita del estimador localmente lineal, al cual denotamos por  $\hat{m}_{LL}$ , se escribe la expresión (4.36) en forma matricial y se usa el estimador de mínimos cuadrados ponderados (ver Wasserman(2006)[99]), llegando a :

$$\hat{m}_{LL}(x) = \frac{S_2(x)T_0(x) - S_1(x)T_1(x)}{S_0(x)S_2(x) - S_1(x)^2}, \quad (4.37)$$

donde

$$S_j(x) = \sum_{i=1}^n K\left(\frac{x_i - x}{h}\right) (x_i - x)^j, \quad j = 0, 1, 2,$$

y

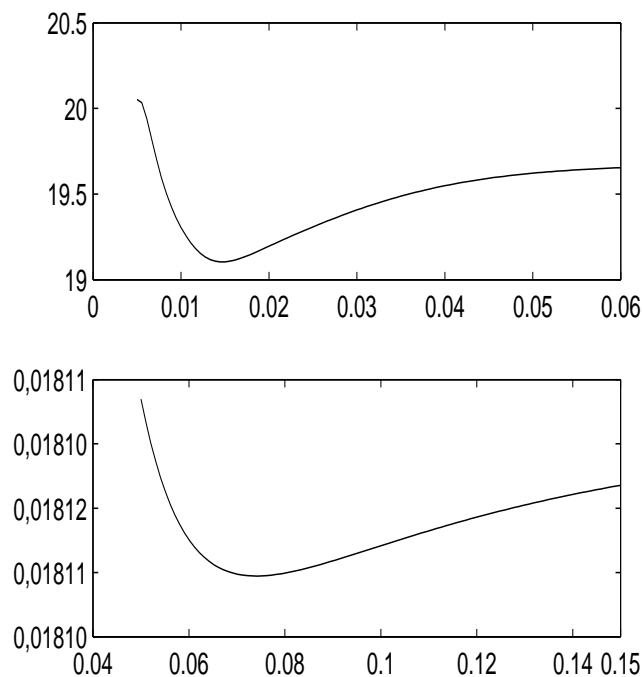
$$T_j(x) = \sum_{i=1}^n K\left(\frac{x_i - x}{h}\right) (x_i - x)^j y_i, \quad j = 0, 1.$$

Para la determinación de la anchura de la ventana de la regresión lineal local mediante el procedimiento de validación cruzada (4.33), en Härdle y Müller (2000)[59] se puede encontrar una expresión análoga a (4.34). En concreto, se tiene que

$$CV(h) = \sum_{i=1}^n [y_i - \hat{m}(x_i)]^2 \cdot \Xi(w_i(x_i)), \quad (4.38)$$

donde  $\Xi(u) = (1 - u)^{-2}$  y  $w_i(x_i) = \frac{K(0) \cdot S_2(x_i)}{S_0(x_i) \cdot S_2(x_i) - S_1(x_i)^2}$ .

En la Figura 4.15 se han representado las funciones  $CV(h)$  dadas por (4.38) para la aproximación de orden uno de la deriva y del cuadrado de la volatilidad usando la regresión lineal local para aproximar las correspondientes esperanzas condicionadas. En concreto, las anchuras óptimas son  $h = 0,0147$  para la deriva y  $h = 0,0740$  para el cuadrado de la volatilidad. Con el valor óptimo así obtenido para la anchura de la ventana y usando un

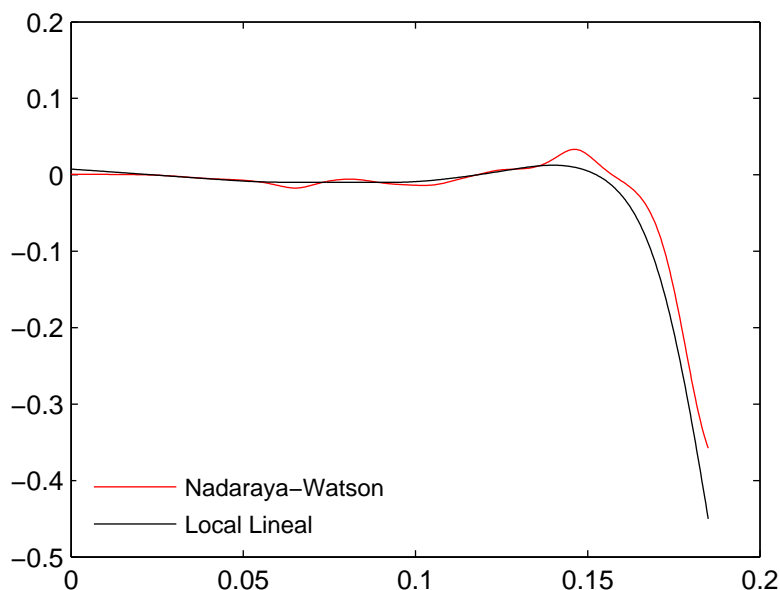


**Figura 4.15:** *Determinación de la anchura de la ventana mediante validación cruzada para la regresión lineal local conducente a la aproximación de orden uno de la deriva (arriba) y del cuadrado de la volatilidad (abajo) de la dinámica de los tipos de interés a corto plazo en España. Los valores mínimos se alcanzan con  $h = 0,0147$  para la deriva y con  $h = 0,0740$  para el cuadrado de la volatilidad.*

núcleo gaussiano, se ha llegado mediante la regresión lineal local (4.37) a una aproximación a la deriva<sup>14</sup> de los tipos a corto plazo en España. Tal aproximación, junto con la obtenida mediante el estimador de Nadaraya-Watson aparecen representadas en la Figura 4.16.

<sup>14</sup>Omitimos la aproximación de orden uno de la volatilidad ya que al estimar su cuadrado se obtienen valores negativos.





**Figura 4.16:** Aproximaciones de primer orden a la deriva de los tipos de interés a corto plazo en España mediante el estimador de Nadaraya-Watson y la regresión lineal local.

Claramente, lo primero que se observa es que ambas aproximaciones a la deriva son muy parecidas y exhiben el mismo comportamiento (lineal con una ligera reversión a la media para tipos medios y bajos y aparentemente no lineal, cayendo bruscamente, para tipos altos). En consecuencia, parece no haber una mejora significativa en la estimación al usar el estimador localmente lineal frente al de Nadaraya-Watson. Por ello, este resultado no nos permite avanzar en la obtención de alguna evidencia respecto a si la no linealidad que muestra la estimación de la deriva es espúrea, debida al proceso de estimación, o implicada por los datos. Este hecho está de acuerdo con Chapman y Pearson (2000)[20], donde se afirma, tras un estudio con datos simulados, que este tipo de no linealidad exhibida se debe más al truncamiento en la frontera superior de cualquier muestra por ser finita, es decir, un sesgo que aparece en la cola derecha de la distribución (donde están los datos mayores) debido al uso de muestras, que al conocido sesgo en la frontera (el cual queda mejorado desde el punto de vista de las propiedades del estimador al usar

la regresión lineal local). Este sesgo de truncamiento viene motivado porque al usar una muestra, aparece una relación negativa entre  $r(t + \Delta) - r(t)$  y  $r(t)$  en el límite superior de los datos observados. Como el estimador de Nadaraya-Watson usa sólo información local al estimar la función de regresión, acomoda esta relación negativa observada en la cola de la distribución.

Dedicamos la sección siguiente a profundizar en este aspecto de la aparente no linealidad de la deriva del modelo planteado.

#### 4.4. Contraste de la no linealidad de la deriva

En la sección anterior se ha obtenido una estimación no lineal de la deriva de los tipos de interés a corto plazo en España (Figura 4.14). Esta peculiar forma estimada de la deriva, con una muy leve reversión a la media para tipos de interés bajos y medios y cayendo bruscamente para tipos de interés altos, ha llevado a algunos autores, como por ejemplo Stanton (1997)[94], Jiang (1998)[64], Fernández (2001)[45] y Corzo y Gómez (2001)[22] a afirmar que la deriva de la difusión que modeliza los tipos instantáneos libres de riesgo en cada caso que analizan es no lineal. Sin embargo, parece algo precipitado afirmar tal no linealidad usando únicamente la estimación, y más teniendo en cuenta que sus bandas de confianza son más anchas (indicando una menor precisión en las estimaciones) para los tipos de interés más altos, que es precisamente donde se apoya la afirmación de no linealidad de la deriva. Además, Chapman y Pearson (2000)[20] sugieren que la no linealidad podría ser espúrea, debida a los sesgos ya mencionados en la sección anterior: El sesgo en la frontera de la regresión núcleo y el sesgo de truncamiento en el límite

superior que aparece al usar muestras finitas. En este sentido, con los datos considerados de la Economía española, hemos comprobado que el uso de una regresión lineal local para estimar la deriva, estimador que tiene mejores propiedades teóricas respecto al del sesgo en la frontera, no aporta más información sobre la linealidad o no de la deriva que el estimador más simple de Nadaraya-Watson (ver Figura 4.16).

Todas las afirmaciones realizadas en las secciones previas sobre la deriva de los tipos de interés a corto plazo en España y las que hacen al respecto los autores anteriormente mencionados a partir de otras series de datos, se apoyan fundamentalmente en estimaciones y representaciones gráficas. Nosotros consideramos que para justificar si la deriva estimada se aparta significativamente de la linealidad, es conveniente disponer de un test de hipótesis adecuado. Recientemente, se han desarrollado dos test de especificación de modelos no paramétricos, los cuales se pueden usar para contrastar la linealidad de la deriva en el modelo unifactorial continuo (4.1): El test de razón de verosimilitud generalizado (conocido como test GLR) de Fan, Zhang y Zhang (2001)[40] y el test de especificación de modelos no paramétricos (conocido como test  $J_n$ ) de Li (1994)[72] y Zheng (1996)[102].

La versión bootstrap del test GLR es considerada por Fan y Zhang (2003)[41], quienes argumentan tras diferentes experimentaciones que es un método potente para contrastar la especificación de varios modelos unifactoriales continuos. Sin embargo, Kim y Wang (2006)[68] estudian las versiones bootstrap de los dos test anteriores y encuentran, mediante experimentos de simulación, que la versión bootstrap del test GLR no proporciona tamaños estables sobre un rango de anchuras de ventana para llevar a cabo el test sobre la deriva en algunos modelos continuos, lo que dificulta el análisis de la potencia del test. Además, aplicar, como se hace en el test  $J_n$ , el método bootstrap salvaje (ver nota al pie de la página 169) en el test GLR lleva a resultados similares a los de el bootstrap ordinario.

Los propios Fan y Zhang (2003)[41] (pág. 127) señalan que la justificación rigurosa de que el test GLR se puede aplicar a datos de series temporales no ha sido aún proporcionada.

Por otra parte, el test  $J_n$  no adolece de ninguno de estos defectos ya que:

- Fan y Li (1997)[43] y Li (1999)[73] prueban que el test es aplicable a datos de series temporales y
- Kim y Wang (2006)[68] demuestran mediante experimentos de simulación con los modelos CIR y CKLS<sup>15</sup> que el tamaño del test es estable sobre un rango de anchuras de ventana.

Teniendo en cuenta todos los comentarios anteriores respecto la situación actual de los contrastes sobre la deriva aplicables a los modelos unifactoriales, descartamos el test GLR y pasamos a describir brevemente y a aplicar a los datos considerados el test  $J_n$ .

#### 4.4.1. El test $J_n$ . Resultados de su aplicación

En la sección anterior comentamos la conveniencia de trabajar con las aproximaciones de orden uno a la deriva y al cuadrado de la volatilidad del modelo unifactorial continuo planteado en (4.1). Tales aproximaciones vienen dadas por las ecuaciones (4.17) y (4.20) las cuales volvemos a reproducir aquí:

$$\mu(r(t)) = \frac{1}{\Delta} \mathbb{E}_t [r(t + \Delta) - r(t)] + O(\Delta), \quad (4.39)$$

$$\sigma^2(r(t)) = \frac{1}{\Delta} \mathbb{E}_t [(r(t + \Delta) - r(t))^2] + O(\Delta). \quad (4.40)$$

---

<sup>15</sup>Ver Tabla 4.1. Este segundo es uno de los modelos unifactoriales paramétricos más generales que se han planteado para la dinámica de los tipos de interés a corto plazo.

En base a la anterior aproximación a la deriva (4.39), se puede considerar el siguiente modelo de regresión:

$$Y(t) = \mu(X(t)) + v(t), \quad t = 1, \dots, n-1, \quad (4.41)$$

donde  $Y(t) = \frac{r(t+\Delta) - r(t)}{\Delta}$ ,  $X(t) = r(t)$  y  $\{v(t)\}$  son los términos de error, los cuales verifican que  $\mathbb{E}(v(t)|X(t)) = 0$ .

Para contrastar la hipótesis de un modelo paramétrico en la regresión (4.41), se consideran las siguientes hipótesis nula y alternativa:

$$\begin{aligned} H_0 : \quad \mu(x) &= g(x, \beta), \\ H_1 : \quad \mu(x) &\neq g(x, \beta), \end{aligned} \quad (4.42)$$

donde  $g(x, \beta)$  es una función conocida con un parámetro (o vector de parámetros) desconocido  $\beta$ . En concreto, estamos especialmente interesados en contrastar la linealidad de la deriva, es decir, en la especificación  $g(x, \beta) = \beta_1 + \beta_2 x$ .

El test  $J_n$  usa una estimación núcleo de la expresión  $\mathbb{E} \{ \mathbb{E} [e(t)|X(t)]^2 f(X(t)) \} \geq 0$ , siendo  $e$  el término error en la regresión  $Y(t) = g(X(t), \beta) + e(t)$  y  $f$  la función de densidad de  $X(t)$ . En la desigualdad anterior se da la igualdad cuando  $H_0$  es cierta. En la práctica, se usa  $\hat{e}(t)$  en lugar de  $e(t)$ , donde  $\hat{e}(t) = Y(t) - g(X(t), \hat{\beta})$  es un residuo y  $\hat{\beta}$  es un estimador de  $\beta$  obtenido por mínimos cuadrados ordinarios. Usando el estimador núcleo que deja una observación fuera (*leave one out*),  $\frac{1}{nh} \sum_{u \neq t} \hat{e}(u) \cdot K\left(\frac{X(u) - X(t)}{h}\right)$ , el test

estadístico se basa en la siguiente expresión:

$$\begin{aligned} I_n &= \frac{1}{n} \sum_{t=1}^n \hat{\varepsilon}(t) \left\{ \frac{1}{nh} \sum_{u=1, u \neq t}^n \hat{\varepsilon}(u) \cdot K \left( \frac{X(u)-X(t)}{h} \right) \right\} \\ &= \frac{1}{n^2 h} \sum_{t=1}^n \sum_{u=1, u \neq t}^n \hat{\varepsilon}(t) \hat{\varepsilon}(u) \cdot K \left( \frac{X(u)-X(t)}{h} \right), \end{aligned} \quad (4.43)$$

donde  $K$  es una función de densidad acotada y simétrica.

Bajo algunas condiciones de regularidad, Li y Wang (1998)[74] prueban que:

- Se tiene una distribución asintótica normal; más concretamente, bajo  $H_0$

$$J_n = \frac{n\sqrt{h}I_n}{\sqrt{\hat{\Omega}^2}} \rightarrow N(0, 1)$$

donde  $\hat{\Omega}^2 = \frac{2}{n^2 h} \sum_{t=1}^n \sum_{u=1, u \neq t}^n \hat{\varepsilon}(t)^2 \hat{\varepsilon}(u)^2 \cdot K \left( \frac{X(u)-X(t)}{h} \right)^2$  es un estimador consistente de  $\Omega^2 = 2\mathbb{E} \left\{ \mathbb{E}(e(1)^2 | X(1))^2 f(X(1)) \right\} \left\{ \int K^2(s) ds \right\}$ .

- El test  $J_n$  es un test consistente; más concretamente, bajo  $H_1$  se tiene que  $\mathbb{P}(J_n > c_n) \rightarrow 1$  para cualquier sucesión no estocástica  $c_n = o(n\sqrt{h})$ .

Li y Wang (1998)[74] recomiendan no usar la distribución normal estándar anterior, ya que comprueban mediante estudios de simulación que incluso con  $n = 1000$  el estadístico  $J_n$  aun no está cerca de dicha distribución asintótica. En su lugar, recomiendan usar el método bootstrap, especialmente el método bootstrap salvaje. Estos autores prueban que el método bootstrap salvaje lleva a aproximaciones válidas a la distribución de  $J_n$  bajo  $H_0$  incluso cuando los términos de error en el modelo de regresión son heteroscedásticos.

En definitiva, las etapas para obtener el  $p$ -valor del test  $J_n$  son las siguientes:

*Etapa 1:* Estimar la regresión  $Y(t) = g(X(t), \beta) + e(t)$  por mínimos cuadrados ordinarios y obtener los residuos  $\hat{e}(t) = Y(t) - g(X(t), \hat{\beta})$  y el estadístico  $J_n$ .

*Etapa 2:* Generar los errores bootstrap  $e^*(t)$  usando el método bootstrap salvaje, es decir,

$$e^*(t) = \begin{cases} a \cdot \hat{e}(t) & \text{con probabilidad } p = \frac{1+\sqrt{5}}{2\sqrt{5}} \\ b \cdot \hat{e}(t) & \text{con probabilidad } 1 - p \end{cases},$$

donde  $a = \frac{1-\sqrt{5}}{2}$  y  $b = \frac{1+\sqrt{5}}{2}$  (ver nota al pie<sup>16</sup>).

*Etapa 3:* Obtener  $Y^*(t) = g(X(t), \hat{\beta}) + e^*(t)$ . La muestra resultante  $(\mathcal{X}, \mathcal{Y}^*)$  se llama muestra bootstrap. Calcular los residuos bootstrap  $\hat{e}^*(t) = Y^*(t) - g(X(t), \hat{\beta}^*)$ , donde  $\hat{\beta}^*$  es el estimador por mínimos cuadrados de  $\beta$  usando la muestra bootstrap  $(\mathcal{X}, \mathcal{Y}^*)$ .

*Etapa 4:* Usar los residuos bootstrap para calcular el estadístico  $J_n^* = \frac{n\sqrt{h}I_n^*}{\sqrt{\hat{\Omega}^{*2}}}$ , donde  $I_n^*$  y  $\hat{\Omega}^{*2}$  son los anteriores  $I_n$  y  $\hat{\Omega}^2$  reemplazando los residuos  $\hat{e}(t)$  por los residuos bootstrap  $\hat{e}^*(t)$  obtenidos en la etapa anterior.

*Etapa 5:* Repetir las etapas 2–4 un número suficientemente grande de veces para obtener la distribución empírica del estadístico  $J_n^*$ . Esta distribución se llama distribución bootstrap y se usa para aproximar la distribución en muestras finitas del estadístico  $J_n$  bajo  $H_0$ . Por tanto, la proporción de veces que ocurre que  $J_n < J_n^*$  es el  $p$ -valor aproximado de  $J_n$ .

---

<sup>16</sup> Se llama bootstrap salvaje porque en cierto sentido se puede ver como un intento de reconstruir la distribución de cada residuo a través del uso de una sola observación, en lugar de remuestrear de todo el conjunto de residuos. De esta manera, se conserva mejor las características distribucionales del residuo (ver Härdle (1997)[58]). Con estas elecciones de los valores  $a$ ,  $b$  y  $r$  se consigue que  $\mathbb{E}e^*(t) = 0$ ,  $\mathbb{E}e^*(t)^2 = \hat{e}(t)^2$  y  $\mathbb{E}e^*(t)^3 = \hat{e}(t)^3$ . Resulta curioso que el valor  $b$  sea precisamente la razón áurea o límite de las razones de la sucesión de Fibonacci.

Para aplicar el test  $J_n$  a los datos semanales de los tipos interbancarios a un mes en la economía española, necesitamos únicamente elegir la anchura de la ventana. Kim y Wang (2006)[68], usan los valores  $h_j = j \cdot S \cdot n^{-1/5}$ , con  $j = 0,5, 1$  y  $1,5$ , siendo  $n$  el número de datos y  $S$  su desviación estándar. Tales valores tratan de recoger ventanas de diferentes tamaños: pequeño, mediano y grande (una especie de análisis de sensibilidad muy primitivo). Estas anchuras de ventana *ad hoc*, que son asintóticamente óptimas, se suelen usar al investigar los tamaños y potencias de los test, evitando otras como la validación cruzada mínimo-cuadrática por cuestiones de coste computacional. En este trabajo, también se usa esta cuarta anchura de ventana en todos los contrastes que se realizan. Los resultados de aplicar este test para contrastar la no linealidad de la deriva que modeliza la dinámica de los tipos de interés a corto plazo en España, usando los datos semanales de los tipos de interés interbancarios a un mes, las cuatro anchuras de ventana antes mencionadas y 5000 replicaciones bootstrap de la muestra original, aparecen recogidas en la Tabla 4.5. La deriva lineal estimada en este caso viene dada por  $g(x, \hat{\beta}) = 0,001807 - 0,101022 \cdot x$ .

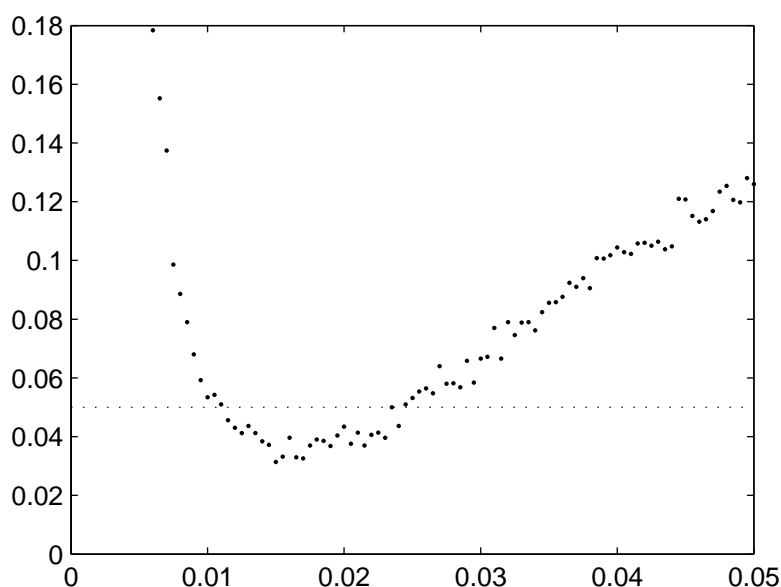
Anchura de Ventana	Estadístico $J_n$	$p$ -valor
$h_1 = 0,0063$	0,7992	0,1566
$h_2 = 0,0127$	1,0451	0,0412
$h_3 = 0,0190$	0,6044	0,0396
$h_{CV} = 0,0068$	0,8647	0,1315

**Tabla 4.5:** Resultados del test  $J_n$  con 5000 replicaciones bootstrap para contrastar la linealidad de la deriva en la dinámica de los tipos de interés a corto plazo en España.

Al contrario de lo que sería deseable, se observa en la Tabla 4.5 cierta variación del  $p$ -valor del contraste para las diferentes anchuras de ventana consideradas. Esto nos lleva a hacer un análisis más profundo, efectuando los cálculos anteriores para todo un rango de anchuras de ventana que incluye a las cuatro anteriormente mencionadas. La Figura 4.17 muestra los  $p$ -valores obtenidos para anchuras de ventana en el intervalo  $[0'006, 0'050]$ , la línea horizontal corresponde al usual 5% de significatividad. Observamos que la mayoría de los  $p$ -valores se sitúan por encima de la línea del 5% de significatividad. Sólo unas



pocas elecciones (entre las que se encuentran  $h_2$  y  $h_3$ ) conducen a  $p$ -valores inferiores a 0,05, y por tanto a rechazar la hipótesis de linealidad en la deriva. Además todos los  $p$ -valores se sitúan por encima del 3% (el más pequeño observado es 0,0314). Como resultado importante teniendo todo lo anterior en cuenta, se puede concluir que no existe una evidencia contundente que conduzca a rechazar la hipótesis de linealidad de la deriva del modelo propuesto para describir los tipos de interés a corto plazo en España.



**Figura 4.17:**  $P$ -valores del test  $J_n$  para contrastar la linealidad de la deriva de los tipos de interés a corto plazo en España en función del valor de la anchura de la ventana. La línea horizontal representa el usual 5% de significatividad.

Por otro lado, en la sección anterior se afirma que la volatilidad estimada induce a pensar que se ajusta a la forma potencial especificada en el modelo paramétrico CKLS (ver Figura 4.14). Podemos contrastar formas específicas del cuadrado de la volatilidad usando el test  $J_n$  sin más que tener en cuenta su aproximación de orden uno de la ecuación (4.40). En concreto, tomando ahora  $Y(t) = \frac{[r(t+\Delta)-r(t)]^2}{\Delta}$  y  $X(t) = r(t)$ , la regresión (4.41) se convierte en

$$Y(t) = \sigma^2(X(t)) + e(t), \quad t = 1, \dots, n-1. \quad (4.44)$$

Los resultados del contraste de un cuadrado de la volatilidad constante como en el modelo de Vasicek (1977)[97] y de la forma paramétrica potencial especificada por el modelo CKLS aparece en la Tabla 4.6. En este caso, hay un claro rechazo de una volatilidad constante y no se rechaza, a los niveles usuales de significatividad, una forma paramétrica potencial para la volatilidad, en acuerdo con lo que aparentemente se deduce a partir de las estimaciones obtenidas en la sección anterior. La volatilidad estimada mediante mínimos cuadrados no lineales resulta ser, para los datos considerados  $\hat{\sigma}(x) = 0,241450 \cdot x^{1,430828}$ .

Anchura de Ventana	$H_0 : \sigma^2(x) = cte$		$H_0 : \sigma^2(x) = \sigma^2 x^{2\gamma}$	
	Estadístico $J_n$	$p$ -valor	Estadístico $J_n$	$p$ -valor
$h_1 = 0,0063$	4,4245	0,0000	2,5043	0,137
$h_2 = 0,0127$	5,1567	0,0000	0,2548	0,201
$h_3 = 0,0190$	5,5250	0,0000	-0,3810	0,232
$h_{CV} = 0,0372$	4,7032	0,0000	-0,3598	0,184

**Tabla 4.6:** Resultados del test  $J_n$  con 1000 replicaciones bootstrap para contrastar un cuadrado de la volatilidad constante y potencial.

## 4.5. Aplicación a la valoración de bonos: Comparaciones entre de diferentes modelos

Una vez establecido el modelo unifactorial dado por la ecuación diferencial estocástica (4.1) para el tipo de interés instantáneo libre de riesgo en España y estimado mediante técnicas no paramétricas, obteniendo como resultado las estimaciones expuestas en la Figura 4.14, podemos plantearnos usarlo para una de sus aplicaciones fundamentales: La valoración de activos derivados de los tipos de interés. Para ello, según la Proposición 3.15 solo queda estimar el parámetro  $\lambda = \lambda_0 \cdot \sigma$ , donde  $\lambda_0$  es precio del riesgo de mercado, y persiguiendo este fin, vamos a usar la metodología no paramétrica propuesta en Jiang

(1998)[64] para estimar parámetros de la forma  $\lambda(r(t)) = \lambda_0(r(t)) \cdot \sigma(r(t))$ , que conduce a un modelo libre de arbitraje.

Con este objetivo, y teniendo en cuenta que el precio del riesgo de mercado  $\lambda_0(r(t))$  está relacionado con el exceso de rendimiento de los activos dependientes de los tipos de interés (Proposición 3.14), Jiang (1998)[64] considera los rendimientos de dos bonos cupón cero, denotados aquí por  $R_{T_i}(r(t))$ ; siendo  $r(t)$  el nivel de los tipos de interés a corto plazo y  $T_i$  el instante de vencimiento del bono  $i$ , con  $i = 1, 2$ . Supongamos que las dinámicas de tales rendimientos vienen también dadas por ecuaciones diferenciales estocásticas homogéneas, en concreto

$$dR_{T_i}(r(t)) = \xi_{T_i}(r(t)) dt + \nu_{T_i}(r(t)) dW(t). \quad (4.45)$$

Teniendo en cuenta la relación entre los precios y los rendimientos en tiempo continuo (3.15), el Lema de Itô junto con la dinámica de los precios dadas por las ecuaciones (3.18) y (3.19) llevan, para  $i = 1, 2$ , a que

$$\begin{aligned} \alpha_{T_i}(r(t)) &= R_{T_i}(r(t)) - (T_i - t) \xi_{T_i}(r(t)) + \frac{1}{2} (T_i - t)^2 \nu_{T_i}^2(r(t)), \\ \sigma_{T_i}(r(t)) &= - (T_i - t) \nu_{T_i}(r(t)). \end{aligned} \quad (4.46)$$

Como además el precio del riesgo de mercado no depende del instante de vencimiento, se deduce a partir de la ecuación (3.22) particularizada en  $T_1$  y  $T_2$  que

$$\lambda_0(r(t)) = \frac{\alpha_{T_1}(r(t)) - \alpha_{T_2}(r(t))}{\sigma_{T_1}(r(t)) - \sigma_{T_2}(r(t))}.$$

Sustituyendo ahora en la igualdad anterior las expresiones alternativas de (4.46) se obtiene

finalmente

$$\lambda_0(r(t)) = \frac{R^d(r(t), \tau_1, \tau_2) + \tau_2 \xi_{T_2}(r(t)) - \tau_1 \xi_{T_1}(r(t)) + \frac{1}{2} [\tau_1^2 \nu_{T_1}^2(r(t)) - \tau_2^2 \nu_{T_2}^2(r(t))]}{\tau_2 \nu_{T_2}(r(t)) - \tau_1 \nu_{T_1}(r(t))}, \quad (4.47)$$

donde  $\tau_i = T_i - t$ ,  $i = 1, 2$  son los tiempos hasta los vencimientos de cada activo y  $R^d(r(t), \tau_1, \tau_2)$  es el diferencial del rendimiento entre ambos activos. En consecuencia, para obtener una aproximación al precio del riesgo de mercado, basta elegir dos instantes de vencimiento cualquiera  $T_1$  y  $T_2$  y estimar la deriva y la volatilidad de los procesos de los rendimientos a vencimiento  $R_{T_i}(r(t))$ , para  $i = 1, 2$ , mediante la regresión núcleo.

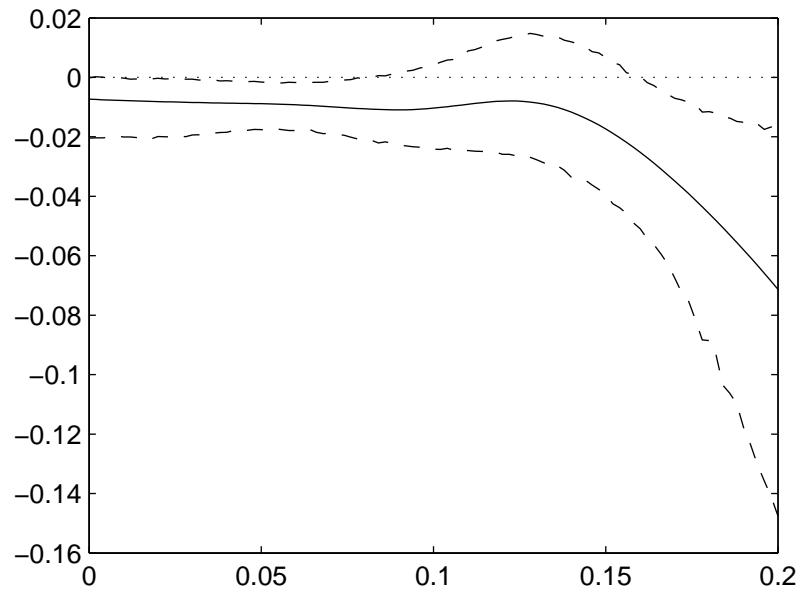
Para realizar tal estimación con datos de los tipos de interés interbancarios en España, se toman datos semanales de los rendimientos a un mes ( $T_1 = \frac{1}{12}$ ) y de los rendimientos a una semana ( $T_2 = \frac{1}{50}$ ) y se estiman los procesos de los rendimientos a vencimiento (4.45) mediante una aproximación de orden uno, usando la regresión núcleo con un núcleo gaussiano<sup>17</sup>. En la Figura 4.18 aparece dicha estimación, junto con una banda de confianza puntual al 95 % obtenida por el método bootstrap de bloques móviles.

Respecto a las estimaciones obtenidas y representadas en la Figura, 4.18 cabe hacer las siguientes apreciaciones:

- Es no nulo globalmente, siendo cercano a cero para tipos de interés bajos y medios y decreciendo fuertemente conforme los tipos de interés aumentan. Por tanto, exhibe un comportamiento bastante diferente del que cabría esperar bajo algunas funciones paramétricas propuestas con anterioridad (por ejemplo, constante

---

<sup>17</sup>Respecto a la anchura de la ventana, después de considerar varias posibilidades (validación cruzada en cada regresión y algunas elecciones *ad hoc* de la literatura), el mejor resultado se obtiene, en consonancia con los resultados alcanzados en Arapis y Gao (2006)[2] para otros conjuntos de datos similares, tomando una anchura de ventana global para todas las regresiones de  $h = 0,10 \cdot n^{-\frac{1}{5}}$  donde  $n$  es el número de observaciones.



**Figura 4.18:** *Estimación del precio del riesgo de mercado mediante la ecuación (4.47), usando una aproximación de orden uno y regresiones núcleo con núcleo gaussiano y  $h = 0,10 \cdot n^{-\frac{1}{5}}$  para estimar las correspondientes derivas y volatilidades. Se añade además, banda puntual de confianza al 95 % obtenida por el método bootstrap de bloques móviles.*

en Vasicek (1977)[97], nulo en el modelo CKLS (1992)[19], lineal en el modelo CIR (1985)[24, 25]).

- La estimación resultante es negativa, lo que corresponde a una prima positiva por soportar el riesgo de tipo de interés.
- No sorprende que la prima de riesgo sea menor para tipos de interés menores, ya que los tipos de interés bajos están asociados con una menor volatilidad que los altos.
- De una manera similar a la deriva y a la volatilidad estimada del proceso, la banda de confianza puntual es más ancha conforme los tipos son más altos, ya que hay menos observaciones con tipos de interés altos.

Una vez estimada la deriva, la volatilidad y el precio del riesgo de mercado del modelo

---

(4.1) propuesto para la evolución de los tipos de interés a corto plazo en España, se puede usar para valorar activos derivados de los tipos de interés. Para ello, disponemos principalmente de dos enfoques (ver Capítulo 5 de Kwok (1998)[71]):

- Resolver, mediante algún algoritmo numérico, la ecuación en derivadas parciales en (3.25) con las condiciones adicionales (inicial, final, de frontera) adecuadas a cada caso.
- Obtener simulaciones Monte Carlo del proceso de los tipos de interés a corto plazo bajo la medida de riesgo neutral, es decir, del proceso

$$dr(t) = [\mu(r(t)) - \lambda(r(t))] dt + \sigma(r(t)) dW^*(t),$$

donde  $W^*$  es un movimiento browniano estándar bajo tal medida de probabilidad. Con esas trayectorias simuladas, todas comenzando en el tipo de interés  $r(0) = r$  observado en el instante inicial  $t = 0$  y terminando en el instante final  $T$ , podemos estimar el precio del activo derivado aproximando el valor esperado en (3.26) por la correspondiente media.

Respecto al método Monte Carlo, introducido en Economía Financiera por Boyle (1977)[13], hoy día es una técnica poderosa y versátil para valorar activos derivados. Este método, que parece estar en desventaja cuando hay un algoritmo numérico estable que no requiere demasiado tiempo de cómputo para resolver la ecuación en derivadas parciales, se suele usar cuando se quieren calcular los precios de los activos derivados sólo en unos pocos puntos o cuando el enfoque basado en resolver una ecuación en derivadas parciales es difícil de implementar debido a la complejidad del problema.

La principal ventaja del método Monte Carlo para valorar activos derivados es que

resulta fácil de acomodar a funciones de pagos terminales complicadas. Así cuando los profesionales se enfrentan a un nuevo activo derivado que acaba de ser diseñado, pueden obtener un precio estimado usando la simulación Monte Carlo, en lugar de arriesgarse ellos mismos a construir un modelo analítico para la valoración del nuevo activo. Sin embargo, su principal desventaja es la demanda de un gran número de simulaciones para alcanzar un alto nivel de precisión. Aunque, este aspecto se ha visto enormemente mejorado con los recientes avances en el campo de la Informática, propiciando un nuevo impulso en el desarrollo y las aplicaciones de esta metodología. Además, la eficiencia de la simulación Monte Carlo se puede aumentar mediante alguna técnica de reducción de la varianza (Boyle *et al* (1997)[14]).

A continuación, en vista de la importancia de este enfoque, explicamos brevemente cómo usar la simulación Monte Carlo para valorar activos derivados de los tipos de interés y cómo reducir la variabilidad de las estimaciones usando un procedimiento que resulta adecuado en nuestro contexto: el método de las variaciones antitéticas.

En primer lugar, se ha visto en el capítulo anterior (Proposición 3.15) que el valor en el instante  $t < T$  de un  $T$ -activo derivado de la forma  $X = \Phi(r(T))$  viene dado por el siguiente valor esperado, calculado bajo una medida equivalente de martingalas:

$$p(t, T) = \mathbb{E}_{t,r}^* \left[ X \cdot \exp \left\{ - \int_t^T r(u) du \right\} \right]. \quad (4.48)$$

Teniendo en cuenta que la simulación Monte Carlo es básicamente un procedimiento numérico para estimar el valor esperado de una variable aleatoria, resulta ser una herramienta natural para valorar ciertos activos derivados. El procedimiento para ello, se lleva a cabo mediante las siguientes etapas:

1. Simular, en el intervalo temporal  $[t, T]$ , una cantidad  $M$  suficientemente grande de trayectorias del proceso de los tipos de interés a corto plazo  $\{r(t)\}$  bajo la medida equivalente de martingalas. Dicho proceso, bajo tal medida de probabilidad obedece, en nuestro contexto, a la siguiente dinámica (ecuación (3.27)):

$$dr(t) = [\mu(r(t)) - \lambda(r(t))] dt + \sigma(r(t)) dW^*(t), \quad \text{con } r(t) = r. \quad (4.49)$$

2. Para cada camino simulado, evaluar la expresión en el interior de la esperanza matemática (4.48). Así obtenemos unos precios simulados  $p_i(t, T)$  con  $i = 1, \dots, M$ .
3. Estimar el precio del activo derivado mediante la media muestral de los valores obtenidos en la etapa anterior:

$$\hat{p}(t, T) = \frac{\sum_{i=1}^M p_i(t, T)}{M}. \quad (4.50)$$

Centrándonos en la etapa 1, para simular una trayectoria del proceso descrito por la ecuación (4.49) se suele usar una versión discreta del mismo, la más habitual es la aproximación de orden uno o esquema de Euler, el cual viene dado por:

$$\hat{r}(t + \Delta) = \hat{r}(t) + [\mu(\hat{r}(t)) - \lambda(\hat{r}(t))] \Delta + \sigma(\hat{r}(t)) \sqrt{\Delta} \cdot Z,$$

donde  $\{\hat{r}(t)\}$  denota una aproximación en tiempo discreto al proceso  $\{r(t)\}$ ,  $Z$  es una variable aleatoria normal estándar y  $\Delta$  el tiempo entre dos observaciones consecutivas en la trayectoria simulada. En definitiva, para obtener una aproximación discreta  $\{\hat{r}(t)\}$  en un conjunto de tiempos equidistantes  $t_j = t_0 + j \cdot \Delta$  para  $j = 1, \dots, N$ , basta fijar el valor en el instante inicial  $r(0) = \hat{r}(0)$  y aplicar la siguiente ecuación recursiva

$$\hat{r}(t_j) = \hat{r}(t_{j-1}) + [\mu(\hat{r}(t_{j-1})) - \lambda(\hat{r}(t_{j-1}))] \Delta + \sigma(\hat{r}(t_{j-1})) \sqrt{\Delta} \cdot Z_j, \quad (4.51)$$



donde  $Z_j$  ( $j = 1, \dots, N$ ) son variables aleatorias normales estándar independientes, las cuales se pueden generar sin dificultad con el software adecuado.

La aproximación de Euler, aunque es la más sencilla proporciona muy buenos resultados<sup>18</sup>. Sin embargo, con un pequeño ajuste se obtiene una aproximación con un mejor comportamiento en cuanto a la convergencia del proceso discretizado  $\hat{r}$  al proceso de tiempo continuo  $r$  (Talay (1996)[95]). En concreto, la aproximación de Milstein, usada por ejemplo en los trabajos de Jiang (1998)[64] y Fernández (2001)[45], se obtiene añadiendo un término adicional en (4.51), que proviene de considerar un término más en un desarrollo de Taylor-Itô, llegando a la siguiente ecuación recursiva

$$\begin{aligned} \hat{r}(t_j) = & \hat{r}(t_{j-1}) + [\mu(\hat{r}(t_{j-1})) - \lambda(\hat{r}(t_{j-1}))] \Delta + \sigma(\hat{r}(t_{j-1})) \sqrt{\Delta} \cdot Z_j + \\ & \frac{1}{2} \sigma^2(\hat{r}(t_{j-1})) \cdot \Delta (Z_j^2 - 1). \end{aligned} \quad (4.52)$$

Un aspecto importante a tener en cuenta al usar el enfoque Monte Carlo basado en la ecuación (4.50) es que si las replicaciones son independientes, además de aproximar el precio del activo derivado de los tipos de interés se puede dar una estimación de su desviación estándar mediante

$$\frac{\hat{S}_p}{\sqrt{M}}, \quad \text{donde} \quad \hat{S}_p = \sqrt{\frac{\sum_{i=1}^M [p_i(t, T) - \hat{p}_i(t, T)]^2}{M - 1}}.$$

En consecuencia, la precisión del estimador se puede incrementar de dos formas diferentes:

- Aumentando el número de simulaciones  $M$ . La aparición del factor  $\frac{1}{\sqrt{M}}$  implica que la reducción de la desviación estándar por ejemplo a 10 veces menos requiere multiplicar por 100 el número de simulaciones.

---

<sup>18</sup>Stanton (1997)[94] comprueba, mediante la simulación de varios procesos, que cuando la frecuencia de observación es al menos mensual los errores de discretización son pequeños.

- Usar algún método de reducción de la varianza para disminuir el término  $\hat{S}_p$ . Una técnica comúnmente usada para ello es el método de las variaciones antitéticas, con el que se intenta reducir la varianza introduciendo dependencia negativa entre pares de replicaciones.

A continuación, explicamos el método de las variaciones antitéticas: Hemos visto en (4.51) y (4.52) que para simular una trayectoria del proceso  $\{r(t)\}$  bajo  $\mathbb{P}^*$  se necesita generar valores de las variables normales estándar  $Z_j$  para  $j = 1, \dots, N$ . Una vez simulada la trayectoria, en base a (4.48) se genera un precio y repitiendo el proceso un número  $M$  grande de veces se consigue un estimador insesgado del precio del activo derivado mediante (4.50). Ahora bien, nótese que como las variables  $Z_j$  son normales estándar, las variables  $-Z_j$  también lo son. Entonces, se puede usar los valores  $-Z_j$  ( $j = 1, \dots, N$ ) para generar una nueva trayectoria del proceso  $\{r(t)\}$  bajo  $\mathbb{P}^*$  y un nuevo precio  $p^*(t, T)$ . En consecuencia, la repetición del procedimiento conduce a la construcción de un nuevo estimador insesgado del precio del activo derivado, denotado por  $p^*(t, T)$ , mediante

$$\hat{p}^*(t, T) = \frac{\sum_{i=1}^M p_i^*(t, T)}{M}. \quad (4.53)$$

Es razonable esperar que  $p_i(t, T)$  y  $p_i^*(t, T)$  estén negativamente correlados, es decir, si uno sobreestima el verdadero valor, el otro lo subestima. Por tanto, parece adecuado estimar el precio del activo derivado tomando la media de esos dos estimadores, construyendo así el estimador mediante las variaciones antitéticas, denotado por  $\hat{p}(t, T)$ , como

$$\hat{p}(t, T) = \frac{\hat{p}(t, T) + \hat{p}^*(t, T)}{2}.$$

La eficiencia de este estimador mejora la del Monte Carlo ordinario supuesto que

$Cov(p_i(t, T), p_i^*(t, T)) \leq 0$  (ver por ejemplo Boyle (1977)[13], Glasserman (2004)[52], Platen y Heath (2006)[87]). Además, es importante observar el ahorro en el esfuerzo computacional al cambiar de signo las variables  $Z_j$  ya generadas en lugar de simular nuevas variables aleatorias normales estándar.

Por tanto, estamos ya en condiciones de aplicar el método Monte Carlo para valorar bonos cupón cero con diferentes vencimientos y niveles iniciales de los tipos de interés a corto plazo. Para estimar cada precio, se generan 2000 trayectorias mediante la aproximación de Milstein con 100 datos diarios en cada una de ellas y usando las estimaciones de orden uno de la deriva, la volatilidad y el precio del riesgo de mercado obtenidas con una anchura de ventana común (ver nota al pie de la página 174). Además, la variabilidad de las estimaciones es reducida usando el método de las variaciones antitéticas<sup>19</sup>. Los resultados obtenidos aparecen recogidos en la Tabla 4.7.

$T \setminus r$	3 %	6 %	9 %	12 %
1	0,9681	0,9405	0,9122	0,8829
2	0,9329	0,8824	0,8284	0,7727
3	0,8951	0,8255	0,7490	0,6701
4	0,8556	0,7704	0,6738	0,5763
5	0,8148	0,7171	0,6027	0,4945

**Tabla 4.7:** Precios estimados de bonos cupón cero para la Economía española mediante el método Monte Carlo. Se generan, para cada precio, 2000 trayectorias con 100 datos diarios mediante la aproximación de Milstein y reduciendo la variabilidad de las estimaciones mediante el método de las variaciones antitéticas.

Con el objetivo de ahondar un poco más sobre las repercusiones que tiene imponer en el modelo (4.1) de la dinámica del tipo de interés a corto plazo una deriva lineal, valoramos de nuevo los mismos bonos cupón cero de la Tabla 4.7 pero imponiendo la deriva lineal

---

<sup>19</sup>Con estas elecciones, Jiang (1998)[64] encuentra que el método Monte Carlo es de gran precisión. Para ello, usando datos de la economía estadounidense, compara las curvas de rendimiento obtenidas por este método con la que se obtiene resolviendo los correspondientes problemas de contorno (3.25) y encuentra que las diferencias son pequeñas. Además, comprueba que el método de las variaciones antitéticas reduce en gran parte la varianza de las estimaciones.

para la que no se encontró una evidencia contundente para su rechazo usando el test  $J_n$  (ver §4.4) y estimando el resto de los parámetros (volatilidad y precio del riesgo de mercado) de manera no paramétrica. Los precios estimados en esta nueva situación y siguiendo la misma metodología y elecciones que en la Tabla 4.7 se presentan en la Tabla 4.8; en la cual se han resaltado en **negrita** los precios que difieren menos del 1% del correspondiente no paramétrico general (sin imponer la linealidad de la deriva), en **azul** los que difieren entre un 1% y un 2% y en **rojo** los que difieren entre un 2% y un 2'5% (la mayor diferencia observada es de un 2,4022%). Comparando las dos tablas anteriores

$T \setminus r$	3%	6%	9%	12%
1	<b>0,9670</b>	<b>0,9395</b>	<b>0,9124</b>	<b>0,8872</b>
2	<b>0,9290</b>	<b>0,8784</b>	<b>0,8303</b>	<b>0,7847</b>
3	<b>0,8870</b>	<b>0,8178</b>	<b>0,7534</b>	<b>0,6840</b>
4	<b>0,8421</b>	<b>0,7578</b>	<b>0,6804</b>	<b>0,5885</b>
5	<b>0,7952</b>	<b>0,6995</b>	<b>0,6118</b>	<b>0,5064</b>

**Tabla 4.8:** Precios estimados de bonos cupón cero en España imponiendo en la ecuación (4.1) una deriva lineal. Se sigue el método Monte Carlo generando, para cada precio, 2000 trayectorias con 100 datos diarios mediante la aproximación de Milstein y reduciendo la variabilidad de las estimaciones mediante el método de las variaciones antitéticas.

se observa que:

- En general, las diferencias entre los precios estimados mediante el modelo completamente no paramétrico y el que impone una deriva lineal aumentan con el nivel inicial de los tipos de interés y con el periodo del vencimiento del bono cupón cero.
- Todos los precios estimados bajo la hipótesis de una deriva lineal difieren de los correspondientes no paramétricos menos de un 2,5%. Además, para tipos de interés iniciales inferiores al 9% (rango en el cual la deriva estimada de forma no paramétrica se aproxima a la linealidad) y periodos de vencimiento de 3 años o menos, estas diferencias son inferiores al 1%.

En conclusión, la adopción de una deriva lineal lleva a precios de bonos cupón cero bastante similares a los del modelo general, más aún si el tipo de interés inicial es bajo o medio (no superior al 9%) y el periodo de vencimiento no supera los 3 años.

Según los resultados obtenidos en §4.4, a un nivel de significatividad del 3%, no existen evidencias para rechazar una deriva lineal y aún menos una volatilidad potencial. Teniendo en cuenta que precisamente éstas son las formas funcionales propuestas para la deriva y la volatilidad en el modelo CKLS [19], parece que éste es un modelo paramétrico razonable para modelizar el tipo de interés a corto plazo en España mediante la difusión homogénea (4.1). Para ahondar un poco más en este aspecto, nos proponemos valorar con este modelo paramétrico los mismos bonos cupón cero que antes con el modelo no paramétrico.

Para estimar la deriva y la volatilidad de tal proceso paramétrico de tiempo continuo en base a las observaciones discretas  $\{r(t), t = 1, \dots, n\}$  seguimos el procedimiento propuesto en Fan y Zhang (2003)[41]. Este proceso de estimación se basa en la siguiente discretización de primer orden del modelo paramétrico:

$$Y(t) \approx (\alpha + \beta X(t)) \Delta + \sigma X(t) \sqrt{\Delta} Z(t), \quad (4.54)$$

donde  $Y(t) = r(t+1) - r(t)$ ,  $X(t) = r(t)$ ,  $\Delta$  es la separación temporal entre dos observaciones consecutivas y  $Z(t)$  son variables aleatorias normales estándar independientes, con  $t = 1, \dots, n - 1$ . Las etapas para realizar la estimación son las siguientes:

**Etapas I:** Estimar el modelo (4.54) como si fuera homoscedástico y mediante mínimos cuadrados ordinarios obtener  $(\hat{\alpha}^{(1)}, \hat{\beta}^{(1)})$  estimaciones de  $(\alpha, \beta)$ .

**Etapas II:** Sea la variable  $\hat{e}(t) = [Y(t) - (\hat{\alpha}^{(1)} + \hat{\beta}^{(1)} X(t)) \Delta] \Delta^{-\frac{1}{2}}$ , la cual transforma

el modelo (4.54) en

$$\log(\hat{\varepsilon}(t)^2) \approx \log(\sigma^2) + \gamma \log(X(t)^2) + \log(Z(t)^2), \quad t = 1, \dots, n-1. \quad (4.55)$$

Obtener por mínimos cuadrados  $(\hat{\sigma}^{(1)}, \hat{\gamma}^{(1)})$  estimaciones de  $(\sigma, \gamma)$ , una vez restado  $\mathbb{E}[\log(Z^2)] \approx -1,270362845$  de cada lado del modelo (4.55), donde  $Z$  es una variable aleatoria normal estándar.

**Etapa III**<sup>20</sup>: Sustituir  $(\hat{\sigma}^{(1)}, \hat{\gamma}^{(1)})$  en el modelo (4.54) y obtener por mínimos cuadrados ponderados  $(\hat{\alpha}^{(2)}, \hat{\beta}^{(2)})$ , para tener en cuenta la heteroscedasticidad en (4.54), nuevas estimaciones de  $(\alpha, \beta)$ . Finalmente, usando esta nueva estimación, obtener  $(\hat{\sigma}^{(2)}, \hat{\gamma}^{(2)})$  estimaciones de  $(\sigma, \gamma)$  como en la etapa II.

Los resultados de aplicar el procedimiento anterior a los datos semanales de los tipos de interés interbancarios a un mes en España aparecen recogidos en la Tabla 4.9.

	$\alpha$	$\beta$	$\sigma$	$\gamma$
Etapas I y II	0,001808	-0,101022	0,102850	1,135093
Etapa III	0,002355	-0,123824	0,104994	1,139133

**Tabla 4.9:** *Estimación del modelo CKLS usando datos semanales de los tipos de interés interbancarios a un mes en España y siguiendo el método propuesto en Fan y Zhang (2003)[41].*

Finalmente, para poder valorar activos derivados con un modelo totalmente paramétrico, necesitamos especificar y estimar una forma paramétrica para el precio del riesgo de mercado  $\lambda(r)$ . A la vista de la estimación no paramétrica de la Figura 4.18 no parece adecuado suponer, como por ejemplo hacen Chan *et al* (1992)[19], que es nulo. Tras

---

<sup>20</sup>Fan y Zhan (2003)[41] proponen esta etapa como opcional al no observar, con sus datos simulados, diferencias apreciables con los resultados obtenidos siguiendo sólo las dos primeras etapas.

considerar algunas formas funcionales propuestas en la literatura (nula, constante, proporcional a  $r, \dots$ ), la que ofrece un mejor resultado al comparar la valoración de bonos cupón cero con el caso no paramétrico es  $\lambda(r) = q \cdot r$ . Una estimación basada en la aproximación mínimo-cuadrática a los valores obtenidos en el caso no paramétrico es  $\hat{\lambda}(r) = -0,131863 \cdot r$ .

Una vez estimado el modelo paramétrico se valoran los mismos bonos cupón cero que antes con el modelo no paramétrico y mediante el mismo procedimiento. Los resultados obtenidos aparecen recogidos en la Tabla 4.10, en la cual se han resaltado en **negrita** los precios que difieren menos del 1% del correspondiente no paramétrico, en **azul** los que difieren entre un 1% y un 2% y en **rojo** los que difieren más de un 2%.

$T \setminus r$	3%	6%	9%	12%
1	<b>0,9692</b>	<b>0,9404</b>	<b>0,9125</b>	<b>0,8855</b>
2	<b>0,9369</b>	<b>0,8819</b>	<b>0,8302</b>	<b>0,7815</b>
3	<b>0,9032</b>	<b>0,8246</b>	<b>0,7530</b>	<b>0,6875</b>
4	<b>0,8685</b>	<b>0,7690</b>	<b>0,6808</b>	<b>0,6031</b>
5	<b>0,8330</b>	<b>0,7149</b>	<b>0,6139</b>	<b>0,5272</b>

**Tabla 4.10:** Precios estimados de bonos cupón cero usando el modelo CKLS y con  $\hat{\lambda}(r) = -0,131863 \cdot r$ . Se sigue el método Monte Carlo, generando, para cada precio, 2000 trayectorias con 100 datos diarios mediante la aproximación de Milstein y reduciendo la variabilidad de las estimaciones mediante el método de las variaciones antitéticas.

Comparando la Tabla 4.7 de la valoración no paramétrica de los bonos cupón cero con la Tabla 4.10 de la valoración de tales bonos con el modelo paramétrico se pueden hacer prácticamente los mismos comentarios que antes al comparar la valoración no paramétrica con la que se obtiene en la Tabla 4.8 de imponer sólo una deriva lineal. Concretamente la adopción del modelo paramétrico CKLS lleva a precios de bonos cupón cero bastante similares a los del modelo no paramétrico general, más aún si el tipo de interés inicial

es bajo o medio (no superior al 9%) y el periodo de vencimiento no supera los 3 años. Las diferencias aumentan con el nivel inicial de los tipos de interés y con el periodo del vencimiento del bono cupón cero.

En consecuencia, teniendo todo lo anterior en cuenta, y recordando además las conclusiones de aplicar el test  $J_n$  en §4.4 a la deriva y a la volatilidad, se tiene que el modelo paramétrico CKLS es adecuado para representar la evolución de los tipos de interés a corto plazo en España. Su estimación tiene un comportamiento bastante similar a la estimación no paramétrica del modelo general planteado en la ecuación (4.1), al menos para tipos de interés bajos o medios y periodos de vencimiento no muy alejados en el tiempo. Además, esto mismo es cierto para el modelo planteado en la ecuación (4.1) imponiéndole una deriva lineal.



# Capítulo 5

## Conclusiones

### 5.1. Conclusiones

Este trabajo termina con una breve exposición de las principales conclusiones que se han obtenido y proponiendo, desde el estudio aquí realizado, la continuidad de futuras líneas de investigación.

La aportación fundamental, aunque no la única, de esta tesis es el estudio empírico de los tipos de interés a corto plazo en España llevado a cabo en el capítulo 4. Es importante señalar que el trabajo de recopilación, síntesis y presentación de las ideas fundamentales realizado en los tres primeros capítulos tiene por objetivo presentar un trabajo lo más autocontenido posible respecto al cálculo estocástico y su uso en la teoría general de modelos de mercados financieros. Queremos destacar también el diseño de una batería de

programas, presentados en el Apéndice, los cuales permiten realizar los cálculos específicos a los que nos enfrentamos a la hora de realizar el estudio empírico del capítulo 4.

La gran cantidad de literatura existente respecto al estudio de los tipos de interés deja claro, por una parte, su enorme importancia en Economía y, por otra parte, la ausencia de consenso en la elección de un modelo adecuado. Los trabajos empíricos (Dybvig (1989)[36], Litterman y Scheinkman (1991)[75]) muestran que es razonable usar únicamente el tipo de interés a corto plazo como variable endógena para la descripción de los movimientos de toda la estructura temporal de los tipos de interés, lo que justifica el uso de los modelos unifactoriales. En este marco de trabajo y con el objetivo de evitar una mala especificación del modelo, han surgido (Aït-Sahalia (1996)[3], Jiang y Knight (1997)[65], Stanton (1997)[94]) técnicas de estimación no paramétricas que permiten hacer un planteamiento general de la modelización. En consecuencia, el punto de partida en el estudio empírico que se ha realizado es una ecuación diferencial estocástica homogénea (evitando así el gran esfuerzo computacional que requiere el planteamiento y constante reestimación que requiere el modelo no homogéneo) unidimensional para recoger los movimientos de los tipos de interés a corto plazo en España. Una vez planteado el modelo, y realizado un primer análisis de los datos, se pasa a estimar sus parámetros mediante la metodología propuesta en Stanton (1997)[94], ya que ésta es la más general. De este estudio inicial de los datos y de la estimación de los parámetros cabe destacar las siguientes conclusiones:

- La estimación de la densidad estacionaria, presentada en la Figura 4.8, no invita a pensar que los modelos simples, como por ejemplo los gaussianos, sean adecuados para representar la evolución de los tipos de interés a corto plazo en España.
- Las aproximaciones de los tres primeros órdenes tanto de la deriva como de la volatilidad (Figuras 4.12 y 4.13) son muy similares, que es lo que se espera que ocurra cuando los errores de aproximación son pequeños.

- Las bandas de confianza de las aproximaciones de orden uno a la deriva y la volatilidad (que son las que se dan como definitivas teniendo en cuenta el enorme crecimiento de la varianza asintótica de los estimadores propuestos con el orden de la aproximación, Fan y Zhang (2003)[41]) son más estrechas para tipos de interés medios y bajos y más anchas, reflejando una menor precisión en la estimación, para tipos altos.
- La deriva estimada (Figura 4.14) exhibe un comportamiento no lineal, con una leve reversión a la media para tipos de interés medios y bajos, cayendo fuertemente cuando los tipos se sitúan por encima del 15 %.
- La volatilidad estimada (Figura 4.14) aumenta con el nivel de los tipos de interés, estando en consonancia con la planteada en el modelo paramétrico de Chan *et al.*(1992) [19].

La forma particular de la deriva estimada, es similar a la de otros trabajos donde se usan diferentes conjuntos de datos para modelizar la dinámica de los tipos de interés a corto plazo en diversos países; entre ellos, destacamos a Stanton (1997)[94], Jiang (1998)[64], Fernández (2001)[45] y Corzo y Gómez (2001)[22]. Esto ha motivado un análisis más profundo de la situación, no estando actualmente cerrado el problema de la afirmación de la no linealidad de la verdadera deriva a partir de la no linealidad de su estimación (Chapman y Pearson (2000)[20]). Con el objetivo de avanzar algo en esta controversia, consideramos la estimación de la deriva mediante la regresión lineal local (método que extiende al de Nadaraya-Watson y que no presenta algunos de sus inconvenientes). Sin embargo, la estimación obtenida es completamente similar a la antes comentada (Figura 4.16). En esta situación, y dado que lo que se pretende investigar es si hay evidencia suficiente en los datos como para rechazar una deriva lineal, se considera oportuno aplicar algún contraste no paramétrico de los existentes en la literatura. A este respecto,

encontramos que se han propuesto recientemente las versiones bootstrap de dos test de especificación de modelos no paramétricos: El test GLR y el test  $J_n$ . Teniendo en cuenta que el test GLR no proporciona tamaños estables sobre un rango de anchuras de ventana y que la justificación rigurosa de que se puede aplicar a datos de series temporales aún no ha sido establecida, decidimos aplicar el test  $J_n$ , el cual no presenta estos inconvenientes (Kim y Wang (2006)[68]). La conclusión principal que se alcanza tras la aplicación del test es que no existe una evidencia contundente para rechazar la linealidad de la deriva (a un nivel de significatividad del 3% el test no conduce a rechazar la hipótesis de linealidad en la deriva para ninguna anchura de ventana en un amplio rango, Tabla 4.5). En este contexto, se aplica el test para contrastar dos formas funcionales para la volatilidad, llegando a los siguientes resultados: Hay evidencia para rechazar una volatilidad constante y se acepta que la volatilidad es potencial, como la propuesta en el modelo de Chan *et al.*(1992)[19], Tabla 4.6.

Finalmente, con el objetivo de usar el modelo para valorar activos derivados de los tipos de interés (bonos concretamente), es necesario previamente aproximar el precio del riesgo de mercado. Esto se hace mediante el enfoque no paramétrico propuesto en Jiang (1998)[64], llegando a la estimación que aparece en la Figura 4.18, sobre la que se pueden destacar las siguientes características: Al igual que para la deriva y la volatilidad, la banda de confianza es más ancha conforme los tipos de interés son más altos. La estimación es negativa (correspondiente a una prima positiva por soportar el riesgo de tipo de interés), siendo cercana a cero para tipos de interés medios y bajos y decreciendo fuertemente cuando éstos son altos, llevando a una prima de riesgo mayor cuando los tipos de interés son más altos (y más volátiles).

Una vez estimados la deriva, la volatilidad y el precio del riesgo de mercado para el

modelo propuesto, se usa el método Monte Carlo para valorar bonos cupón cero, esto requiere simular trayectorias, bajo la medida equivalente de martingalas, del proceso que describe los movimientos de los tipos de interés a corto plazo en España. Además, con el objetivo de disminuir la varianza de las estimaciones, se aplica el método de las variaciones antitéticas. En este contexto, y con el propósito de ahondar en las consecuencias de una especificación lineal de la deriva y de un modelo completamente paramétrico, se valoran bonos para diferentes niveles de los tipos de interés y vencimientos y para tres modelos diferentes: El modelo completamente no paramétrico, el modelo anterior con la especificación de una deriva lineal y el modelo completamente paramétrico propuesto en Chan *et al.*(1992)[19]. La principal conclusión a la que se llega es que los tres modelos llevan a precios muy similares cuando el nivel inicial de los tipos de interés es medio o bajo (inferior al 9 %) y el vencimiento es inferior a 3 años (Tablas 4.7, 4.8 y 4.10). Además, en general, las diferencias aumentan con el nivel inicial de los tipos de interés y el periodo de vencimiento.

Como consecuencia de los resultados del test  $J_n$  y de la valoración de bonos, una conclusión destacable es que el modelo paramétrico de Chan *et al.*(1992)[19] parece adecuado para describir los movimientos de los tipos de interés a corto plazo en España, al menos cuando éstos se mantienen en niveles medios o bajos (como ocurre ahora).

En cuanto a las futuras líneas de investigación que pretendemos desarrollar, manteniendo siempre la generalidad que ofrece el planteamiento no paramétrico de los modelos, efectuando comparaciones con otros modelos existentes (paramétricos y no paramétricos) y tomando como punto de partida el estudio realizado en este trabajo, destacamos las siguientes:

- Llevar a cabo un estudio similar en modelos más generales (modelos multifactoriales y modelos con trayectorias discontinuas).
- Usar los modelos para valorar otros activos derivados de los tipos de interés y aplicarlos también al mercado de renta variable.

# Apéndice A

## Programas Elaborados en MATLAB

En este apéndice mostramos los programas que hemos elaborado para llevar a cabo los cálculos, gráficas y ejemplos que se realizan en este trabajo. Todos ellos se han construido usando como soporte informático el paquete MATLAB en su versión 7.2 (R2006a), bajo el sistema operativo Windows XP.

Básicamente, en este apéndice exponemos las líneas de código de cada programa, ya que en cada uno de ellos se describe, al principio en líneas de comentarios no ejecutables, su objetivo, funcionamiento e instrucciones de uso.

## A.1. El programa bono\_ckls

```
% Esta función estima, por el método Monte Carlo, el precio (p) y su
% desviación estándar (ee), en un instante inicial t0=0 de un bono cupón
% cero cuando los tipos de interés a corto plazo siguen el modelo CKLS:
%
%           dr=(a+br)dt+s(r^g)dW.
%
% Se supone un precio del riesgo de mercado proporcional a r (l=l0·r), esto
% se puede modificar fácilmente a otras formas paramétricas. Con objeto de
% reducir la varianza del estimador se usa el método de las variaciones
% antitéticas.
%
% Argumentos de entrada:
% (param): vector 4x1 con los parámetros del modelo (a),(b),(s),(g) en
%         este orden.
% (l0): Para el precio del riesgo de mercado.
% (T): Vencimiento (en años) del bono cupón cero.
% (ns): Número de trayectorias simuladas, del proceso r bajo la medida de
%       riesgo neutral, siguiendo el esquema de Milstein (debe ser un numero
%       par).
% (r0): Valor inicial del tipo de interés a corto plazo.
% (N): Número de puntos (incluido r0) que componen la partición de [0,T]
%       donde se simula el proceso r.
%
% Uso: [p ee] = bono_ckls(param,l0,T,ns,r0,N)
```



```

function [p ee] = bono_ckls(param,l0,T,ns,r0,N)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Control de los argumentos de entrada %%%%%%%%%
if (nargin ~= 6); error('número de argumentos erróneo en
bono_ckls');end

[k basura]=size(param); if ((basura ~= 1) | k ~= 4);
    error('param debe ser un vector columna de orden 4 en bono_ckls')
end

if (isscalar(T)~=1|isscalar(r0)~=1|isscalar(l0)~=1|T<=0 |r0<=0...
    | isscalar(N)~=1 | isscalar(ns)~=1);
    error('T,r0,l0,N y ns deben ser escalares con T>0 y r0>0 en bono_ckls')
end

if (ns ~= floor(ns) | N ~= floor(N) | N<=0 | ns<=0 | ns/2 ~=
floor(ns/2));
    error('N y ns deben ser naturales con ns par en bono_ckls')
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Programa %%%%%%%%%

d=T/(N-1); % diámetro de la rejilla.
rd=sqrt(d);

```

```

S=zeros(ns,1); rini=r0*ones(ns,1);

for j=1:N-2
    mu=param(1)+param(2)*rini;          % deriva
    sigma=param(3)*(rini.^param(4)); % volatilidad
    lambda=l0*rini;                      % precio del riesgo de mercado

    z=normrnd(0,rd,ns/2,1);
    z=[z;-z];

    % etapa método Milstein
    rfin=rini+d.*(mu-lambda)+sigma.*z+0.5*sigma.^2.*(z.^2-d);

    S=S+rfin; rini=rfin;
end

p=exp(-d*S); ee=std(p); p=mean(p);

```

## A.2. El programa bono\_tabla

```

% Esta función estima el precio de un bono cupón cero (p) y su desviación
% típica (ee) a partir de una tabla donde, por columnas, aparecen
% discretizaciones de la deriva, volatilidad y precio del riesgo de mercado
% estimados del modelo de los tipos de interés a corto plazo. Se usa la
% simulación Monte Carlo y el método de las variaciones antitéticas para

```

## A.2. El programa bono\_tabla

---

```
% reducir la varianza del estimador. Se aplica el esquema de Milstein para
% generar las trayectorias. En cada trayectoria se toman 100 puntos
% diarios. Se aplica interpolación lineal para las estimaciones que no
% aparecen en la tabla.
%
% Argumentos de entrada:
% (tabla): tabla de 5 columnas. En cada columna deben aparecer las
%         las siguientes cantidades en el orden que se indica: tipos a
%         corto plazo (equiespaciados), deriva, deriva si se fuerza a que
%         sea lineal, volatilidad, precio del riesgo de mercado.
% (r0): Valor inicial del tipo a corto plazo.
% (T): Tiempo a vencimiento, en años.
% (opc): si es 1 se calcula bajo una deriva lineal, si es 0 se calcula sin
%         esa restricción.
% (ns): Número de trayectorias simuladas para cada precio (debe ser par).
%
% Uso:          [p ee]=bono_tabla(tabla,r0,T,opc,ns)

function [p ee]=bono_tabla(tabla,r0,T,opc,ns)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Control de los argumentos de entrada %%%%%%%%%
if (nargin ~= 5); error('número de argumentos erróneo en
bono_ckls');end

[basura temp]=size(tabla); if (temp ~= 5);
    error('tabla debe tener 5 columnas en bono_tabla')
end
```

```

if (min(diff(tabla(:,1)))-max(diff(tabla(:,1)))> 1e-015);
    error('la primera columna debe ser equiespaciada en bono_tabla')
end

if (isscalar(T)~=1|isscalar(r0)~=1| T<=0 | r0<=0 | isscalar(ns)~=1);
    error('T,r0 y ns deben ser escalares con T>0 y r0>0 en bono_tabla')
end

if (ns ~= floor(ns) | ns<=0 | ns/2 ~= floor(ns/2));
    error('ns debe ser natural y par en bono_tabla')
end

if (opc ~= 0 & opc ~= 1);
    error('opc debe ser 0 ó 1 en bono_tabla')
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Programa %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

N=100*250*T;    % Puntos en cada trayectoria
incr=tabla(2,1)-tabla(1,1); % incremento entre dos r consec.
primer=tabla(1,1); % primer r

d=T/(N-1); rd=sqrt(d); S=zeros(ns,1);

rini=r0*ones(ns,1);

```

```
for i=1:N-2
    z=normrnd(0,rd,ns/2,1);
    z=[z;-z];

    posicion=1+(1/incr)*(rini-primer);

    I=floor(posicion); % se interpola linealmente en la tabla
    D=posicion-floor(posicion);

    if opc==0 mu=tabla(I,2)+(tabla(I+1,2)-tabla(I,2)).*D;
    else mu=tabla(I,2+opc)+(tabla(I+1,2+opc)-tabla(I,2+opc)).*D; end
    lambda=tabla(I,5)+(tabla(I+1,5)-tabla(I,5)).*D;
    sigma=tabla(I,4)+(tabla(I+1,4)-tabla(I,4)).*D;

    % etapa método Milstein
    rfin=rini+d.*(mu-lambda)+sigma.*z+0.5*sigma.^2.*(z.^2-d);

    S=S+rfin; rini=rfin;
end
p=exp(-d*S); ee=std(p); p=mean(p);
```

### A.3. El programa BonoCuponCero

% Esta función estima, de forma no paramétrica, el precio (p), y su

---

```

% desviación típica (ee) de un bono cupón cero con vencimiento en el
% instante (T) años. Para ello, usando la discretización de Milstein, se
% simula un número par (ns), de trayectorias del proceso de los tipos de
% interés instantáneos (r1) bajo la medida equivalente de martingalas. Para
% estimar el precio del riesgo de mercado, siguiendo a Jiang (1998), se
% necesitan los rendimientos de dos activos dependientes de los tipos de
% interés: (r1) con plazo (t1) años y (r2) con plazo (t2) años. El primer
% activo se usa también como proxy de los tipos instantáneos. Todas las
% estimaciones necesarias para obtener aproximaciones de los parámetros de
% este proceso se realizan por el método del núcleo con una anchura de
% ventana común (h). (delta) es el tiempo entre dos observaciones
% consecutivas de los datos r1 y r2. (r0) es el valor inicial del tipo de
% interés. (N) es el número de puntos (incluido r0) que componen la
% partición de [0,T] donde se simula el proceso, y (opc) permite usar una
% deriva de los tipos a corto lineal (opc=1), en el caso (opc=0) se estima
% sin restricciones. Para reducir la varianza del estimador se usa el
% método de las variaciones antitéticas.

%
% Uso: [p ee] = BonoCuponCero(T,ns,r1,t1,r2,t2,h,delta,r0,N,opc)
%
function [p ee] = BonoCuponCero(T,ns,r1,t1,r2,t2,h,delta,r0,N,opc)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Control de los argumentos de entrada %%%%%%%%%%%%%%%
if (nargin ~= 11);
    error('número de argumentos erróneo en BonoCuponCero');end

```

### A.3. El programa BonoCuponCero

---

```
if (sum(size(r1)~=size(r2))~=0 | size(r1,2)~=1);
    display('r1 y r2 deben ser vectores columna de la misma');
    display('dimensión en BonoCuponCero');error(' ');
end

if (isscalar(T)~=1 | isscalar(r0)~=1 | isscalar(delta)~=1 |...
    isscalar(N)~=1 | isscalar(ns)~=1 | isscalar(t1)~=1 |...
    isscalar(t2)~=1 | isscalar(h)~=1 | T<=0 | r0<=0 | delta<=0 |...
    t1<=0 | t2<=0 | h<=0);
    display('T,r0,delta,N,ns,t1,t2 deben ser escalares');
    display('con T>0,r0>0,delta>0,t1>0,t2>0 en BonoCuponCero');
    error(' ');
end

if (ns~=floor(ns) | N~=floor(N) | N<=0 | ns<=0 | ns/2~=floor(ns/2));
    error('N y ns deben ser naturales con ns par en BonoCuponCero')
end

if (opc ~= 0 & opc ~= 1);
    error('opc debe ser 0 ó 1 en BonoCuponCero')
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Programa %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
d=T/(N-1); % diámetro de la rejilla.
rd=sqrt(d);

y1=(r1(2:end)-r1(1:end-1))/delta; yv1=y1.*(r1(2:end)-r1(1:end-1));
```

```
y2=(r2(2:end)-r2(1:end-1))/delta; yv2=y2.*(r2(2:end)-r2(1:end-1));
x=r1(1:end-1); yd=r1-r2;
```

```
if opc==1; X=[ones(length(x),1) x]; b=regress(y1,X); end
```

```
S=zeros(ns,1); rini=r0*ones(ns,1);
```

```
for j=1:N-2
```

```
    % deriva lineal
```

```
    if opc==1 mu=b'*[ones(1,ns); rini']; end
```

```
    for i=1:ns
```

```
        k=normpdf((x-rini(i))/h);
```

```
        kt=[k ; normpdf((r1(end)-rini(i))/h)];
```

```
        w=k/sum(k);
```

```
        wt=kt/sum(kt);
```

```
    % deriva primer activo
```

```
    a1(i)=sum(y1.*w);
```

```
    % volatilidad primer activo
```

```
    sigma(i)=sqrt(sum(yv1.*w));
```

```
    % deriva segundo activo
```

```
    a2(i)=sum(y2.*w);
```



```
% volatilidad segundo activo
k2(i)=sqrt(sum(yv2.*w));

% diferencia rendimientos
yd_mu(i)=sum(yd.*wt);

end

% precio del riesgo de mercado
num=yd_mu+0.5.*(((t1^2)*(sigma.^2))-((t2^2)*(k2.^2)))+((t2*a2)-(t1*a1));
den=(t2*k2)-(t1*sigma);
lambda=sigma.*(num./den);

z=normrnd(0,rd,ns/2,1);
z=[z;-z];

if opc==0 mu=a1; end
% etapa método Milstein
rfin=rini'+d.*(mu-lambda)+sigma.*z'+0.5*sigma.^2.*(z'.^2-d);

S=S+rfin'; rini=rfin';

end

p=exp(-d*S); ee=std(p); p=mean(p);
```

## A.4. El programa cambrow

```
function [t,w]=cambrow(n);  
%  
% uso:          [t,w]=cambrow(n);  
%  
% Dado un vector columna n, de dimensión k, de números naturales  
% n=(n(1),...,n(k)', se generan k trayectorias aproximadas del movimiento  
% browniano unidimensional estándar en el intervalo [0,1] basadas en el  
% TCLF o principio de invarianza de Donsker. Todas las aproximaciones se  
% construyen con una misma muestra de tamaño max{n(i)} de variables  
% aleatorias independientes normales estándar. Las cantidades devueltas por  
% la función son:  
% (t) que es un array de k celdas donde en cada celda se almacena los  
% diferentes instantes de tiempo de una trayectoria, y  
% (w) que es otro array de k celdas donde en cada celda se almacena los  
% valores de una trayectoria browniana aproximada para los correspondientes  
% instantes t.  
%  
% A cada trayectoria se le añade además el punto (0,0), con lo que la  
% trayectoria i-ésima tiene n(i)+1 puntos, i=1,...,k.  
%  
% Dada una trayectoria en [0,1], basta escalar el tiempo por el factor T y  
% escalar la trayectoria por el factor sqrt(T) para obtener una  
% trayectoria en [0,T].  
%
```

#### A.4. El programa cambrow

---

```
% Por ejemplo:
% [T B]=cambrow([10000;50000]); devuelve (y dibuja) dos trayectorias del MB
% construidas a partir de una MAS de tamaño 50000 de una N(0,1):
% (T{1},B{1}) formada por 10000 puntos, añadiendo después el (0,0) y
% (T{2},B{2}) formada por 50000 puntos, añadiendo después el (0,0).
%
% La modificación (10*T{1},sqrt(10)*B{1}) sería una trayectoria simulada
% del MB estándar formada por 10001 puntos equiespaciados en el intervalo [0,10].
% Y análogamente con la otra trayectoria.
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Control de los argumentos de entrada %%%%%%%%%%%%%%%
%
%comprobando el número de argumentos en la función
if (nargin ~= 1);%
error('número de argumentos erróneo en cambrow');end
%
%comprobando que n es un vector de números naturales
[k basura]=size(n);%
if ((basura ~= 1) | (n ~= floor(n)) | (min(n)<=0));
    error('n debe ser un vector columna de números naturales en cambrow')
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Programa %%%%%%%%%%%%%%%
nmax=max(n);
e=normrnd(0,1,nmax,1); % Generación de nmax variables iid normales estándar

for i=1:k
```

```

t{i}(1)=1/n(i);
w{i}(1)=e(1)/sqrt(n(i));
for j=2:n(i)
    w{i}(j)=w{i}(j-1)+(e(j)/sqrt(n(i)));
    t{i}(j)=j/n(i);
end
t{i}=[0 t{i}];
w{i}=[0 w{i}];
end

% dibujo de las trayectorias
hold on;
for i=1:k
    plot(t{i},w{i},'k-');
end
hold off;

```

## A.5. El programa ckl\_s\_fan\_zhang

```

% Dados un vector columna (r) conteniendo los datos de los tipos de interés
% instantáneos libres de riesgo y un escalar (delta) siendo el tiempo entre
% dos observaciones consecutivas, esta función estima el modelo CKLS
%
%          dr=(alpha+beta*r)+sigma*(r^gamma),
% usando el procedimiento propuesto en Fan y Zhang (2003).
%

```

## A.5. El programa ckls\_fan\_zhang

---

```
% Uso: [alpha beta sigma gamma]= ckls_fan_zhang(r,delta)

function [alpha beta sigma gamma]= ckls_fan_zhang(r,delta)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Control de los argumentos de entrada %%%%%%%%%%%%%
if (nargin ~= 2);
    error('error de número de argumentos en ckls_fan_zhang');end

[n temp] = size(r); if (temp ~= 1);
    error('r debe ser un vector columna en ckls_fan_zhang');end;

if (isscalar(delta)~=1 | delta<=0);
    error('delta debe un escalar positivo en ckls_fan_zhang'); end;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Programa %%%%%%%%%%%%%

% variables
y=r(2:end,1)-r(1:end-1,1); x=r(1:end-1,1);%
X=[ones(size(x,1),1) x];%
logX=[ones(size(x,1),1) log(x.*x)];

% etapa 1
etapa1=regress(y,X.*delta);%
alpha=etapa1(1);%
beta=etapa1(2);

% etapa 2
```

```

e=(y-(alpha+beta.*x).*delta)./sqrt(delta);
etapa2=regress(log(e.*e)+1.270362845,logX);%
sigma2=exp(etapa2(1));%
gamma=etapa2(2);

% etapa 3 (parte 1)
w=x.^gamma;%
etapa3a=regress(y./w,(X.*delta)./[w w]);%
alpha(2)=etapa3a(1);%
beta(2)=etapa3a(2);%

% etapa 3 (parte 2)
e=(y-(alpha(2)+beta(2).*x).*delta)./sqrt(delta);
etapa3b=regress(log(e.*e)+1.270362845,logX);
sigma2(2)=exp(etapa3b(1));%
gamma(2)=etapa3b(2);

sigma=sqrt(sigma2);

% impresión en pantalla de los resultados
display('          alpha          beta          sigma          gamma');
display('-----')
fprintf('Estim. inicial  %9.6f  %9.6f  %9.6f  %9.6f\n',...
        alpha(1),beta(1),sigma(1),gamma(1));
fprintf('Estim. final    %9.6f  %9.6f  %9.6f  %9.6f\n',...
        alpha(2),beta(2),sigma(2),gamma(2));

```

## A.6. El programa densidadMBB

```
% Dado un vector de datos x, esta función calcula la densidad estimada
% (fest) en los puntos xx (N puntos equiespaciados en el intervalo
% [min,max]), usando un kernel Gaussiano y con anchura de ventana h, y los
% límites inferior (inf) y superior (sup) de un intervalo de confianza a
% nivel 90%. La construcción del intervalo se basa en aproximar a la
% distribución muestral del estimador kernel de la densidad mediante el
% método bootstrap de los bloques móviles (Moving Block Bootstrap). El
% número de datos en cada bloque es (b) y la cantidad de replicaciones
% bootstrap (nb).
%
% Uso: [xx,inf,fest,sup]= densidadMBB(x,min,max,N,h,b,nb)

function [xx,inf,fest,sup]= densidadMBB(x,min,max,N,h,b,nb)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Control de los argumentos de entrada %%%%%%%%%%%%%%%
if (nargin ~= 7);
    error('error de número de argumentos en densidadMBB');end

[n temp] = size(x);%
if (temp ~= 1);%
error('x debe ser un vector columna en densidadMBB');end;

if (size(min,1)~=1 | size(min,2)~=1 | size(max,1)~=1 | ...
    size(max,2)~=1 | max<min);
```

```

error('min y max deben ser escalares con min<=max en densidadMBB'); end;

if checktyp('N',N,'int',mfilename);return;end%
if N<1 error('b debe ser un natural en densidadMBB');end;

if (h<=0); error('la ventana debe ser positiva en densidadMBB');
end;

if checktyp('b',b,'int',mfilename);return;end%
if b<1 error('b debe ser un natural en densidadMBB');end;

if checktyp('nb',nb,'int',mfilename);return;end%
if nb<1 error('nb debe ser un natural en densidadMBB');end;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Programa %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% puntos donde se obtendrá f estimada y los intervalos de confianza.
xx=linspace(min,max,N)';

temp=n*h; k=round(n/b); l=k*b;
xb=zeros(1,1); % vector donde se va a guardar cada muestra bootstrap
tempb=l*h;
a=0.10; % Intervalos al 90% de confianza

for i=1:N % cálculos para el punto xx(i)
    fest(i,1)=sum(normpdf((xx(i)-x)/h))/temp; % densidad estimada en xx(i)

```



```
% Obtención de la banda de confianza puntual Bootstrap.

% Elección de los k bloques móviles que compondrán la j-ésima
% muestra bootstrap

for j=1:nb
    bloque=unidrnd(n-b+1,1,k);

% Construcción de la j-ésima muestra bootstrap con los bloques elegidos
    for m=1:k
        xb((m-1)*b+1:m*b)=x(bloque(m):bloque(m)+b-1);
    end

% Estimación de la densidad con la muestra bootstrap
    festb(j,1)=sum(normpdf((xx(i)-xb)/h))/tempb;
    home;fprintf('%2.2f, %2.2f.\n',100*i/N,j*100/nb); %control tiempo
end

% Intervalo de confianza basado en los percentiles bootstrap
inf(i,1)=prctile(festb,100*a/2);
sup(i,1)=prctile(festb,100*(1-a/2));

end; plot(xx,fest','k-',xx,inf','k:',xx,sup,'k:');
```

## A.7. El programa generaCIR

```

% Este programa simula una trayectoria de (n) puntos del modelo CIR, el
% cual viene dado por la ecuación diferencial estocástica
%
%           dr=k(theta-r)dt+sigma*sqrt(r)*dW,
% usando que la densidad condicionada de r es proporcional a una
% distribución ji-cuadrado no centrada. (k, theta, sigma) son los
% parámetros del modelo y (delta) el paso de tiempo entre dos observaciones
% consecutivas generadas (con delta=1 se generan observaciones anuales).
%
% Uso:      [t, r] = generaCIR(n,k,theta,sigma,delta)

function [t, r] = generaCIR(n,k,theta,sigma,delta)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Control de los argumentos de entrada %%%%%%%%%%%%%%%
if (nargin ~= 5);
    error('error de número de argumentos en generaCIR');end

if (isscalar(n)~=1 | isscalar(k)~=1 | isscalar(theta)~=1 | ...
    isscalar(sigma)~=1 | isscalar(delta)~=1);
    error('todos las entradas deben ser escalares en generaCIR')
end

if checktyp('n',n,'int',mfilename);return;end%
if n<1 error('n debe ser un natural en generaCIR');end;

```

```

if (sigma<=0 | delta<=0);
    error('sigma y delta deben ser positivos en generaCIR'); end;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Programa %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

k4=4*k; d=(k4*theta)/(sigma^2);

% el primer valor se toma de la distribución asintótica
r(1)=gamrnd(d/2,(sigma^2)/(2*k));

temp1=(sigma^2)*(exp(k*delta)-1);
temp2=((sigma^2)*(1-exp(-k*delta)))/(k4); for i=2:n
lambda=(k4*r(i-1))/temp1; r(i)=ncx2rnd(d,lambda)*temp2; end
t=[0:delta:(n-1)*delta]'; r=r'; % devolviendo vectores columna
plot(t,r,'k-');

```

## A.8. El programa histograma

```

function []=histograma(y,y0,h)
% Dados un vector de datos (y), un origen (y0) y una anchura para los
% intervalos (h), esta función dibuja el histograma a partir de las
% densidades de frecuencias relativas de cada uno de los intervalos
% [y0,y0+h), [y0+h,y0+2h), [y0+2h,y0+3h),...
%
% Uso:          histograma(y,y0,h)

```

```
%%%%%%%%%% Control de los argumentos de entrada %%%%%%%%%%%
if (nargin ~= 3);%
error('número de argumentos erróneo en histograma');end

if (size(y,2) ~= 1);
    error('y debe ser un vector columna en histograma');end

if (isscalar(y0)~=1 | isscalar(h)~=1 | h<=0);
    error('y0 y h deber ser escalares con h>0 en histograma ');
end

if (y0>min(y));
    error('y0 debe ser menor o igual que el menor dato en histograma');
end

%%%%%%%%%% Programa %%%%%%%%%%%

n=size(y,1); % cantidad de datos

% Encontrando en número de ventanas M
M=1; xf=y0+M*h; while xf<=max(y)
    M=M+1; xf=y0+M*h;
end

x=y0:h:xf; % Puntos x que delimitan las ventanas (=intervalos)
```

```
d=[0; histc(y,x)/(n*h)]; % densidades de frecuencia de cada ventana

% Construcción de los puntos (xd,dh) para dibujar el histograma
for i=2:2:2*(M+1)
    xd(i-1)=x(i/2); dh(i-1)=d(i/2);
    xd(i)=x(i/2); dh(i)=d(i/2+1);
end

% dibujo del histograma
plot(xd,dh,'k-');
```

## A.9. El programa LLR

```
% En el contexto de la regresión no paramétrica simple de Y sobre X, dados
% dos vectores de datos (x) e (y) y una anchura de ventana (h), esta
% función proporciona la curva de regresión estimada (mhat) en (N) puntos
% equiespaciados en el intervalo [xmin,xmax] (vector xx). Si (opc=0) se
% obtiene el estimador de Nadaraya-Watson y si (opc=1) se aplica la
% regresión local lineal.
%
% Uso:      [xx mhat] = LLR(y,x,opc,xmin,xmax,N,h)
%
function [xx mhat] = LLR(y,x,opc,xmin,xmax,N,h)
```

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Control de los argumentos de entrada %%%%%%%%%
if (nargin ~= 7);
    error('error de número de argumentos en LLR');end

if (sum(size(x)~=size(y))~=0 | size(x,2)~=1);%
error('x e y deben ser vectores columna de igual dimensión en LLR');
end;

if (opc ~= 0 & opc~= 1);
    error('opc debe ser 0 (NW) ó 1 (LLR) en LLR');end;

if (isscalar(xmax)~=1 | isscalar(xmin)~=1 | xmax<xmin);%
error('xmin y xmax deben ser escalares con xmin<=xmax en LLR'); end;

if checktyp('N',N,'int',mfilename);return;end%
if (N<1 | isscalar(N)~=1)
    error('N debe ser un escalar natural en LLR');end;

if (isscalar(h)~=1 | h<=0 );
    error('la ventana debe ser un escalar positivo en LLR'); end;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Programa %%%%%%%%%
xx=linspace(xmin,xmax,N)'; % vector donde se estima la regresión

temp=zeros(N,1);
y = meshgrid(y,temp)'; % Matriz en cuyas N columnas hay copias de y.

```

```
[difs temp] = meshgrid(xx,x);
difs = (temp - difs); % matriz de diferencias xx(i)-x(j)

difsh = difs/h;

% matriz de valores kernels (gaussiano)
kernvals = exp(-((difsh.*difsh)/2)) ;

s0 = sum(kernvals); t0 = sum(kernvals.*y);
if opc==0; mhat=t0./s0; % estimador Nadaraya-Watson
else % regresión lineal local
    temp = difs.*kernvals;
    s1 = sum(temp);
    t1 = sum(temp.*y);
    temp = difs.*temp;
    s2 = sum(temp);
    num=s2.*t0 - s1.*t1;
    den=s0.*s2 - s1.*s1;
    mhat = num ./ den;
end plot(xx,mhat,'r-');
```

## A.10. El programa LLRCV

```
function [hopt,hh,CV] = LLRCV(y,x,opc,hmin,hmax,N)
% En el contexto de la regresión no paramétrica simple de Y sobre X, dados
```

---

```

% dos vectores de datos (x) e (y), esta función proporciona la anchura de
% la ventana óptima (hopt) por el método de validación cruzada mínimo-
% cuadrática (LSCV), cual es equivalente al enfoque de la función
% penalizadora (concretamente a la validación cruzada generalizada, GCV).
% Para ello se obtiene el valor de la función de validación cruzada (CV) en
% un intervalo de (N) posibles anchuras de ventana (vector hh), que se
% construyen de forma equiespaciada entre (hmin) y (hmax). Si (opc=0) los
% cálculos se hacen para el estimador de Nadaraya-Watson y si (opc=1) se
% hacen para la regresión local lineal.
%
% Uso:      [hopt,hh,CV] = LLRCV(y,x,opc,hmin,hmax,N)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Control de los argumentos de entrada %%%%%%%%%%
if (nargin ~= 6);
    error('error de número de argumentos en LLRCV');end

if (sum(size(x)~=size(y))~=0 | size(x,2)~=1);%
error('x e y deben ser vectores columna de igual dimensión en
LLRCV'); end;

if (opc ~= 0 & opc~= 1);
    error('opc debe ser 0 (NW) ó 1 (LLR) en LLRCV');end;

if (isscalar(hmax)~=1 | isscalar(hmin)~=1 | hmax<hmin | hmin<=0);
error('hmin y hmax deben ser escalares positivos con hmin<=hmax en
LLRCV'); end;

```



```

if checktyp('N',N,'int',mfilename);return;end%
if (N<1 | isscalar(N)~=1)
    error('N debe ser un escalar natural en LLRCV');end;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Programa %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

hh=linspace(hmin,hmax,N); % vector de posibles anchuras de ventana
y = meshgrid(y,y)';      % matriz en cuyas columnas hay n copias
                          % del vector y

difs = meshgrid(x,x);
difs = (difs' - difs);   % matriz de las diferencias xi-xj

p=0; % porcentaje de obs a quitar en total
    % se podría poner como otro argumento de entrada
omega=(x>=prctile(x,p/2) & x<=prctile(x,100-p/2))';

% Cálculo de la función CV en cada valor de hh
for i=1:N
    difsh = difs/hh(i);

    % matriz de valores kernels (gaussiano)
    kernvals = exp(-((difsh.*difsh)/2));

    s0 = sum(kernvals);
    t0 = sum(kernvals.*y);
    if opc==0; % Caso Nadaraya-Watson

```

```

    mhat=t0./s0;
    ponder=1-1./s0;
    temp=(y(:,1)')-mhat)./ponder;
    CV(i)=sum(temp.*temp.*omega);
else      % Caso regresión lineal local
    temp = difs.*kernvals;
    s1 = sum(temp);
    t1 = sum(temp.*y);
    temp = difs.*temp;
    s2 = sum(temp);
    num=s2.*t0 - s1.*t1;
    den=s0.*s2 - s1.*s1;
    mhat = num ./ den;
    ponder=1 - s2 ./ den;
    temp=(y(:,1)')-mhat)./ponder;
    CV(i)=sum(temp.*temp.*omega);
end
end

% Obtención de la anchura óptima
[CVmin,i]=min(CV), hopt=hh(i); plot(hh,CV,'k-');

```

## A.11. El programa LLRMBB

```
function [xx minf mhat msup] = LLRMBB(y,x,opc,xmin,xmax,N,h,b,nb)
```

## A.11. El programa LLRMBB

---

```
% En el contexto de la regresión no paramétrica simple de Y sobre X, dados
% dos vectores de datos (x) e (y) y una anchura de ventana (h), esta
% función proporciona la curva de regresión estimada (mhat) en (N) puntos
% equiespaciados en el intervalo [xmin,xmax] (vector xx) así como bandas
% de confianza puntuales al 95% (minf,msup) obtenida por el método
% bootstrap de los bloques móviles (MBB) con (b) el tamaño de cada bloque
% y (nb) el número de replicaciones bootstrap. Si (opc=0) usa el estimador
% de Nadaraya-Watson y si (opc=1) se aplica la regresión local lineal.
%
% Uso: [xx minf mhat msup] = LLRMBB(y,x,opc,xmin,xmax,N,h,b,nb)
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Control de los argumentos de entrada %%%%%%%%%%%%%%%
if (nargin ~= 9);
    error('error de número de argumentos en LLRMBB');end

if (sum(size(x)~=size(y))~=0 | size(x,2)~=1);%
error('x e y deben ser vectores columna de igual dimensión en
LLRMBB'); end;

if (opc ~= 0 & opc~= 1);
    error('opc debe ser 0 (NW) ó 1 (LLR) en LLRMBB');end;

if (isscalar(xmax)~=1 | isscalar(xmin)~=1 | xmax<xmin);
error('xmin y xmax deben ser escalares con xmin<=xmax en LLRMBB');
end;

if checktyp('N',N,'int',mfilename);return;end%
```

```

if (N<1 | isscalar(N)~=1)
    error('N debe ser un escalar natural en LLRMBB');end;

if (isscalar(h)~=1 | h<=0 );
    error('la ventana debe ser un escalar positivo en LLRMBB'); end;

if checktyp('b',b,'int',mfilename);return;end%
if checktyp('nb',nb,'int',mfilename);return;end

if (b<1 | isscalar(b)~=1 | nb<1 | isscalar(nb)~=1 )
    error('b y nb deben ser escalares naturales en LLRMBB');end;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Programa %%%%%%%%%%%%%+
n=size(x,1); k=round(n/b);
l=k*b;          % tamaño de cada muestra bootstrap

xb=zeros(1,1);
yb=zeros(1,1); % vectores donde se van a guardar cada muestra bootstrap

a=0.05;        % Intervalos al 95% de confianza

% Obtención de m estimada
[xx mhat] = LLR(y,x,opc,xmin,xmax,N,h);

% Obtención de la banda de confianza puntual Bootstrap al 95 en xx(i)
for i=1:N

```

```

% Elección de los k bloques móviles que compondrán la j-ésima
% muestra bootstrap

for j=1:nb
    bloque=unidrnd(n-b+1,1,k);

% Construcción de la j-ésima muestra bootstrap con los bloques elegidos
    for m=1:k
        xb((m-1)*b+1:m*b)=x(bloque(m):bloque(m)+b-1);
        yb((m-1)*b+1:m*b)=y(bloque(m):bloque(m)+b-1);
    end

% Estimación de la regresión con la muestra bootstrap
[temp mhatb(j)] = LLR(yb,xb,opc,xx(i),xx(i),1,h);

    home;fprintf('%2.2f, %2.2f.\n',100*i/N,j*100/nb); %control tiempo
end

% Intervalo de confianza basado en los percentiles bootstrap
minf(i,1)=prctile(mhatb,100*a/2);
msup(i,1)=prctile(mhatb,100*(1-a/2));

end;%

plot(xx,mhat','k-',xx,minf','k.',xx,msup','k.');
```

## A.12. El programa PRMjiang

```
% Esta función estima el precio del riesgo de mercado (lambda=lambda0·sigma)
% a partir de los rendimientos (Y1,Y2) de dos activos (que no pagan dividendos)
% dependientes de los tipos de interés. (t1) y (t2) los periodos de vencimiento de cada activo. (h) es la anchura de la ventana que se usa para
% efectuar las regresiones pertinentes (que se hacen por el método kernel usando un núcleo gaussiano). El precio del riesgo de mercado se estima
% para diferentes valores de los tipos de interés (rr); en concreto para
% (N) valores equiespaciados en el intervalo [rmin,rmax]. (delta) es la
% diferencia en el tiempo (constante) entre dos observaciones
% consecutivas). El rendimiento (Y1) se usa como proxy de los tipos de
% interés instantáneos. El método seguido se encuentra en Jiang (1998).
%
% Uso: [rr lambda] = PRMjiang(Y1,t1,Y2,t2,h,delta,rmin,rmax,N)

function [rr lambda] = PRMjiang(Y1,t1,Y2,t2,h,delta,rmin,rmax,N)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Control de los argumentos de entrada %%%%%%%%%
if (nargin ~= 9);
    error('error de número de argumentos en PRMjiang');end

if (sum(size(Y1)~=size(Y2))~=0 | size(Y1,2)~=1);%
error('Y1 e Y2 deben ser vectores columna de igual dimensión en
PRMjiang'); end;
```

```

if (isscalar(rmax)~=1 | isscalar(rmin)~=1 | rmax<rmin);%
error('rmin y rmax deben ser escalares con rmin<=rmax en PRMjiang');
end;

if checktyp('N',N,'int',mfilename);return;end%
if (N<1 | isscalar(N)~=1)
    error('N debe ser un escalar natural en PRMjiang');end;

if (isscalar(t1)~=1 | t1<=0 | isscalar(t2)~=1 | t2<=0 | ...
    isscalar(h)~=1 | h<=0 | isscalar(delta)~=1 | delta<=0);
    error('t1,t2,h y delta deben ser escalares positivos en PRMjiang');
end;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Programa %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Obtención de la deriva y la volatilidad con el primer rendimiento, que
% además se toman como los tipos de interés instantáneos.
n=size(Y1,1); Y1d=(Y1(2:n)-Y1(1:n-1)); x1=Y1(1:n-1);
y1=(1/delta)*Y1d; y1v=(1/delta)*(Y1d.^2);

[rr a1] = LLR(y1,x1,0,rmin,rmax,N,h);    % deriva
[rr k1] = LLR(y1v,x1,0,rmin,rmax,N,h);
k1=sqrt(k1);                            % volatilidad

% Obtención de la deriva y la volatilidad con el 2º rendimiento.

```

```

Y2d=(Y2(2:n)-Y2(1:n-1));
Y2_mu=(1/delta).*Y2d;%
Y2_sigma=(1/delta)*(Y2d.^2);%

[rr a2] = LLR(Y2_mu,x1,0,rmin,rmax,N,h);      % deriva
[rr k2] = LLR(Y2_sigma,x1,0,rmin,rmax,N,h);
k2=sqrt(k2);                                  % volatilidad

% Obtención de la diferencia de rendimientos en función del nivel
% de los tipos de interés.
Yd_mu=Y1-Y2; [rr Yd] = LLR(Yd_mu,Y1,0,rmin,rmax,N,h);

% Precio del riesgo de mercado.
num=Yd+0.5.*(((t1^2)*(k1.^2))-((t2^2)*(k2.^2)))+((t2*a2)-(t1*a1));
den=(t2*k2)-(t1*k1); lambda=k1.*(num./den);
plot(rr,lambda,'k-',rr,zeros(N,1),'r--');

```

### A.13. El programa PRMjiangMBB

```

function [r0 lambdainf lambdasup] = ...
    PRMjiangMBB(Y1,t1,Y2,t2,h,delta,r0,b,nb)
% Esta función aplica la técnica Bootstrap de los bloques móviles (MBB) al
% procedimiento descrito en Jiang (1998) para obtener un intervalo al 95%

```



### A.13. El programa PRMjiangMBB

---

```
% de confianza [lambdainf,lambdasup] para el precio del riesgo de mercado,  
% calculado como en el programa PRMjiang, en el punto (r0). Se debe  
% proporcionar además: los rendimientos (Y1,Y2) de dos activos (que no  
% pagan dividendos) dependientes de los tipos de interés, (t1) y (t2) los  
% periodos de vencimiento de cada activo, (h) la anchura de la ventana que  
% se usa para efectuar las regresiones pertinentes (que se hacen por el  
% método kernel usando un núcleo gaussiano), (delta) la diferencia en el  
% tiempo (constante) entre dos observaciones consecutivas), (b) el tamaño  
% de los bloques bootstrap y (nb) el número de replicaciones bootstrap. El  
% rendimiento (Y1) se usa como proxy de los tipos de interés instantáneos.  
%  
% Uso: [r0 lambdainf lambdasup]=PRMjiangMBB(Y1,t1,Y2,t2,h,delta,r0,b,nb)
```

```
%%%%%%%%%% Control de los argumentos de entrada %%%%%%%%%%%
```

```
if (nargin ~= 9);
```

```
    error('error de número de argumentos en PRMjiangMBB');end
```

```
if (sum(size(Y1)~=size(Y2))~=0 | size(Y1,2)~=1); error('Y1 e Y2  
deben ser vectores columna de igual dimensión en PRMjiangMBB'); end;
```

```
if (isscalar(t1)~=1 | t1<=0 | isscalar(t2)~=1 | t2<=0 | ...
```

```
    isscalar(h)~=1 | h<=0 | isscalar(delta)~=1 | delta<=0);
```

```
    error('t1,t2,h y delta deben ser escalares positivos en PRMjiangMBB');
```

```
end;
```

```
if (isscalar(r0)~=1);
```

```

    error('r0 debe ser escalar en PRMjiangMBB');
end;

if checktyp('b',b,'int',mfilename);return;end%
if checktyp('nb',nb,'int',mfilename);return;end

if (b<1 | isscalar(b)~=1 | nb<1 | isscalar(nb)~=1 )
    error('b y nb deben ser escalares naturales en PRMjiangMBB');end;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Programa %%%%%%%%%
a=0.05;    % 95% de confianza
n=size(Y1,1);

% Obtención de la deriva y la volatilidad con el primer rendimiento, que
% además se toman como los tipos de interés instantáneos.
Y1d=(Y1(2:n)-Y1(1:n-1));%
x1=Y1(1:n-1);%
Y2d=(Y2(2:n)-Y2(1:n-1));%
Yd_mu=Y1-Y2;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% MBB de estas tres variables %%%%%%%%%

k=round((n-1)/b);
l=k*b;          % tamaño de cada muestra bootstrap

Y1db=zeros(l,1);%
Y2db=zeros(l,1);%

```

```

x1b=zeros(1,1); % vectores donde se van a guardar cada muestra bootstrap

kd=round(n/b);%
ld=kd*b;

Yd_mub=zeros(ld,1);%
Y1b=zeros(ld,1);

for j=1:nb bloque=unidrnd(n-b,1,k);

    for m=1:k
        Y1db((m-1)*b+1:m*b)=Y1d(bloque(m):bloque(m)+b-1);
        Y2db((m-1)*b+1:m*b)=Y2d(bloque(m):bloque(m)+b-1);
        x1b((m-1)*b+1:m*b)=x1(bloque(m):bloque(m)+b-1);
    end

y1b=(1/delta)*Y1db; y1vb=(1/delta)*(Y1db.^2);
Y2_mub=(1/delta).*Y2db;%
Y2_sigmab=(1/delta)*(Y2db.^2);%

[r0 a1b(j)] = LLR(y1b,x1b,0,r0,r0,1,h);      % deriva
[r0 k1b(j)] = LLR(y1vb,x1b,0,r0,r0,1,h);
k1b(j)=sqrt(k1b(j));                        % volatilidad
[r0 a2b(j)] = LLR(Y2_mub,x1b,0,r0,r0,1,h);  % deriva
[r0 k2b(j)] = LLR(Y2_sigmab,x1b,0,r0,r0,1,h);
k2b(j)=sqrt(k2b(j));                        % volatilidad

```

```
% Obtención de la diferencia de rendimientos en función del nivel
% de los tipos de interés.
```

```
bloqued=unidrnd(n-b+1,1,kd);
```

```
for m=1:kd
```

```
    Yd_mub((m-1)*b+1:m*b)=Yd_mu(bloqued(m):bloqued(m)+b-1);
```

```
    Y1b((m-1)*b+1:m*b)=Y1(bloqued(m):bloqued(m)+b-1);
```

```
end
```

```
[r0 Ydb(j)] = LLR(Yd_mub,Y1b,0,r0,r0,1,h);
```

```
% Precio del riesgo de mercado.
```

```
num=Ydb(j)+0.5.*(((t1^2)*(k1b(j).^2))-((t2^2)*(k2b(j).^2)))+...
```

```
    ((t2*a2b(j))-(t1*a1b(j)));
```

```
den=(t2*k2b(j))-(t1*k1b(j)); lambda(j)=k1b(j).*(num./den);
```

```
end
```

```
% Intervalo de confianza basado en los percentiles bootstrap
```

```
    lambdainf=prctile(lambda,100*a/2);
```

```
    lambdasup=prctile(lambda,100*(1-a/2));
```

## A.14. El programa testJn

```
% En el contexto de la regresión no paramétrica simple de Y sobre X, dados
% dos vectores de datos (x) e (y), una anchura de ventana (h) y una
% cantidad (nb) de replicaciones bootstrap, esta función contrasta si la
% especificación paramétrica (H0) (lineal si H0=1, constante si H0=2,
% potencial si H0=3) es aceptable a partir de los datos. El test aplicado
% es la versión bootstrap salvaje del denominado test Jn, Kim y Wang(2006).
% Argumentos de salida:
% (Jn) Valor del estadístico.
% (pvalue) p-valor del contraste.
%
% Uso:          [Jn, pvalue] = testJn (y,x,h,nb,H0)
%
function [Jn, pvalue] = testJn (y,x,h,nb,H0)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Control de los argumentos de entrada %%%%%%%%%%%%%%%

if (nargin ~= 5);
    error('error de número de argumentos en testJn');end

[n temp] = size(x); if (sum(size(x)~=size(y))~=0 | temp~=1);
error('x e y deben ser vectores columna de igual dimensión en
testJn'); end;

if (isscalar(H0)~=1 | (H0 ~= 1 & H0 ~= 2 & H0 ~= 3));%
```

```

error('H0 debe ser 0 (lineal), 1 (constante) ó 2 (potencial) en
testJn'); end;

if checktyp('nb',nb,'int',mfilename);return;end%
if (nb<1 | isscalar(nb)~=1)
    error('nb debe ser un escalar natural en testJn');end;

if (isscalar(h)~=1 | h<=0 );
    error('la ventana h debe ser un escalar positivo en testJn'); end;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Programa %%%%%%%%%%%%%%%
r=(1+sqrt(5))/(2*sqrt(5));%
a=(1-sqrt(5))/2;%
b=(1+sqrt(5))/2;%
c=1/sqrt(2);

% etapa 1

if H0==1      % regresión lineal
    X=[ones(n,1) x];
    betahat=regress(y,X);
    yhat=X*betahat;
    ehat=y-yhat;
elseif H0==2  % regresión constante
    yhat=mean(y);
    ehat=y-yhat;

```

```
else          % ajuste potencial
    betahat = nlinfit(x,y,@potencial,[1e+10 15]);
    yhat=potencial(betahat,x);
    ehat=y-yhat;
end

difs = meshgrid(x,x);
difs = (difs' - difs);    % matriz de las diferencias xi-xj
difsh = difs/h;
% matriz de valores kernels (gaussiano)
kernvals = exp(-((difsh.*difsh)/2));

diagonal=1:n+1:n*n;
kernvals(diagonal)=0; % leave one out (0's en diagonal)

num=ehat'*kernvals* ehat;%
denom=(ehat.*ehat)'+ (kernvals.*kernvals)* (ehat.*ehat);
Jn=c*num/sqrt(denom);

for i=1:nb;
% etapa 2
temp=rand([n,1]);%
temp=temp<=r;%
temp=b+(a-b)*temp;
estar=temp.*ehat; %errores bootstrap

% etapa 3
```

```
ystar=yhat+estar;

if H0==1      % regresión lineal
    betahatstar=regress(ystar,X);
    ehatstar=ystar-X*betahatstar; % residuos bootstrap
elseif H0==2  % regresión constante
    ehatstar=ystar-mean(ystar); % residuos bootstrap
else
    betahatstar = nlinfit(x,ystar,@potencial,[1e+5 10]); %h1, h2
    %betahatstar = nlinfit(x,ystar,@potencial,[1e+6 10]); % h3 hcv
    ehatstar=ystar-potencial(betahatstar,x);
end

% etapa 4
num=ehatstar'*kernvals* ehatstar;%
denom=(ehatstar.*ehatstar)'*(kernvals.*kernvals)*(ehatstar.*ehatstar);
Jnstar(i)=c*num/sqrt(denom);%
end

pvalue=mean(Jn<Jnstar);
```



# Bibliografía

- [1] Andersen, T.G. y Lund, J. (1977). Estimating continuous-time stochastic volatility models of the short-term interest rate. *Journal of Econometrics*, **77**, 343–377.
- [2] Arapis, M. y Gao, J. (2006). Empirical Comparisons in Short-Term Interest Rate Models Using Nonparametric Methods. *Journal of Financial Econometrics*, **4**, 310–345.
- [3] Aït-Sahalia, Y. (1996). Nonparametric Pricing of Interest Rate Derivative Securities. *Econometrica*, **64**, 527–560.
- [4] Aït-Sahalia, Y. (1996). Testing Continuous-Time Models of the Spot Rate. *Review of Financial Studies*, **9**, 385–426.
- [5] Bachelier, L. (1900). Théorie de la spéculation. *Annales scientifiques de l'École Normale Supérieure*, **17**, 21–86.
- [6] Billingsley, P. (1968). *Convergence of Probability Measures*. John Wiley & Sons, New York.
- [7] Björk, T. (1997). Interest Rate Theory. *Lecture Notes in Mathematics*, **1656**, 53–122. Springer-Verlag, Berlin.

- [8] Björk, T. (2004). *Arbitrage Theory in Continuous Time*. Oxford University Press, 2<sup>a</sup> edición.
- [9] Black, F., Derman, E., y Toy, W. (1990). A one-factor model of interest rates and its applications to treasury bond options. *Financial Analysts Journal*, **46**, 33–39.
- [10] Black, F. y Karasinski, P. (1991). Bond and option pricing when short rates are lognormal. *Financial Analysts Journal*, **47**, 52–59.
- [11] Black, F. y Scholes, M. (1973). The pricing of options and corporate liabilities. *The Journal of Political Economy*, **81**, 637–659.
- [12] Box, G., Jenkins, G.M., y Reinsel, G. (1994). *Time Series Analysis: Forecasting and Control*. Prentice–Hall, 3 edición.
- [13] Boyle, P. (1977). Options: a Monte–Carlo approach. *Journal of Financial Economics*, **4**, 323–338.
- [14] Boyle, P., Broadie, M., y Glasserman, P. (1997). Monte–Carlo methods for security pricing. *Journal of Economic Dynamics and Control*, **21**, 1267–1322.
- [15] Brennan, M. y Schwartz, E. (1980). Analyzing convertible bonds. *Journal of Financial and Quantitative Analysis*, **15** (4), 907–929.
- [16] Brigo, D. y Mercurio, F. (2006). *Interest Rate Models. Theory and Practice*. Springer–Verlag, 2 edición.
- [17] Cairns, A.G.J. (2004). *Interest Rate Models: An Introduction*. Princeton University Press.
- [18] Canabarro, E. (1995). Where do one-factor interest rate models fail? *The Journal of Fixed Income*, **5**, 31–52.

- [19] Chan, K.C., Karolyi, G.A., Longstaff, F.A., y Sanders, A.B. (1992). An Empirical Comparison of Alternative Models of the Term Structure of Interest Rates. *Journal of Finance*, **47**, 1209–1228.
- [20] Chapman, D.A. y Pearson, N.D. (2000). Is the Short Rate Drift Actually Nonlinear? *Journal of Finance*, **55** (1), 355–388.
- [21] Constantinides, G. (1992). A Theory of the Nominal Term Structure of Interest Rates. *Review of Financial Studies*, **5**, 531–552.
- [22] Corzo, T. y Gómez, J. (2001). Nonparametric Estimation of Interest Rates Processes in Europe. *Revista de Economía Aplicada*, **9** (27), 83–101.
- [23] Cox, J. (1975). Notes on Option Pricing I: Constant Elasticity of Variance Diffusions. Working paper, Stanford University, Graduate School of Business. Reimpreso en *Journal of Portfolio Management* 1996, 22, 15-17.
- [24] Cox, J.C., Ingersoll, J.E., y Ross, S. (1985). An intemporal general equilibrium model of asset prices. *Econometrica*, **53**, 363–384.
- [25] Cox, J.C., Ingersoll, J.E., y Ross, S. (1985). A theory of the term structure of interest rates. *Econometrica*, **53**, 385–407.
- [26] Delbaen, F. y Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, **300**, 463–520.
- [27] Delbaen, F. y Schachermayer, W. (2006). *The Mathematics of Arbitrage*. Springer–Verlag.
- [28] Dellacherie, C. (1972). *Capacités et processus stochastiques*. Springer–Verlag, Berlin.

- [29] Doléans-Dade, C. (1970). Quelques applications de la formule de changement de variables pour les semimartingales. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, **16**, 181–194.
- [30] Donsker, M.D. (1951). An invariance principle for certain probability limit theorems. *Memoirs of the American Mathematical Society*, **6**, 1–12.
- [31] Doob, J.L. (1953). *Stochastic Processes*. John Wiley & Sons, New York.
- [32] Dothan, L.U. (1978). On the Term Structure of Interest Rates. *Journal of Financial Economics*, **6**, 59–69.
- [33] Dothan, M. (1990). *Prices in Financial Markets*. Oxford University Press, New York.
- [34] Duffie, D. y Kan, R. (1996). A yield-factor model of interest rates. *Mathematical Finance*, **6**, 379–406.
- [35] Durrett, R. (1996). *Stochastic Calculus: A practical introduction*. CRC Press.
- [36] Dybvig, P.H. (1997). Bond and Bond Option Pricing Based on the Current Term Structure. En M.A.H. Dempster y S. Pliska, eds., *Mathematics of Derivative Securities*. Cambridge University Press.
- [37] Dynkin, E.B. (1965). *Markov Processes*. Springer-Verlag. Dos volúmenes.
- [38] Dynkin, E.B. y Yushkevich, A.A. (1969). *Markov Processes: Theorems and Problems*. Plenum Press, New York.
- [39] Einstein, A. (1905). On the Movement of Small Particles Suspended in Stationary Liquids Required by the Molecular-Kinetic Theory of Heat. *Annalen der Physik*, **17**, 549–560.

- [40] Fan, J., Zhang, C., y Zhang, J. (2001). Generalized likelihood ratio statistics and Wilks phenomenon. *Annals of Statistics*, **29**, 153–193.
- [41] Fan, J. y Zhang, C. (2003). A reexamination of diffusion estimators with applications to financial model validation. *Journal of the American Statistical Association*, **98**, 118–134.
- [42] Fan, J. y Gijbels, I. (1995). Data-Driven Bandwidth Selection in Local Polynomial Fitting: Variable Bandwidth and Spatial Adaptation. *Journal of the Royal Statistical Society. Series B (Methodological)*, **57** (2), 371–394.
- [43] Fan, Y. y Li, Q. (1997). A Consistent Nonparametric Test for Linearity of AR(p) Models. *Economics Letters*, **55**, 53–59.
- [44] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications, Volumen 2*. John Wiley & Sons.
- [45] Fernandez, V. (2001). A nonparametric approach to model the term structure of interest rates: The case of Chile. *International Review of Financial Analysis*, **10**, 99–122.
- [46] García, J. (1998). Tipos de interés a corto plazo en España. *Revista de Economía Aplicada*, **18** (6), 5–26.
- [47] Gavilán Ruiz, J.M. y Alba Riesco, J.M. (2003). Un Análisis no Paramétrico de los Tipos de Interés a Corto Plazo en España. VI Encuentro de Economía Aplicada.
- [48] Gavilán Ruiz, J.M. y González Abril, L. (1999). Construcción y Suavización de Histogramas: Método Cum y Splines Parabólicos. Actas de la I Jornada Asepelt. Nuevas Tecnologías para la Enseñanza de Economía Aplicada.

- [49] Gavilán Ruiz, J.M. y González Abril, L. (2000). Una Metodología para la Construcción de Histogramas. Aplicación a los Ingresos de los Hogares Andaluces. XIV Reunión ASEPELT-España. Anales de Economía Aplicada.
- [50] Gavilán Ruiz, J.M. (1999). Una Introducción al Cálculo Estocástico. Aplicaciones a la Economía Financiera. Tesina. Facultad de Ciencias Económicas y Empresariales. Universidad de Sevilla.
- [51] Girsanov, I.V. (1960). On Transforming a Certain Class of Stochastic Processes by Absolutely Continuous Substitution of Measures. *Theory of Probability and its Applications*, **7**, 325–331.
- [52] Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering*. Springer.
- [53] Hale, J. (1969). *Ordinary Differential Equations*. John Wiley & Sons Interscience, New York.
- [54] Hamilton, J.D. (1994). *Time Series Analysis*. Princeton University Press.
- [55] Harrison, J.M. y Pliska, S.R. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and Appl.*, **11**, 215–260.
- [56] Hille, E. y Phillips, R.S. (1957). *Functional Analysis and Semigroups*. American Mathematical Society, Providence, RI.
- [57] Ho, T. y Lee, S. (1986). Term Structure Movements and Pricing Interest Rate Contingent Claims. *Journal of Finance*, **41**, 1011–1029.
- [58] Härdle, W. (1997). *Applied Nonparametric Regression*. Cambridge University Press.
- [59] Härdle, W. y Müller, M. (2000). Multivariate and Semiparametric Kernel Regression. En M.G. Schimek, ed., *Smoothing and Regression: Approaches, Computation, and Application*. Wiley-Interscience.

- [60] Hull, J. y White, A. (1990). Pricing interest rate derivative securities. *Review of Financial Studies*, **3**, 573–592.
- [61] Itô, K. (1942). Differential equations determining Markov Processes. *Zenkoku Shijo Sugaku Danwakai*, **1077**, 1352–1400.
- [62] Itô, K. (1944). Stochastic integral. *Proceedings of the Imperial Academy Tokyo*, **20**, 519–524.
- [63] James, J. y Webber, N. (2000). *Interest Rate Modelling*. John Wiley & Sons.
- [64] Jiang, G.J. (1998). Nonparametric Modeling of U.S. Interest Rate Term Structure Dynamics and Implications on the Prices of Derivative Securities. *Journal of Financial and Quantitative Analysis*, **33** (4), 465–497.
- [65] Jiang, G.J. y Knight, J.L. (1997). A Nonparametric Approach to the Estimation of Diffusion Processes, with an Application to a Short-Term Interest Rate Model. *Econometric Theory*, **13**, 615–645.
- [66] Karatzas, I. y Shreve, S.E. (1991). *Brownian Motion and Stochastic Calculus*. Springer–Verlag, New York, 2 edición.
- [67] Karatzas, I. y Shreve, S.E. (1998). *Methods of Mathematical Finance*. Springer.
- [68] Kim, M.S. y Wang, S. (2006). Sizes of two bootstrap-based nonparametric specification tests for the drift function in continuous time models. *Computational Statistics and Data Analysis*, **50**, 1793–1806.
- [69] Knight, F.B. (1981). *Essentials of Brownian Motion*. American Mathematical Society.
- [70] Künsch, H. (1989). The Jackknife and the Bootstrap for General Stationary Observations. *Annals of Statistics*, **17** (3), 1217–1241.

- [71] Kwok, Y.K. (1998). *Mathematical Models of Financial Derivatives*. Springer-Verlag Singapore.
- [72] Li, Q. (1994). A Consistent Test for Linearity in Partially Linear Regression Models. Universidad de Guelph. Disponible en <http://ideas.repec.org/p/fth/guelph/1994-7.html>.
- [73] Li, Q. (1999). Consistent model specification tests for time series econometric models. *Journal of Econometrics*, **92**, 101–147.
- [74] Li, Q. y Wang, S. (1998). A simple consistent bootstrap test for a parametric regression function. *Journal of Econometrics*, **87**, 145–165.
- [75] Litterman, R. y Scheinkman, J.A. (1991). Common Factors Affecting Bond Returns. *Journal of Fixed Income*, **1**, 54–61.
- [76] Liu, R.Y. y Singh, K. (1992). Moving Blocks Jackknife and Bootstrap Capture Weak Dependence. En: *Exploring the Limits of Bootstrap*. Editado por LePage, R, y Billard, L.
- [77] Longstaff, F.A. y Schwartz, E.S. (1992). Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model. *Journal of Finance*, **47**, 1259–1282.
- [78] McKean, H.P.J. (1969). *Stochastic Integrals*. Academic Press, New York.
- [79] Merton, R.C. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science*, **4**, 141–183.
- [80] Meyer, P. (1976). *Un cours sur les intégrales stochastiques*, capítulo en Séminaire de Probabilités X, no. 511 in *Lecture notes in mathematics*. Springer–Verlag, págs. 245–400.



- [81] Mikosch (2004). *Elementary Stochastic Calculus with Finance in view*. World Scientific, 5 edición.
- [82] Moreno, M. (2000). Modelización de la estructura temporal de los tipos de interés: valoración de activos derivados y comportamiento empírico. *Revista Española de Financiación y Contabilidad*, **29** (104), 345–376.
- [83] Nowman, K. (1997). Gaussian Estimation of Single-Factor Continuous Time models of the Term Structure of Interest Rates. *Journal of Finance*, **52** (4), 1695–1706.
- [84] Oksendal, B. (1998). *Stochastic Differential Equations: An Introduction with Applications*. Springer-Verlag, 5 edición.
- [85] Paley, R.E.A.C., Wiener, N., y Zygmund, A. (1933). Note on random functions. *Mathematische Zeitschrift*, **37**, 647–668.
- [86] Pelsser, A. (2000). *Efficient Methods for Valuing Interest Rate Derivatives*. Springer-Verlag.
- [87] Platen, E. y Heath, D. (2006). *A Benchmark Approach to Quantitative Finance*. Springer-Verlag.
- [88] Protter, P. (2004). *Stochastic integration and differential equations*. Springer-Verlag, 2 edición.
- [89] Rebonato, R. (1998). *Interest-Rate Option Models*. John Wiley & Sons, 2 edición.
- [90] Revuz, D. y Yor, M. (1991). *Continuous Martingales and Brownian motion*. Springer-Verlag, New York.
- [91] Samuelson, P.A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, **6**, 41–50.
- [92] Samuelson, P.A. (1973). Mathematics of speculative prices. *SIAM Review*, **15**, 1–39.

- [93] Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall.
- [94] Stanton, R. (1997). A nonparametric model of term structure dynamics and the market price of interest rate risk. *Journal of Finance*, **52** (5), 1973–2002.
- [95] Talay, D. (1996). Probabilistic numerical methods for PDEs: Elements of analysis. En D. Talay y L. Tubaro, eds., *Probabilistic Methods for Nonlinear PDEs*. Springer.
- [96] Uhlenbeck, G.E. y Ornstein, L.S. (1930). On the Theory of the Brownian Motion. *Physical Review*, **36**, 823–841.
- [97] Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, **5**, 177–188.
- [98] Vetzal, K.R. (1997). Stochastic Volatility, Movements in Short-Term Interest Rates and Bond Option Values. *Journal of Banking and Finance*, **21**, 169–196.
- [99] Wasserman, L. (2006). *All of Nonparametric Statistics*. Springer.
- [100] Wiener, N. (1923). Differential space. *Journal of Mathematical Physics*, **2**, 131–174.
- [101] Wiener, N. (1924). Un problème de probabilités dénombrables. *Bulletin de la Société Mathématique de France*, **52**, 569–578.
- [102] Zheng, J.X. (1996). A consistent test of functional form via non-parametric estimation techniques. *Journal of Econometrics*, **75**, 263–289.