Universidad de Cantabria

Facultad
de
Ciencias

Teoría de Martingalas y Cálculo Estocástico para la derivación de la fórmula de Black-Scholes
(Martingale Theory and Stochastic Calculus for the Derivation of the Black-Scholes Formula)

Trabajo de Fin de Grado
para acceder al

GRADO EN MATEMATICAS

Autor: Ander Bingen Ibarrondo Bilbao

Director: Luis González De La Fuente

Junio - 2025

Resumen

Este trabajo tiene como objetivo la obtención de la fórmula de Black-Scholes-Merton, empleada en la valoración de opciones financieras. Para ello, se desarrolla el marco matemático necesario, centrado principalmente en la teoría de martingalas y el cálculo estocástico, así como en una breve aplicación y consecuencias de la fórmula en la práctica. Se abordan conceptos fundamentales como el movimiento browniano, los procesos estocásticos, las integrales y procesos de Itô, así como el teorema de Girsanov, entre otros. La metodología seguida no solo permite obtener la fórmula de Black-Scholes de manera formal, sino que también proporciona una base teórica sólida para abordar la valoración de derivados financieros más complejos.

Palabras clave: Black-Scholes, cálculo estocástico, martingalas, Itô, teorema de Girsanov, derivados financieros.

Abstract

This work aims to derive the Black-Scholes-Merton formula, which is used in the valuation of financial options. For this purpose, the necessary mathematical framework is developed, focusing primarily on martingale theory and stochastic calculus, as well as a brief application and practical consequences of the formula. Key concepts such as Brownian motion, stochastic processes, Itô integrals and processes, as well as Girsanov's theorem, are discussed. The methodology presented not only leads to a formal derivation of the Black-Scholes formula, but also provides a solid theoretical foundation for tackling the valuation of more complex financial derivatives.

Keywords: Black-Scholes, stochastic calculus, martingales, Itô, Girsanov's theorem, financial derivatives.

Table of Contents

Capítulo 1

Introducción

La valoración de productos financieros, especialmente los derivados, es una de las cuestiones más relevantes y complejas en el ámbito de las finanzas modernas. El auge de los mercados financieros durante el siglo XX, especialmente a partir de los años 70, puso de manifiesto la necesidad de herramientas matemáticas capaces de modelizar con precisión el comportamiento incierto de los precios y de establecer criterios objetivos y coherentes para su valoración.

Dentro de estos instrumentos financieros, los derivados ocupan un lugar destacado tanto por su utilidad práctica como por su creciente complejidad. A medida que los productos derivados se han vuelto más sofisticados, se ha hecho imprescindible construir una base teórica sólida que permita abordar su estudio con garantías. En este trabajo se ha decidido comenzar por uno de los modelos fundamentales: la fórmula de Black-Scholes-Merton, que constituye uno de los resultados más emblemáticos dentro de la teoría matemática de las finanzas, al proporcionar una expresión cerrada para la valoración de opciones europeas bajo ciertas hipótesis de mercado. Su comprensión detallada es esencial para poder avanzar hacia la valoración de derivados más complejos en el futuro.

La deducción de la fórmula de Black-Scholes-Merton que se presenta en el trabajo sigue de cerca el enfoque original de Black y Scholes (1973), pero se apoya también de forma esencial en el principio de no arbitraje desarrollado en profundidad por Merton (1973). Esta fórmula supuso una revolución en los mercados financieros, al introducir un enfoque basado en modelos continuos de tiempo y técnicas del análisis estocástico. El impacto fue tal que, en 1997, Scholes y Merton recibieron el Premio Nobel de Economía (Black falleció antes y no pudo ser premiado).

En el trabajo se empieza explicando algunos conceptos financieros para ubicar al lector en este contexto. Luego se exponen resultados matemáticos a modo de preliminares para empezar a construir los principios del marco teórico del trabajo, en este caso, Martingalas y el movimiento Browniano, junto a los resultados que acometen.

Después, en el capítulo 3, se aborda el campo del cálculo estocástico. En este caso la integral, procesos y fórmula o lema de Itô.

Por último, en el cuarto capítulo se trata la formulación de la ecuación de Black-Scholes y su resolución, pasando por el importante teorema de Girsanov o cambio de medida. Y se acaba con un apartado más práctico con el objetivo de cerrar de manera completa los conceptos que rodean el modelo de Black-Scholes.

Capítulo 2

Preliminares

El trabajo sigue principalmente, a modo de síntesis, el contenido del libro de Steven E.Shreve "Stochastic Calculus for Finance II: Continuous-Time Models", [9]. Donde se explica de manera extendida la mayoría de los conceptos tratados a lo largo del trabajo. Para introducir estas primeras definiciones se utiliza también la mencionada obra de E.Shreve cuando es posible y se recurre a ([1], [6], [7]) para explicar contenido complementario que se da por sabido en el libro. Y para los conceptos financieros se ha recurrido principalmente a la obra de John C. Hull, [5].

2.1. Preliminares financieros

Los conceptos financieros fundamentales para el estudio del trabajo se presentan a continuación.

Normalmente, se asume que el *mercado monetario* o de dinero es el conjunto de mercados financieros, en los que se intercambian activos financieros que tienen como denominador común un plazo de amortización corto, un bajo riesgo y una elevada liquidez.

Aunque, en sí, en la definición explícitamente venga que es un conjunto de mercados financieros y que solo hay inversiones a corto plazo, seguiremos el uso común de muchos autores ([5], [9]) y nos referiremos, simplemente, a invertir en el mercado monetario como la inversión en un único activo con una rentabilidad r de riesgo cero (llamada rentabilidad libre de riesgo), ya sea en una posición larga o corta. Se puede imaginar esto como conceder un préstamo (posición larga) que el comprador va a tener que devolver con un interés r, o pedir un préstamo (posición corta) que se va a tener que devolver con un interés r.

De hecho, este último mecanismo de banco que presta o al que podemos prestar dinero es lo único necesario que tenemos que tener en mente para todo lo que sigue.

A veces, aparte de mercado monetario, también se suele utilizar el término de "mercado de bonos". Pero toda esta aclaración es para recalcar que el instrumento financiero al que nos referimos no es exactamente un bono, sino una simplificación teórica, porque el precio del bono es afectado de manera instantánea por los cambios de intereses (mirar "Warning" de [9, p. 215]).

Pero vamos a profundizar a lo que nos referimos con intereses. En el ámbito financiero, el concepto de interés representa la ganancia o el costo asociado al uso del dinero a lo largo del tiempo. Se utiliza comúnmente para calcular el rendimiento de inversiones o el costo de préstamos. Existen dos formas fundamentales de calcular el interés: el interés simple y el interés compuesto.

Si pedimos un préstamo por una cantidad P, llamado principal invertido, al banco y nos dicen que hay que devolverlo el año siguiente con un interés r, es que hay que pagar la cantidad prestada P más los intereses rP, es decir, $P + rP = P(1 + r)$. Y nos podemos preguntar, el banco me pide unos intereses de r al año, entonces ¿Y si solo quiero un préstamo de medio año? Pues entonces, como tenía que devolver rP por tener el dinero ese año, por tenerlo medio año, ahora tengo que devolver $\frac{r}{2}P$, es decir, al final de medio año devuelvo $P + \frac{r}{2}P = P(1 + \frac{r}{2})$. Todos estos intereses son *intereses simples*.

El interés compuesto considera no solo el capital inicial, sino también los intereses acumulados en periodos anteriores, generando un crecimiento exponencial. Cada periodo, el monto total se reinvierte, lo que produce intereses sobre intereses. Ponemos un ejemplo para que se vea mejor y deducir la fórmula.

Supongamos que ahora eres tú quien invierte dinero en el banco. Depositas una cantidad inicial, P, y el banco te ofrece una tasa de interés r compuesta **semianual**. Esto significa que al cabo de medio año (6 meses), el banco te paga la mitad del interés anual, es decir, $\frac{r}{2}P$, pero todavía no te devuelve el capital inicial.

Ahora bien, tú no decides quedártelo porque quieres seguir generando dinero con esa cantidad, por eso lo reinviertes, tu nuevo capital para la segunda mitad del año ya no es solo P, sino:

$$P + \frac{r}{2}P = P\left(1 + \frac{r}{2}\right)$$

Durante los siguientes seis meses, el banco vuelve a aplicar la misma tasa semianual de $\frac{r}{2}$, pero ahora sobre el nuevo capital. Por lo tanto, al final del año, el capital se multiplica otra vez por $\left(1 + \frac{r}{2}\right)$, como hemos visto en el ejemplo de antes, y se obtiene:

$$P\left(1 + \frac{r}{2}\right)\left(1 + \frac{r}{2}\right) = P\left(1 + \frac{r}{2}\right)^2$$

Si el banco, en lugar de pagar semestralmente, compusiera cada trimestre, entonces repetirías el proceso cuatro veces al año, y el resultado sería $P\left(1 + \frac{r}{3}\right)^3$.

Si en vez de tres veces, el interés se compone n veces al año, entonces en cada periodo se aplica una tasa de $\frac{r}{n}$, y se repite el proceso n veces. Al final del año, tendrías:

$$P\left(1 + \frac{r}{n}\right)^n$$

Y si el número de periodos de composición crece indefinidamente, entonces en el **límite** se obtiene lo que se conoce como *interés compuesto continuo*:

$$\lim_{n\to\infty} P\left(1 + \frac{r}{n}\right)^n = Pe^r$$

Por eso, en condiciones idealizadas en las que siempre exista un servicio como el mercado monetario, decimos que el potencial de P monedas a 1 año con interés compuesto continuo es Pe^r, a dos años $Pe^r \cdot e^r = Pe^{2r}$, y en general para un tiempo t el que sea Pe^{tr}. Ese es el valor potencial, es decir, el valor al que sabemos que puede crecer P. Invirtiendo el pensamiento podemos pensar cuanto estoy dispuesto a pagar a alguien por una cantidad P que me va a dar en el tiempo t. Más de Pe^{-tr} por supuesto que no, porque sé que si invierto más que Pe^{-tr}, puedo ganar más de $Pe^{-tr} \cdot e^{tr} = P$, ese precio justo, que es Pe^{-tr} por el que pagaría P en el futuro, sin salir ni perdiendo ni ganando, se le llama *valor presente* o *valor actualizado* de P con interés compuesto continuo. Y supondremos que el interés continuo compuesto es el caso a menos que se especifique lo contrario.

Ya hemos explicado los mecanismos que nos ofrecen renta fija y como lo hacen, ahora vemos los elementos de renta variable, en este caso la bolsa y derivados.

La bolsa de valores es un mercado organizado donde se compran y venden activos financieros, siendo los más comunes las acciones de empresas. Un *activo financiero* es un instrumento que representa un derecho económico: por ejemplo, una acción representa una parte proporcional de la propiedad de una empresa.

A diferencia de los activos de renta fija, como los bonos, donde el rendimiento está pactado de antemano, en la bolsa se negocian principalmente activos de renta variable, cuyo valor puede subir o bajar de forma impredecible. ¿Por qué? Porque el precio de un activo en la bolsa no se determina únicamente por su valor "real" o contable, sino por lo que las personas están dispuestas a pagar por él, es decir, por las expectativas del mercado.

Cuando alguien compra una acción, lo hace no solo por lo que vale hoy esa empresa, sino por lo que cree que valdrá mañana. Si muchas personas piensan que la empresa crecerá, obtendrá más beneficios o ganará importancia en su sector, estarán dispuestas a pagar más por sus acciones hoy. Esa demanda hace que el precio suba. Por el contrario, si se cree que la empresa tendrá pérdidas, que el sector entrará en crisis o simplemente se genera incertidumbre, los inversores querrán vender y pocos querrán comprar: el precio bajará.

Así como en la bolsa se negocian activos cuyo valor depende de lo que se espera que ocurra, existen otros instrumentos financieros más sofisticados llamados derivados financieros. Su nombre proviene del hecho de que su valor se "deriva" del valor de otro activo, al que llamamos *activo subyacente*.

A diferencia de una acción, que representa una participación directa en una empresa, un derivado es esencialmente un contrato entre dos partes que acuerdan pagar o recibir una cantidad de dinero dependiendo del activo subyacente en el futuro. Aquí nos disponemos a definir los de mayor importancia para el trabajo.

Definición 2.1. ([5]) Una *opción de compra Europea* es un contrato que da al poseedor el derecho pero no la obligación de comprarle al vendedor del contrato, el activo subyacente acordado, por un precio acordado en una fecha acordada. También es normal usar el término en inglés *call Europea*.

A este precio y fecha se les llama como *precio* y *fecha del ejercicio* respectivamente.

Ejemplo 2.1. Supongamos que adquirimos una opción de compra (call) europea por 5\$, al lo que se le denomina como *la prima*. Esta opción nos otorga el derecho, pero no la obligación, de comprar una acción de Google a un precio de ejercicio de 180\$ el 1 de enero del año siguiente.

Si en esa fecha el precio de mercado de la acción de Google se encuentra por debajo del precio de ejercicio, por ejemplo, en 150\$, no ejerceríamos nuestro derecho, ya que sería más conveniente adquirir la acción directamente en el mercado a un precio inferior. En este caso, la pérdida incurrida sería equivalente al valor de la prima pagada por la opción.

En cambio, si el precio de mercado de la acción en la fecha de vencimiento supera el precio de ejercicio, digamos que alcanza los 200$, entonces ejerceríamos la opción, comprando la acción a 180$. Posteriormente, podríamos venderla en el mercado a 200$, obteniendo así un beneficio bruto de 20$, sin considerar el costo inicial de la prima.

Cabe destacar que, en la práctica, este intercambio físico de acciones rara vez se lleva a cabo. En su lugar, el vendedor de la opción suele liquidar la diferencia directamente, abonando al comprador el valor correspondiente al beneficio obtenido.

Desde un punto de vista matemático, si denotamos por $S(T)$ el precio del activo subyacente en la fecha de vencimiento T, y por K el precio de ejercicio, la ganancia obtenida por el comprador de la opción de compra europea, sin considerar la prima, se expresa como:

$$(S(T) - K)^+ = \max(S(T) - K, 0).$$

Por su parte, el vendedor de la opción recibe como ganancia la prima, pero, en caso de que la opción resulte favorable al comprador, incurre en una pérdida equivalente a $(S(T) - K)^+$.

El objetivo último será establecer principalmente el precio para este contrato.

Definición 2.2. ([5]) Una *opción de venta Europea* es un contrato que da al poseedor el derecho pero no la obligación de venderle al vendedor del contrato, el activo subyacente acordado, por un precio acordado en una fecha acordada. También es normal usar el término en inglés *put Europea*.

Siguiendo la misma lógica del ejemplo para la opción de compra y la misma notación, el pago para una opción de venta a fecha del ejercicio sería $(K - S(T))^+$.

Observación 2.1. Notar que hay cuatro tipos de posiciones que se pueden tomar en un mercado de opciones:

1. Comprar una call
2. Vender una call
3. Comprar put
4. Vender una put

A comprar opciones lo llamamos operar en largo y vender operar en corto.

Uno de los pilares fundamentales de la valoración de instrumentos derivados es el principio de *ausencia de arbitraje*. Este principio establece que en un mercado financiero eficiente no deben existir oportunidades que permitan obtener una ganancia segura sin necesidad de realizar una inversión inicial y sin asumir ningún tipo de riesgo.

La existencia persistente de tales oportunidades llevaría a los agentes del mercado a explotarlas inmediatamente, lo cual provocaría ajustes en los precios de los activos hasta que dichas oportunidades desaparezcan. Por tanto, en condiciones de mercado competitivo, se espera que los precios se ajusten de forma tal que no sea posible construir estrategias de arbitraje.

En la práctica, este enfoque lleva al concepto de *precio de no arbitraje*, entendido como aquel valor para un derivado que impide que cualquier participante en el mercado pueda generar ganancias libres de riesgo mediante estrategias de compra y venta de los activos disponibles.

La formulación matemática de este principio, así como su utilización para valorar opciones, serán abordadas en las secciones siguientes.

2.2. Fundamentos matemáticos y probabilisticos

El uso del no arbitraje para valorar derivados financieros se basa en la idea de planes de contingencia, es decir, en cómo se debe actuar dependiendo de la información disponible en cada momento. Para poder modelar esto matemáticamente, se utiliza el concepto de información como un conjunto de resultados posibles que se pueden considerar "resueltos" en función de lo que se sabe hasta un instante determinado.

Por ejemplo, imaginemos que tenemos un activo que empieza valiendo $S_0 = 4$ dólares cuya subida o bajada de precio viene dado aleatoriamente por el resultado de tirar una moneda tres veces, es decir, para la primera tirada, si sale cara (la denominaremos H por su inicial heads en inglés) se duplica el precio anterior y alcanza un precio al que llamamos $S_1(H) = 8$, si sale cruz (la denominaremos T por su inicial tails en inglés) baja a la mitad de precio y alcanza un precio al que llamamos $S_1(T) = 2$. Y así recursivamente con las dos tiradas restantes [1]. Es decir, matemáticamente nuestro espacio muestral tendría a Ω como la tríada de posibles combinaciones de cara o cruz (HHH,HHT,...). Ahora imaginemos que realizamos el experimento de lanzar tres monedas, pero no conocemos la terna resultante ω, pero si nos dicen el resultado del primer lanzamiento ya podemos dividir el espacio de resultados en dos conjuntos: aquellos que comienzan con H y los que comienzan con T. Esto nos da dos subconjuntos:

$$A_H = \{HHH, HHT, HTH, HTT\} \quad A_T = \{THH, THT, TTH, TTT\}$$

Aunque no sepamos el resultado completo, la información parcial (como saber lo que ha salido en el primer lanzamiento) nos permite hacer una lista de conjuntos tales que podemos asegurar que el resultado final pertenece o no a esos conjuntos. A estos, les decimos que están resueltos por la información actual y forman una σ-álgebra.

Si conocemos el resultado de los dos primeros lanzamientos, la información es aún más detallada, y los subconjuntos se vuelven más pequeños, por ejemplo, $A_{HH} = \{HHH, HHT\}$, $A_{HT} = \{HTH, HTT\}$, $A_{TH} = \{THH, THT\}$, $A_{TT} = \{TTH, TTT\}$ están resueltos y, al igual que su unión y complementarios también.

Cada vez que se obtiene más información (por ejemplo, conocer más lanzamientos), la resolución del espacio de resultados mejora, y las σ-álgebras correspondientes contienen más subconjuntos. Estas -álgebras están organizadas jerárquicamente en lo que se llama una filtración, que modela cómo la información crece con el tiempo.

[1] Este es un modelo binomial de activos, mirar [8]

Definición 2.3. (Filtración) ([9]) Sea Ω un conjunto no vacío. Sea T un número positivo fijo, y supongamos que para cada $t \in [0, T]$ existe una σ-álgebra $\mathcal{F}(t)$. Supongamos además que si $s \leq t$, entonces cada conjunto en $\mathcal{F}(s)$ también pertenece a $\mathcal{F}(t)$. Entonces, llamamos *filtración* a la colección de σ-álgebras $\mathcal{F}(t)$, $0 \leq t \leq T$.

En el caso anterior, sin haber lanzado la moneda, el único conjunto resuelto es el total, es decir, $\mathcal{F}(0) = \{\emptyset, \Omega\}$. Para el primer lanzamiento ya hemos visto que los conjuntos resueltos son $\emptyset, \Omega$ (que además pertenecen a $\mathcal{F}(0)$) y $A_H = \{HHH, HHT, HTH, HTT\}, A_T = \{THH, THT, TTH, TTT\}$ que formarían $\mathcal{F}(1)$. Del mismo modo, por lo visto antes y siguiendo el mismo proceso

$$\mathcal{F}_2 = \begin{Bmatrix} \emptyset, \Omega, A_H, A_T, A_{HH}, A_{HT}, A_{TH}, A_{TT}, A_{HH}^c, A_{HT}^c, A_{TH}^c, A_{TT}^c, \\ A_{HH} \cup A_{TH}, A_{HH} \cup A_{TT}, A_{HT} \cup A_{TH}, A_{HT} \cup A_{TT} \end{Bmatrix}$$

Y finalmente $\mathcal{F}(3)$ sería las partes de Ω.

Ahora, imaginemos que la información que recibimos no es el resultado de la moneda, sino el valor del activo, es decir, la variable aleatoria $S_1(\omega)$, $S_2(\omega)$..., para el cual tenemos la formula en la que pones ω y nos da el valor:

$$S_i(\omega) = S_{i-1} \cdot \begin{cases} 2 & \text{si } \omega_i = H \\ \frac{1}{2} & \text{si } \omega_i = T \end{cases}$$

Y si nos dicen que $S_2 = 16$ la información que podemos sacar es que ω puede ser $\{HHH\}$ o $\{HHT\}$ por lo que $\{S_2 = 16\} = A_{HH}$. Todos estos conjuntos que podemos formar dependiendo de la variable aleatoria se formaliza matemáticamente con la siguiente definición.

Definición 2.4. ([9]) Sea X una variable aleatoria definida sobre un espacio muestral no vacío Ω. La σ-álgebra generada por X, denotada $\sigma(X)$, es la colección de todos los subconjuntos de Ω de la forma $\{X \in B\} = \{\omega \in \Omega : X(\omega) \in B\}$, donde B es cualquier subconjunto de Borel en $\mathbb{R}$.

En la discusión previa a la definición solo se tomaba $B = 16$. Vamos a ver a modo de ejemplo cuál sería $\sigma(S_2)$, mirando los posibles valores de S_2:

$$\begin{aligned} &S_2(HHH) = S_2(HHT) = 16 \\ &S_2(HTH) = S_2(HTT) = S_2(THH) = S_2(THT) = 4 \\ &S_2(TTH) = S_2(TTT) = 1 \end{aligned}$$

Tomando $B = 16$ hemos visto que tenemos el conjunto A_{HH}, con $B = 4$ tenemos $A_{HT} \cup A_{TH}$, con $B = 1$ tenemos A_{TT}, con $B = [4, 16]$ tenemos $A_{HH} \cup A_{HT} \cup A_{TH}$, y en resumen, tomando B cualquier subconjunto de Borel, tendremos en $\sigma(S_2)$ todas las uniones y complementarios de $\emptyset$, Ω, A_{HH}, A_{TT}, $A_{HT} \cup A_{TH}$.

Todo conjunto en $\sigma(S_2)$ está en la σ-álgebra $\mathcal{F}_2$, que contiene la información de los dos primeros lanzamientos de moneda. Por otro lado, A_{HT} y A_{TH} aparecen por separado en $\mathcal{F}_2$ y solo su unión aparece en $\sigma(S_2)$. Esto se debe a que ver los dos primeros lanzamientos de moneda nos permite distinguir una cara seguida de una cruz frente a una cruz seguida de una cara, pero conocer solo el valor de S_2 no lo permite. Por eso afirmamos que hay suficiente información en $\mathcal{F}_2$ para determinar el valor de S_2. Y decimos que S_2 es $\mathcal{F}_2$-medible.

Definición 2.5. ([9]) **($\mathcal{G}$-medibilidad)** Sea X una variable aleatoria definida en un espacio muestral no vacío Ω. Sea $\mathcal{G}$ una σ-álgebra de subconjuntos de Ω. Si cada conjunto en $\sigma(X)$ también pertenece a $\mathcal{G}$, decimos que X es $\mathcal{G}$-medible.

Y con todos estos conceptos previos definimos el principal elemento con el que vamos a trabajar, los procesos estocásticos adaptados.

Definición 2.6. ([6]) **(Proceso estocástico)** sea $X(t)$ una colección de variables aleatorias en $(\Omega, \mathcal{F}, \mathbb{P})$ indexadas por $t \in [0, T]$ llamamos a dicha colección $\{X(t) : t \in T\}$ un proceso estocástico.

La definición anterior se puede generalizar tomando un subconjunto cualquiera de $\mathbb{R}$ en vez del intervalo $[0, T]$. Pero en el contexto de series temporales y aplicaciones financieras suele ser suficiente restringirse al caso $t \in [0, T]$, con $T < \infty$, lo cual facilita tanto el análisis como la interpretación temporal del proceso.

Una vez definido qué entendemos por proceso estocástico, es natural preguntarse cómo se relaciona este con la información disponible en cada instante de tiempo. Para abordar esta cuestión, introducimos el concepto de proceso estocástico adaptado, el cual impone una condición de compatibilidad entre la evolución del proceso y una filtración que representa el flujo de información a través del tiempo.

Definición 2.7. ([9]) **(Proceso estocástico adaptado)** Sea Ω un espacio muestral no vacío equipado con una filtración $\mathcal{F}(t), 0 \leq t \leq T$ y sea $X(t)$ una colección de variables aleatorias indexadas por $t \in [0, T]$, es decir, un proceso estocástico. Decimos que esta colección de variables aleatorias es un *proceso estocástico adaptado* si, para cada t, la variable aleatoria $X(t)$ es $\mathcal{F}(t)$-medible.

Ahora nos damos cuenta de que el experimento de la moneda y el activo es un ejemplo de proceso estocástico adaptado, en concreto la colección de variables aleatorias $S(t)$ es el proceso estocástico y está adaptado a la filtración del lanzamiento de las monedas descrita justo después de la definición 2.3.

Por último, continuando con el ejemplo de la moneda, suponemos que esta vez solo se realiza un lanzamiento, suponemos que el interés del mercado monetario es $r = \frac{1}{4}$ y que el precio de ejercicio de la opción de compra europea es $K = 5$. Supongamos que el dinero del que disponemos en un principio es de $X_0 = 1{,}20$ dólares y compramos $\Delta_0 = \frac{1}{2}$ acciones al tiempo cero, dicha cantidad de acciones a tiempo cero suponemos que cuestan dos dólares.

Luego debemos usar nuestra riqueza inicial $X_0 = 1{,}20$ y pedir prestado un adicional de 0.80 para hacer esta compra. Esto nos deja con una posición de efectivo de $X_0 - \Delta_0 S_0 = -0{,}80$ (es decir, una deuda de 0.80 en el mercado monetario). En el tiempo uno, teniendo en cuenta que estamos en un caso discreto, es decir, los intereses generados por una cantidad de dinero en el siguiente periodo los consideramos en este caso como (1+r) por esa cantidad nuestra posición de efectivo será:

$$(1 + r)(X_0 - \Delta_0 S_0) = 1{,}25 \cdot (-0{,}8) = -1$$

Es decir, tendremos una deuda de 1 en el mercado monetario.

Por otro lado, en el tiempo uno, el valor de la acción será o bien $\frac{1}{2}S_1(H) = 4$ o $\frac{1}{2}S_1(T) = 1$. En particular:

- Si el lanzamiento de la moneda resulta en cara, teniendo en cuenta que poseemos media acción, el valor de nuestra cartera será:

$$X_1(H) = \frac{1}{2}S_1(H) + (1 + r)(X_0 - \Delta_0 S_0) = 4 - 1 = 3$$

- Del mismo modo si el lanzamiento de la moneda resulta en cruz, el valor de nuestra cartera será:

$$X_1(T) = \frac{1}{2}S_1(T) + (1 + r)(X_0 - \Delta_0 S_0) = 1 - 1 = 0$$

En cualquiera de los casos, el valor de la cartera coincide con el valor de la opción en el tiempo uno, que es:

$$(S_1(H) - 5)^+ = 3 \quad \text{o} \quad (S_1(T) - 5)^+ = 0$$

Hemos replicado la opción comerciando acciones y operando con el mercado monetario. Lo que da la riqueza inicial de 1.20 dólares necesarios para configurar la cartera replicante descrita es el *precio de no arbitraje de la opción en el tiempo cero* de la que hablábamos al final de la sección de preliminares financieros.

De esta manera, formalmente:

Definición 2.8. ([5]) Una cartera replicante es una cartera de activos financieros que reproduce exactamente los flujos de caja de otro activo o estrategia de inversión.

Dicha cartera, en el contexto que nos acomete más adelante (modelo de Black-Scholes) siempre existe, mirar sección 8.2 Completeness in the Black-Scholes Model de ([2]).

Una vez terminado con el ejemplo, volvemos a las definiciones matemáticas que nos ocupan.

Definición 2.9. ([9]) **(σ-álgebras independientes)** Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, y sean $\mathcal{G}$ y $\mathcal{H}$ sub-σ-álgebras de $\mathcal{F}$ (es decir, los conjuntos en $\mathcal{G}$ y los conjuntos en $\mathcal{H}$ también están en $\mathcal{F}$). Decimos que estas dos σ-álgebras son independientes si

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B) \quad \text{para todo } A \in \mathcal{G},\ B \in \mathcal{H}.$$

Definición 2.10. ([9]) **(Variable aleatoria independiente de una σ-álgebra)** Decimos que la variable aleatoria X es independiente de la σ-álgebra $\mathcal{G}$ si $\sigma(X)$ y $\mathcal{G}$ son independientes.

Observación 2.2. Además, si X y Y son variables aleatorias en $(\Omega, \mathcal{F}, \mathbb{P})$. Sin cambiar lo que significa que dos variables sean independientes en el sentido más clásico, podemos decir que estas dos variables aleatorias son independientes si las σ-álgebras que generan, $\sigma(X)$ y $\sigma(Y)$, son independientes.

En lo que resta de sección nos disponemos a recordar algunos resultados y definiciones ya conocidos:

Definición 2.11. ([7]) Dado $p > 0$ y la variable aleatoria real X se llama momento de orden p de X a la cantidad $\mathbb{E}[X^p]$ suponiendo que exista.

Definición 2.12. ([7])La función generadora de momentos $M(t)$ de una variable aleatoria continua X , con función de densidad $f(x)$, se define como:

$$M(t) = \mathbb{E}[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx$$

Se le llama generadora de momentos porque todos los momentos de X pueden obtenerse derivando sucesivamente $M(t)$ y evaluando en $t = 0$. Por ejemplo, derivando formalmente:

$$M'(t) = \frac{d}{dt}\mathbb{E}[e^{tX}] = \mathbb{E}\left[\frac{d}{dt}e^{tX}\right] = \mathbb{E}[Xe^{tX}]$$

donde se ha supuesto que es válido intercambiar el operador de derivación con el de esperanza. Es decir, se asume que:

$$\frac{d}{dt}\left[\int_{-\infty}^{\infty} e^{tx} f(x)\, dx\right] = \int_{-\infty}^{\infty} \frac{d}{dt}\left(e^{tx}\right) f(x)\, dx$$

Esta suposición suele estar justificada en la mayoría de distribuciones usuales.

Definición 2.13. ([9]) Dado un espacio probabilístico $(\Omega, \mathcal{F}, \mathbb{P})$ y una variable aleatoria no negativa Z que cumpla que $\mathbb{E}[Z] = 1$. Podemos definir una nueva probabilidad $\tilde{\mathbb{P}}$:

$$\tilde{\mathbb{P}}(A) = \int_A Z(\omega) d\mathbb{P}(\omega)$$

Y decimos que Z es la derivada de Radon-Nikodým de $\tilde{\mathbb{P}}$ respecto de $\mathbb{P}$

Teorema 2.1. *([9]) **(Teorema de Radon-Nikodým)** Sean $\mathbb{P}$ y $\widetilde{\mathbb{P}}$ medidas de probabilidad equivalentes definidas en $(\Omega, \mathcal{F})$. Entonces existe una variable aleatoria Z casi seguramente positiva tal que $\mathbb{E}[Z] = 1$ y*

$$\widetilde{\mathbb{P}}(A) = \int_A Z(\omega)\, d\mathbb{P}(\omega) \quad \textit{para todo } A \in \mathcal{F}$$

Se puede ver el anterior teorema con una condición más débil, basta con que $\widetilde{\mathbb{P}}$ sea absolutamente continua respecto a $\mathbb{P}$, que quiere decir que $\forall E \in \mathcal{F}$ tal que $\mathbb{P}(E) = 0$ se tiene que $\widetilde{\mathbb{P}}(E) = 0$. Mientras que la condición que nosotros hemos impuesto es un si y solo si. La demostración se puede encontrar en [1].

2.2.1. Resultados de convergencia

Recordamos también algunos resultados básicos que utilizaremos. Los cuales no demostraremos.

Definición 2.14. ([6]) **(Espacio L^p)** En un espacio muestral $(\Omega, \mathcal{F}, \mu)$, siendo u una función y donde por su clase de equivalencia nos referimos a la relación de ser equivalentes en casi todo punto respecto a μ, se tiene que:

$$L^p = \{\text{clases de equivalencia de } u : \int_\Omega |u|^p \, d\mu < \infty\}$$

Más en nuestro contexto, y para aclarar la notación, si nuestro espacio muestral es $(\Omega, \mathcal{F}, \mathbb{P})$, y X es una variable aleatoria, la definición anterior equivale a decir que $L^2 = \{\text{clases de equivalencia de } X : \mathbb{E}[|X|^p] < \infty\}$. Que es lo que vamos a emplear a partir de ahora.

Definición 2.15. ([6]) **(Convergencia en L^p)** Decimos que la sucesión $(u_n,\, n \geq 1)$ converge a u en L^p (en concreto, nos referimos a la convergencia fuerte) si todas las variables están en L^p y

$$\lim_{n\to\infty} \int |u_n - u|^p \, d\mu = 0$$

De nuevo en el contexto que se indica previo a esta definición, si u se trata de una variable aleatoria que denotamos como X, etc. La condición del límite la podemos escribir como $\lim_{n\to\infty} \mathbb{E}[|X_n - X|^p] = 0$.

En concreto, prestamos interés a L^2 donde dicha convergencia también se la conoce como convergencia en media cuadrática.

Esta forma de convergencia en L^p es más fuerte que la convergencia en probabilidad, ya que implica que los X_n se acercan a X no solo "probabilísticamente", sino también en promedio. Lo que se ve en el siguiente resultado.

Proposición 2.1. *([6]) Convergencia en L^p implica convergencia en probabilidad ($p < \infty$)*

Demostración. Mirar Teorema 7.3 de [6] ■

Observación 2.3. Por la identidad de la varianza $\mathbb{E}[(X_n - X)^2] = Var(X_n) + (\mathbb{E}[X_n] - X)^2$, para demostrar que X_n converge a X en media cuadrática, es decir,$\lim_{n\to\infty} \mathbb{E}[(X_n - X)^2] = 0$, basta con demostrar que $Var(X_n)$ tiende a 0 y $\mathbb{E}[X_n]$ tiende a X.

Teorema 2.2. *([6]) Si la sucesión X_n converge en probabilidad a X en Ω, existe una subsucesión (X_{n_k}) que converge casi seguro a X.*

Demostración. Mirar Corolario 6.13 de ([6]) ■

2.3. Esperanza condicionada

Hasta ahora hemos introducido la noción de información disponible en cada instante a través del concepto de filtración, y cómo los procesos estocásticos se adaptan a esta evolución de la información. También hemos visto que algunas variables aleatorias pueden ser medibles respecto a cierta sub-σ-álgebra, lo que significa que su comportamiento puede determinarse con la información contenida en dicha colección de subconjuntos.

En este contexto, surge la necesidad de redefinir el concepto clásico de esperanza de una variable aleatoria, adaptándolo a situaciones en las que no disponemos de toda la información del espacio muestral, sino solo de una parte de ella, representada por una σ-álgebra.

Definición 2.16. ([9]) **(Esperanza condicionada)** Consideremos el espacio probabilístico $(\Omega, \mathcal{F}, \mathbb{P})$ y sea $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$. Si X es una variable aleatoria y existe una variable aleatoria Y, que es $\mathcal{G}$-medible, tal que

$$\int_A Y d\mathbb{P} = \int_A X d\mathbb{P}, \quad \text{para todo } A \in \mathcal{G}.$$

Entonces decimos que Y es la *esperanza condicionada* dada una σ-álgebra de X y escribimos:

$$Y = \mathbb{E}[X \mid \mathcal{G}].$$

Si $\mathcal{G}$ es la σ-álgebra generada por alguna otra variable aleatoria W (es decir, $\mathcal{G} = \sigma(W)$), generalmente escribimos $\mathbb{E}[X \mid W]$ en lugar de $\mathbb{E}[X \mid \sigma(W)]$.

Teorema 2.3. *([9]) **(Existencia y unicidad)** Siempre existe $\mathbb{E}[X \mid \mathcal{G}]$ descrita en la definición 2.16 y es única casi seguro.*

Demostración.

- **Existencia:** Sea $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$ y sea X una variable aleatoria, la cual podemos escribirla como combinación de dos términos positivos $X = X^+ - X^-$. Definimos $\mathbb{Q}^+$ y $\mathbb{Q}^- : \mathcal{G} \to [0, \infty)$ tal que:

$$\mathbb{Q}^+(A) := \int_A X^+ d\mathbb{P}$$
$$\mathbb{Q}^-(A) := \int_A X^- d\mathbb{P}$$

  Que son medidas (mirar Teorema 1.4 de [1]). Además, estas medidas son equivalentes con $\mathbb{P}$, luego por el Teorema de Radon-Nikodým(2.1) existen Y_1 y Y_2 funciones $\mathcal{G}$-medibles tal que:

$$\mathbb{Q}^+(A) = \int_A Y_1 \, d\mathbb{P}$$
$$\mathbb{Q}^-(A) = \int_A Y_2 \, d\mathbb{P}$$

  Ahora escribiendo $Y = Y_1 - Y_2$, tengo que:

$$\begin{aligned}\int_A X \, d\mathbb{P} &= \int_A X^+ - X^- \, d\mathbb{P} = \int_A X^+ \, d\mathbb{P} - \int_A X^- \, d\mathbb{P} \\ &= \int_A Y_1 \, d\mathbb{P} - \int_A Y_2 \, d\mathbb{P} = \int_A Y_1 - Y_2 \, d\mathbb{P} = \int_A Y \, d\mathbb{P}\end{aligned}$$

  Probando así la existencia de tal Y.

- **Unicidad:** Suponemos que existen dos variables aleatorias Y y Z cumpliendo la definición 2.16 y veremos que son iguales. Notar primero que por definición Y y Z son $\mathcal{G}$-medibles luego su diferencia Y-Z también lo es y entonces $B = \{Y - Z > 0\}$ está en $\mathcal{G}$. Como hemos supuesto que Y y Z cumplen la definición, también tenemos:

$$\int_A Y d\mathbb{P} = \int_A X d\mathbb{P} = \int_A Z d\mathbb{P}, \quad \text{para todo } A \in \mathcal{G}.$$

En concreto para B:

$$\int_B Y d\mathbb{P} = \int_B X d\mathbb{P} = \int_B Z d\mathbb{P}$$

Teniendo así,

$$\int_B (Y - Z) d\mathbb{P} = 0$$

En B el integrando Y-Z es siempre estrictamente mayor que 0, así que la única posibilidad para que se cumpla esta última igualdad es que B tenga probabilidad 0 (es decir, $Y \leq Z$ casi seguro). Haciendo lo mismo simétricamente tenemos la otra desigualdad llegando a que $Y = Z$ casi seguro.

■

Proposición 2.2. *([9]) Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y sea $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$.*

(i) ***Linealidad*** *Si X e Y son variables aleatorias integrables*[2] *y c_1 y c_2 son constantes, entonces*

$$\mathbb{E}[c_1 X + c_2 Y \mid \mathcal{G}] = c_1 \mathbb{E}[X \mid \mathcal{G}] + c_2 \mathbb{E}[Y \mid \mathcal{G}].$$

Esta ecuación también es válida si asumimos que X e Y son no negativas (en lugar de integrables) y que c_1 y c_2 son positivos, aunque ambos lados podrían ser $+\infty$.

(ii) ***Sacar lo que se conoce:*** *Si X e Y son variables aleatorias integrables, Y y XY son integrables, y X es $\mathcal{G}$-medible, entonces*

$$\mathbb{E}[XY \mid \mathcal{G}] = X \mathbb{E}[Y \mid \mathcal{G}].$$

Esta ecuación también es válida si asumimos que X es positivo e Y es no negativo (en lugar de integrables), aunque ambos lados podrían ser $+\infty$.

(iii) ***Iteración de la esperanza:*** *Si $\mathcal{H}$ es una sub-σ-álgebra de $\mathcal{G}$ ($\mathcal{H}$ contiene menos información que $\mathcal{G}$) y X es una variable aleatoria integrable, entonces*

$$\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}] \mid \mathcal{H}] = \mathbb{E}[X \mid \mathcal{H}].$$

(iv) ***Independencia:*** *Si X es integrable e independiente de $\mathcal{G}$, entonces*

$$\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[X].$$

Esta ecuación también es válida si asumimos que X es no negativo (en lugar de integrable), aunque ambos lados podrían ser $+\infty$.

[2] Cuando decimos que una variable aleatoria es integrable nos referimos a que $\int_\Omega |X| \, d\mathbb{P} < \infty$ o lo que es lo mismo $\mathbb{E}[|X|] < \infty$

Observación 2.4. En nuestro contexto, se suele dar que $\mathcal{F}(0) = \{\emptyset, \Omega\}$, en cuyo caso, si condicionamos a dicha σ-álgebra, es como no condicionar a nada, y entonces la propiedad tres de la esperanza condicionada, que es: $\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}] \mid \mathcal{F}(0)] = \mathbb{E}[X \mid \mathcal{F}(0)]$. Se puede ver como $\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]] = \mathbb{E}[X]$

2.4. Martingalas

Una vez introducida la esperanza condicionada, estamos preparados para definir una clase especial de procesos estocásticos que juegan un papel central: las martingalas.

Las martingalas modelan situaciones en las que, dado todo lo que se sabe hasta un cierto momento, la mejor predicción del valor futuro de un proceso es simplemente su valor actual. Es decir, el proceso no tiene "memoria"ni "tendencia", lo que las convierte en herramientas ideales para modelar mercados eficientes, juegos justos o precios de activos bajo ciertas condiciones.

Además de martingalas, también introduciremos las nociones de submartingala y supermartingala, que permiten modelar procesos con tendencia al alza o a la baja, respectivamente.

Definición 2.17. ([8]) **(Martingala)** Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, sea T un número positivo fijo, y sea $\mathcal{F}(t), 0 \leq t \leq T$ una filtración de sub-σ-álgebras de $\mathcal{F}$. Consideremos un proceso estocástico adaptado $M(t)$, $0 \leq t \leq T$.

Si

$$\mathbb{E}[M(t) \mid \mathcal{F}(s)] = M(s) \quad \text{para todo } 0 \leq s \leq t \leq T,$$

decimos que este proceso es un *martingala.*

Proposición 2.3. *Si* $\{X(t) : 0 \leq t \leq T\}$ *es una martingala con respecto a la filtración* $\{\mathcal{F}(t) : 0 \leq t \leq T\}$ *para la cual* $\mathcal{F}(0) = \{\emptyset, \Omega\}$*, entonces* $\mathbb{E}[X(t)] = \mathbb{E}[X(0)]$ *para todo* $0 \leq t \leq T$*.*

Demostración. Para $0 \leq s \leq t$ tenemos por la observación 2.4 que:

$$\mathbb{E}[X(t)] = \mathbb{E}[\mathbb{E}[X(t) \mid \mathcal{F}(s)]]$$

Y como X es martingala:

$$\mathbb{E}[X(t)] = \mathbb{E}[\mathbb{E}[X(t) \mid \mathcal{F}(s)]] = \mathbb{E}[X(s)]$$

Si tomamos $s = 0$, obtenemos el resultado deseado. ■

En definiciones anteriores hemos repasado lo que era la derivada de Radon-Nykodým, ahora vemos un proceso resultante de ello y que resulta ser una martingala.

Definición 2.18. ([9]) Dado un espacio probabilístico $(\Omega, \mathcal{F}, \mathbb{P})$, una filtración $\{\mathcal{F}(t) : 0 \leq t \leq T$ y Z una derivada de Radon-Nikodým de $\tilde{\mathbb{P}}$ respecto de $\mathbb{P}$ (Definida de igual manera que en la definición 2.13), en particular Z es positiva casi seguro ($\mathbb{P}(Z > 0) = 1$). Llamamos proceso de la derivada de Radon-Nikodým a:

$$Z(t) = \mathbb{E}[Z|\mathcal{F}(t)] \quad 0 \leq t \leq T$$

Proposición 2.4. *([9]) Dado un espacio probabilístico* $(\Omega, \mathcal{F}, \mathbb{P})$*. Sea* Z *una variable aleatoria positiva casi seguro y que cumpla que* $\mathbb{E}[Z] = 1$ *y* $\tilde{\mathbb{P}}$ *definida como en la definición 2.18. Sea* Y *una función medible. Entonces:*

$$\tilde{\mathbb{E}}[Y] = \mathbb{E}[YZ]$$

Demostración. Suponemos que Y es una función indicatriz $Y = \mathbb{I}_A$. Entonces:

$$\tilde{\mathbb{E}}[Y] = \tilde{\mathbb{E}}[\mathbb{I}_A] = \tilde{\mathbb{P}}(A) = \int_\Omega \mathbb{I}_A Z d\mathbb{P}(\omega) = \mathbb{E}[\mathbb{I}_A Z] = \mathbb{E}[YZ]$$

Luego, por linearidad de la esperanza, tenemos el resultado para funciones simples (suma de funciones indicatrices). Conociendo que cualquier función medible no negativa se puede aproximar como una secuencia de funciones simples crecientes (mirar teorema 1.30 de [1]), se concluye por el teorema de convergencia monótona el resultado cuando Y es no negativa. Por último, considerando $Y = Y^+ - Y^-$, como ambas Y^+ e Y^-, son no negativas $\mathcal{G}$-medibles, de nuevo por linearidad concluimos el resultado para toda función $\mathcal{G}$-medible, como queríamos. ■

Proposición 2.5. *([9]) Dado un espacio probabilístico* $(\Omega, \mathcal{F}, \mathbb{P})$*, una filtración* $\{\mathcal{F}(t) : 0 \leq t \leq T$ *y un proceso de la derivada de Radon-Nikodým* $Z(t)$ *definida como en la definición 2.18 y con la misma* $\tilde{\mathbb{P}}$*. Sea* Y *una variable aleatoria* $\mathcal{F}(t)$*-medible. Entonces:*

1. $Z(t) : 0 \leq t \leq T$ *es una martingala*
2. $\tilde{\mathbb{E}}[Y] = \mathbb{E}[YZ(t)]$
3. $\tilde{\mathbb{E}}[Y|\mathcal{F}(s)] = \frac{1}{Z(s)}\mathbb{E}[YZ(t)|\mathcal{F}(s)]$

Demostración.

1. Queremos ver que el proceso de Radon-Nikodým $Z(t) : 0 \leq t \leq T$ es una martingala:

   Sea $0 \leq s \leq t \leq T$, primero por la propia definición y después utilizando la propiedad de iteración de la esperanza condicionada:

   $$\mathbb{E}[Z(t)|\mathcal{F}(s)] = \mathbb{E}[\mathbb{E}[Z|\mathcal{F}(t)]|\mathcal{F}(s)] = \mathbb{E}[Z|\mathcal{F}(s)] = Z(s),$$

   Luego $Z(t)$ es una martingala.

2. Queremos ver $\tilde{\mathbb{E}}[Y] = \mathbb{E}[YZ(t)]$:

   Por la proposición 2.4 y dado que Y es $\mathcal{F}(t)$-medible:

   $$\tilde{\mathbb{E}}[Y] = \mathbb{E}[YZ] = \mathbb{E}[\mathbb{E}[YZ|\mathcal{F}(t)]] = \mathbb{E}[Y\mathbb{E}[Z|\mathcal{F}(t)]] = \mathbb{E}[YZ(t)].$$

3. Queremos ver que $\tilde{\mathbb{E}}[Y|\mathcal{F}(s)] = \frac{1}{Z(s)}\mathbb{E}[YZ(t)|\mathcal{F}(s)]$, para ello, recurrimos a la definición 2.16 y el resultado previo del teorema 2.3 que nos asegura la unicidad de esta. Por lo que solo hace falta probar que

   $$\int_A \frac{1}{Z(s)}\mathbb{E}[YZ(t)|\mathcal{F}(s)]\, d\tilde{\mathbb{P}} = \int_A Y\, d\tilde{\mathbb{P}}, \quad \forall A \in \mathcal{F}.$$

   Podemos escribir el lado izquierdo como:

   $$\int_A \frac{1}{Z(s)}\mathbb{E}[YZ(t)|\mathcal{F}(s)]\, d\tilde{\mathbb{P}} = \int_\Omega \mathbb{I}_A \frac{1}{Z(s)}\mathbb{E}[YZ(t)|\mathcal{F}(s)]\, d\tilde{\mathbb{P}} = \tilde{\mathbb{E}}\left[\mathbb{I}_A \frac{1}{Z(s)}\mathbb{E}[YZ(t)|\mathcal{F}(s)]\right]$$

Y está claro que $\mathbb{I}_A \frac{1}{Z(s)} \mathbb{E}[YZ(t)|\mathcal{F}(s)]$ es $\mathcal{F}(s)$-medible, luego utilizaremos el resultado 1 ya probado de esta proposición, para el caso de una función $\mathcal{F}(s)$-medible.

$$\begin{aligned}\tilde{\mathbb{E}}\left[\mathbb{I}_A \frac{1}{Z(s)} \mathbb{E}[YZ(t)|\mathcal{F}(s)]\right] &= \tilde{\mathbb{E}}\left[Z(s)\mathbb{I}_A \frac{1}{Z(s)} \mathbb{E}[YZ(t)|\mathcal{F}(s)]\right] = \mathbb{E}\left[\mathbb{I}_A \mathbb{E}[YZ(t)|\mathcal{F}(s)]\right] \\ &= \mathbb{E}\left[\mathbb{E}[\mathbb{I}_A YZ(t)|\mathcal{F}(s)]\right]\end{aligned}$$

Considerando que la esperanza externa se puede ver como condicionada a $\mathcal{F}(0)$ (observación 2.4), es decir, $\mathbb{E}\left[\mathbb{E}[\mathbb{I}_A YZ(t)|\mathcal{F}(s)]\right] = \mathbb{E}\left[\mathbb{E}[\mathbb{I}_A YZ(t)|\mathcal{F}(s)]|\mathcal{F}(0)\right]$ y usamos la propiedad de iteración, para después volver a utilizar el resultado 1 de esta proposición:

$$\mathbb{E}\left[\mathbb{E}[\mathbb{I}_A YZ(t)|\mathcal{F}(s)]|\mathcal{F}(0)\right] = \mathbb{E}\left[\mathbb{I}_A YZ(t)|\mathcal{F}(0)\right] = \mathbb{E}\left[\mathbb{I}_A YZ(t)\right] = \tilde{\mathbb{E}}\left[\mathbb{I}_A Y\right] = \int_A Y d\tilde{\mathbb{P}}$$

Que siguiendo las cadenas de igualdades es lo que queríamos probar.

■

Otro proceso que nos interesa es el siguiente.

Definición 2.19. ([9]) **(Proceso de Markov)** Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, sea T un número positivo fijo, y sea $\{\mathcal{F}(t) : 0 \leq t \leq T\}$, una filtración de sub-$\sigma$-álgebras de $\mathcal{F}$. Consideremos un proceso estocástico adaptado $\{X(t) : 0 \leq t \leq T\}$. Supongamos que para todo $0 \leq s \leq t \leq T$ y para toda función f no negativa y medible en el sentido de Borel, existe otra función medible en el sentido de Borel g tal que:

$$\mathbb{E}[f(t, X(t))|\mathcal{F}(s)] = g(s, X(s)).$$

Entonces decimos que X es un proceso de Markov.

2.5. Movimiento Browniano

El movimiento browniano es una de las piedras angulares de la teoría moderna de los procesos estocásticos y desempeña un papel fundamental en la modelización matemática de fenómenos aleatorios continuos en el tiempo.

Su aparición se remonta al estudio del movimiento errático de partículas microscópicas en un fluido, observado por primera vez por el botánico Robert Brown en 1827, quedando documentado en 1828. Aunque algunos trabajos preliminares, como los de Thiele en 1880, anticiparon ciertas ideas, la primera construcción matemática rigurosa del proceso fue realizada por Norbert Wiener en 1923. Por esta razón, el movimiento browniano también es conocido como proceso de Wiener.

En esta sección se formalizan sus propiedades matemáticas y su papel como martingala, fundamental para el desarrollo posterior de la teoría de Itô y la valoración de derivados financieros.

Definición 2.20. **(Movimiento Browniano)** Un proceso estocástico $\{W(t) : t \geq 0\}$ se llama *movimiento browniano* si satisface las siguientes propiedades:

1. $P[W(0) = 0]$ y $W(t)$, $t \geq 0$, es continuo con probabilidad 1.

2. Para $0 \leq s \leq t$, se tiene que $W(t) - W(s) \sim \mathcal{N}(0, t - s)$, es decir, tiene una distribución normal con media 0 y varianza $t - s$.

3. $\{W(t) : t \geq 0\}$ tiene incrementos estacionarios e independientes. Para todo tiempo $0 < t_1 < t_2 < t_n$,las variables aleatorias $W(t_1), W(t_2) - W(t_1), ..., W(t_n) - W(t_{n-1})$ son independientes.

Definición 2.21. ([9]) **(Filtración para el movimiento Browniano)**. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad en el cual está definido un movimiento Browniano $W(t)$, $t \geq 0$. Una *filtración para el movimiento Browniano* es una colección de σ-álgebras $\mathcal{F}(t)$, $t \geq 0$, que satisface:

(i) **(Acumulación de información)** Para $0 \leq s < t$, cada conjunto en $\mathcal{F}(s)$ también está en $\mathcal{F}(t)$. En otras palabras, hay al menos tanta información disponible en $\mathcal{F}(t)$ como en $\mathcal{F}(s)$.

(ii) **(Adaptabilidad)** Para cada $t \geq 0$, el movimiento Browniano $W(t)$ en el tiempo t es $\mathcal{F}(t)$-medible. Es decir, la información disponible en el tiempo t es suficiente para evaluar el movimiento Browniano $W(t)$ en ese instante.

(iii) **(Independencia de incrementos futuros)** Para $0 \leq t < u$, el incremento $W(u) - W(t)$ es independiente de $\mathcal{F}(t)$. En otras palabras, cualquier incremento del movimiento Browniano después del tiempo t es independiente de la información disponible en el tiempo t.

Teorema 2.4. *([9]) El movimiento Browniano es una martingala.*

Demostración. Sea $0 \leq s \leq t$. Entonces:

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

■

Definición 2.22. ([9]) **(Variación cuadrática)**

Sea $f(t)$ una función definida para $0 \geq t \geq T$, y sea $\Pi = \{t_0, t_1, \ldots, t_n\}$ una partición de $[0, T]$ tal que $0 < t_1 < \ldots < t_n = T$. Consideremos $\|\Pi\| := \text{máx}_{0 \leq i \leq (n-1)} |t_{i+1} - t_i|$, entonces la *variación cuadrática de* $f(t)$ es:

$$[f, f](T) := \lim_{\|\Pi\| \to 0} \sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)|^2,$$

Vemos ahora un lema para la demostración del teorema próximo.

Lema 2.5.1. *([9]) Dado una variable normal X de media cero y varianza σ^2, se tiene que* $\mathbb{E}[X^4] = 3\sigma^4$

Demostración. El desarrollo de la siguiente prueba está presente en el ejercicio 3.3 de [9] cuando se prueba un resultado para la kurtosis de una normal. En nuestro caso nos basta con probar lo expuesto en el enunciado. Dando por sabido que la función generadora de momentos de una normal de media cero es $\varphi(u) = \mathbb{E}[e^{uX}]$, para hallar el cuarto momento central, es decir, $\mathbb{E}[X^4]$, calculamos $\varphi^{(4)}(0)$. Derivando se tiene que

$$\varphi^{(4)}(0) = \left(3\sigma^4 + 6\sigma^6 u^2 + \sigma^8 u^4\right) e^{\frac{1}{2}u^2\sigma^2}\Big|_{u=0} = 3\sigma^4$$

Como se quería probar. ∎

Teorema 2.5. *([9]) Sea W un movimiento Browniano. Entonces $[W,W](T) = T, \forall T \geq 0$ casi seguro.*

Demostración. Sea $\Pi = \{t_0, t_1, ..., t_n\}$ una partición de [0,T] y definimos:

$$Q_\Pi = \sum_{j=0}^{n-1} (W(t_{j+1}) - W(t_j))^2$$

Esto es la variación cuadrática empírica del movimiento Browniano, la cual converge a $[W,W](T)$. Vamos a ver ahora que esta variable aleatoria Q_Π, converge en media cuadrática a T cuando $\|\Pi\| \longrightarrow 0$. Bastará ver por la observación 2.3 que su esperanza es T y que su varianza es 0; procedemos a calcularlas:

- **(Esperanza)** Por la segunda propiedad del movimiento Browniano tenemos:

$$\mathbb{E}[(W(t_{j+1}) - W(t_j))^2] = Var[W(t_{j+1}) - W(t_j)] = t_{j+1} - t_j \tag{2.1}$$

  Y por la linealidad de la esperanza:

$$\mathbb{E}[Q_\Pi] = \mathbb{E}[\sum_{j=0}^{n-1}(W(t_{j+1}) - W(t_j))^2] = \sum_{j=0}^{n-1}\mathbb{E}[(W(t_{j+1}) - W(t_j))^2] \overset{(2,1)}{=} \sum_{j=0}^{n-1}(t_{j+1} - t_j) = T$$

  Como queríamos

- **(Varianza)** Empezamos como en el caso previo calculando:

$$\begin{aligned} Var[(W(t_{j+1}) - W(t_j))^2] &= \mathbb{E}[((W(t_{j+1}) - W(t_j))^2 - (t_{j+1} - t_j))^2] \\ &= \mathbb{E}[((W(t_{j+1}) - W(t_j))^4] - 2(t_{j+1} - t_j)\mathbb{E}[(W(t_{j+1}) - W(t_j))^2] \\ &+ (t_{j+1} - t_j)^2 \end{aligned}$$

  Por el lema 2.5.1, tenemos que:

$$\mathbb{E}[((W(t_{j+1}) - W(t_j))^4] = 3Var[(W(t_{j+1}) - W(t_j)]^2 = 3(t_{j+1} - t_j)^2 \tag{2.2}$$

Luego continuando con la última expresión última utilizando (2.1) en el segundo término y (2.2) en el primero obtenemos:

$$Var[(W(t_{j+1}) - W(t_j))^2] = 3(t_{j+1} - t_j)^2 - 2(t_{j+1} - t_j)^2 + 3(t_{j+1} - t_j)^2$$
$$2(t_{j+1} - t_j)^2$$

Y parecido a lo que hicimos con la esperanza:

$$Var(Q_\Pi) = Var[\sum_{j=0}^{n-1}(W(t_{j+1}) - W(t_j))^2] = \sum_{j=0}^{n-1} Var[(W(t_{j+1}) - W(t_j))^2]$$
$$= \sum_{j=0}^{n-1} 2(t_{j+1} - t_j)^2 \leq \sum_{j=0}^{n-1} 2\,\|\Pi\|\,(t_{j+1} - t_j)$$

Esto nos dice que , $lim_{\|\Pi\| \longrightarrow 0} Var(Q_\Pi) = 0$ como queríamos

Concluyendo así que efectivamente converge en media cuadrática a T, y por la proposición 2.1 también en probabilidad, y por el teorema 2.2 existe una subsucesión convergente casi seguro a T, es decir: $[W(\omega), W(\omega)](T) = lim_{\|\Pi\| \longrightarrow 0} Q_\Pi(\omega) = T$, $\forall \omega \in \Omega$ tal que $\mathbb{P}[\omega] > 0$, probando así la propiedad en casi todas partes como queríamos, en otras palabras, para todo camino con posibilidad no nula se cumple la propiedad.

■

Observación 2.5. Recalcar que del teorema anterior podemos escribir informalmente $dW(t)dW(t) = dt$. Es decir, intuitivamente, el movimiento Browniano acumula variación cuadrática a proporción uno por unidad de tiempo. Para una aclaración más formal de la adecuada interpretación y trasfondo de esta afirmación, mirar la observación 3.4.4 de [9].

Definición 2.23. (Movimiento Browniano Geométrico) Sea $\alpha > 0$ y σ constantes, y sea $W(t)$ un movimiento browniano. Definimos el *movimiento Browniano geométrico* como:

$$S(t) = S(0)e^{\sigma W(t) + (\alpha - \frac{1}{2}\sigma^2)t}$$

Por último, exponemos dos resultados que no demostramos por la relevancia de las pruebas y su extensión.

Teorema 2.6. *([9]) **(Caracterización de Lévy del Movimiento Browniano)** Sea $\{W(t) : t \geq 0\}$ una martingala con respecto a una filtración $\{\mathcal{F}(t) : t \geq 0\}$. Si $W(0) = 0$, $W(t)$ tiene trayectorias continuas y $[W, W](t) = t$, $\forall t \geq 0$, entonces $W(t)$ es un movimiento browniano.*

Demostración. Mirar teorema 4.6.4 de [9] ■

Teorema 2.7. *([9]) Sea $W(t)$, $t \geq 0$, un movimiento browniano y sea $\mathcal{F}(t)$, $t \geq 0$, una filtración para este movimiento browniano. Entonces, $W(t)$, $t \geq 0$, es un proceso de Markov.*

Demostración. Mirar teorema 3.5.1 de [9] ■

Capítulo 3

Integral y procesos de Itô

En este capítulo, definiremos la integral de Itô, es decir, una integral con respecto a un movimiento Browniano. Inicialmente, lo haremos para un proceso estocástico adaptado simple, estableciendo la definición formal y demostrando algunas propiedades esenciales como la martingalidad y la isometría. Posteriormente, ampliaremos este marco a procesos estocásticos generales. También introduciremos los procesos de Itô, los cuales permiten describir una amplia clase de dinámicas aleatorias. Finalmente, culminamos el capítulo con la fórmula de Itô-Doeblin, o a veces llamado lema de Itô, que es una extensión del teorema de Taylor al caso estocástico.

3.1. Integral de Itô

Definición 3.1. ([9]) **(Proceso Simple)** Sea $\Pi = \{t_0, t_1, \ldots, t_n\}$ una partición de $[0, T]$. Supongamos que $\{\Delta(t) : 0 \leq t \leq T\}$ es constante en t en el intervalo $[t_i, t_{i+1})$, para todo i. Llamamos a $\{\Delta(t) : 0 \leq t \leq T\}$ un *proceso simple.*

Definición 3.2. ([9]) **(Integral de Itô para Procesos Simples)** Sea $\{W(t) : t \geq 0\}$ un movimiento Browniano con la filtración $\{\mathcal{F}(t) : t \geq 0\}$ y sea $\{\Delta(t) : 0 \leq t \leq T\}$ un proceso simple adaptado con partición $\Pi = \{t_0, t_1, \ldots, t_n\}$ de $[0, T]$. Entonces, si $t_k \leq t \leq t_{k+1}$, definimos la *Integral de Itô para procesos simples* como:

$$I(t) = \sum_{i=0}^{k-1} \Delta(t_i)[W(t_{i+1}) - W(t_i)] + \Delta(t_k)[W(t) - W(t_k)] = \int_0^t \Delta(t) dW(t)$$

Observación 3.1. Siguiendo con la integral definida anteriormente, también podemos definir una integral respecto a un diferencial en tiempo en vez de dW, esto sería, siguiendo la notación precedente:

$$\sum_{i=0}^{k-1} \Delta(t_i)(t_{i+1} - t_i) + \Delta(t_k)(t - t_k) = \int_0^t \Delta(t) dt$$

Pero, para mayor claridad, en la mayoría de los casos denotaremos lo de dentro de la integral de diferente forma $\int_0^t \Delta(u) du$

Teorema 3.1. *[9] La Integral de Itô para procesos simples es una Martingala con respecto a la filtración* $\{\mathcal{F}(t) : t \geq 0\}$.

Demostración. Sea s y t tal que $0 \leq s < t \leq T$ y podemos asumir sin pérdida de generalidad que pertenecen a intervalos distintos de Π (partición del proceso simple), i.e., $\exists l < k$ tal que $s \in [t_l, t_{l+1})$ y $t \in [t_k, t_{k+1})$. Tenemos que ver que:

$$\mathbb{E}[I(t)|\mathcal{F}(s)] = I(s)$$

Notar ahora que podemos separar $I(t)$ en dos sumatorios:

$$\begin{aligned} I(t) = & \sum_{i=0}^{l-1} \Delta(t_i)[W(t_{i+1}) - W(t_i)] + \Delta(t_l)[W(t_{l+1}) - W(t_l)] + \\ & \sum_{i=l+1}^{k-1} \Delta(t_i)[W(t_{i+1}) - W(t_i)] + \Delta(t_k)[W(t) - W(t_k)] \end{aligned}$$

Ahora, por la linealidad de la esperanza condicionada, miramos la de cada término de los anteriores por separado. Para el primero:

$$\mathbb{E}\left[\sum_{i=0}^{l-1} \Delta(t_i)[W(t_{i+1}) - W(t_i)] \middle| \mathcal{F}(s)\right]$$

Como $t_l \leq s$ tenemos que la expresión del interior es $\mathcal{F}(s)$-medible, y por la propiedad (ii) de (2.2) nos queda:

$$\mathbb{E}\left[\sum_{i=0}^{l-1} \Delta(t_i)[W(t_{i+1}) - W(t_i)] \middle| \mathcal{F}(s)\right] = \sum_{i=0}^{l-1} \Delta(t_i)[W(t_{i+1}) - W(t_i)]$$

De forma similar para el segundo término, como $t_l \leq s$ y utilizando en el último paso que el movimiento Browniano es una martingala (teorema 2.4):

$$\begin{aligned} \mathbb{E}\left[\Delta(t_l)[W(t_{l+1}) - W(t_l)] \middle| \mathcal{F}(s)\right] &= \Delta(t_l)\mathbb{E}\left[W(t_{l+1}) - W(t_l) \middle| \mathcal{F}(s)\right] \\ &= \Delta(t_l)(\mathbb{E}[W(t_{l+1})|\mathcal{F}(s)] - W(t_l)) \\ &= \Delta(t_l)(W(t_s) - W(t_l)) \end{aligned}$$

Es decir la expresión resultante de los dos primeros términos es:

$$\sum_{i=0}^{l-1} \Delta(t_i)[W(t_{i+1}) - W(t_i)] + \Delta(t_l)(W(t_s) - W(t_l)) = I(s)$$

Y solo nos queda ver que los términos restantes son cero. Para el tercer término, como $t_i \geq s$ para $i \in \{l+1, l+2, ..., k-1\}$ entonces $\mathcal{F}_s \subset \mathcal{F}_{t_i}$, por la linealidad (propiedad i) e iteración de la esperanza (propiedad iii) de las propiedades de la esperanza condicionada:

$$\begin{aligned} \mathbb{E}\left[\sum_{i=l+1}^{k-1} \Delta(t_i)[W(t_{i+1}) - W(t_i)] \middle| \mathcal{F}(s)\right] &\overset{(i)}{=} \sum_{i=l+1}^{k-1} \mathbb{E}\left[\Delta(t_i)[W(t_{i+1}) - W(t_i)] \middle| \mathcal{F}(s)\right] \\ &\overset{(iii)}{=} \sum_{i=l+1}^{k-1} \mathbb{E}\left[\mathbb{E}\left[\Delta(t_i)[W(t_{i+1}) - W(t_i)] \middle| \mathcal{F}_{t_i}\right] \middle| F(s)\right] \\ &\overset{(i+ii)}{=} \sum_{i=l+1}^{k-1} \mathbb{E}\left[\Delta(t_i)\left(\mathbb{E}[W(t_{i+1})|\mathcal{F}(t_i)] - W(t_i)\right) \middle| \mathcal{F}(s)\right] \\ &\overset{Teorema\ 2{,}4}{=} \sum_{i=l+1}^{k-1} \mathbb{E}\left[\Delta(t_i)(W(t_i) - W(t_i)) \middle| \mathcal{F}(s)\right] = 0 \end{aligned}$$

Para el último término, repitiendo lo que acabamos de hacer $\mathbb{E}[\Delta(t_k)[W(t) - W(t_k)]] = 0$

■

Teorema 3.2. *([9]) **(Isometría de Itô)** La integral de Itô para un proceso simple cumple que:*

$$\mathbb{E}[I^2(t)] = \mathbb{E}\left[\int_0^t \Delta^2(u)du\right]$$

Demostración. Vamos a denominar $\Delta W_i = W(t_{i+1}) - W(i)$ y $\Delta W_k = W(t) - W(t_k)$, de tal forma que podemos escribir la integral de Itô como:

$$I(t) = \sum_{i=0}^{k} \Delta(t_i)\Delta W_i$$

Proseguimos calculando:

$$I^2(t) = (\sum_{i=0}^{k} \Delta(t_i)\Delta W_i)^2 = \sum_{i=0}^{k} \Delta^2(t_i)\Delta W_i^2 + 2\sum_{i<j} \Delta(t_i)\Delta(t_j)\Delta W_i \Delta W_j$$

Tomando la esperanza y usando linealidad:

$$\begin{aligned}\mathbb{E}\left[I^2(t)\right] &= \mathbb{E}\left[\sum_{i=0}^{k} \Delta^2(t_i)\Delta W_i^2 + 2\sum_{i<j} \Delta(t_i)\Delta(t_j)\Delta W_i \Delta W_j\right] \\ &= \sum_{i=0}^{k} \mathbb{E}\left[\Delta^2(t_i)\Delta W_i^2\right] + 2\sum_{i<j} \mathbb{E}\left[\Delta(t_i)\Delta(t_j)\Delta W_i \Delta W_j\right]\end{aligned}$$

Procedemos a calcular $\mathbb{E}[\Delta(t_i)\Delta(t_j)\Delta W_i \Delta W_j]$, utilizando para cada igualdad de la expresión a continuación respectivamente:

- La iteración de la esperanza (como se explica en la observación 2.4)
- $\Delta(t_i)\Delta(t_j)\Delta W_i$ es $\mathcal{F}(t_j)$-medible para $i < j$, que es el caso en el sumatorio que nos ocupa, por lo que utilizaremos la segunda propiedad de la esperanza.
- ΔW_j es independiente de $\mathcal{F}(t_j)$, por definición de filtración para un movimiento Browniano (definición 2.21), por lo que podemos usar la propiedad (iv) de independencia de la esperanza condicionada
- $\mathbb{E}[\Delta W_j] = 0$ por el punto dos de la definición de movimiento Browniano (definición 2.20.

Se tiene que:

$$\begin{aligned}\mathbb{E}[\Delta(t_i)\Delta(t_j)\Delta W_i \Delta W_j] &\overset{(iii)}{=} \mathbb{E}[\mathbb{E}[\Delta(t_i)\Delta(t_j)\Delta W_i \Delta W_j|\mathcal{F}(t_j)]] \\ &\overset{(ii)}{=} \mathbb{E}[\Delta(t_i)\Delta(t_j)\Delta W_i \mathbb{E}[\Delta W_j|\mathcal{F}(t_j)]] \overset{(iv)}{=} \mathbb{E}[\Delta(t_i)\Delta(t_j)\Delta W_i \mathbb{E}[\Delta W_j]] \\ &= \mathbb{E}[\Delta(t_i)\Delta(t_j)\Delta W_i \cdot 0] = \mathbb{E}[0] = 0\end{aligned}$$

Por el otro lado, del mismo modo utilizando que $\Delta^2(t_i)$ es $\mathcal{F}(t_i)$-medible y que ΔW_i^2 es independiente de $\mathcal{F}(t_i)$:

$$\mathbb{E}[\Delta^2(t_i)\Delta W_i^2] \overset{(iii)}{=} \mathbb{E}[\mathbb{E}[\Delta^2(t_i)\Delta W_i^2|\mathcal{F}(t_i)]] \overset{(ii)}{=} \mathbb{E}[\Delta^2(t_i)\mathbb{E}[\Delta W_i^2|\mathcal{F}(t_i)]] \overset{(iv)}{=} \mathbb{E}[\Delta^2(t_i)\mathbb{E}[\Delta W_i^2]]$$

Por la ecuación 2.1 de la demostración del teorema 2.5, $\mathbb{E}[\Delta W_i^2] = t_{i+1} - t_i$, luego:

$$\mathbb{E}[\Delta^2(t_i)\mathbb{E}[\Delta W_i^2]] = \mathbb{E}[\Delta^2(t_i)(t_{i+1} - t_i)]$$

Y lo mismo ocurre para y $\mathbb{E}\left[\Delta^2(t_k)\Delta W_k^2\right]$ que resulta ser tras el mismo procedimiento $\mathbb{E}\left[\Delta^2(t_k)(t-t_k)\right]$, todo junto:

$$\begin{aligned}
I^2(t) &= \sum_{i=0}^{k} \mathbb{E}\left[\Delta^2(t_i)\Delta W_i^2\right] = \sum_{i=0}^{k-1} \mathbb{E}\left[\Delta^2(t_i)\Delta W_i^2\right] + \mathbb{E}\left[\Delta^2(t_k)\Delta W_k^2\right] \\
&\overset{(i)}{=} \sum_{i=0}^{k-1} \mathbb{E}\left[\Delta^2(t_i)(t_{i+1}-t_i)\right] + \mathbb{E}\left[\Delta^2(t_k)(t-t_k)\right] \\
&= \mathbb{E}\left[\sum_{i=0}^{k-1} \Delta^2(t_i)(t_{i+1}-t_i) + \Delta^2(t_k)(t-t_k)\right] \\
&= \mathbb{E}\left[\int_0^t \Delta^2(u)du\right]
\end{aligned}$$

Es decir, hemos probado como queríamos que:

$$\mathbb{E}\left[I^2(t)\right] = \mathbb{E}\left[\int_0^t \Delta^2(u)du\right].$$

■

Teorema 3.3. *([9]) La variación cuadrática de la integral de Itô para un proceso simple está dada por:*

$$[I,I](t) = \int_0^t \Delta^2(u)du.$$

Demostración. Vamos a probarlo mirando a los incrementos. Sea $[t_i, t_{i+1}]$ un subintervalo donde $\{\Delta(t) : 0 \leq t \leq T\}$ es constante. Y consideramos una partición de ese intervalo:

$$t_i = s_0 < s_1 < ... < s_n = t_{i+1}$$

Siendo $\|\Pi\| := \text{máx}_{0\leq i\leq(n-1)} |t_{i+1} - t_i|$, calculamos la variación en dicho intervalo:

$$\begin{aligned}
\lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} [I(s_{j+1}) - I(s_j)]^2 &= \lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} [\Delta(t_i)W(s_{j+1}) - \Delta(t_i)W(s_j)]^2 \\
&= \lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} [\Delta(t_i)\left(W(s_{j+1}) - W(s_j)\right)]^2 \qquad (3.1) \\
&= \Delta^2(t_i) \lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} [W(s_{j+1}) - W(s_j)]^2
\end{aligned}$$

Notar que el sumatorio es la variación cuadrática del movimiento Browniano entre t_i y t_{i+1}, por lo que por la observación 2.5 dicha variación es $t_{i+1} - t_i$. Y teniendo en cuenta que $\Delta(u) = \Delta(t_j)$ para $u \in [t_i, t_{i+1}]$, tenemos:

$$\Delta^2(t_i) \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} [W(s_{j+1}) - W(s_j)]^2 = \Delta^2(t_i)(t_{i+1} - t_i) = \int_{t_i}^{t_{i+1}} \Delta^2(u) du$$

Para los tiempos t_k y t haciéndolo de forma análoga quedaría $\int_{t_k}^{t} \Delta^2(u)du$. Y sumando todas las variaciones de los distintos intervalos obtenemos como queríamos que:

$$[I, I](t) = \int_0^t \Delta^2(u) du$$

■

Observación 3.2. Notar que al igual que ya hemos hecho antes y siguiendo de la misma manera podemos escribir $dIdI = \Delta^2(t)dt$. Fijarse que ya desde la ecuación (3.1) podíamos proceder de una manera más informal para llegar a la misma conclusión $dIdI = \Delta^2(t)dW(t)dW(t) \overset{2,5}{=} \Delta^2(t)dt$.

3.1.1. Integral de Itô en el caso general

Recordamos que en lo anterior se estaba tratando con funciones $\Delta(t)$ simples. Ahora $\Delta(t)$ tomará el papel de un proceso estocástico cualquiera, es decir, no tiene por qué ser constante a trozos, simplemente asumimos que $\{\Delta(t) : 0 \leq t \leq T\}$ es un proceso estocástico adaptado a la filtración $\{\mathcal{F}(t) : 0 \leq t \leq T\}$, y también asumiremos por razones de convergencia que pertenece a $L^2(\Omega, \mathcal{F}, \mathbb{P})$, donde nos referimos a la convergencia en media cuadrática. Es decir, asumimos que $\mathbb{E}\left[\int_0^T \Delta^2(t)dt\right] < \infty$. Vamos a construir ahora una sucesión de procesos adaptados simples que los denotaremos como $\Delta_n(t)$ tales que aproximen a $\{\Delta(t) : 0 \leq t \leq T\}$. Definimos la partición $\Pi_1 = \{t_0, t_1, ..., t_n\}$ y escribimos el proceso a trozos para que se tome el valor de la izquierda en cada intervalo de la partición, es decir, $\Delta_1(t) = \Delta(t_i)$ $\forall t \in [t_i, t_{i+1})$. Siguiendo el mismo procedimiento con una secuencia creciente de particiones $(\Pi_n)_{n\in\mathbb{N}}$, es decir, con $\Pi_n \subset \Pi_{n+1}$, es fácil demostrar que la serie creada converge en $L^2(\Omega, \mathcal{F}, \mathbb{P})$, es decir:

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^T |\Delta_n(t) - \Delta(t)|^2 \, dt\right] = 0 \tag{3.2}$$

Ahora, sea $I_n(t) = \int_0^T \Delta_n(u)dW(u)$ una integral de Itô para procesos simples definida anteriormente en 3.2, demostraremos que la serie $(I_n)_{n\in\mathbb{N}}$ es de Cauchy (usando la norma L^2). Es decir, queremos demostrar que $\mathbb{E}[(I_n(t) - I_m(t))^2]$ tiende a cero. Notar que $I_n(t) - I_m(t) = \int_0^T \Delta_n(u) - \Delta_m(u)dW(u)$ que sigue siendo una integral de Itô, ya que $\Delta_n(u) - \Delta_m(u)$ sigue siendo un proceso simple. Luego, por la isometría de Itô (teorema 3.2) tenemos que:

$$\mathbb{E}[(I_n(t) - I_m(t))^2] = \mathbb{E}\left[\int_0^t |\Delta_n(t) - \Delta_m(t)|^2 \, dt\right]$$

Cuya parte de la derecha tiende a cero, debido a que la serie $(\Delta_n(t))_{n\in\mathbb{N}}$ converge como se ha mencionado antes, concluyendo así que $(I_n)_{n\in\mathbb{N}}$ es una sucesión de Cauchy, y como $L^2(\Omega, \mathcal{F}, \mathbb{P})$ es completo, esta sucesión converge, y a lo que converge es lo que definimos como integral de Itô para un proceso adaptado cualquiera.

Definición 3.3. ([9]) Sea $\{W(t) : t \geq 0\}$ un movimiento Browniano con la filtración $\{\mathcal{F}(t) : t \geq 0\}$ y sea $\{\Delta(t) : 0 \leq t \leq T\}$ un proceso simple adaptado, que cumpla que $\mathbb{E}\left[\int_0^T \Delta^2(t)dt\right] < \infty$. Y sea la sucesión de procesos simples $(\Delta_n(t))_{n\in\mathbb{N}}$ tal y como se ha definido en lo anterior. Definimos la *Integral de Itô* como:

$$\int_0^t \Delta(u)dW(u) = \lim_{n\to\infty} \int_0^t \Delta_n(u)dW(u).$$

Por lo visto en esta sección, a partir de ahora, cada vez que hagamos referencia a la integral de Itô de un proceso adaptado cualquiera, vamos a tener que suponer que dicho proceso cumple con la hipótesis de esperanza cuadrática finita $\left(\mathbb{E}\left[\int_0^T \Delta^2(t)dt\right] < \infty\right)$

Observación 3.3. A uno le puede recordar la manera en que hemos construido la serie $(\Delta_n(t))_{n\in\mathbb{N}}$ y cómo hemos definido la integral pasando al límite, a lo que se suele explicar cuando se enseña la integral de Riemann. Sin embargo, con Riemann para construir la serie que aproxima da igual tomar los valores de la izquierda (como hemos hecho nosotros), los del medio, los de la derecha o cualquier otro. En nuestro caso podríamos haber seguido otro camino y haber tomado el punto medio, por ejemplo, lo que se conoce entonces con el nombre de la integral de Stratonovich, y que tiene ventajas como que se le puedan aplicar las reglas más ordinales de cálculo. Sin embargo, es inapropiado para las finanzas. En finanzas, el integrando representa una posición en un activo y el integrador representa el precio de ese activo. No podemos decidir a la 1:00 p.m. qué posición tomamos a las 9:00 a.m. Debemos decidir la posición al comienzo de cada intervalo de tiempo, y la integral de Itô es el límite de la ganancia obtenida por ese tipo de operaciones a medida que el tiempo entre operaciones se aproxima a cero. En nuestro caso, para funciones que tengan derivada, daría el mismo límite, como es el caso con el símil con Riemann, pero en cuanto el proceso sobre el que integremos no tenga varianza cuadrática nula la integral de Itô y la de Stratonovich difieren en general, y de hecho esta última no tiene por qué cumplir que sea una martingala cosa que con la de Itô sí que se tiene.

3.2. Procesos de Itô

Definición 3.4. ([9]) Sea $\{W(t) : t \geq 0\}$ un movimiento Browniano y $\{\mathcal{F}(t) : t \geq 0\}$ una filtración para $\{W(t) : t \geq 0\}$. Un *proceso de Itô* es un proceso estocástico de la forma

$$X(t) = X(0) + \int_0^t \Delta(u)dW(u) + \int_0^t \Theta(u)du,$$

donde $X(0)$ es no aleatorio y $\{\Delta(u) : u \geq 0\}$, $\{\Theta(u) : u \geq 0\}$ son procesos estocásticos adaptados. Suponemos que las integrales existen y, en particular, que $\mathbb{E}\left[\int_0^T \Delta^2(t)dt\right] < \infty$ (Por lo visto anteriormente sobre convergencia)

Nótese que un movimiento Browniano es un proceso de Itô, debido a que $W(0) = 0$ y tomamos $\Delta(u)$ constantemente uno y $\Theta(u)$ constantemente cero.

Definición 3.5. ([9]) *(Integral con respecto a un Proceso de Itô)*

Sea $\{X(t) : t \geq 0\}$ un proceso de Itô, y sea $\{\Gamma(t) : t \geq 0\}$ un proceso adaptado. Entonces, la *integral con respecto al proceso de Itô* se define como

$$\int_0^t \Gamma(u)dX(u) = \int_0^t \Gamma(u)\Delta(u)dW(u) + \int_0^t \Gamma(u)\Theta(u)du.$$

Suponemos que las integrales anteriores existen y, en particular, que $\mathbb{E}\left[\int_0^T \Delta^2(t)dt\right] < \infty$.

3.3. Fórmula de Itô-Doeblin

Teorema 3.4. *([9]) **(Fórmula de Itô-Doeblin)***

Sea $f(t,x)$ una función con derivadas parciales continuas $\frac{\partial}{\partial t}f(t,x)$, $\frac{\partial}{\partial x}f(t,x)$ y $\frac{\partial^2}{\partial x^2}f(t,x)$. Y sea $\{X(t) : t \geq 0\}$ un proceso de Itô como en la Definición 3.4, se tiene que:

$$\begin{aligned} f(T,X(T)) =& f(0,X(0)) + \int_0^T \frac{\partial}{\partial t}f(t,X(t))dt + \int_0^T \frac{\partial}{\partial x}f(t,X(t))dX(t) \\ &+ \frac{1}{2}\int_0^T \frac{\partial^2}{\partial x^2}f(t,X(t))d[X,X](t). \end{aligned}$$

Demostración. Consideramos el intervalo $[0,T]$ y $\Pi = \{t_0, t_1, \cdots, t_n\}$ una partición de dicho intervalo y recordamos que $\|\Pi\| := \max_{0\leq i\leq(n-1)} |t_{i+1} - t_i|$. Podemos escribir:

$$f(T,X(T)) - f(0,X(0)) = \sum_{i=0}^{n-1} f\left(t_{i+1}, X(t_{i+1})\right) - f\left(t_i, X(t_i)\right)$$

Usando ahora el desarrollo de Taylor en dos dimensiones, desarrollando cada $f\left(t_{i+1}, X(t_{i+1})\right)$ en $(t_i, X(t_i))$, es decir

$$\begin{aligned} f(t,x) = f(t_i,X(t_i)) &+ \left[\frac{\partial}{\partial t}f\left(t,x\right)\left(t_{i+1}-t_i\right) + \frac{\partial}{\partial x}f\left(t,x\right)\left(X(t_{i+1}) - X(t_i)\right)\right] \\ &+ \frac{1}{2}\left[\frac{\partial^2}{\partial x^2}f\left(t,x\right)\left(X(t_{i+1}) - X(t_i)\right)^2 + \frac{\partial^2}{\partial t\partial x}f\left(t,x\right)\left(t_{i+1}-t_i\right)\left(X(t_{i+1}) - X(t_i)\right)\right. \\ &+ \left.\frac{\partial^2}{\partial t^2}f\left(t,x\right)\left(t_{i+1}-t_i\right)^2\right] + \text{términos de orden superior.} \end{aligned}$$

Que sustituyendo en el punto $(t_i, X(t_i))$, donde por abreviar notación indicamos $\frac{\partial}{\partial t}f(t_i, X(t_i))$ en vez de $\frac{\partial}{\partial t}f(t,x)\Big|_{t=t_i;\, x=X(t_i)}$ tenemos:

$$\begin{aligned} &f(t_i, X(t_i) - f(t_i, X(t_i)) = \\ &= \left[\frac{\partial}{\partial t}f\left(t_{i+1}, X(t_{i+1})\right)\left(t_{i+1}-t_i\right) + \frac{\partial}{\partial x}f\left(t_{i+1}, X(t_{i+1})\right)\left(X(t_{i+1}) - X(t_i)\right)\right] \\ &+ \frac{1}{2}\left[\frac{\partial^2}{\partial x^2}f\left(t_{i+1}, X(t_{i+1})\right)\left(X(t_{i+1}) - X(t_i)\right)^2\right. \\ &+ \frac{\partial^2}{\partial t\partial x}f\left(t_{i+1}, X(t_{i+1})\right)\left(t_{i+1}-t_i\right)\left(X(t_{i+1}) - X(t_i)\right) \\ &+ \left.\frac{\partial^2}{\partial t^2}f\left(t_{i+1}, X(t_{i+1})\right)\left(t_{i+1}-t_i\right)^2\right] + \text{términos de orden superior.} \end{aligned}$$

Sustituyendo ahora en el sumatorio:

$$\begin{aligned}\sum_{i=0}^{n-1} f\left(t_{i+1}, X(t_{i+1})\right) - f\left(t_i, X(t_i)\right) = &\sum_{i=0}^{n-1} \frac{\partial}{\partial t} f\left(t_i, X(t_i)\right)\left(t_{i+1}-t_i\right)\\
+&\sum_{i=0}^{n-1} \frac{\partial}{\partial x} f\left(t_i, X(t_i)\right)\left(X(t_{i+1})-X(t_i)\right)\\
+&\frac{1}{2}\sum_{i=0}^{n-1} \frac{\partial^2}{\partial x^2} f\left(t_i, X(t_i)\right)\left(X(t_{i+1})-X(t_i)\right)^2\\
+&\frac{1}{2}\sum_{i=0}^{n-1} \frac{\partial^2}{\partial t\partial x} f\left(t_i, X(t_i)\right)\left(t_{i+1}-t_i\right)\left(X(t_{i+1})-X(t_i)\right)\\
+&\frac{1}{2}\sum_{i=0}^{n-1} \frac{\partial^2}{\partial t^2} f\left(t_i, X(t_i)\right)\left(t_{i+1}-t_i\right)^2 + \text{términos de orden superior.}\end{aligned}$$

Ahora, si hacemos tender $\|\Pi\| \to 0$ tenemos:

- Primer termino:

$$\lim_{\|\Pi\|\to 0} \sum_{i=0}^{n-1} \frac{\partial}{\partial t} f\left(t_i, X(t_i)\right)\left(t_{i+1}-t_i\right) = \int_0^T \frac{\partial}{\partial t} f(t, X(t))dt$$

- Segundo termino:

$$\lim_{\|\Pi\|\to 0} \sum_{i=0}^{n-1} \frac{\partial}{\partial x} f\left(t_i, X(t_i)\right)\left(X(t_{i+1})-X(t_i)\right) = \int_0^T \frac{\partial}{\partial x} f(t, X(t))dX(t)$$

- Tercer termino:

$$\lim_{\|\Pi\|\to 0} \frac{1}{2}\sum_{i=0}^{n-1} \frac{\partial^2}{\partial x^2} f\left(t_i, X(t_i)\right)\left(X(t_{i+1})-X(t_i)\right)^2 = \frac{1}{2}\int_0^T \frac{\partial^2}{\partial x^2} f(t, X(t))d[X,X](t)$$

- Cuarto termino:

$$\begin{aligned}&\lim_{\|\Pi\|\to 0} \sum_{i=0}^{n-1} \frac{\partial^2}{\partial t\partial x} f\left(t_i, X(t_i)\right)\left(t_{i+1}-t_i\right)\left(X(t_{i+1})-X(t_i)\right)\\
\leq &\lim_{\|\Pi\|\to 0} \max_{0\leq i\leq (n-1)} \left(X(t_{i+1})-X(t_i)\right) \sum_{i=0}^{n-1} \frac{\partial^2}{\partial t\partial x} f\left(t_i, X(t_i)\right)\left(t_{i+1}-t_i\right)\\
= &0 \cdot \int_0^T \frac{\partial^2}{\partial t\partial x} f(t, X(t))dt = 0\end{aligned}$$

- Quinto termino:

$$\begin{aligned}&\lim_{\|\Pi\|\to 0} \frac{1}{2}\sum_{i=0}^{n-1} \frac{\partial^2}{\partial t^2} f(t_i, X(t_i))(t_{i+1}-t_i)^2\\
\leq &\lim_{\|\Pi\|\to 0} \frac{1}{2}\|\Pi\| \sum_{i=0}^{n-1} \frac{\partial^2}{\partial t^2} f(t_i, X(t_i))(t_{i+1}-t_i)\\
= &0 \cdot \int_0^T \frac{\partial^2}{\partial t^2} f(t, X(t))dt\end{aligned}$$

Del mismo modo nos podemos deshacer de los términos de mayor órdenes sacando diferenciales de t $(t_{i+1} - t_i)$ o de X $(X(t_{i+1}) - X(t_i))$ según corresponda.

Luego:

$$\begin{aligned} f(T, X(T)) - f(0, X(0)) = &\int_0^T \frac{\partial}{\partial t} f(t, X(t))dt + \int_0^T \frac{\partial}{\partial x} f(t, X(t))dX(t) \\ &+ \frac{1}{2}\int_0^T \frac{\partial^2}{\partial x^2} f(t, X(t))d[X, X](t). \end{aligned}$$

Como queríamos. ■

Observación 3.4. En su forma diferencial también se puede escribir como:

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

Y aún más, si queremos dejarlo en función de dW y dt,de manera informal solo hay que usar la expresión diferencial del proceso de Itô ($dX(t) = \Delta(t)dW + \Theta(t)dt$) y sustituir la variación cuadrática ($dX(t)dX(t) = \Delta(t)dt$), quedando así:

$$\begin{aligned} df(t, X(t)) = &\frac{\partial}{\partial t} f(t, X(t))\, dt + \frac{\partial}{\partial x} f(t, X(t))\Delta(t)\, dW(t) + \frac{\partial}{\partial x} f(t, X(t))\Theta(t)\, dt \\ &+ \frac{1}{2}\frac{\partial^2}{\partial x^2} f(t, X(t))\Delta^2(t)\, dt \end{aligned}$$

Observación 3.5. Recalcar que del teorema anterior podemos escribir informalmente $dtdt = 0$ y $dtdX(t) = dX(t)dt = 0$

Teorema 3.5. *([9]) **(Regla del producto)** Sea $X(t)$ y $Y(t)$ dos procesos de Itô, entonces:*

$$d(X(t)Y(t)) = X(t)dY(t) + Y(t)dX(t) + dX(t)dY(t)$$

Demostración. De manera similar al anterior teorema, aunque ahora lo haremos con la notación de diferenciales para abreviar, utilizamos la expansión de Taylor en 3 dimensiones. Dado $f(t, x, y)$ con derivadas parciales continuas $\frac{\partial}{\partial t} f(t, x, y)$, $\frac{\partial}{\partial x} f(t, x, y)$, $\frac{\partial}{\partial y} f(t, x, y)$, $\frac{\partial^2}{\partial x^2} f(t, x, y)$, $\frac{\partial^2}{\partial y^2} f(t, x, y)$, $\frac{\partial^2}{\partial x \partial y} f(t, x, y)$ y $\frac{\partial^2}{\partial y \partial x} f(t, x, y)$, se cumple que:

$$\begin{aligned} df(t, X(t), Y(t)) = &\frac{\partial}{\partial t} f(t, X(t), Y(t))dt + \frac{\partial}{\partial x} f(t, X(t), Y(t))dX(t) + \frac{\partial}{\partial y} f(t, X(t), Y(t))dY(t) \\ &+ \frac{1}{2}\frac{\partial^2}{\partial x^2} f(t, X(t), Y(t))dX(t)X(t) + \frac{1}{2}\frac{\partial^2}{\partial y^2} f(t, X(t), Y(t))dY(t)Y(t) \\ &+ \frac{\partial^2}{\partial x \partial y} f(t, X(t), Y(t))dX(t)dY(t) \cdots \end{aligned} \tag{3.3}$$

Ahora, sea $f(t, x, y) = xy$, utilizamos (3.3):

$$\begin{aligned} d(X(t)Y(t)) = df(t, X(t), Y(t)) &= 0 + Y(t)dX(t) + X(t)dY(t) + 0 + 0 + dX(t)dY(t) \\ &= X(t)dY(t) + Y(t)dX(t) + dX(t)dY(t). \end{aligned}$$

■

Capítulo 4

Ecuación y fórmula de Black-Scholes

4.1. La ecuación de Black-Scholes

En esta sección obtendremos la ecuación diferencial parcial usada para poder valorar una opción basada en un activo financiero modelado por un movimiento Browniano geométrico, a esta EDP se la conoce con el nombre de ecuación de Black-Scholes o Black-Scholes-Merton (de forma abreviada también llamada ecuación de B-S). Para en posteriores secciones obtener la fórmula de Black-Scholes(-Merton).

La idea subyacente con lo que vamos a proceder es la misma que en el ejemplo anterior a la definición 2.8, consistente básicamente en determinar el capital inicial necesario para cubrir/asegurar perfectamente una posición corta en la opción.

Suponemos que tenemos una cartera replicante cuyo valor es $X(t)$ para cada instante de tiempo t, donde en él se invierte (ya sea en una posición larga o corta) en el mercado monetario con un interés constante r, y en un activo financiero cuyo precio en el instante t es S(t), que viene modelado por el movimiento Browniano geométrico. Recordamos de la definición 2.23 que entonces:

$$S(t) = S(0)e^{\sigma W(t)+(\alpha-\frac{1}{2}\sigma^2)t}$$

Donde $\alpha > 0$, σ son constantes y $\{W(t) : t \geq 0\}$ es un movimiento Browniano. Usando la fórmula de Itô-Doeblin (3.4) en su forma diferencial (observación 3.4), siendo $f(t,x) = S(0)e^{\sigma x+(\alpha-\frac{1}{2}\sigma^2)t}$ obtenemos la expresión diferencial para el movimiento Browniano geométrico:

$$dS(t) = df(t,W(t)) = (\alpha - \frac{1}{2}\sigma^2)S(t)dt + \sigma S(t)dW(t) + \frac{1}{2}\sigma^2 S(t)dt = \alpha S(t)dt + \sigma S(t)dW(t) \tag{4.1}$$

Lo que será utilizado más tarde.

Vamos ahora a recordar como se formaba la cartera $X(t)$, en cada tiempo t tenemos $\Delta(t)$ acciones, y el dinero restante (o el que nos haga falta, recordar que podemos tomar posiciones cortas) en el mercado monetario con un interés constante r. Luego el cambio en el valor de la cartera viene dado por el cambio en el valor de las acciones que poseemos $\Delta(t)dS(t)$ y en las ganancias (o intereses del préstamo) $r(X(t) - \Delta(t)S(t))dt$, luego:

$$dX(t) = \Delta(t)dS(t) + r(X(t) - \Delta(t)S(t))dt$$

Y utilizando la expresión diferencial del movimiento Browniano (4.1), sustituimos en lo anterior:

$$
\begin{aligned}
dX(t) &= \Delta(t)dS(t) + r(X(t) - \Delta(t)S(t))dt \\
&= \Delta(t)(\alpha S(t)dt + \sigma S(t)dW(t)) + r(X(t) - \Delta(t)S(t))dt \\
&= \Delta(t)(\alpha - r)S(t)dt + \Delta(t)\sigma S(t)dW(t) + rX(t)dt
\end{aligned} \tag{4.2}
$$

Otra expresión que vamos a necesitar es la expresión diferencial de los precios de las acciones $S(t)$ y de la cartera $X(t)$ actualizados (en valor presente, no en moneda corriente), es decir, $e^{-rt}S(t)$ y $e^{-rt}X(t)$. Esto lo obtenemos usando la fórmula de Itô-Doeblin, con $f(t,x) = e^{-rt}x$

$$
\begin{aligned}
d\left(e^{-rt}S(t)\right) &= \frac{\partial}{\partial t}e^{-rt}x\bigg|_{x=S(t)} dt + \frac{\partial}{\partial x}e^{-rt}x\bigg|_{x=S(t)} dS(t) + \frac{1}{2}\,\frac{\partial^2}{\partial x^2}e^{-rt}x\bigg|_{x=S(t)} dS(t)dS(t) \\
&= -re^{-rt}S(t)dt + e^{-rt}dS(t) \\
&\overset{(4,1)}{=} -re^{-rt}S(t)dt + e^{-rt}\left(\alpha S(t)dt + \sigma S(t)dW(t)\right) \\
&= (\alpha - r)e^{-rt}S(t)dt + \sigma e^{-rt}S(t)dW(t)
\end{aligned} \tag{4.3}
$$

Del mismo modo:

$$
\begin{aligned}
d\left(e^{-rt}X(t)\right) &= \frac{\partial}{\partial t}e^{-rt}x\bigg|_{x=X(t)} dt + \frac{\partial}{\partial x}e^{-rt}x\bigg|_{x=X(t)} dX(t) + \frac{1}{2}\,\frac{\partial^2}{\partial x^2}e^{-rt}x\bigg|_{x=X(t)} dX(t)dX(t) \\
&= -re^{-rt}X(t)dt + e^{-rt}dX(t) \\
&\overset{(4,2)}{=} -re^{-rt}X(t)dt + e^{-rt}\left(\Delta(t)(\alpha - r)S(t)dt + \Delta(t)\sigma S(t)dW(t) + rX(t)dt\right) \\
&= \Delta(t)(\alpha - r)e^{-rt}S(t)dt + \Delta(t)\sigma e^{-rt}S(t)dW(t) \\
&\overset{(4,3)}{=} \Delta(t)d\left(e^{-rt}S(t)\right)
\end{aligned} \tag{4.4}
$$

Ahora tenemos en mente que queremos valorar una opción de compra Europea (definición 2.1) con precio de ejercicio K, es decir, en el tiempo T paga $(S(T) - K)^+$. Black, Scholes y Merton argumentaron que el valor de esta opción de compra en cualquier momento debería depender del tiempo (más precisamente, del tiempo hasta el vencimiento) y del valor del precio de la acción en ese momento. Por supuesto, también debería depender de los parámetros del modelo, r, σ y α, así como del precio de ejercicio K.

Sin embargo, solo dos de estas, el tiempo y el precio de la acción, son variables. Siguiendo este razonamiento, definimos $c(t,x)$ como el valor de la opción de compra en el momento t, si el precio de la acción en ese instante es $S(t) = x$.

No hay nada aleatorio en la función $c(t,x)$. Sin embargo, el valor de la opción sí es aleatorio, es decir, al sustituir $S(t)$ como x en $c(t,S(t))$ obtenemos un proceso estocástico. En el tiempo inicial, no conocemos los precios futuros de las acciones $S(t)$ y, por lo tanto, no conocemos los valores futuros de la opción $c(t,S(t))$. Nuestro objetivo es determinar la función $c(t,x)$ para que al menos tengamos una fórmula para los valores futuros de la opción en términos de los precios futuros de las acciones.

Comenzamos calculando el diferencial de $c(t, S(t))$. Según la fórmula de Itô-Doeblin, se tiene:

$$dc(t, S(t)) = \frac{\partial}{\partial t} c(t, x) dt \Big|_{x=S(t)} + \frac{\partial}{\partial x} c(t, x) dS(t) \Big|_{x=S(t)} + \frac{1}{2} \frac{\partial^2}{\partial x^2} c(t, x) dS(t) dS(t) \Bigg|_{x=S(t)} \tag{4.5}$$

Usando (4.1), la observación 3.5 ($dWdt = 0$) y 2.5(dW dW=dt):

$$\begin{aligned} dS(t)dS(t) &= \alpha^2 S^2(t) dt dt + 2\sigma\alpha S(t) dW(t) dt + \sigma^2 S^2(t) dW(t) dW(t) \\ &= \sigma^2 S^2(t) dt \end{aligned} \tag{4.6}$$

Por lo tanto, volviendo a la expresión (4.5) sustituyendo $dS(t)dS(t)$ por lo que acabamos de obtener, tendríamos:

$$\begin{aligned} dc(t, S(t)) &= \frac{\partial}{\partial t} c(t, x) dt \Big|_{x=S(t)} + \frac{\partial}{\partial x} c(t, x) dS(t) \Big|_{x=S(t)} + \frac{1}{2} \frac{\partial^2}{\partial x^2} c(t, x) \sigma^2 S^2(t) dt \Bigg|_{x=S(t)} \\ &\overset{(4,1) \text{ y } (4,6)}{=} \frac{\partial}{\partial t} c(t, x) dt \Big|_{x=S(t)} + \frac{\partial}{\partial x} c(t, x) (\alpha S(t) dt + \sigma S(t) dW(t)) \Big|_{x=S(t)} \\ &+ \frac{1}{2} \frac{\partial^2}{\partial x^2} c(t, x) \sigma^2 S^2(t) dt \Bigg|_{x=S(t)} = \left(\frac{\partial}{\partial t} c(t, x) + \alpha S(t) \frac{\partial}{\partial x} c(t, x) + \frac{1}{2} \sigma^2 S^2(t) \frac{\partial^2}{\partial x^2} c(t, x) \right) dt \Bigg|_{x=S(t)} \\ &+ \sigma S(t) \frac{\partial}{\partial x} c(t, x) dW(t) \Big|_{x=S(t)} \end{aligned}$$

Y por conveniencia, escribimos:

$$\begin{aligned} dc(t, S(t)) = &\left(\frac{\partial}{\partial t} c(t, S(t)) + \alpha S(t) \frac{\partial}{\partial x} c(t, S(t)) + \frac{1}{2} \sigma^2 S^2(t) \frac{\partial^2}{\partial x^2} c(t, S(t)) \right) dt \\ &+ \sigma S(t) \frac{\partial}{\partial x} c(t, S(t)) dW(t) \end{aligned} \tag{4.7}$$

Análogamente, para el precio descontado de la opción:

$$\begin{aligned} d(e^{-rt} c(t, S(t))) = &-re^{-rt} c(t, x) + e^{-rt} \frac{\partial}{\partial t} c(t, x) dt \Big|_{x=S(t)} + e^{-rt} \frac{\partial}{\partial x} c(t, x) dS(t) \Big|_{x=S(t)} \\ &+ \frac{1}{2} e^{-rt} \frac{\partial^2}{\partial x^2} c(t, x) \sigma^2 S^2(t) dt \Bigg|_{x=S(t)} \overset{(4,1)}{=} -re^{-rt} c(t, x) + e^{-rt} \frac{\partial}{\partial t} c(t, x) dt \Big|_{x=S(t)} \\ &+ e^{-rt} \frac{\partial}{\partial x} c(t, x) (\alpha S(t) dt + \sigma S(t) dW(t)) \Big|_{x=S(t)} \\ &+ e^{-rt} \frac{1}{2} \frac{\partial^2}{\partial x^2} c(t, x) \sigma^2 S^2(t) dt \Bigg|_{x=S(t)} \end{aligned}$$

Que reordenando y expresándolo de manera abreviada como en el anterior caso es:

$$\begin{aligned} d(e^{-rt} c(t, S(t))) = e^{-rt} &\left(-rc(t, S(t)) + \frac{\partial}{\partial t} c(t, S(t)) + \alpha S(t) \frac{\partial}{\partial x} c(t, S(t)) \right. \\ &\left. + \frac{1}{2} \sigma^2 S^2(t) \frac{\partial^2}{\partial x^2} c(t, S(t)) \right) dt + e^{-rt} \sigma S(t) \frac{\partial}{\partial x} c(t, S(t)) dW(t) \end{aligned} \tag{4.8}$$

Ahora miramos la cartera replicante comentada al principio del capítulo, que con un capital inicial $X(0)$, el valor en cada instante t, es decir, $X(t)$ tiene que coincidir con el de la opción $c(t, S(t))$.

Pero no vamos a igualar esas expresiones directamente, sino su valor presente calculado en las expresiones (4.3) y (4.4). Esto se debe a que, como puede verse en la fórmula (4.4) podemos expresar $d(e^{-rt}X(t))$ de manera simple y directa en función de $d(e^{-rt}S(t))$, lo que facilitará considerablemente los cálculos. Además, trabajar en términos descontados tiene ventajas conceptuales, ya que permite representar el comportamiento del mercado. Sin embargo, estas consideraciones no las abordaremos aquí. Es decir, igualamos $e^{-rt}X(t) = e^{-rt}c(t, S(t))$ $\forall t \in [0, T]$, o de manera equivalente:

$$\begin{cases} d\left(e^{-rt}X(t)\right) = d\left(e^{-rt}c(t, S(t))\right) \\ X(0) = c(0, S(0)) \end{cases}$$

Ya que al integrar lo anterior nos dice que $e^{-rt}X(t) - X(0) = e^{-rt}c(t, S(t)) - c(0, S(0)), \quad \forall t \in [0, T)$, y si $X(0) = c(0, S(0))$ se cancelan y nos queda la igualdad deseada.

Una expresión de $d\left(e^{-rt}c(t, S(t))\right)$ la tenemos en la ecuación (4.8), con ello, desarrollamos ahora $d\left(e^{-rt}X(t)\right)$:

$$d\left(e^{-rt}X(t)\right) \overset{(4,4)}{=} \Delta(t) d\left(e^{-rt}S(t)\right) \overset{(4,3)}{=} \Delta(t)\left((\alpha - r)e^{-rt}S(t)dt + \sigma e^{-rt}S(t)dW(t)\right) \tag{4.9}$$

Como la ecuación (4.9) y (4.8) tienen que ser iguales, igualando los sumandos correspondientes a $dW(t)$ y dt, tenemos, por un lado, que:

$$\Delta(t)\sigma e^{-rt}S(t)dW(t) = e^{-rt}\sigma S(t)\frac{\partial}{\partial x}c(t, S(t))dW(t)$$

Luego

$$\Delta(t) = \frac{\partial}{\partial x}c(t, S(t)) \quad \forall t \in [0, T) \tag{4.10}$$

Por el otro lado,

$$\begin{aligned} \Delta(t)(\alpha - r)e^{-rt}S(t)dt = e^{-rt}\Bigg(&-rc(t, S(t)) + \frac{\partial}{\partial t}c(t, S(t)) + \alpha S(t)\frac{\partial}{\partial x}c(t, S(t)) \\ &+\frac{1}{2}\sigma^2 S^2(t)\frac{\partial^2}{\partial x^2}c(t, S(t))\Bigg)dt \end{aligned}$$

Sustituyendo $\Delta(t)$ por la expresión en (4.10):

$$\begin{aligned} &e^{-rt}\left(\alpha S(t)\frac{\partial}{\partial x}c(t, S(t)) - rS(t)\frac{\partial}{\partial x}c(t, S(t))\right)dt = \\ &e^{-rt}\left(-rc(t, S(t)) + \frac{\partial}{\partial t}c(t, S(t)) + \alpha S(t)\frac{\partial}{\partial x}c(t, S(t)) + \frac{1}{2}\sigma^2 S^2(t)\frac{\partial^2}{\partial x^2}c(t, S(t))\right)dt \end{aligned}$$

Cancelando y reordenando:

$$\frac{\partial}{\partial t}c(t, S(t)) + rS(t)\frac{\partial}{\partial x}c(t, S(t))) + \frac{1}{2}\sigma^2 S^2(t)\frac{\partial^2}{\partial x^2}c(t, S(t)) = rc(t, S(t))$$

A esta última expresión es a la que se la conoce como la ecuación de Black-Scholes-Merton.

Recordamos también por definición de una opción de compra (definición 2.1) también se tiene que cumplir que $c(T, S(T)) = (S(T) - K)^+$. Es decir, que el sistema a resolver sería:

$$\begin{cases} \frac{\partial}{\partial t}c(t,S(t)) + rx\frac{\partial}{\partial x}c(t,S(t))) + \frac{1}{2}\sigma^2 S^2(t)\frac{\partial^2}{\partial x^2}c(t,S(t)) = rc(t,S(t)) \\ c(T,x) = (x-K)^+ \end{cases}$$

4.2. Teorema de Girsanov

Teorema 4.1. *([9]) **(Teorema de Girsanov)***

Consideremos el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. Sea $\{W(t) : 0 \leq t \leq T\}$ un movimiento Browniano, sea $\{\mathcal{F}(t) : 0 \leq t \leq T\}$ una filtración para $\{W(t) : 0 \leq t \leq T\}$ y sea $\{\Theta(t) : 0 \leq t \leq T\}$ un proceso adaptado.

Definimos:

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

Sea $Z = Z(T)$, definimos

$$\tilde{\mathbb{P}}(A) = \int_A Z d\mathbb{P}.$$

Entonces, $\mathbb{E}[Z] = 1$ y $\widetilde{W}(t)$ es un $\tilde{\mathbb{P}}$-movimiento Browniano.

Para que la integral de Itô converja correctamente, asumimos que $\mathbb{E}\left[\int_0^T (\Theta(t)Z(t))^2 dt\right] < \infty$

Demostración. Procederemos a demostrarlo por la caracterización de Levy (2.6). Llamamos $X(t) = -\int_0^t \Theta(u)dW(u) - \frac{1}{2}\int_0^t \Theta^2(u)du$ (el exponente de la expresión de $Z(t)$), que en forma diferencial sería $dX(t) = -\Theta(t)dW(t) - \frac{1}{2}\Theta^2(t)dt$, luego $dX(t)dX(t)$ sería:

$$\begin{aligned} dX(t)dX(t) &= \left(-\Theta(t)dW(t) - \frac{1}{2}\Theta^2(t)dt\right)^2 \\ &= \Theta^2(t)dW(t)dW(t) + \frac{1}{2}\Theta^3(t)dtdW(t) + \frac{1}{4}\Theta^4(t)dtdt \overset{(3,5),(2,5)}{=} \Theta^2(t)dt \end{aligned} \tag{4.11}$$

Tomando ahora $f(t,x) = e^x$ para aplicar la fórmula de Itô-Doeblin tenemos:

$$\begin{aligned} dZ = d(e^{X(t)}) &= \frac{\partial}{\partial x}e^x dX(t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}e^x dX(t)dX(t)\bigg|_{x=X(t)} \\ &\overset{(4,11)}{=} e^{X(t)}\left(-\Theta(t)dW(t) - \frac{1}{2}\Theta^2(t)dt\right) + \frac{1}{2}e^{X(t)}\Theta^2(t)dt \\ &= -\Theta(t)Z(t)dW(t) \end{aligned} \tag{4.12}$$

Lo que nos dice que $Z(t) - Z(0) = -\int_0^t \Theta(t)Z(t)dW(t)$, es decir, $Z(t) = Z(0) - \int_0^t \Theta(t)Z(t)dW(t)$. Notar que la integral de la expresión es una integral de Itô, que ya hemos probado que es una martingala (Teorema 3.1 ampliado a procesos generales). Luego haciendo lo siguiente vemos que $Z(t)$ también es una $\mathbb{P}$-martingala:

$$\begin{aligned}\mathbb{E}[Z(t)|\mathcal{F}(s)] &= \mathbb{E}\left[Z(0) - \int_0^t \Theta(t)Z(t)dW(t)\middle|\mathcal{F}(s)\right] \\ &= \mathbb{E}\left[Z(0)|\mathcal{F}(s)\right] - \mathbb{E}\left[\int_0^t \Theta(t)Z(t)dW(t)\middle|\mathcal{F}(s)\right] \\ &= Z(0) - \int_0^s \Theta(t)Z(t)dW(t) = Z(s)\end{aligned} \tag{4.13}$$

Utilizando esto último, al ser $Z(t)$ una martingala tenemos que $Z(t) = \mathbb{E}[Z(T)|\mathcal{F}(t)] = \mathbb{E}[Z|\mathcal{F}(t)]$. Y aplicando la proposición 2.3 tenemos que $\mathbb{E}[Z] = \mathbb{E}[Z(T)] = \mathbb{E}[Z(0)] = \mathbb{E}[e^0] = 1$. Luego Z(t) es un proceso de Radon-Nykodým derivado como se define en 2.18.

Notar que según la expresión de $\widetilde{W}(t)$ tenemos que $d\widetilde{W}(t) = dW(t) + \Theta(t)dt$. Vamos a sustituir esto y la expresión de $dZ(t)$ de la ecuación (4.12) para el cálculo de $d\left(\widetilde{W}(t)Z(t)\right)$ en el que utilizamos la regla del producto del teorema 3.5.

$$\begin{aligned}&d\left(\widetilde{W}(t)Z(t)\right) = \widetilde{W}(t)dZ(t) + Z(t)d\widetilde{W}(t) + d\widetilde{W}(t)dZ(t) \\ &= -\widetilde{W}\Theta(t)Z(t)dW(t) + Z(t)\left(dW(t) + \Theta(t)dt\right) + \left(dW(t) + \Theta(t)dt\right)\left(-\Theta(t)Z(t)dW(t)\right) \\ &= -\widetilde{W}\Theta(t)Z(t)dW(t) + Z(t)dW(t) + Z(t)\Theta(t)dt - \Theta(t)Z(t)dW(t)dW(t) - \Theta^2(t)Z(t)dtdW(t)) \\ &= -\widetilde{W}\Theta(t)Z(t)dW(t) + Z(t)dW(t) + Z(t)\Theta(t)dt - \Theta(t)Z(t)dt \\ &= \left(-\widetilde{W}\Theta(t) + 1\right)Z(t)dW(t)\end{aligned}$$

Esta expresión que hemos obtenido resulta no tener término dt, luego análogamente al procedimiento utilizado para demostrar que Z(t) era una martingala, se demuestra que $\widetilde{W}(t)Z(t)$ también es una $\mathbb{P}$-martingala. Usando esto y la proposición 2.5

$$\tilde{\mathbb{E}}\left[\widetilde{W}(t)|\mathcal{F}(s)\right] \overset{(2{,}5)}{=} \frac{1}{Z(s)}\mathbb{E}\left[\widetilde{W}(t)Z(t)|\mathcal{F}(s)\right] = \frac{1}{Z(s)}\widetilde{W}(s)Z(s) = \widetilde{W}(s)$$

Luego $\widetilde{W}(t)$ es una $\tilde{\mathbb{P}}$-martingala. Ahora como también tenemos $\widetilde{W}(0) = W(0) = 0$ y que $d\widetilde{W}(t)\widetilde{W}(t) = (dW(t) + \Theta(t)dt)^2 = dW(t)dW(t) + \Theta(t)dW(t)dt + \Theta(t)dtdt = dt$, es decir se cumple que la variación cuadrática también es t. Concluimos por la caracterización de Levy (Teorema 2.6) que $\widetilde{W}(t)$ es un $\tilde{\mathbb{P}}$-movimiento Browniano.

■

4.3. La fórmula de Black-Scholes

Ahora nos disponemos, finalmente, a hallar la fórmula de Black-Scholes para valorar una opción de compra Europea.

Recordamos de la ecuación (4.3) que $d\left(e^{-rt}S(t)\right) = (\alpha - r)e^{-rt}S(t)dt + \sigma e^{-rt}S(t)dW(t)$ que reordenando podemos escribir como:

$$d\left(e^{-rt}S(t)\right) = \sigma e^{-rt}S(t)\left(\frac{(\alpha - r)}{\sigma}dt + dW(t)\right)$$

Llamando $\Theta = \frac{(\alpha - r)}{\sigma}$ podemos aplicar el teorema de Girsanov (4.1), en este caso con Θ constante, para obtener, entre otras cosas el $\widetilde{W}(t)$, el cual recordamos que su diferencial es $d\widetilde{W}(t) = \Theta dt + dW(t)$, con lo que sustituyendo en la expresión anterior tenemos que:

$$d\left(e^{-rt}S(t)\right) = \sigma e^{-rt}S(t)d\widetilde{W}(t) \tag{4.14}$$

De la misma expresión, $d\widetilde{W}(t) = \Theta dt + dW(t) = \frac{(\alpha-r)}{\sigma}dt + dW(t)$, al integrar entre 0 y t, obtenemos $\widetilde{W}(t) = \frac{(\alpha-r)}{\sigma}t + W(t)$ y despejando, llegamos a $\sigma W(t) = \sigma d\widetilde{W}(t) - (\alpha - r)t$, que lo sustituimos en la expresión del movimiento Browniano geométrico $S(t)$:

$$S(t) = S(0)e^{\sigma W(t)+(\alpha-\frac{1}{2}\sigma^2)t} = S(0)e^{\sigma\widetilde{W}(t)+(r-\frac{1}{2}\sigma^2)t} \tag{4.15}$$

Volviendo con la cartera replicante $X(t)$ explicada al principio del capítulo tenemos que:

$$d\left(e^{-rt}X(t)\right) \overset{(4,4)}{=} \Delta(t)d\left(e^{-rt}S(t)\right) \overset{(4,14)}{=} \Delta(t)\sigma e^{-rt}S(t)d\widetilde{W}(t)$$

Que observamos que solo tiene el término acompañando $d\widetilde{W}(t)$ sin dt, luego de la misma forma que vimos que $Z(t)$ era una martingala se ve que $e^{-rt}X(t)$ es una $\tilde{\mathbb{P}}$-martingala.

Recordamos que en secciones anteriores obtuvimos la ecuación de Black-Scholes para opciones de compra europeas donde el valor de la opción c(t, S(t)) nos lo daba el valor de la cartera replicante X(t), es decir, en concreto por ser una opción europea en el instante T el valor es $X(T) = (S(T) - K)^+$ o lo que es lo mismo $e^{-rt}X(T) = e^{-rt}(S(T) - K)^+$, y como por lo anterior, $e^{-rt}X(t)$ es una $\tilde{\mathbb{P}}$-martingala, tenemos:

$$e^{-rt}X(t) = \tilde{\mathbb{E}}\left[e^{-rT}X(T)\middle|\mathcal{F}(t)\right]$$

Y multiplicamos por e^{-rt} a ambos lados y podemos meter ese término dentro de la esperanza, ya que es $\mathcal{F}(t)$-medible. Resultando en:

$$X(t) = \tilde{\mathbb{E}}\left[e^{-r(T-t)}X(T)\middle|\mathcal{F}(t)\right]$$

Aunque se construya $X(t)$ como cartera replicante, no se puede garantizar que $X(t)$ sea una función determinista de t y $S(t)$, ya que únicamente se conoce la evolución de su diferencial.

Como hemos visto en el teorema 2.7 al ser S(t) un movimiento browniano geométrico también es un proceso de Markov y tomando como función adentro de la esperanza $f(T, S(T)) = e^{-r(T-t)}(S(T) - K)^+$, podemos asegurar que existe una función $g(t, S(t))$ tal que:

$$g(t, S(t)) = \tilde{\mathbb{E}}\left[e^{-r(T-t)}(S(T) - K)^+\middle|\mathcal{F}(t)\right] = \tilde{\mathbb{E}}\left[e^{-r(T-t)}X(T)\middle|\mathcal{F}(t)\right] \tag{4.16}$$

A esta función $g(t, S(t))$, cuya existencia acabamos de justificar, la denominamos a partir de ahora $c(t, S(t))$ para unificar la notación.

Esa función $g(t, S(t))$ que acabamos de demostrar que existe, la llamamos a partir de ahora $c(t, S(t))$, para unificar la notación.

Recordamos ahora la expresión de $S(t)$ obtenida en la ecuación (4.15), y denominamos $Y(t) = \sigma\widetilde{W}(t) + (r - \frac{1}{2}\sigma^2)t$, entonces:

$$\begin{aligned}S(T) &= S(0)e^{Y(T)} = S(0)e^{Y(T)+Y(t)-Y(t)} = S(0)e^{Y(t)}e^{Y(T)-Y(t)} = S(t)e^{Y(T)-Y(t)} \\ &= S(t)e^{\sigma\left(\widetilde{W}(T)-\widetilde{W}(t)\right)+\left(r-\frac{1}{2}\sigma^2\right)(T-t)}\end{aligned}$$

Ahora hacemos un cambio de variables y notación. Llamamos $\tau = T - t$ (es el tiempo hasta el vencimiento), y $Z = -\frac{\widetilde{W}(T)-\widetilde{W}(t)}{\sqrt{\tau}}$, que por definición de movimiento browniano $Z \sim \mathcal{N}(0,1)$ (propiedad 2 de la definición 2.20), y podemos escribir $S(T) = S(t)e^{-\sigma\sqrt{\tau}Z+\left(r-\frac{1}{2}\sigma^2\right)\tau}$. Retomando la expresión de (4.16) se tiene en función de x que:

$$c(t,x) = \tilde{\mathbb{E}}\left[e^{-r\tau}\left(xe^{-\sigma\sqrt{\tau}Z+\left(r-\frac{1}{2}\sigma^2\right)\tau} - K\right)^+\right]$$

Y como $Z \sim \mathcal{N}(0,1)$

$$\begin{aligned}c(t,x) &= \tilde{\mathbb{E}}\left[e^{-r\tau}\left(xe^{-\sigma\sqrt{\tau}Z+\left(r-\frac{1}{2}\sigma^2\right)\tau} - K\right)^+\right] \\ &= \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-r\tau}\left(xe^{-\sigma\sqrt{\tau}z+\left(r-\frac{1}{2}\sigma^2\right)\tau} - K\right)^+ e^{-\frac{1}{2}z^2}dz\end{aligned}$$

El integrando es positivo si y solo si $\left(xe^{-\sigma\sqrt{\tau}z+\left(r-\frac{1}{2}\sigma^2\right)\tau} - K\right)^+ > 0$, o lo que es lo mismo, al despejar, si y solo si:

$$z < \frac{1}{\sigma\sqrt{\tau}}\left(\log(\frac{x}{K}) + (r - \frac{1}{2}\sigma^2)\tau\right) = d_-(\tau,x)$$

Luego

$$\begin{aligned}c(t,x) &= \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-r\tau}\left(xe^{-\sigma\sqrt{\tau}z+\left(r-\frac{1}{2}\sigma^2\right)\tau} - K\right)^+ e^{-\frac{1}{2}z^2}dz \\ &= \int_{-\infty}^{d_-(\tau,x)}\frac{1}{\sqrt{2\pi}}e^{-r\tau}\left(xe^{-\sigma\sqrt{\tau}z+\left(r-\frac{1}{2}\sigma^2\right)\tau} - K\right) e^{-\frac{1}{2}z^2}dz \\ &= x\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_-(\tau,x)} e^{-\sigma\sqrt{\tau}z-\frac{1}{2}\sigma^2\tau}e^{-\frac{1}{2}z^2}dz - Ke^{-r\tau}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_-(\tau,x)} e^{-\frac{1}{2}z^2}dz \\ &= x\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_-(\tau,x)} e^{-\frac{1}{2}(z+\sigma\sqrt{\tau})^2}dz - Ke^{-r\tau}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_-(\tau,x)} e^{-\frac{1}{2}z^2}dz \\ &= x\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_-(\tau,x)} e^{-\frac{1}{2}(z+\sigma\sqrt{\tau})^2}dz - Ke^{-r\tau}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_-(\tau,x)} e^{-\frac{1}{2}z^2}dz\end{aligned}$$

Para la primera integral hacemos el cambio de variable $u = z + \sigma\sqrt{\tau}$ y llamamos $d_+(\tau,x) = d_-(\tau,x) + \sigma\sqrt{\tau}$ para que quede:

$$\begin{aligned}c(t,x) &= x\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_-(\tau,x)} e^{-\frac{1}{2}(z+\sigma\sqrt{\tau})^2}dz - Ke^{-r\tau}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_-(\tau,x)} e^{-\frac{1}{2}z^2}dz \\ &= x\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_+(\tau,x)} e^{-\frac{1}{2}u^2}du - Ke^{-r\tau}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d_-(\tau,x)} e^{-\frac{1}{2}z^2}dz\end{aligned}$$

Donde si denotamos la función de distribución acumulativa de la normal estándar como Φ, es decir, $\Phi(z) = \mathbb{P}(Z < z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{1}{2}t^2} dt$, nos queda la expresión final de la fórmula de Black-Scholes:

$$c(t,x) = x\Phi(d_+(\tau,x)) - Ke^{-r\tau}\Phi(d_-(\tau,x)) \tag{4.17}$$

4.3.1. Paridad de call-put

Teniendo ahora el conocimiento de como calcular el precio de una call con la fórmula de Black-Scholes, podemos preguntarnos que ocurre con una opción put.

Suponemos que tenemos una call y una put de las mismas características, es decir, comparte el mismo activo subyacente, fecha de ejercicio T y precio de Strike K. El precio a tiempo t de la call la denominamos como $C(t)$, la de put $P(t)$, y el activo $S(t)$. Notar que la denotamos por $C(t)$ en vez de $c(t,x)$ como solemos hacer para simplificar la notación, y lo mismo para la opción put.

Vamos a comparar dos carteras:

- $\pi_A(t)$: Una opción call y una cantidad de efectivo igual al valor presente del precio de ejercicio K , es decir, $Ke^{-r(T-t)}$, invertido en un banco. La inversión necesaria en el tiempo t para comprar la cartera es $C(t) + Ke^{-r(T-t)}$.
- $\pi_B(t)$: Una opción put y una acción del activo subyacente. La inversión necesaria en el tiempo t para comprar la cartera es $P(t) + S(t)$

Al vencimiento T, el valor de la cartera π_A es el pago de la call al vencimiento, que ya hemos visto en los preliminares que es $\max(S(T) - K, 0)$, y el dinero invertido en el banco más intereses, $Ke^{-r(T-t)} \cdot e^{r(T-t)} = K$, es decir, en total:

$$\max(S(T) - K, 0) + K = \max(S(T), K)$$

Donde la igualdad se puede ver si dividimos casos. Si $K > S(T)$ entonces $\max(S(T)-K, 0) = 0$ y $\max(S(T) - K, 0) + K = K = \max(S(T), K)$. Si $K < S(T)$ entonces $\max(S(T) - K, 0) = S(T) - K$ y $\max(S(T) - K, 0) + K = S(T) - K + K = S(T) = \max(S(T), K)$.

Del mismo modo, el valor de la cartera π_B es, el valor de la put al vencimiento más lo que valga el activo en el tiempo T:

$$\max(K - S(T), 0) + S(T) = \max(S(T), K)$$

Dado que las carteras π_A y π_B tienen el mismo valor en el vencimiento y suponemos que el mercado es eficiente y no hay oportunidades de arbitraje, las carteras deben tener el mismo valor en el presente, es decir $\pi_A = \pi_B$. Esto nos lleva a la relación de paridad put-call:

$$\pi_A = \pi_B \Longrightarrow C(t) + Ke^{-r(T-t)} = P(t) + S(t)$$

Despejando, el precio de la put es:

$$P(t) = S(t) - C(t) - Ke^{-r(T-t)}$$

Así ya tenemos el precio de la put, pues $S(t)$, $Ke^{-r(T-t)}$ y $C(t)$ son conocidos ($C(t)$ mediante B-S). Con cálculos extras similares a los ya vistos en el trabajo se puede obtener una fórmula explícita, pero la omitimos para mantener la exposición concisa.

Capítulo 5

Algunas implicaciones y usos

A lo largo de este trabajo hemos explorado la ecuación de Black-Scholes como uno de los pilares fundamentales en la valoración de opciones, destacando su formulación matemática. Sin embargo, al llegar al corazón práctico de su aplicación, emerge un componente crítico como es la volatilidad del mercado, cuyo valor no se obtiene de manera directa.

Este capítulo final se centra precisamente en ese aspecto, abordando la volatilidad desde dos enfoques complementarios: la volatilidad implícita, en la que partir del precio de mercado de las opciones se utiliza el modelo de Black-Scholes para inferir dicha volatilidad, y la volatilidad histórica, estimada directamente a partir de datos observados de precios del activo subyacente.

Veremos las incógnitas que aparecen con la volatilidad implícita y analizaremos los métodos para estimar la volatilidad histórica. Ajustar la volatilidad del activo nos permitirá valorar según Black-Scholes cuál es el precio adecuado del activo. De lo que podemos sacar conclusiones para operar.

Este cierre busca no solo reforzar la conexión entre teoría y práctica, sino también evidenciar cómo, incluso partiendo de un modelo idealizado como Black-Scholes, la complejidad del comportamiento del mercado obliga a incorporar herramientas estadísticas y econométricas más flexibles y adaptativas para sacar mayor partido al modelo. Las siguientes secciones están basadas principalmente en [4], [3], [10], [5].

5.1. Volatilidad implícita

Dentro del marco del modelo de Black-Scholes, la volatilidad del activo subyacente es uno de los parámetros clave que determinan el precio teórico de una opción. Sin embargo, en la práctica, este valor no siempre es conocido con certeza, a diferencia de los demás que pueden ser observados. En los mercados financieros se utiliza el concepto de volatilidad implícita, que consiste en invertir el modelo: dado un precio de mercado observado para una opción, se busca la volatilidad que, al ser introducida en la fórmula de Black-Scholes, reproduce dicho precio. Y es muy utilizada para predecir o saber la volatilidad esperada por la gente.

Vemos un ejemplo con datos reales, sacados de [11], donde se está graficando la volatilidad implícita y el precio de Strike o precio del ejercicio del mercado para las opciones respectivas al

S&P 500, en concreto para opciones call con fecha de vencimiento en el 7 de Julio de 2025 [1]. Todo el procesamiento y la visualización de los datos se ha realizado utilizando el lenguaje de programación R, en esta sección y las posteriores.

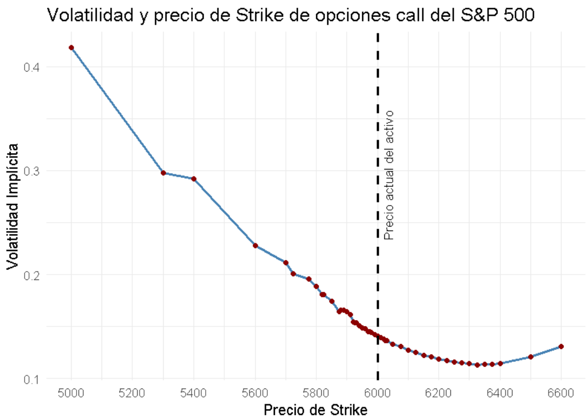

Figura 5.1: Volatilidad implícita para distintas opciones call sobre el mismo activo

Si la fórmula de Black-Scholes fuera totalmente válida, la volatilidad implícita sería igual para todas las opciones europeas sobre un mismo activo, independientemente del precio de ejercicio. Sin embargo, esto no ocurre: las opciones "in the money" (precio de strike de la call menor que el precio del activo) suelen tener mayor valor que las "out of the money" (precio de strike de la call mayor que el precio del activo), con diferencias de volatilidad de hasta un 10 %, y normalmente dibujando una forma de sonrisa, como se suele llamar a este fenómeno. Este comportamiento, común desde la caída del mercado en 1987, evidencia las limitaciones del modelo. Aunque se han propuesto ajustes, ninguno ha logrado reemplazar a Black-Scholes ni en la teoría ni en la práctica.

Este comportamiento observado en la volatilidad implícita no solo pone de manifiesto los límites del modelo Black-Scholes, sino que también apunta hacia una posibilidad interesante: si se pudiera estimar de forma fiable la volatilidad futura del subyacente, cabría utilizar dicha estimación como base para valorar las opciones y detectar potenciales desajustes con los precios de mercado. Aunque este trabajo no desarrolla esa línea, sí se analiza a continuación el problema de estimar la volatilidad, lo cual constituye un paso previo necesario para estrategias de este tipo.

5.2. Volatilidad histórica

Una de las maneras más directas de estimar la volatilidad del activo subyacente es observar su comportamiento en el pasado. A esta aproximación se le denomina volatilidad histórica, y se fundamenta en el análisis estadístico de las variaciones del precio del activo a lo largo del tiempo. Aunque parte de una mirada retrospectiva, este enfoque ofrece una base cuantitativa útil para anticipar escenarios futuros y alimentar modelos como el de Black-Scholes con datos más cercanos a la dinámica real del mercado. Y aunque Black-Scholes supone volatilidad constante,

[1]Cuyo precio utilizado ha sido el del día 6 de junio de 2025

podemos usar las dinámicas y comportamientos del pasado para intentar conocer la volatilidad instantánea en el presente o la que más se aproxime para un futuro.

A continuación, describiremos algunos procesos para estimar la volatilidad a partir de los retornos del activo, así como su implementación práctica sobre datos reales.

Primero, hacemos notar que si tomamos logaritmos en la expresión del movimiento browniano geométrico $S(t) = S(0)e^{\sigma W(t)+(\alpha-\frac{1}{2}\sigma^2)t}$, podemos expresarlo como:

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

Ahora restando los logaritmos de $S(t)$ y de $S(s)$:

$$\log(S(t)) - \log(S(s)) = \sigma(W(t) - W(s)) + (\alpha - \frac{1}{2}\sigma^2)(t-s)$$

Lo que renombramos como una variable aleatoria $U = \log(S(t)) - \log(S(s)) = \log \frac{S(t)}{S(s)}$. Dado que $W(t) - W(s) \sim \mathcal{N}(0, t-s)$, al multiplicarlo por σ obtenemos una normal con varianza $\sigma^2(t-s)$ y al sumarle el término $\left(\alpha - \frac{1}{2}\sigma^2\right)(t-s)$, se concluye que U sigue una distribución normal con media $\left(\alpha - \frac{1}{2}\sigma^2\right)(t-s)$ y varianza $\sigma^2(t-s)$, es decir:

$$U \sim \mathcal{N}\left(\left(\alpha - \frac{1}{2}\sigma^2\right)(t-s),\ \sigma^2(t-s)\right).$$

Ahora suponemos un activo de acciones del cual sabemos el precio de cierre en cada instante hasta la fecha actual, es decir $S(t_i)$, para todo $0 \leq i \leq n$, donde n en este caso es el tiempo presente (con intervalo el que elijamos, donde llamaremos Δt a $t_i - t_{i-1}$). Y denominamos la variable aleatoria $U_i = \log(S(t_i)) - \log(S(t_{i-1})) = \log \frac{S(t_i)}{S(t_{i-1})}$, con realización u_i, llamado el retorno logarítmico o por abreviar simplemente retorno. Por lo anterior, cada una de estas variables U_i son una normal $\mathcal{N}\Big((\alpha - \frac{1}{2}\sigma^2)\Delta t, \sigma^2\Delta t\Big)$.

Por lo visto en el Apéndice A, sea $m = (\alpha - \frac{1}{2}\sigma^2)(t-s)$, $v = \sigma^2(t-s)$, $\Theta = [m, v]$ y f_Θ la función de densidad de una normal de media m y varianza v. La función de máxima verosimilitud es $\mathcal{L}(\Theta) = \prod_{i=1}^n f_\Theta(u_i)$. Que se sabe (aclarado en el apéndice) que al maximizar, los valores estimados son:

$$\hat{m} = \frac{1}{n}\sum_{i=1}^n u_i \qquad \hat{v} = \frac{1}{n}\sum_{i=1}^n (u_i - \hat{m})^2$$

Una vez obtenidas las estimaciones, podemos despejar los parámetros originales del modelo. Recordando que $m = (\alpha - \frac{1}{2}\sigma^2)\Delta t$ y $v = \sigma^2\Delta t$, se obtienen los siguientes estimadores:

$$\hat{\sigma}^2 = \frac{\hat{v}}{\Delta t} \qquad \hat{\alpha} = \frac{\hat{m}}{\Delta t} + \frac{1}{2}\hat{\sigma}^2 \tag{5.1}$$

Esto nos dice cuál es la volatilidad del activo más plausible.

Ahora, cogemos los datos de [11] sobre la cotización del índice del S&P 500 desde hace un año. Notar que debemos suponer que los $U(t_i)$ son independientes, de lo contrario no podemos usar lo expuesto en el Apéndice A. Aunque no podemos saber esto con certeza, una forma de justificar la suposición de independencia es mediante el análisis de la función de autocorrelación: si esta no exhibe retardos significativos (o “lags”, en inglés), la independencia puede considerarse razonable (mirar [3]).

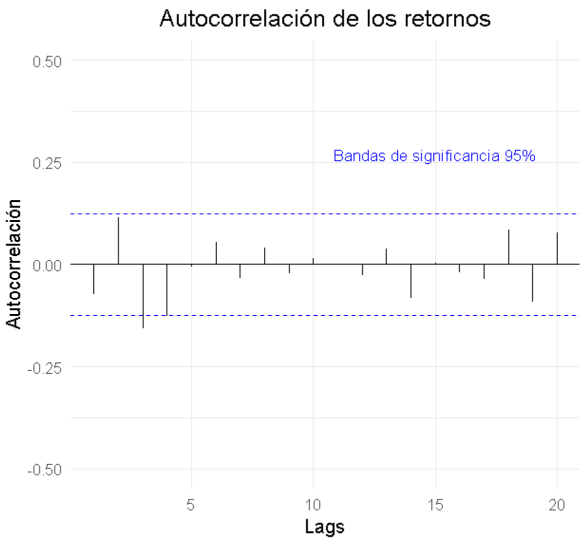

Figura 5.2: Autocorrelación de los retornos

No se observan lags significativos en la serie temporal, lo que significa que la suposición de independencia es aceptable en este ejemplo.

Ahora, dibujamos la volatilidad del activo calculada mediante la estimación de máxima verosimilitud de los retornos y utilizando la ecuación (5.1). Tenemos en cuenta que para nuestros datos $\Delta t = 1/249$, ya que el intervalo de un año de nuestros datos contiene 249 observaciones (hubo 249 días laborales de la bolsa).

Ahora bien, dado que en la práctica la volatilidad no es constante, resulta más realista estimarla de forma local, atendiendo únicamente al comportamiento reciente del activo. En este caso, calculamos la volatilidad en cada instante empleando solo la información de aproximadamente el último mes. Ambas estimaciones se muestran en el siguiente gráfico, junto con la evolución temporal del precio del activo, lo que permite tener una visión conjunta del comportamiento del activo y de su variabilidad.

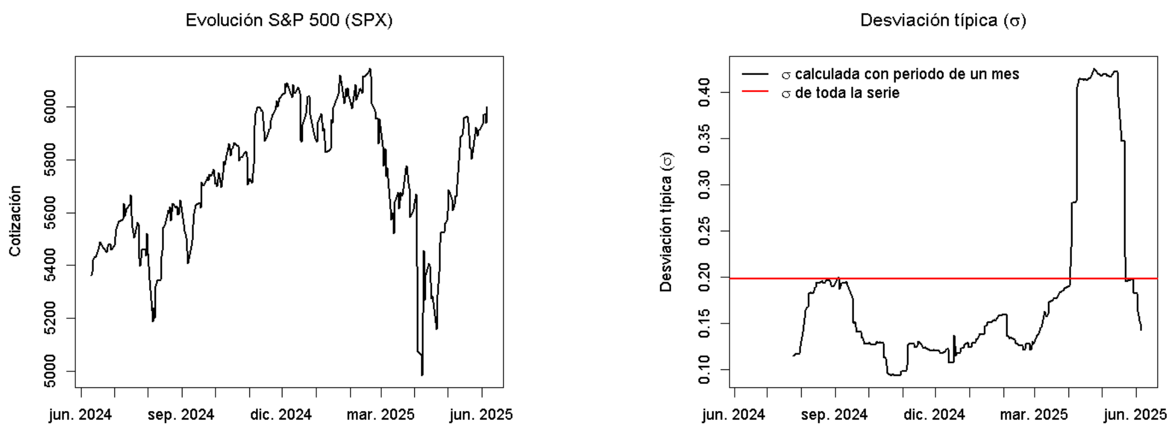

Figura 5.3: Evolución del S&P 500 (izquierda) junto con su volatilidad estimada por máxima verosimilitud (derecha).

Esto implica complicaciones sobre la decisión de las longitudes de las ventanas de tiempo y dan resultados no siempre del todo confiables. Por esto existen aproximaciones más modernas para modelar la volatilidad. En especial usaremos los modelos GARCH, en concreto explicaremos el más utilizado, el GARCH(1,1).

5.3. Modelo GARCH(1,1)

Diversos problemas, como la valoración de opciones en finanzas, han motivado el estudio de la volatilidad, o variabilidad, de una serie temporal. Los modelos ARMA se utilizaron para modelar la media condicional de un proceso cuando la varianza es constante (mirar [10]). En muchos problemas, sin embargo, se violará la suposición de una varianza constante. Modelos como el modelo autorregresivo condicionalmente heterocedástico, o ARCH, introducido por primera vez por Engle (1982), fueron desarrollados para modelar cambios en la volatilidad. Donde ya la volatilidad no era constante, sino que se modela con dependencia de una volatilidad en el largo plazo y los retornos u_i ponderados, es decir, en el caso de un ARCHA(m) (dependencia de los últimos m valores):

$$\sigma_n^2 = \gamma\sigma_L^2 + \sum_{i=1}^{m} \alpha_i u_{n-i}^2$$

Donde $\gamma + \sum_{i=1}^{m} \alpha_i = 1$ y donde hay veces que la volatilidad a largo plazo σ_L se decide omitir.

Estos modelos fueron posteriormente extendidos a modelos ARCH generalizados, o GARCH, por Bollerslev (1986). En concreto trataremos GARCH(1,1), donde el “(1, 1)” indica que la estimación de σ^2 en un instante se basa en la observación más reciente de u^2 y en la estimación más reciente de la tasa de varianza, lo que es lo novedoso del modelo. El GARCH(p, q) más general estima σ^2 a partir de las p observaciones más recientes de u^2 y las q estimaciones más recientes de la tasa de varianza. GARCH(1, 1) es, con diferencia, el más popular de los modelos GARCH. El cual estima σ_n^2 como:

$$\sigma_n^2 = \gamma\sigma_L^2 + \alpha u_{n-1}^2 + \beta\sigma_{n-1}^2$$

Donde $\gamma + \alpha + \beta = 1$. Y los debidos coeficientes se hallan mediante estimación por máxima verosimilitud.

Si no se especifica σ_L lo que se hace es, aparte de estimar α y β, también se estima $\omega = \gamma\sigma_L^2$, que además nos permite, despejando, hallar σ_L^2 como $\sigma_L^2 = \frac{\omega}{\gamma} = \frac{\omega}{1-\alpha-\beta}$ (Esto se hará en la práctica en la observación 5.1).

Ahora, con los mismos datos que el apartado anterior, procedemos esta vez utilizando el GARCH(1,1), aplicándoselo a la serie de los retornos y hallando los estimadores como en la ecuación (5.1):

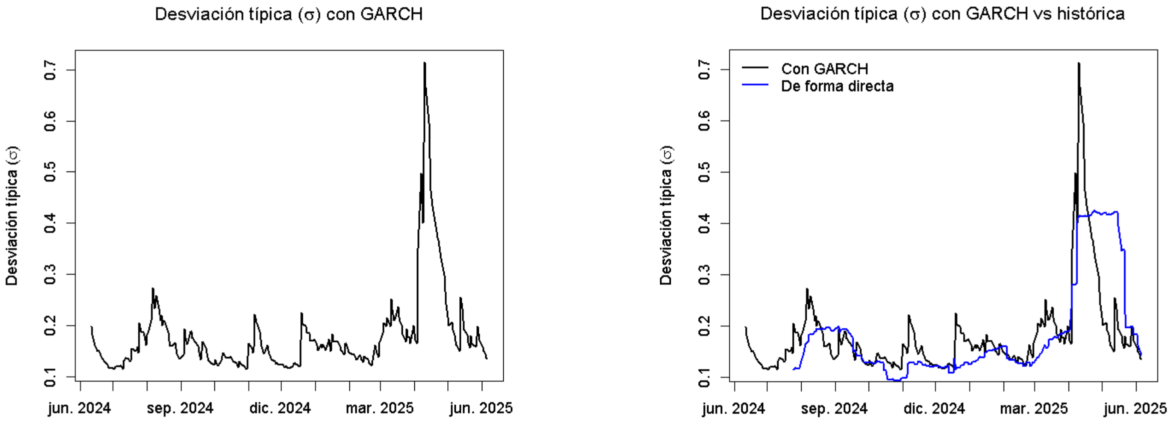

Figura 5.4: Volatilidad modelada con GARCH(1,1) (izquierda) y la comparación con la estimada anteriormente (derecha)

Como se puede observar, con el primer método, el modelo no reacciona tan rápido a los cambios como con GARCH como se puede ver en fechas como septiembre de 2024 o mayo del 2025, entre otras. Y una ventaja, de entre otras muchas, de este proceso es la capacidad de hacer predicciones. Para los mismos datos, el modelo nos permite hacer las siguientes predicciones:

Observación 5.1. Al ejecutar el código incluido en el Apéndice B, obtenemos los siguientes coeficientes para el modelo GARCH(1,1): $\omega = 0{,}0000100857$, $\alpha = 0{,}1704659097$, y $\beta = 0{,}7653818332$.

A partir de ellos, podemos calcular la volatilidad a largo plazo de los retornos calculada como se ha explicado con anterioridad.

$$\sigma_L = \sqrt{\frac{\omega}{1-\alpha-\beta}}.$$

Para expresar esta volatilidad en términos del activo, la multiplicamos por $\sqrt{249}$, la misma conversión que ya hemos hecho más veces.

$$\sigma_L^{\text{Activo}} = \sigma_L \cdot \sqrt{249} = \sqrt{\frac{\omega}{1-\alpha-\beta} \cdot 249} = 0{,}1978551.$$

Este valor coincide con las predicciones de volatilidad a medio-largo plazo obtenidas mediante el modelo, como puede observarse en el siguiente gráfico, donde se representa la evolución esperada para los próximos 60 días junto con la volatilidad a largo plazo de referencia:

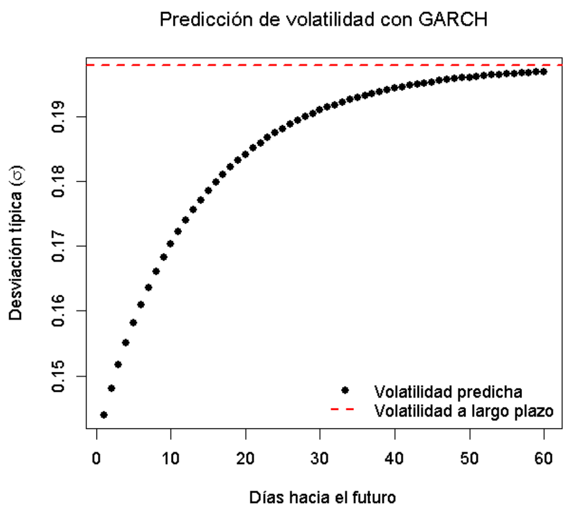

Figura 5.5: Predicciones de volatilidad realizadas con el modelo GARCH(1,1) en los próximos 60 días, junto con la volatilidad a largo plazo

Además, del análisis de los coeficientes obtenidos se puede extraer información sobre la dinámica del modelo. El valor relativamente alto de $\beta = 0{,}765$ frente a $\alpha = 0{,}170$ o de $\gamma = 1 - \alpha - \beta = 0{,}0641522571$ indica que el modelo otorga más peso a la persistencia de la volatilidad que a los shocks recientes en los retornos. Es decir, la evolución futura de la volatilidad depende en mayor medida del comportamiento pasado de esta, que de los retornos recientes.

Bibliografía

[1] Jesús Araujo, *Measure theory*, Apuntes de clase, Universidad de Cantabria, Measury theory course., 2024-2025.

[2] Tomas Björk, *Arbitrage theory in continuous time*, 3 ed., Oxford University Press, Oxford, 2009.

[3] Damiano Brigo, Antonio Dalessandro, Matthias Neugebauer, and Fares Triki, *A stochastic processes toolkit for risk management*, 2007, https://arxiv.org/pdf/0812.4210.pdf.

[4] Desmond J. Higham, *An introduction to financial option valuation: Mathematics, stochastics and computation*, Cambridge University Press, Cambridge, 2004.

[5] John C. Hull, *Options, futures, and other derivatives*, 8 ed., Prentice Hall, Upper Saddle River, NJ, 2012.

[6] Achim Klenke, *Probability theory: A comprehensive course*, 2 ed., Universitext, Springer, London, 2014.

[7] Sheldon M. Ross, *A first course in probability*, 8 ed., Pearson Prentice Hall, Upper Saddle River, NJ, 2010.

[8] Steven E. Shreve, *Stochastic calculus for finance i: The binomial asset pricing model*, Springer Finance, Springer, 2004.

[9] ______, *Stochastic calculus for finance ii: Continuous-time models*, Springer Finance, Springer, 2004.

[10] Robert H. Shumway and David S. Stoffer, *Time series analysis and its applications: With R examples*, 3 ed., Springer Texts in Statistics, Springer, New York, 2011.

[11] Yahoo Finance, *S&P 500 Index (^SPX)*, 2025, Accessed: 2025-06-08. Available at: `https://finance.yahoo.com/quote/%5ESPX/`.

Apéndice A

Consideremos variables aleatorias $X_1, X_2, \ldots, X_n$ con realizaciones $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$. Supongamos que estas variables aleatorias siguen una distribución caracterizada por un parámetro θ, y que tienen una función de densidad o de masa conjunta $f_\theta(x_1, x_2, \ldots, x_n)$. Entonces, definimos la función de máxima verosimilitud como:

$$L(\theta) = f_\theta(x_1, x_2, \ldots, x_n),$$

la cual representa la probabilidad de que ocurran estas realizaciones. El máximo de esta función se llama el estimador de máxima verosimilitud (o MLE por sus siglas en inglés). Así, $\hat{\theta}$ para la cual $L(\theta)$ es máxima, es precisamente el valor con mayor probabilidad de generar estas observaciones.

Supongamos que las variables aleatorias $X_1, \ldots, X_n$ son independientes e idénticamente distribuidas (i.i.d.), entonces podemos escribir:

$$L(\theta) = \prod_{i=1}^{n} f_\theta(x_i),$$

donde $f_\theta(x)$ es la función de densidad de X. A veces esta función puede ser difícil de diferenciar, por lo tanto, a menudo tomamos logaritmos:

$$\log(L(\theta)) = \log\left(\prod_{i=1}^{n} f_\theta(x_i)\right) = \sum_{i=1}^{n} \log(f_\theta(x_i)).$$

Maximizar esta expresión a la que llamamos función de log-verosimilitud y denominamos por $l(\theta)$ es equivalente, ya que el logaritmo natural es una función monótonamente creciente, así que maximizar $l(\theta)$ da el mismo valor que maximizar $L(\theta)$.

Consideremos $X_1, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ i.i.d., con realizaciones $x_1, \ldots, x_n$. Entonces podemos calcular los estimadores de máxima verosimilitud para μ y σ. La función de densidad es:

$$f_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Entonces la función de log-verosimilitud es:

$$l(\mu, \sigma^2) = \sum_{i=1}^{n} \log(f_{\mu,\sigma^2}(x_i)) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2}.$$

Ahora derivamos para hallar el máximo:

$$\frac{\partial l}{\partial \mu} = \frac{\partial}{\partial \mu}\left(-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2}\right) = \sum_{i=1}^{n} \frac{(x_i-\mu)}{\sigma^2}.$$

Igualando a cero

$$\sum_{i=1}^{n} \frac{(x_i-\mu)}{\sigma^2} = 0 \Longrightarrow \sum_{i=1}^{n}(x_i-\mu) = 0 \Longrightarrow \sum_{i=1}^{n} x_i - n\mu = 0.$$

Por lo tanto,

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Dado que $\hat{\mu}$ es el único extremo y la segunda derivada respecto a μ es negativa, esto es un máximo. Por otro lado,

$$\frac{\partial l}{\partial \sigma} = \frac{\partial}{\partial \sigma}\left(-\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^2}\right) = -\frac{n}{\sigma} + \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^3}.$$

Entonces, igualando a cero y resolvemos

$$-n\sigma^2 + \sum_{i=1}^{n}(x_i-\mu)^2 = 0.$$

Lo que resulta en

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(x_i-\hat{\mu})^2}{n}.$$

Vemos que $\hat{\mu}$ y $\hat{\sigma}$ maximizan $l(\mu,\sigma^2)$, entonces, ya que el par $(\hat{\mu},\hat{\sigma}^2)$ resulta en un máximo de $l(\mu,\sigma^2)$, concluimos que las estimaciones de máxima verosimilitud están dadas por

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(x_i-\hat{\mu})^2}{n}.$$

Apéndice B

```r
library(quantmod)
library(forecast)
library(PerformanceAnalytics)
library(xts)

load("S&P500 prices.Rdata")
m <- length(cotizaciones)
retornos <- na.omit(Return.calculate(cotizaciones))

#Autocorrelacion
acf(retornos, main = "Autocorrelacion de los retornos")

#MLE
ventana <- 30
ventana2 <- ventana-1

varianza=c()
for(i in ventana:length(retornos))
{f=var(retornos[(i-ventana2):i])
varianza[i]=(f)}

SD_SP500 <- sqrt(varianza*m)
SD_SP500_Total <- sqrt(var(retornos)*m)


#########GARCH################
install.packages("fGarch")
library(fGarch)

GARCH <- garchFit(~ garch(1,1), cond.dist="norm", data = (retornos),
                  include.mean = FALSE, trace = FALSE)

SD <- GARCH@sigma.t
SD_SP500_GARCH <- SD*sqrt(m)


########Predicciones########
forecast <- (predict(GARCH,n.ahead=30))
vol_pred <- (forecast$standardDeviation)*sqrt(m)

coeficientes <- coef(GARCH)
omega <- coeficientes["omega"]
alpha <- coeficientes["alpha1"]
beta <- coeficientes["beta1"]
vol_largo <- sqrt(omega / (1 - alpha - beta))*sqrt(249)
```