



UNIVERSIDAD TECNOLÓGICA DE LA MIXTECA

Aplicaciones Computacionales de las EDP's Asociadas a
Procesos Estocásticos

TESIS

para obtener el título de

LICENCIADO EN MATEMÁTICAS APLICADAS

presenta

JORGE ZAVALA SÁNCHEZ

Director de Tesis:

M.C. JUAN CARLOS MENDOZA SANTOS

Huajuapán de León, Oaxaca, Noviembre de 2007.

Con cariño especial:

a mi madre,
Yolanda Alicia Sánchez L.;

al recuerdo de mi padre,
Jorge José Zavaleta R.

a mis hermanos,
Carlos Francisco y Yazmín;

y a todos aquellos que me han brindado su amistad sincera.

Dedicatoria.

A mi madre por todo el cariño, apoyo y por todo el esfuerzo y sacrificios hechos durante tantos años para sacarnos adelante a mi hermano y a mí; al recuerdo de mi padre que siempre luchó contra sus enfermedades pero que lamentablemente no pudo ver la culminación de este proyecto; a mi hermano por su amistad y apoyo incondicional y a mi hermana, que a pesar del distanciamiento siempre está en mis pensamientos.

A mis tíos y familia en general, en especial, al recuerdo de mi tío Pedro que lamentablemente murió en meses recientes y también de mi tío Sergio, que gracias a su apoyo y el de su familia, este proyecto se ha completado.

A mis amigos, los que conocí antes de entrar a la universidad y los que conocí durante esta, por su amistad y por los gratos momentos, y que por temor a olvidar mencionar a alguno no pongo sus nombres, pero estoy seguro de que ellos saben quiénes son.

Aprovecho esta parte para agradecerles a todos mis maestros que tuve durante mi estancia en la UTM, ya que todos son parte de mi formación académica, en particular, al M. C. Juan Ramón Tijerina por su amistad y apoyo, al M.C. Vulfrano Tochihuitl, M.C. Gustavo Yáñez y M.C. Adolfo Maceda por sus magníficas clases. A mis sinodales M.C. Norma Edith Alamilla López, M.C. Adolfo Maceda Méndez y al M.C. José del Carmen Jiménez Hernández, por sus observaciones que ayudaron a mejorar este trabajo.

De manera especial, con toda la gratitud posible, a mi asesor M.C. Juan Carlos Mendoza Santos, por todo su apoyo durante la realización de la tesis, su confianza, apoyo y por su amistad. Finalmente, al Dr. Onésimo Hernández-Lerma, por el apoyo que me ofreció durante mis estancias en el CINVESTAV y por haberme dado las bases que me sirvieron para la realización de este trabajo.

Introducción.

Actualmente las ecuaciones diferenciales parciales tienen muchas aplicaciones tanto en matemáticas aplicadas como en física. En esta última, dichas ecuaciones son utilizadas para describir fenómenos lineales [8] y no lineales [3].

Además, por ejemplo, la aplicación de nuevas teorías matemáticas, como el cálculo estocástico (cálculo de Itô), han permitido la aparición de nuevas ecuaciones diferenciales tanto ordinarias como en derivadas parciales, llamadas ecuaciones diferenciales estocásticas [11]. En este sentido, la aparición de ecuaciones diferenciales parciales en el cálculo estocástico, ha dado lugar a la aplicación de algunos métodos de resolución que aparecen en la física matemática y en algunas áreas de las matemáticas aplicadas. En particular, en matemáticas financieras, mediante la generación de modelos matemáticos financieros, como por ejemplo, la valuación de opciones.

Por otro lado, se han generado muchas soluciones de las principales ecuaciones que aparecen en matemáticas financieras; estas soluciones se obtuvieron mediante técnicas utilizadas en la física matemática [14]; pero el problema radica en el hecho de que no todas las soluciones obtenidas mediante dichas técnicas están en correspondencia con algunas aplicaciones de las matemáticas financieras. Por lo tanto, se tienen que hacer determinadas consideraciones para aplicar dichas soluciones a este caso. Como por ejemplo la ecuación en derivadas parciales de Black-Scholes

$$\frac{\partial V(t, x)}{\partial t} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 V(t, x)}{\partial x^2} + rx \frac{\partial V(t, x)}{\partial x} - rV(t, x) = 0, \quad (1) \quad (1)$$

tiene su origen en modelos matemáticos financieros; en particular, en la solución de valuación de opciones.

Pero dicha ecuación ha sido resuelta mediante técnicas de la física matemática, como por ejemplo la mecánica cuántica y la mecánica clásica. En el caso de la mecánica cuántica, esta ecuación tiene una solución análoga a la fórmula de Feynman-Kac [2]; y en el caso de la mecánica clásica se utilizan técnicas del método de separación de variables y teorías de grupos de Lie, mediante la introducción de simetrías y leyes de conservación. Todas estas soluciones obtenidas de (1) son analíticas; las cuales presentan los siguientes problemas, entre otros:

1. Algunas soluciones no se ajustan a los problemas de valor inicial y de frontera que aparecen en los mercados financieros [14].

2. Algunas soluciones no se pueden calcular mediante técnicas computacionales; impidiendo su aplicación en la solución de algunos problemas matemáticos que aparecen en estos mercados.

Así, el objetivo de esta tesis, es utilizar el cálculo de Itô para obtener la ecuación diferencial parcial (1), la cual es utilizada para calcular el valor de las opciones en matemáticas financieras. Además, dado que dicha ecuación se puede reducir a un proceso de difusión estocástico, mediante una transformación adecuada, esto nos permite modelar la valuación de opciones utilizando el método de diferencias finitas [6] y [9], para de esta manera poder comparar dicho cálculo con los valores obtenidos por la fórmula de Black-Scholes, que es la solución analítica del problema de valor a la frontera asociado a la ecuación (1):

$$\frac{\partial C(t, x)}{\partial t} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 C(t, x)}{\partial x^2} + rx \frac{\partial C(t, x)}{\partial x} - rC(t, x) = 0, \quad (2)$$

con las siguientes condiciones de frontera

$$\begin{aligned} C(T, x) &= \text{máx}(x - K, 0), \\ C(t, x) &= x - Ke^{-r(T-t)} \quad \text{cuando } x \rightarrow \infty, \\ C(t, 0) &= 0 \quad \forall t, \end{aligned}$$

Adicionalmente, comparar las nuevas soluciones de la ecuación diferencial parcial (1), obtenidas por Sukhomlin [14] mediante simetrías con la fórmula de Black-Scholes.

La estructura del trabajo es la siguiente. En el Capítulo 1, se da una breve introducción a la teoría de probabilidad y a procesos estocásticos. Aquí se presentan nociones básicas necesarias para la definición de las integrales estocásticas.

En el Capítulo 2 se define un importante proceso estocástico, el movimiento Browniano, el cual tiene un papel relevante en capítulos posteriores. Además, se estudia la ecuación de difusión para la cual se obtiene su solución analítica que será de gran utilidad en la obtención de la solución analítica de la ecuación (1).

En el Capítulo 3 se introduce la definición de integral estocástica de Itô. En este capítulo se obtiene la herramienta más importante en el cálculo de Itô, el lema de Itô.

En el Capítulo 4 se da la definición de ecuación diferencial estocástica la cual se contrasta con las ecuaciones diferenciales determinísticas. Por otro lado, se

utilizará el lema de Itô para obtener las soluciones de algunos ejemplos de ecuaciones diferenciales estocásticas.

En el Capítulo 5 se darán nociones básicas de finanzas, a partir de las cuales, se formulará un modelo matemático para el problema de valuación de opciones, obteniendo de este, la ecuación diferencial parcial (1) o también conocida como la ecuación de Black-Scholes. Dada la ecuación de Black-Scholes, se formularán problemas de valor a la frontera específicos, como por ejemplo (2), y de ellos se derivará la solución analítica de (1) llamada la fórmula de Black-Scholes.

Para el Capítulo 6 se estudiarán diferentes métodos de diferencias finitas para aproximar la solución de la ecuación de difusión, en particular el método de Crank-Nicolson, que nos servirá para modelar la valuación de opciones y para comparar los cálculos realizados con los valores obtenidos mediante la fórmula de Black-Scholes.

Finalmente en el Capítulo 7 se expondrá brevemente las nuevas soluciones de la ecuación de Black-Scholes obtenidas por Sukhomlin mediante simetrías y se compararán con la fórmula de Black-Scholes.

Índice general

Dedicatoria	I
Introducción	III
1. Introducción a los Procesos Estocásticos.	1
1.1. Conceptos básicos de la teoría de la probabilidad.	1
1.1.1. Variables aleatorias.	1
1.1.2. Vectores aleatorios.	5
1.1.3. Independencia y dependencia.	8
1.2. Procesos estocásticos.	9
1.3. La esperanza condicional.	13
1.3.1. La esperanza condicional bajo condiciones discretas.	13
1.3.2. σ -Álgebras.	17
1.3.3. La esperanza condicional general.	21
1.3.4. Reglas para el cálculo de las esperanzas condicionales.	24
2. Movimiento Browniano.	27
2.1. Definición y propiedades.	27
2.2. Procesos derivados del movimiento Browniano.	31
2.3. Movimiento Browniano de una partícula libre.	33
2.3.1. Deducción de la ecuación de difusión.	33
2.3.2. Solución de la ecuación de difusión.	33
3. Cálculo Estocástico de Itô.	37
3.1. Integración estocástica.	37
3.1.1. Las integrales de Riemann y Riemann-Stieltjes.	37
3.2. La integral de Itô.	40
3.2.1. Un caso especial.	40
3.2.2. Integración estocástica de Itô para procesos simples.	43
3.2.3. La integral estocástica general de Itô.	46
3.3. El lema de Itô.	49
3.3.1. La regla de la cadena clásica de diferenciación.	49
3.3.2. Una versión simple del lema de Itô.	50

3.3.3. Versiones extendidas del lema de Itô.	51
4. Ecuaciones Diferenciales Estocásticas.	55
4.1. Ecuaciones diferenciales determinísticas.	55
4.2. Ecuaciones diferenciales estocásticas de Itô.	57
4.2.1. Definición de una ecuación diferencial estocástica.	57
4.2.2. Resolución de ecuaciones diferenciales estocásticas de Itô.	60
4.3. La ecuación diferencial general lineal.	63
4.3.1. Ecuaciones lineales con ruido aditivo.	63
4.3.2. Ecuaciones homogéneas con ruido multiplicativo.	64
4.3.3. El caso general.	65
5. Aplicaciones del Cálculo Estocástico a Matemáticas Financieras.	67
5.1. Conceptos y notación de matemáticas financieras.	67
5.1.1. Conceptos de finanzas.	67
5.1.2. Opciones de compra y venta europeas.	69
5.2. Análisis de la ecuación de Black-Scholes.	70
5.2.1. El problema de valuación de opciones.	70
5.3. Problemas de valor a la frontera con la ecuación de Black-Scholes.	72
5.3.1. Problema de valor a la frontera para una opción de compra.	72
5.3.2. Problema de valor a la frontera para una opción de venta.	78
6. Métodos en Diferencias Finitas.	83
6.1. Ecuaciones diferenciales parabólicas.	83
6.2. Aplicación de la fórmula de Black-Scholes.	91
7. Nuevas Soluciones de la Ecuación de Black-Scholes.	95
7.1. Simetría de la ecuación de Black-Scholes y leyes de conservación.	95
7.2. Grupo de equivalencia de la ecuación de Black-Scholes.	97
7.3. Clasificación de los operadores de simetría de la ecuación de B-S.	101
7.4. Nuevas soluciones de la ecuación de Black-Scholes.	102
7.5. Comparación de las soluciones.	107
Conclusiones	111
Apéndice	113
A. Martingalas.	113
B. Modos de convergencia.	114
C. Algoritmos.	114
D. Reseña Histórica.	118
Bibliografía	119

Capítulo 1

Introducción a los Procesos Estocásticos.

1.1. Conceptos básicos de la teoría de la probabilidad.

1.1.1. Variables aleatorias.

En este capítulo introduciremos algunas nociones básicas de la teoría de la probabilidad necesarias para la definición de los procesos estocásticos y posteriormente para la definición de las integrales estocásticas.

En la práctica ocurren dos tipos de fenómenos o experimentos: Los primeros que podemos nombrar son los fenómenos determinísticos los cuales son determinadas acciones en las cuales se puede predecir que es lo que va a ocurrir con exactitud. A diferencia de estos, están los fenómenos aleatorios los cuales son acciones en las que no sabemos lo que ocurrirá al ser realizados, ya que pueden obtenerse cualquiera de varias alternativas, aún bajo las mismas condiciones en que se realicen estas acciones. Con ayuda de la teoría de la probabilidad, de la cual estamos interesados en abundar, se podrá realizar el estudio de los fenómenos aleatorios. Enseguida daremos algunas definiciones que nos permitirán adentrarnos en el desarrollo de esta teoría.

Definición 1.1. *Dado un fenómeno aleatorio, al conjunto de todos sus posibles resultados se le llama **espacio muestral** y lo denotaremos por Ω .*

Definición 1.2. *Sea un fenómeno aleatorio con espacio muestral Ω , una **variable aleatoria** es una función X cuyo dominio es Ω y contradominio es \mathbb{R} :*

$$X : \Omega \rightarrow \mathbb{R}.$$

En adición a estas dos definiciones, llamaremos evento a cualquier subconjunto del espacio muestral y definiremos a la clase \mathcal{F} como la colección de todos los eventos de Ω , la cual se describirá posteriormente con mayor exactitud.

Con estas definiciones, la medida de probabilidad se puede ver como sigue. Sea \mathcal{F} la clase de eventos de Ω , entonces para cada evento $A \in \mathcal{F}$ se le asigna un número $P(A) \in [0, 1]$. Este número es la fracción esperada de ocurrencia del evento A , a lo largo de una serie de experimentos donde A o A^c son observados.

Algunas propiedades elementales de las medidas de probabilidad son resumidas a continuación.

i) Se cumple que:

$$P(\Omega) = 1 \quad \text{y} \quad P(\emptyset) = 0.$$

ii) Para eventos $A, B \in \mathcal{F}$,

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

iii) Si A y B son disjuntos, es decir $A \cap B = \emptyset$, entonces

$$P(A \cup B) = P(A) + P(B).$$

Observese además que de la propiedad (iii) se sigue que $P(A^c) = 1 - P(A)$.

La relación entre variables aleatorias y probabilidad puede ser caracterizada por ciertas cantidades numéricas. Ahora consideremos algunas de estas.

Definición 1.3. *La colección de las probabilidades*

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

es la **función de distribución de probabilidad acumulada** F_X de X .

Ejemplo 1.1. Utilizando la definición [1.3] podemos determinar probabilidades en términos de la función de distribución de probabilidad acumulada, como:

1. $P(\{\omega : a < X(\omega) \leq b\}) = F_X(b) - F_X(a)$ con $a < b$.
2. $P(X = x) = F_X(x) - \lim_{h \rightarrow 0} F_X(x - h)$.

Con estas probabilidades se puede aproximar la probabilidad del evento $\{\omega : X(\omega) \in B\}$ para cualquier conjunto complicado $B \subset \mathbb{R}$. ■

Definición 1.4. *La colección de las probabilidades*

$$P_X(\mathcal{B}) = P(X \in \mathcal{B}) = P(\{\omega : X(\omega) \in \mathcal{B}\}),$$

es la **distribución** de X para $\mathcal{B} \subset \mathbb{R}$ adecuado.

Nota 1.1. Denotaremos por \mathcal{B} a los subconjuntos adecuados de \mathbb{R}^n . Estos subconjuntos son los conjuntos de Borel, los cuales son obtenidos por un número numerable de operaciones \cap , \cup o c actuando sobre los intervalos de \mathbb{R}^n .

La distribución P_X y la función de distribución de probabilidad acumulada F_X son nociones equivalentes en el sentido de que ambas pueden ser usadas para calcular la probabilidad de cualquier evento $\{X \in \mathcal{B}\}$.

Una función de distribución de probabilidad es continua o discreta (tiene saltos). Consideremos primero el caso especial cuando la función de distribución F_X es una función discreta.

Definición 1.5.

$$F_X(x) = \sum_{k:x_k \leq x} p_k, \quad x \in \mathbb{R}, \quad (1.1)$$

donde

$$0 \leq p_k \leq 1 \text{ para toda } k \text{ y } \sum_{k=1}^{\infty} p_k = 1,$$

y

$$p_k = P(X = x_k).$$

La función de distribución de probabilidad acumulada (1.1) y la correspondiente distribución se dice que son discretas. Una variable aleatoria con función de distribución de probabilidad acumulada (1.1) es una **variable aleatoria discreta**.

En contraste a las distribuciones y variables aleatorias discretas, la función de distribución de una variable aleatoria continua no tiene saltos, por lo tanto $P(x) = 0$ para toda x , o equivalentemente,

$$\lim_{h \rightarrow 0} F_X(x+h) = F_X(x) \quad \forall x, \quad (1.2)$$

es decir, tal variable aleatoria asume cualquier valor con probabilidad 0. Una variable aleatoria continua obtiene su nombre de la propiedad de continuidad (1.2) de la función de distribución de probabilidad acumulada F_X .

Definición 1.6. Sea X una variable aleatoria continua con función de densidad de probabilidad f_X , se define su función de distribución de probabilidad acumulada por:

$$F_X(x) = \int_{-\infty}^x f_X(y) dy, \quad x \in \mathbb{R},$$

donde, la función de densidad de probabilidad cumple con:

$$f_X(x) \geq 0 \quad \forall x \in \mathbb{R}, \quad \int_{-\infty}^{\infty} f_X(x) dx = 1.$$

Ejemplo 1.2. (La distribución normal.)

Una importante distribución continua es la *distribución normal o gaussiana* $N(\mu, \sigma^2)$ con parámetros $\mu \in \mathbb{R}$, $\sigma > 0$. Esta tiene la función de densidad de probabilidad siguiente:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \quad x \in \mathbb{R}.$$

Si $X \sim N(0, 1)$, diremos que la variable aleatoria X tiene una distribución normal estándar, y se denotará con φ su función de densidad de probabilidad f_X y con Φ su función de distribución de probabilidad acumulada F_X . ■

Por otro lado, algunas características interesantes de una variable aleatoria X son la esperanza $E(X)$, la varianza $\text{var}(X)$ y los momentos $E(X^l)$.

A continuación definiremos estas características tanto para variables aleatorias discretas como continuas.

Definición 1.7. *Sea X una variable aleatoria continua con función de densidad de probabilidad f_X .*

i) La **esperanza o valor medio** de X esta dada por:

$$\mu_X = E(X) = \int_{-\infty}^{\infty} x f_X(x) dx.$$

ii) La **varianza** de X esta definida como:

$$\sigma_X^2 = \text{var}(X) = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx.$$

iii) El **l -ésimo momento** de X para $l \in \mathbb{N}$ esta definido como:

$$E(X^l) = \int_{-\infty}^{\infty} x^l f_X(x) dx.$$

iv) Para una función real g la esperanza de $g(X)$ esta dada por:

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

Similarmente se define:

Definición 1.8. Sea X una variable aleatoria discreta con probabilidades $p_k = P(X = x_k)$.

i) La **esperanza o valor medio** de X esta dada por:

$$\mu_X = E(X) = \sum_{k=1}^{\infty} x_k p_k.$$

ii) La **varianza** de X esta definida como:

$$\sigma_X^2 = \text{var}(X) = \sum_{k=1}^{\infty} (x_k - \mu_X)^2 p_k.$$

iii) El **l -ésimo momento** de X para $l \in \mathbb{N}$ esta definido como:

$$E(X^l) = \sum_{k=1}^{\infty} x_k^l p_k.$$

iv) Para una función real g la esperanza de $g(X)$ esta dada por:

$$E[g(X)] = \sum_{k=1}^{\infty} g(x_k) p_k.$$

Podemos considerar la esperanza μ_X como el *centro de gravedad* de la variable aleatoria X , es decir, los valores aleatorios $X(\omega)$, con $\omega \in \Omega$, que están concentrados alrededor del valor no aleatorio μ_X . La extensión o dispersión de los valores aleatorios $X(\omega)$ alrededor de la esperanza μ_X es descrita por la varianza σ_X^2 y la desviación estándar σ_X .

1.1.2. Vectores aleatorios.

Definición 1.9. Decimos que $\mathbf{X} = (X_1, \dots, X_n)$ es un **vector aleatorio** de dimensión n si sus componentes X_1, \dots, X_n son variables aleatorias reales de dimensión uno.

Análogamente a las variables aleatorias de una dimensión se puede introducir la función de distribución de probabilidad, la esperanza, los momentos, la matriz de varianza-covarianza de un vector aleatorio, con el objetivo de describir su distribución y su estructura de dependencia.

Como se hizo anteriormente, se considera una colección \mathcal{F} de subconjuntos de Ω y se define una medida de probabilidad en este, es decir, se asigna un número $P(A) \in [0, 1]$ para cada $A \in \mathcal{F}$.

Definición 1.10. *La colección de las probabilidades*

$$\begin{aligned} F_{\mathbf{X}}(\mathbf{x}) &= P(X_1 \leq x_1, \dots, X_n \leq x_n) \\ &= P(\{\omega : X_1(\omega) \leq x_1, \dots, X_n(\omega) \leq x_n\}), \\ \mathbf{x} &= (x_1, \dots, x_n) \in \mathbb{R}^n, \end{aligned}$$

es la **función de distribución de probabilidad conjunta** $F_{\mathbf{X}}$ de \mathbf{X} .

Esto nos provee con la probabilidad del evento que \mathbf{X} asume valores en el rectángulo

$$(\mathbf{a}, \mathbf{b}] = \{\mathbf{x} : a_i < x_i \leq b_i, i = 1, \dots, n\}.$$

Como en el caso de variables aleatorias de una dimensión, estas probabilidades aproximan $P(\mathbf{X} \in B)$ para conjuntos B muy generales.

Definición 1.11. *La colección de las probabilidades*

$$P_{\mathbf{X}}(\mathcal{B}) = P(\mathbf{X} \in \mathcal{B}) = P(\{\omega : \mathbf{X}(\omega) \in \mathcal{B}\}) \quad \mathcal{B} \subset \mathbb{R}^n,$$

constituyen la **distribución de \mathbf{X}** .

En un sentido matemático, la distribución y la función de distribución de probabilidad conjunta de un vector aleatorio \mathbf{X} son nociones equivalentes. Ambas, $F_{\mathbf{X}}$ y $P_{\mathbf{X}}$, pueden ser usadas para calcular la probabilidad de cualquier evento $\{\mathbf{X} \in \mathcal{B}\}$.

Análogamente a las variables aleatorias se puede introducir vectores aleatorios y distribuciones de probabilidad discretas y continuas. Para nuestros propósitos, los vectores aleatorios continuos con densidad van a ser relevantes, y por tanto restringiremos nuestra atención a estos.

Definición 1.12. *Si la distribución de un vector \mathbf{X} tiene función de densidad de probabilidad $f_{\mathbf{X}}$, se puede representar la función de distribución de probabilidad conjunta $F_{\mathbf{X}}$ de \mathbf{X} como*

$$\begin{aligned} F_{\mathbf{X}}(x_1, \dots, x_n) &= \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{\mathbf{X}}(y_1, \dots, y_n) dy_1 \cdots dy_n, \\ &(x_1, \dots, x_n) \in \mathbb{R}^n, \end{aligned}$$

donde la densidad es una función que satisface

$$f_{\mathbf{X}}(\mathbf{x}) \geq 0 \quad \forall \mathbf{x} \in \mathbb{R}^n, \quad \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{X}}(x_1, \dots, x_n) dx_1 \cdots dx_n = 1.$$

Si un vector \mathbf{X} tiene función de densidad de probabilidad $f_{\mathbf{X}}$, todas sus componentes X_i , los vectores de los pares (X_i, X_j) , tripletas (X_i, X_j, X_k) , etc., tienen una función de densidad de probabilidad. Estas son llamadas *funciones de densidad de probabilidad marginales*.

Además, la esperanza de un vector aleatorio tiene una función similar como el valor medio de una variable aleatoria. Los valores $\mathbf{X}(\omega)$ están concentrados alrededor de este. Así consideramos las siguientes definiciones.

Definición 1.13. Sea \mathbf{X} un vector aleatorio.

i) La **esperanza o media** de \mathbf{X} está dado por:

$$\mu_{\mathbf{X}} = E(\mathbf{X}) = (E(X_1), \dots, E(X_n)).$$

ii) La **matriz de varianza-covarianza** de \mathbf{X} está definida por:

$$\Sigma_{\mathbf{X}} = (\text{cov}(X_i, X_j); i, j = 1, \dots, n),$$

donde

$$\text{cov}(X_i, X_j) = E[(X_i - \mu_{X_i})(X_j - \mu_{X_j})] = E(X_i X_j) - \mu_{X_i} \mu_{X_j},$$

es la **covarianza de X_i y X_j** . Notese que $\text{cov}(X_i, X_i) = \sigma_{X_i}^2$.

Es conveniente el estandarizar covarianzas de las variables aleatorias mediante la división por sus desviaciones estándar correspondientes. La cantidad resultante

$$\text{corr}(X_1, X_2) = \frac{\text{cov}(X_1, X_2)}{\sigma_{X_1} \sigma_{X_2}} = \frac{E[(X_1 - \mu_{X_1})(X_2 - \mu_{X_2})]}{\sigma_{X_1} \sigma_{X_2}},$$

es la *correlación* de X_1 y X_2 . Como resultado de esta estandarización, la correlación de dos variables aleatorias está siempre entre -1 y 1.

Ejemplo 1.3. (Vector aleatorio gaussiano.)

Un vector aleatorio normal o gaussiano tiene función de distribución de probabilidad acumulada normal o gaussiana. La función de distribución de probabilidad normal o gaussiana de dimensión n está dada por su función de densidad de probabilidad

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} (\det \Sigma)^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mu) \Sigma^{-1} (\mathbf{x} - \mu)' \right\}, \quad \mathbf{x} \in \mathbb{R}^n, \quad (1.3)$$

con parámetros $\mu \in \mathbb{R}^n$ y Σ , una matriz simétrica definida positiva de tamaño $n \times n$, donde Σ^{-1} es su inversa y $\det \Sigma$ su determinante. El parámetro μ es la esperanza $\mu_{\mathbf{X}}$ de \mathbf{X} y Σ es la matriz de varianza-covarianza $\Sigma_{\mathbf{X}}$. Así la función de densidad de probabilidad del vector gaussiano, y por tanto su función de distribución de probabilidad, están completamente determinadas a través de su esperanza y matriz de varianza-covarianza.

1.1.3. Independencia y dependencia.

Intuitivamente, si tenemos dos experimentos, la independencia significa que el primer experimento no influye en el resultado del segundo y viceversa. En esta sección daremos algunas definiciones esenciales y propiedades de eventos independientes y variables aleatorias independientes.

Definición 1.14. *Dos eventos A_1 y A_2 son **independientes** si*

$$P(A_1 \cap A_2) = P(A_1)P(A_2).$$

Definición 1.15. *Dos variables aleatorias X_1 y X_2 son **independientes** si*

$$P(X_1 \in \mathcal{B}_1, X_2 \in \mathcal{B}_2) = P(X_1 \in \mathcal{B}_1)P(X_2 \in \mathcal{B}_2), \quad \mathcal{B}_1, \mathcal{B}_2 \subset \mathbb{R}.$$

Esto significa que los eventos $\{X_1 \in \mathcal{B}_1\}$ y $\{X_2 \in \mathcal{B}_2\}$ son independientes.

Alternativamente, se puede definir independencia mediante funciones de distribución de probabilidad conjunta y densidad de probabilidad. Las variables aleatorias X_1 y X_2 son independientes si, y sólo si

$$F_{X_1, X_2}(x_1, x_2) = F_{X_1}(x_1)F_{X_2}(x_2), \quad x_1, x_2 \in \mathbb{R}.$$

Supóngase que (X_1, X_2) tiene función de densidad de probabilidad conjunta f_{X_1, X_2} con funciones de densidad de probabilidad marginales f_{X_1} y f_{X_2} . Entonces se puede probar que las variables aleatorias X_1 y X_2 son independientes si, y sólo si

$$f_{X_1, X_2}(x_1, x_2) = f_{X_1}(x_1)f_{X_2}(x_2), \quad x_1, x_2 \in \mathbb{R}.$$

La definición de independencia puede ser extendida a un número arbitrario finito de eventos y vectores aleatorios. La independencia de los componentes de un vector aleatorio implica la independencia de cada par de sus componentes, pero lo inverso generalmente no es cierto.

Definición 1.16. *Los eventos A_1, \dots, A_n son **independientes** si, para cualquier elección de índices $1 \leq i_1 < \dots < i_k \leq n$ y enteros $1 \leq k \leq n$, se cumple*

$$P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k}).$$

Definición 1.17. *Las variables aleatorias X_1, \dots, X_n son **independientes** si, para cualquier elección de índices $1 \leq i_1 < \dots < i_k \leq n$, enteros $1 \leq k \leq n$ y subconjuntos $\mathcal{B}_1, \dots, \mathcal{B}_n$ de \mathbb{R} , se cumple*

$$P(X_{i_1} \in \mathcal{B}_{i_1}, \dots, X_{i_k} \in \mathcal{B}_{i_k}) = P(X_{i_1} \in \mathcal{B}_{i_1}) \cdots P(X_{i_k} \in \mathcal{B}_{i_k}).$$

Esto significa que los eventos $\{X_1 \in \mathcal{B}_1\}, \dots, \{X_n \in \mathcal{B}_n\}$ son independientes.

Las variables aleatorias X_1, \dots, X_n son independientes si, y sólo si su función de distribución de probabilidad conjunta puede ser escrita como sigue:

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n), \quad (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Si el vector aleatorio $\mathbf{X} = (X_1, \dots, X_n)$ tiene función de densidad de probabilidad $f_{\mathbf{X}}$, entonces se puede probar que X_1, \dots, X_n son independientes si, y sólo si

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n), \quad (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Una consecuencia importante de las variables aleatorias es la siguiente propiedad:

Propiedad 1.1. *Si X_1, \dots, X_n son independientes, entonces para cualesquiera funciones reales g_1, \dots, g_n , se cumple*

$$E[g_1(X_1) \cdots g_n(X_n)] = E[g_1(X_1)] \cdots E[g_n(X_n)]$$

siempre y cuando las esperanzas consideradas esten bien definidas.

En particular, podemos concluir que las variables aleatorias independientes X_1 y X_2 son no correlacionadas, es decir, $\text{corr}(X_1, X_2) = \text{cov}(X_1, X_2) = 0$. Lo inverso generalmente no se cumple.

En lo subsiguiente, con frecuencia trataremos con colecciones infinitas $(X_t, t \in T)$ de variables aleatorias X_t , es decir, T es un conjunto infinito de índices. En este arreglo, incluso se introducirá la independencia.

Definición 1.18. *La colección de variables aleatorias $(X_t, t \in T)$ es **independiente** si para cada elección distinta de índices $t_1, \dots, t_n \in T$ y $n \geq 1$ las variables aleatorias X_{t_1}, \dots, X_{t_n} son independientes. Esta colección es **independiente e idénticamente distribuida** (iid) si esta es independiente y todas las variables aleatorias X_t tienen la misma distribución.*

1.2. Procesos estocásticos.

Definición 1.19. *Un **proceso estocástico** X es una colección de variables aleatorias*

$$(X_t, t \in T) = (X_t(\omega), t \in T, \omega \in \Omega),$$

definida en algún espacio Ω .

Para nuestros propósitos, T es con frecuencia un intervalo, por ejemplo $T = [a, b]$, $[a, b)$ o $[a, \infty)$ para $a < b$. Entonces llamamos a X un proceso de tiempo continuo en contraste a un proceso de tiempo discreto. En el último caso, T es un conjunto finito o infinito numerable. Por obvias razones, el índice t de una variable aleatoria X_t es frecuentemente referida al tiempo, y seguiremos esta convención.

Definición 1.20. *Un proceso estocástico X es una función de dos variables. Para un instante fijo de tiempo t , esta es una variable aleatoria:*

$$X_t = X_t(\omega), \quad \omega \in \Omega$$

Para una salida aleatoria fija $\omega \in \Omega$, esta es una función del tiempo:

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

*Esta función es llamada una **realización**, una **trayectoria** o una **trayectoria de la muestra** del proceso X .*

Tenemos que los conceptos de variable aleatoria X y de proceso estocástico $(X_t, t \in T)$ no son muy diferentes. Ambos tienen trayectorias, pero la trayectoria $X(\omega)$ con $\omega \in \Omega$ de una variable aleatoria es un número, mientras que la trayectoria $X_t(\omega)$, $t \in T$, de un proceso estocástico es una función en T . Así sería correcto entender un proceso estocástico como un elemento aleatorio tomando funciones como valores. Más aún, podemos interpretar una variable aleatoria y un vector aleatorio como un proceso estocástico especial con un conjunto finito de índices T .

En analogía a las variables aleatorias y los vectores aleatorios se desea introducir características no aleatorias de un proceso estocástico tales como su distribución, esperanza, etc. y describir su estructura de dependencia. Esta es una tarea mucho más complicada que la descripción de un vector aleatorio. En efecto, un proceso estocástico no trivial $X = (X_t, t \in T)$ con un conjunto infinito de índices T es un objeto de dimensión infinita; esto puede ser entendido como una colección infinita de variables aleatorias X_t , $t \in T$. Ya que los valores de X son funciones sobre T , la distribución de X puede ser definida sobre subconjuntos de un cierto *espacio de funciones*, es decir

$$P(X \in A), \quad A \in \mathcal{F}, \tag{1.4}$$

donde \mathcal{F} es una colección de subconjuntos adecuados de este espacio de funciones.

La observación clave es que un proceso estocástico puede ser interpretado como una colección de vectores aleatorios.

Definición 1.21. *Las **distribuciones finito-dimensionales** (*fidis*) de un proceso estocástico X son las distribuciones de vectores de dimensión finita*

$$(X_{t_1}, \dots, X_{t_n}), \quad t_1, \dots, t_n \in T,$$

para todas las posibles elecciones de tiempos $t_1, \dots, t_n \in T$ y cada $n \geq 1$.

Se puede conceptualizar las fidis mucho más fácilmente que la complicada distribución (1.4) de un proceso estocástico. Se puede mostrar que las fidis determinan la distribución de X . En este sentido, nos referimos a la colección de las fidis como la *distribución de los procesos estocásticos*.

Los procesos estocásticos pueden ser clasificados de acuerdo a diferentes criterios. Uno de ellos es por el tipo de sus fidis

Ejemplo 1.4. (Proceso gaussiano.)

Un proceso estocástico es llamado gaussiano si todas sus fidis son gaussianas multivariadas con función de densidad de probabilidad (1.3). Del ejemplo 1.3 se tiene que los parámetros μ y Σ de un vector gaussiano son su esperanza y matriz de varianza-covarianza, respectivamente. Por lo tanto, la distribución de un proceso estocástico gaussiano es determinada solamente por la colección de las esperanzas y matrices de varianza-covarianza de sus fidis.

Un proceso gaussiano simple sobre $T = [0, 1]$ consiste de variables aleatorias $N(0, 1)$ iid. En este caso las fidis están caracterizadas por las funciones de distribución

$$\begin{aligned} P(X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n) &= P(X_{t_1} \leq x_1) \cdots P(X_{t_n} \leq x_n) \\ &= \Phi(x_1) \cdots \Phi(x_n), \end{aligned}$$

$$0 \leq t_1 \leq \dots \leq t_n \leq 1, \quad (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Para estos procesos, las trayectorias son muy irregulares. ■

Para un vector aleatorio $\mathbf{X} = (X_1, \dots, X_n)$ definimos la esperanza $\mu_{\mathbf{X}} = (E(X_1), \dots, E(X_n))$ y la matriz de varianza-covarianza $\Sigma_{\mathbf{X}} = (\text{cov}(X_i, X_j))$, $i, j = 1, \dots, n$. Un proceso estocástico $X = (X_t, t \in T)$ puede ser considerado como la colección de todos los vectores aleatorios $(X_{t_1}, \dots, X_{t_n})$ para $t_1, \dots, t_n \in T$ y $n \geq 1$. Para cada uno de ellos podemos determinar la esperanza y la matriz de varianza-covarianza. Alternativamente podemos considerar estas cantidades como funciones de $t \in T$:

Definición 1.22. Sea $X = (X_t, t \in T)$ un proceso estocástico.

i) La **función de esperanza** de X está dada por:

$$\mu_X(t) = \mu_{X_t} = E(X_t), \quad t \in T.$$

ii) La **función de covarianza** de X se define como:

$$c_X(t, s) = \text{cov}(X_t, X_s) = E[(X_t - \mu_X(t))(X_s - \mu_X(s))], \quad t, s \in T.$$

iii) La **función de varianza** de X esta definida por:

$$\sigma_x^2(t) = c_X(t, t) = \text{var}(X_t), \quad t \in T.$$

Como para un vector aleatorio, la función de esperanza $\mu_X(t)$ es una cantidad determinística alrededor de la cual las trayectorias de X están concentradas. La función de covarianza $c_X(t, s)$ es una medida de dependencia en el proceso X . La función de varianza $\sigma_X^2(t)$ puede ser considerada como la medida de dispersión de una trayectoria de X alrededor de $\mu_X(t)$.

Otra forma de clasificación de procesos estocásticos consiste de la imposición de una estructura especial de dependencia.

Definición 1.23. El proceso $X = (X_t, t \in T)$, $T \subset \mathbb{R}$, es **estrictamente estacionario** si la fidis son invariantes bajo cambios de los índices t

$$(X_{t_1}, \dots, X_{t_n}) \stackrel{d}{=} (X_{t_1+h}, \dots, X_{t_n+h}), \quad (1.5)$$

para todas las posibles elecciones de índices $t_1, \dots, t_n \in T$, $n \geq 1$ y h tal que $t_1 + h, \dots, t_n + h \in T$.

Nota 1.2. Aquí $\stackrel{d}{=}$ significa la identidad de las distribuciones. Para los vectores aleatorios en (1.5), esto significa que sus funciones de distribución son idénticas.

Si se describe un proceso de la vida real (estrictamente o en un sentido amplio) mediante un proceso estocástico estacionario, entonces se creería que las propiedades características de este proceso no cambian cuando el tiempo pasa. La estructura de dependencia descrita por las fidis o la función de covarianza es invariante mediante cambios en el tiempo, la cual es una restricción relativamente más fuerte sobre el proceso subyacente, sin embargo, esta es una suposición estándar en muchos campos relacionados con la probabilidad, tales como la estadística y el análisis de series de tiempo.

La propiedad estacionaria puede también ser impuesta sobre los incrementos del proceso. El proceso en si mismo es entonces, no necesariamente estacionario.

Definición 1.24. Sean $X = (X_t, t \in T)$ un proceso estocástico y $T \subset \mathbb{R}$ un intervalo. Se dice que X tiene **incrementos estacionarios** si

$$X_t - X_s \stackrel{d}{=} X_{t+h} - X_{s+h} \text{ para todo } s, t \in T \text{ y } h \text{ con } t+h, s+h \in T.$$

Se dice que X tiene **incrementos independientes** si para cada elección de $t_i \in T$ con $t_1 < \dots < t_n$ y $n > 1$,

$$X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}},$$

son variables aleatorias independientes.

1.3. La esperanza condicional.

1.3.1. La esperanza condicional bajo condiciones discretas.

Sabemos que la probabilidad condicional de que ocurra A dado B , es:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P \neq 0.$$

Claramente,

$$P(A|B) = P(A) \text{ si, y sólo si } A \text{ y } B \text{ son independientes.}$$

La probabilidad $P(A|B)$ puede ser interpretada como sigue. Supóngase que el evento B ocurre. Esta es información adicional lo cual substancialmente cambia la medida de probabilidad subyacente. En particular, se asigna la probabilidad 0 a B^c (sabemos que B^c no ocurrirá) y 1 a B . El evento B se convierte en nuestro nuevo espacio de probabilidad Ω' . Todos los eventos de interés son ahora subconjuntos de $\Omega' : A \cap B \subset \Omega'$. Con el objetivo de obtener una nueva medida de probabilidad en Ω' se tiene que normalizar las viejas probabilidades $P(A \cap B)$ por $P(B)$. En resumen, la ocurrencia de B hace que nuestro espacio original Ω se reduzca a Ω' y las probabilidades originales $P(A)$ tengan que ser reemplazadas con $P(A|B)$.

Dado que $P(B) > 0$, podemos definir la función de distribución de probabilidad condicional de una variable aleatoria X dado B , como:

$$F_X(x|B) = \frac{P(X \leq x, B)}{P(B)}, \quad x \in \mathbb{R},$$

e incluso la esperanza condicional de X dado B

$$E(X|B) = \frac{E(XI_B)}{P(B)}, \tag{1.6}$$

donde

$$I_B(\omega) = \begin{cases} 1 & \text{si } \omega \in B, \\ 0 & \text{si } \omega \notin B, \end{cases}$$

denota la *función indicadora del evento* B . Con el objeto de discutir la ecuación (1.6), asumiremos por el momento que $\Omega = \mathbb{R}$. Si X es una variable aleatoria discreta con valores x_1, x_2, \dots , entonces (1.6) se convierte en

$$E(X|B) = \sum_{k=1}^{\infty} x_k \frac{P(\{\omega : X(\omega) = x_k\} \cap B)}{P(B)} = \sum_{k=1}^{\infty} x_k P(X = x_k|B).$$

Si X tiene función de densidad de probabilidad f_X , entonces (1.6) se transforma en

$$E(X|B) = \frac{1}{P(B)} \int_{-\infty}^{\infty} x I_B(x) f_X(x) dx = \frac{1}{P(B)} \int_B x f_X(x) dx.$$

Usualmente se escribe $\int_B g(x) dx$ por $\int_{-\infty}^{\infty} g(x) I_B(x) dx$.

Ejemplo 1.5. (La esperanza condicional de una variable aleatoria uniforme.)

Consideremos la variable aleatoria $X(\omega) = \omega$ en el espacio $[0, 1]$ dotado con la medida de probabilidad P tal que

$$P((a, b]) = b - a, \quad (a, b] \subset (0, 1].$$

Claramente, X tiene una distribución uniforme en $(0, 1]$, así su función de distribución de probabilidad acumulada esta dada por

$$F_X(x) = P(\{\omega : X(\omega) = \omega \leq x\}) = \begin{cases} P(\emptyset) = 0 & \text{si } x \leq 0, \\ P((0, x]) = x & \text{si } x \in (0, 1], \\ P((0, 1]) = 1 & \text{si } x > 1. \end{cases}$$

Ahora, supongamos que uno de los eventos

$$A_i = \left(\frac{i-1}{n}, \frac{i}{n} \right], \quad i = 1, \dots, n$$

ocurre. Teniendo en cuenta que $f_X(x) = 1$ sobre $(0, 1]$ y además que $P(A_i) = 1/n$, entonces

$$E(X|A_i) = \frac{1}{P(A_i)} \int_{A_i} x f_X(x) dx = n \int_{(i-1)/n}^{i/n} x dx = \frac{1}{2} \frac{2i-1}{n}. \quad (1.7)$$

Las esperanzas condicionales $E(X|A_i)$ se muestran en la Figura 1.1. El valor $E(X|A_i)$ es la esperanza actualizada en el nuevo espacio A_i , dada la información que A_i ocurrió. ■

Ahora considérese una variable aleatoria discreta Y en Ω que toma valores distintos y_i en los conjuntos A_i , es decir,

$$A_i = \{\omega : Y(\omega) = y_i\} \quad i = 1, 2, \dots$$

Claramente, (A_i) es una partición de Ω , es decir,

$$A_i \cap A_j = \emptyset \text{ para } i \neq j \text{ y } \bigcup_{i=1}^{\infty} A_i = \Omega. \quad (1.8)$$

Se supone además por conveniencia que $P(A_i) > 0$ para toda i .

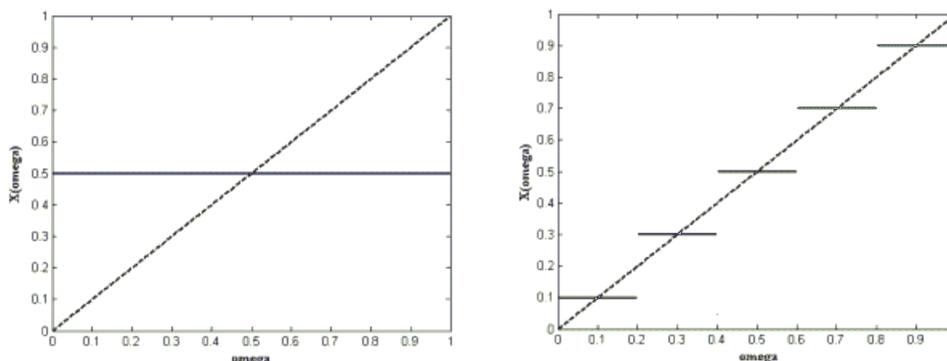


Figura 1.1: A la izquierda una variable aleatoria X en $(0, 1]$ (línea punteada) y su esperanza (línea continua). A la derecha la variable aleatoria X (línea punteada) y sus esperanzas condicionales $E(X|A_i)$ (líneas continuas), donde $A_i = ((i - 1)/5, i/5]$, $i = 1, \dots, 5$.

Definición 1.25. Para una variable aleatoria X en Ω con $E|X| < \infty$, se define la esperanza condicional de X dado Y , una variable aleatoria discreta, como

$$E(X|Y)(\omega) = E(X|A_i) = E(X|Y = y_i), \quad \forall \omega \in A_i, \quad i = 1, 2, \dots \quad (1.9)$$

Si sabemos que A_i ocurre, podemos restringirnos a los ω 's en A_i . Para esos ω 's, $E(X|Y)(\omega)$ coincide con la esperanza condicional $E(X|A_i)$.

Ejemplo 1.6. (Continuación del ejemplo 1.5.)

Interpretemos las $E(X|A_i)$'s en (1.7) como los valores de una variable aleatoria discreta $E(X|Y)$, donde Y es constante sobre los conjuntos $A_i = ((i - 1)/n, i/n]$. Véase la Figura 1.1 para una ilustración de esta variable aleatoria. En este sentido, $E(X|Y)$ no es nada más que una versión aproximada de la variable aleatoria original X , dada la información que cualquiera de los A_i 's ocurren. ■

En lo que sigue se darán algunas propiedades elementales de la variable aleatoria $E(X|Y)$.

Teorema 1.1. La esperanza condicional es lineal, es decir, para variables aleatorias X_1, X_2 y constantes c_1, c_2 se tiene que:

$$E([c_1X_1 + c_2X_2]|Y) = c_1E(X_1|Y) + c_2E(X_2|Y).$$

Demostración. Aplicando la propiedad definida por la ecuación (1.6) se sigue que

$$\begin{aligned} E([c_1X_1 + c_2X_2]|Y) &= \frac{E([c_1X_1 + c_2X_2]I_Y)}{P(Y)} = \frac{E(c_1X_1I_Y + c_2X_2I_Y)}{P(Y)} \\ &= c_1 \frac{E(X_1I_Y)}{P(Y)} + c_2 \frac{E(X_2I_Y)}{P(Y)} = c_1E(X_1|Y) + c_2E(X_2|Y). \end{aligned}$$

□

Teorema 1.2. *Las esperanzas de X y $E(X|Y)$ son las mismas, es decir, $E(X) = E[E(X|Y)]$.*

Demostración. Esto se sigue de una aplicación directa de las propiedades, definidas por las ecuaciones (1.6), (1.9) y por la observación de que $E(X|Y)$ es una variable aleatoria discreta.

$$E(E(X|Y)) = \sum_{i=1}^{\infty} E(X|A_i)P(A_i) = \sum_{i=1}^{\infty} E(XI_{A_i}) = E\left(X \sum_{i=1}^{\infty} I_{A_i}\right) = E(X).$$

Aquí, incluso usamos (1.8), así que

$$\sum_{i=1}^{\infty} I_{A_i} = I_{\cup_{i=1}^{\infty} A_i} = I_{\Omega} = 1.$$

□

Teorema 1.3. *Si X y Y son independientes, entonces $E(X|Y) = E(X)$.*

Demostración. De la definición [1.15] se sigue que la independencia de X y Y , implica

$$P(X \in A, Y = y_i) = P(X \in A)P(Y = y_i) = P(X \in A)P(A_i), \quad (1.10)$$

considerando la variable aleatoria I_{A_i} y notando que

$$\{\omega : I_{A_i}(\omega) = 1\} = A_i = \{\omega : y(\omega) = y_i\},$$

podemos reescribir (1.10) como sigue

$$P(X \in A, I_{A_i} = 1) = P(X \in A)P(I_{A_i} = 1).$$

La relación análoga, donde $\{I_{A_i} = 1\}$ es reemplazada con $\{I_{A_i} = 0\}$ incluso se cumple.

Por lo tanto las variables aleatorias X y I_{A_i} , son independientes y para $\omega \in A_i$, se tiene que:

$$E(X|Y)(\omega) = E(X|A_i) = \frac{E(XI_{A_i})}{P(A_i)} = \frac{E(X)E(I_{A_i})}{P(A_i)} = E(X),$$

donde usamos que

$$E(I_{A_i}) = 0 \cdot P(A_i^c) + 1 \cdot P(A_i) = P(A_i).$$

□

Resumiendo los puntos más importantes de esta sección:

- *La esperanza condicional $E(X|Y)$ de X dado una variable aleatoria discreta Y es una variable aleatoria discreta.*
- *Esto coincide con la esperanza condicional clásica $E(X|Y = y_i)$ en los conjuntos $A_i = \{\omega : Y(\omega) = y_i\}$.*
- *En este sentido esta es una versión aproximada de X .*
- *Cuantos menos valores tenga Y , es más burda la variable aleatoria. En particular, si Y es igual a una constante, entonces $E(X|Y) = E(X)$, si Y asume dos distintos valores, así también $E(X|Y)$, etc.*
- *La esperanza condicional $E(X|Y)$ no es una función de X , es simplemente una función de Y . La variable aleatoria X solo determina el tipo de función. En efecto, podemos escribir*

$$E(X|Y)g(Y), \text{ donde } g(y) = \sum_{i=1}^{\infty} E(X|Y = y_i)I_{\{y_i\}}(y).$$

1.3.2. σ -Álgebras.

En la sección 1.3.1 se introdujo la esperanza condicional $E(X|Y)$ de una variable aleatoria X bajo la condición discreta (es decir, la variable aleatoria discreta) Y . Recordando de (1.8), que los valores de Y realmente no importan para la definición de $E(X|Y)$; pero era crucial que Y asumiera los distintos valores y_i en los ω -conjuntos A_i . Así, la esperanza condicional $E(X|Y)$ puede ser realmente entendida como una variable aleatoria construida a partir de una colección $\sigma(Y)$, de subconjuntos de Ω . Por tanto se podría, de una manera simplificada, escribir $E(X|\sigma(Y))$.

Obviamente, la colección $\sigma(Y)$ nos proporciona la información acerca de la estructura de la variable $Y(\omega)$, como función de $\omega \in \Omega$.

En lo que sigue, queremos hacer más preciso que significa la *colección* $\sigma(Y)$ de subconjuntos de Ω . La llamaremos σ -campo o una σ -álgebra. Su definición es la siguiente:

Definición 1.26. Una σ -álgebra \mathcal{F} (en Ω) es una colección de subconjuntos de Ω que satisfacen las siguientes condiciones:

- i) No es vacía; $\emptyset \in \mathcal{F}$ y $\Omega \in \mathcal{F}$.
- ii) Si $A \in \mathcal{F}$, entonces $A^c \in \mathcal{F}$.
- iii) Si $A_1, A_2, \dots \in \mathcal{F}$ entonces

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{F} \quad \text{y} \quad \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}.$$

Ejemplo 1.7. (Algunas σ -álgebras elementales.)

Las siguientes colecciones de subconjuntos de Ω son ejemplos de σ -álgebras:

1. $\mathcal{F}_1 = \{\emptyset, \Omega\}$.
2. $\mathcal{F}_2 = \{\emptyset, \Omega, A, A^c\}$ para algún $A \neq \emptyset$ y $A \neq \Omega$.
3. $\mathcal{F}_3 = \mathcal{P}(\Omega) = \{A : A \subset \Omega\}$.

\mathcal{F}_1 es la σ -álgebra más pequeña en Ω y \mathcal{F}_3 , el conjunto potencia de Ω , es la más grande, dado que este contiene todos los posibles subconjuntos de Ω . ■

Ahora supóngase que \mathcal{C} es una colección de subconjuntos de Ω , pero no necesariamente una σ -álgebra. Agregando más conjuntos a \mathcal{C} , siempre se puede tener una σ -álgebra, por ejemplo, el conjunto potencia $\mathcal{P}(\Omega)$. Sin embargo, existen razones matemáticas que muestran que $\mathcal{P}(\Omega)$ es en general demasiado grande. Pero se puede probar incluso que:

Observación 1.1. Para una colección \mathcal{C} de subconjuntos de Ω , existe una σ -álgebra $\sigma(\mathcal{C})$ más pequeña en Ω conteniendo a \mathcal{C} .

Llamaremos a $\sigma(\mathcal{C})$, la σ -álgebra generada por \mathcal{C} .

Ejemplo 1.8. (σ -Álgebras generadas a partir de colecciones de subconjuntos de Ω .)

Recordando las σ -álgebras del ejemplo 1.7 tenemos que $\mathcal{F}_i = \sigma(\mathcal{C}_i)$

1. Si dado $\mathcal{C}_1 = \{\emptyset\}$, se agrega Ω a \mathcal{C}_1 .

2. Si dado $\mathcal{C}_2 = \{A\}$, se agregan A^c , \emptyset y Ω a \mathcal{C}_2 .
3. Si dado $\mathcal{C}_3 = \mathcal{F}_3$, se cumple por la definición de \mathcal{F}_3 .

■

En general, es difícil, sino imposible, dar una descripción constructiva de los elementos en $\sigma(\mathcal{C})$. La σ -álgebra $\sigma(Y)$ generada por una variable aleatoria discreta Y es una excepción como se muestra a continuación.

Ejemplo 1.9. (La σ -álgebra generada por una variable aleatoria discreta.)

Considerando una variable aleatoria discreta Y con valores distintos y_i y definimos los subconjuntos $A_i = \{\omega : Y(\omega) = y_i\}$, los cuales constituyen una partición disjunta de Ω . Escogemos

$$\mathcal{C} = \{A_1, A_2, \dots\}.$$

Ya que $\sigma(\mathcal{C})$ es una σ -álgebra, esta debe contener todos los conjuntos de la forma

$$A = \bigcup_{i \in I} A_i, \tag{1.11}$$

donde I es cualquier subconjunto de $\mathbb{N} = \{1, 2, \dots\}$, incluyendo $I = \emptyset$ (dando $A = \emptyset$) e $I = \mathbb{N}$ (dando $A = \Omega$). Ya que los conjuntos (1.11) necesariamente pertenecen a $\sigma(\mathcal{C})$, la σ -álgebra más pequeña que contiene a \mathcal{C} , incluso tenemos que $\sigma(Y) = \sigma(\mathcal{C})$. Después llamaremos a $\sigma(Y)$, la σ -álgebra generada por Y .

Nótese que $\sigma(Y)$ contiene todos los conjuntos de la forma

$$A_{a,b} = \{Y \in (a, b]\} = \{\omega : a < Y(\omega) \leq b\}, \quad -\infty < a < b < \infty.$$

De hecho, el conjunto $I = \{i : a < y_i \leq b\}$ es un subconjunto de \mathbb{N} , así

$$A_{a,b} = \bigcup_{i \in I} \{\omega : Y(\omega) = y_i\} \in \sigma(Y).$$

■

Ejemplo 1.10. (Los conjuntos de Borel.)

Tomamos $\Omega = \mathbb{R}$ y

$$\mathcal{C}^{(1)} = \{(a, b) : -\infty < a < b < \infty\}.$$

La σ -álgebra $\mathcal{B}_1 = \sigma(\mathcal{C}^{(1)})$ contiene muchos subconjuntos generales de \mathbb{R} . Esta es llamada la σ -álgebra de Borel, sus elementos son los conjuntos de Borel. Por ejemplo, es un hecho, pero no es fácil de verificar que \mathcal{B}_1 es un subconjunto

propio del conjunto potencia $\mathcal{P}(\mathbb{R})$. Se puede introducir incluso las σ -álgebras de los conjuntos n -dimensionales de Borel $\mathcal{B}_n = \sigma(\mathcal{C}^{(n)})$ donde $\Omega = \mathbb{R}^n$ y

$$\mathcal{C}^{(n)} = \{(\mathbf{a}, \mathbf{b}) : -\infty < a_i < b_i < \infty, i = 1, \dots, n\}.$$

Los conjuntos $(\mathbf{a}, \mathbf{b}]$ son llamados rectángulos. Cualquier conjunto *razonable* de \mathbb{R} es un conjunto de Borel. Por ejemplo, bolas, esferas, curvas suaves, superficies, etc., son conjuntos de Borel, así también son los conjuntos abiertos y cerrados. Con el objetivo de mostrar que dado un subconjunto $C \subset \mathbb{R}^n$ es un conjunto de Borel, es necesario obtener C mediante un número numerable de operaciones $\cap, \cup, ^c$ actuando sobre los rectángulos. ■

Ahora recordemos el ejemplo 1.9, donde generamos la σ -álgebra $\sigma(Y)$ de los conjuntos $A_i = \{\omega : Y(\omega) = y_i\}$ de una variable aleatoria discreta Y con valores y_i . Si Y es una variable aleatoria tomando valores continuos, la σ -álgebra generada a partir de los conjuntos de la forma $\{\omega : Y(\omega) = y\}$, $y \in \mathbb{R}$, no es, desde un punto de vista matemático lo suficiente rico. Por ejemplo, los conjuntos de la forma $\{\omega : a < Y(\omega) \leq b\}$ no pertenecen a tal σ -álgebra, pero es deseable tenerlos en $\sigma(Y)$.

Se vio en el ejemplo 1.9 que, para una variable aleatoria discreta Y , los conjuntos $\{\omega : a < Y(\omega) \leq b\}$ pertenecen a $\sigma(Y)$. Por consiguiente requerimos de una suposición mínima que estos conjuntos pertenezcan a $\sigma(Y)$ cuando Y es cualquier variable aleatoria. Ya que incluso estamos interesados en vectores aleatorios \mathbf{Y} , inmediatamente definimos las σ -álgebras $\sigma(\mathbf{Y})$ para el caso multivariado.

Definición 1.27. Sea $\mathbf{Y} = (Y_1, \dots, Y_n)$ un vector aleatorio de dimensión n , es decir, Y_1, \dots, Y_n son variables aleatorias. La σ -álgebra $\sigma(\mathbf{Y})$ es la menor σ -álgebra que contiene todos los conjuntos de la forma

$$\{\mathbf{Y} \in (\mathbf{a}, \mathbf{b}]\} = \{\omega : a_i < Y_i(\omega) \leq b_i, i = 1, \dots, n\},$$

$$-\infty < a_j < b_j < \infty, j = 1, \dots, n.$$

Llamaremos a $\sigma(\mathbf{Y})$ la σ -álgebra generada por el vector aleatorio \mathbf{Y} .

Un elemento de $\sigma(\mathbf{Y})$ nos dice para que $\omega \in \Omega$ el vector aleatorio \mathbf{Y} toma valores en un rectángulo $(\mathbf{a}, \mathbf{b}]$ o en un conjunto de Borel más general.

Nota 1.3. La σ -álgebra $\sigma(\mathbf{Y})$ generada por \mathbf{Y} contiene la información esencial acerca de la estructura del valor aleatorio \mathbf{Y} como una función de $\omega \in \Omega$. Esta contiene todos los conjuntos de la forma $\{\omega : \mathbf{Y}(\omega) \in C\}$ para todos los conjuntos de Borel $C \subset \mathbb{R}^n$.

Como se ve, es difícil de imaginar la σ -álgebra $\sigma(Y)$, y es incluso más difícil para un proceso estocástico.

Definición 1.28. Para un proceso estocástico $Y = (Y_t, t \in T, \omega \in \Omega)$, la σ -álgebra $\sigma(Y)$ es la menor σ -álgebra que contiene todos los conjuntos de la forma

$$\{\omega : \text{la trayectoria } (Y_t(\omega), t \in T) \text{ pertenece a } C\},$$

para todo conjunto adecuado C de funciones en T . Entonces $\sigma(Y)$ es llamada **la σ -álgebra generada por Y** .

Hemos aprendido que la σ -álgebra $\sigma(Y)$ es un objeto no trivial. En lo que sigue, queremos usar la siguiente regla con el objetivo de ser capaces de imaginar a $\sigma(Y)$.

Definición 1.29. Para una variable aleatoria, vector aleatorio o proceso estocástico Y en Ω , la σ -álgebra generada por Y contiene la información esencial acerca de la estructura de Y como una función de $\omega \in \Omega$. Esta consiste de todos los subconjuntos $\{\omega : Y(\omega) \in C\}$ para conjuntos C adecuados.

Dado que Y genera una σ -álgebra, incluso decimos que Y **contiene información representada por $\sigma(Y)$ o Y lleva la información de $\sigma(Y)$** .

Concluiremos esta sección con una observación de utilidad. Sea f una función actuando sobre Y , y considerese el conjunto

$$\{\omega : f(Y(\omega)) \in C\},$$

para conjuntos adecuados C . Para funciones f adecuadas, este conjunto pertenece a $\sigma(Y)$, es decir,

$$\sigma(f(Y)) \subset \sigma(Y).$$

Esto significa que una función f actuando sobre Y no provee de nueva información acerca de la estructura de Y . Decimos que la información contenida por $f(Y)$ está contenida en la información $\sigma(Y)$.

1.3.3. La esperanza condicional general.

En la sección 1.3.1 se definió la esperanza condicional $E(X|Y)$ de una variable aleatoria X , dada una variable aleatoria discreta Y . Esta definición no hace explícito el uso de los valores y_i de Y , pero este depende de los subconjuntos $A_i = \{\omega : Y(\omega) = y_i\}$ de Ω . Se mostró en el ejemplo 1.9 que la colección de los subconjuntos A_i generan la σ -álgebra $\sigma(Y)$. Este es el punto de inicio para la definición de la esperanza condicional $E(X|\mathcal{F})$ dada una σ -álgebra \mathcal{F} en Ω . En aplicaciones siempre tomaremos $\mathcal{F} = \sigma(Y)$ para variables aleatorias, vectores

aleatorios o procesos estocásticos Y apropiados. En la sección 1.3.2 se dejó claro que la información esencial acerca de la estructura de Y esta contenida en la σ -álgebra $\sigma(Y)$, generada por Y . En este sentido, decimos que Y tiene la información $\sigma(Y)$.

Definición 1.30. Sean Y, Y_1, Y_2 variables aleatorias, vectores aleatorios o procesos estocásticos en Ω y \mathcal{F} una σ -álgebra en Ω .

Decimos que:

- i) La información de Y está contenida en \mathcal{F} o Y no contiene más información que la que esta contenida en \mathcal{F} si $\sigma(Y) \subset \mathcal{F}$, y
- ii) Y_2 contiene más información que Y_1 si $\sigma(Y_1) \subset \sigma(Y_2)$.

Ahora estamos preparados para dar una definición rigurosa de la esperanza condicional $E(X|Y)$ bajo una σ -álgebra \mathcal{F} abstracta.

Definición 1.31. Una variable aleatoria Z es llamada la **esperanza condicional** de X dada la σ -álgebra \mathcal{F} (escribimos $Z = E(X|\mathcal{F})$) si

- i) Z no contiene más información que la que está contenida en \mathcal{F} :
 $\sigma(Z) \subset \mathcal{F}$.
- ii) Z satisface la relación

$$E(XI_A) = E(ZI_A) \quad \forall A \in \mathcal{F}. \quad (1.12)$$

La propiedad definida por la ecuación (1.12), muestra que las variables aleatorias X y $E(X|\mathcal{F})$ son *cercanas* unas a otras, no en el sentido de que ellas coinciden para cualquiera $\omega \in \Omega$, pero aproximaciones (esperanzas) de X y $E(X|\mathcal{F})$ sobre conjuntos adecuados A son las mismas. Esto apoya nuestra conceptualización acerca de la variable aleatoria $E(X|\mathcal{F})$, la cual obtuvimos en la sección 1.3.1.

Nota 1.4. La *esperanza condicional* es una versión aproximada de la variable aleatoria original X .

La propiedad definida en (1.12) deja cierta libertad para la construcción de la variable aleatoria $Z = E(X|\mathcal{F})$. Esto significa que pueden existir versiones Z' de $E(X|\mathcal{F})$ las cuales pueden diferir de Z sobre conjuntos de probabilidad 0. Por esta razón, todas las relaciones que envuelven a $E(X|\mathcal{F})$ deberían en realidad ser interpretadas en un sentido más seguro, e incluso deberíamos indicar esta incertidumbre en toda fórmula relevante con la etiqueta *c.s.* (véase apéndice B). Sin embargo, una vez que hemos señalado este hecho, encontraremos conveniente el suprimir esta etiqueta donde quiera.

Ejemplo 1.11. (La esperanza condicional bajo condiciones discretas.)

Queremos mostrar que la definición de $E(X|Y)$ en la sección 1.3.1 se puede ver como un caso especial de la definición [1.31] de $E(X|\mathcal{F})$ donde $\mathcal{F} = \sigma(Y)$. Del ejemplo 1.9 tenemos que cada elemento A de $\sigma(Y)$ es de la forma,

$$A = \bigcup_{i \in I} A_i = \bigcup_{i \in I} \{\omega : Y(\omega) = y_i\} \quad I \subset \mathbb{N}. \quad (1.13)$$

En (1.9) definimos $E(X|Y)$ como la variable aleatoria Z tal que

$$Z(\omega) = E(X|A_i) \text{ para } \omega \in A_i.$$

Recordando que Z es solamente una función de Y , no de X , por lo tanto $\sigma(Z) \subset \sigma(Y)$. Más aún, para A dado por (1.13),

$$E(XI_A) = E\left(X \sum_{i \in I} I_{A_i}\right) = \sum_{i \in I} E(XI_{A_i}).$$

Por otro lado, vimos que ZI_A es una variable aleatoria discreta con esperanza

$$E(ZI_A) = \sum_{i \in I} E(X|A_i)P(A_i) = \sum_{i \in I} E(XI_{A_i}).$$

Así Z satisface la relación definida en (1.12) y esta no contiene más información que Y (es una función de Y). Por consiguiente esta es en efecto la esperanza condicional de X dada la σ -álgebra $\sigma(Y)$. ■

En el caso de una variable aleatoria discreta Y hemos visto que $E(X|Y)$ y $E(X|\sigma(Y))$ representan la misma variable aleatoria. Esto sugiere la siguiente definición:

Definición 1.32. *Sea Y una variable aleatoria, un vector aleatorio o un proceso estocástico en Ω y $\sigma(Y)$ una σ -álgebra generada por Y . La esperanza condicional de una variable aleatoria X dada Y es definida por*

$$E(X|Y) = E(X|\sigma(Y)).$$

Ejemplo 1.12. (La probabilidad condicional y la esperanza condicional.)

La probabilidad condicional y la esperanza condicional son casos especiales de la noción general de la esperanza condicional definidos en la definición [1.31]. En efecto, sea B tal que $P(B) \neq 0$, $P(B^c) \neq 0$ y definimos $\mathcal{F}_B = \sigma(\{B\})$. Del ejemplo 1.8 tenemos que $\mathcal{F}_B = \{\emptyset, \Omega, B, B^c\}$ y considerando el ejemplo 1.11 tenemos que

$$E(X|\mathcal{F}_B)(\omega) = E(X|B) \text{ para } \omega \in B.$$

Esta es la noción clásica de la esperanza condicional. Si especificamos que $X = I_A$ para algún evento A , obtenemos para $\omega \in B$,

$$E(I_A|\mathcal{F}_B)(\omega) = E(I_A|B) = \frac{P(A \cap B)}{P(B)}.$$

El término del lado derecho es la probabilidad condicional de A dado B . ■

1.3.4. Reglas para el cálculo de las esperanzas condicionales.

Teorema 1.4. *Si $E|X| < \infty$, la esperanza condicional $E(X|\mathcal{F})$ existe y es única.*

La demostración de este teorema está fuera del alcance de las herramientas utilizadas, por lo que solo se usará como un hecho para saber que la esperanza condicional está bien definida. Con el objetivo de utilizar más adelante este concepto importante, se darán a continuación algunas reglas para el cálculo de las esperanzas condicionales.

Regla 1. *La esperanza condicional es lineal. Para variables aleatorias X_1 , X_2 y constantes c_1 , c_2 , se tiene*

$$E([c_1X_1, c_2X_2]|\mathcal{F}) = c_1E(X_1|\mathcal{F}) + c_2E(X_2|\mathcal{F}).$$

Regla 2. *Las esperanzas de X y $E(X|\mathcal{F})$ son las mismas:*

$$E(X) = E[E(X|\mathcal{F})].$$

Regla 3. *Si X y la σ -álgebra \mathcal{F} son independientes, entonces $E(X|\mathcal{F}) = E(X)$. En particular, si X y Y son independientes, entonces $E(X|Y) = E(X)$.*

Regla 4. *Si la σ -álgebra $\sigma(X)$, generada por la variable aleatoria X , está contenida en \mathcal{F} , entonces*

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

En particular si X es una función de Y , $\sigma(X) \subset \sigma(Y)$, así $E(X|Y) = X$.

Regla 5. *Si la σ -álgebra $\sigma(X)$, generada por la variable aleatoria X , está contenida en \mathcal{F} , entonces para toda variable aleatoria G*

$$E(XG|\mathcal{F}) = XE(G|\mathcal{F}).$$

En particular, si X es una función de Y , $\sigma(X) \subset \sigma(Y)$, así $E(XG|Y) = XE(G|Y)$.

Regla 6. Si \mathcal{F} y \mathcal{F}' son dos σ -álgebras con $\mathcal{F} \subset \mathcal{F}'$, entonces

$$E(X|\mathcal{F}) = E(E(X|\mathcal{F}')|\mathcal{F}),$$

$$E(X|\mathcal{F}) = E(E(X|\mathcal{F})|\mathcal{F}').$$

Regla 7. Si X es independiente de \mathcal{F} , y la información dada por la variable aleatoria, vector aleatorio o proceso estocástico G está contenida en \mathcal{F} , entonces para cualquier función $h(x, y)$

$$E(h(X, G)|\mathcal{F}) = E(E_X[h(X, G)]|\mathcal{F}),$$

donde $E_X[h(X, G)]$ significa que fijamos G y tomamos la esperanza con respecto a X .

Capítulo 2

Movimiento Browniano.

2.1. Definición y propiedades.

El movimiento Browniano juega un papel central en la teoría de la probabilidad, la teoría de los procesos estocásticos, física, finanzas, etc. Se comenzará con la definición de este importante proceso para después continuar con algunas de sus propiedades elementales.

Definición 2.1. *Un proceso estocástico $B = (B_t, t \in [0, \infty))$ es llamado un **movimiento Browniano (estándar)** o un **proceso de Wiener** si las siguientes condiciones se cumplen:*

- i) Comienza en cero: $B_0 = 0$.*
- ii) Tiene incrementos estacionarios e independientes.*
- iii) Para cada $t > 0$, B_t tiene una distribución normal $N(0, t)$.*
- iv) Tiene trayectorias continuas.*

Las fidsis (véase definición [1.21]) de un movimiento Browniano son gaussianas multivariadas, dado que B es un proceso gaussiano. Se puede verificar esta aseveración observando que el movimiento Browniano tiene incrementos gaussianos independientes y usando la fórmula para transformaciones lineales de un vector aleatorio gaussiano.

Observación 2.1. *Las variables aleatorias $B_t - B_s$ y B_{t-s} tienen una distribución $N(0, t - s)$ para $s < t$.*

Esto se sigue de los incrementos estacionarios. En efecto, $B_t - B_s$ tienen la misma distribución como $B_{t-s} - B_0 = B_{t-s}$, el cual es normal con media cero y varianza $t - s$. Así, la varianza es proporcional a la longitud del intervalo $[s, t]$. Esto significa intuitivamente que a un intervalo mayor, mayores serán las fluctuaciones del movimiento Browniano en este intervalo.

Nota 2.1. La identidad distribucional $B_t - B_s \stackrel{d}{=} B_{t-s}$ no implica la identidad de la forma de la trayectoria: en general,

$$B_t(\omega) - B_s(\omega) \neq B_{t-s}(\omega).$$

Ahora se tratará de determinar las funciones de esperanza y covarianza para el movimiento Browniano. Es inmediato de la definición (2.1) que el movimiento Browniano tiene función de esperanza

$$\mu_B(t) = EB_t = 0, \quad t \geq 0,$$

y dado que los incrementos $B_s - B_0 = B_s$ y $B_t - B_s$ son independientes para $t > s$, este tiene función de covarianza

$$\begin{aligned} \text{cov}_B(t, s) &= E\{[(B_t - B_s) + B_s]B_s\} = E[(B_t - B_s)B_s] + E(B_s^2) \\ &= E(B_t - B_s)E(B_s) + s = 0 + s = s, \quad 0 \leq s < t. \end{aligned}$$

Ya que un proceso gaussiano es caracterizado por sus funciones de esperanza y covarianza, podemos dar una definición alternativa:

Definición 2.2. Un movimiento Browniano es un proceso gaussiano con

$$\mu_B(t) = 0 \quad \text{y} \quad \text{cov}_B(t, s) = \text{mín}(s, t).$$

En lo que sigue, se fijará una trayectoria $B_t(\omega)$, $t \geq 0$ y $\omega \in \Omega$, y consideraremos sus propiedades. De la definición de movimiento Browniano [2.1] se tiene que sus trayectorias son continuas. Sin embargo, al simular estas trayectorias se observa que estas funciones que dependen de t son extremadamente irregulares. La razón principal es que los incrementos de B son independientes. En particular, los incrementos del movimiento Browniano sobre intervalos adyacentes son independientes cualquiera que sea la longitud de los intervalos.

Antes de ver que tan irregulares son las trayectorias del movimiento Browniano se definirá una clase de procesos estocásticos, los cuales contienen al movimiento Browniano como caso especial y donde todos los miembros de esta clase tienen trayectorias irregulares.

Definición 2.3. Un proceso estocástico $(X_t, t \in [0, \infty])$ es *H-auto-similar* para algún $H > 0$ si sus fids satisfacen la condición

$$(T^H X_{t_1}, \dots, T^H X_{t_n}) \stackrel{d}{=} (X_{Tt_1}, \dots, X_{Tt_n}),$$

para cada $T > 0$, y cualquier elección de $t_i \geq 0$, $i = 1, \dots, n$ y $n \geq 1$.

Nota 2.2. *La auto-similitud es una propiedad distribucional, no de la forma de la trayectoria, por lo que $\stackrel{d}{=}$ no puede ser reemplazado por $=$.*

En términos generales, la auto-similitud significa que patrones propiamente escalados de una trayectoria en cualquier intervalo grande o pequeño de tiempo tienen una forma semejante, pero no son idénticos. Ahora se probará que las trayectorias de procesos auto-similares son diferenciables en ninguna parte.

Proposición 2.1. *Suponga que (X_t) es un proceso H -auto-similar con incrementos estacionarios para algún $H \in (0, 1)$. Entonces para cada punto fijo t_0*

$$\limsup_{t \rightarrow t_0} \frac{|X_t - X_{t_0}|}{t - t_0} = \infty,$$

es decir, las trayectorias de procesos H -auto-similares son diferenciables en ninguna parte con probabilidad 1.

Demostración. Sin pérdida de generalidad se elige $t_0 = 0$. Sea (t_n) una sucesión tal que $t_n \rightarrow 0$. Entonces, por la H -auto-similitud, $X_0 = 0$ casi seguramente o con probabilidad 1, y por lo tanto

$$\begin{aligned} P \left(\limsup_{n \rightarrow \infty} \sup_{0 \leq s \leq t_n} \left| \frac{X_s}{s} \right| > x \right) &= \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq s \leq t_n} \left| \frac{X_s}{s} \right| > x \right) \\ &\geq \limsup_{n \rightarrow \infty} P \left(\left| \frac{X_{t_n}}{t_n} \right| > x \right) \\ &= \limsup_{n \rightarrow \infty} P (t_n^{H-1} |X_1| > x) \\ &= 1, \quad x > 0. \end{aligned}$$

Por lo tanto, con probabilidad 1, $\limsup_{n \rightarrow \infty} \left| \frac{X_{t_n}}{t_n} \right| = \infty$ para cualquier sucesión $t_n \rightarrow 0$. □

Observación 2.2. *El movimiento Browniano es $\frac{1}{2}$ -auto-similar, es decir*

$$(T^{\frac{1}{2}}B_{t_1}, \dots, T^{\frac{1}{2}}B_{t_n}) \stackrel{d}{=} (B_{Tt_1}, \dots, B_{Tt_n})$$

para cada $T > 0$, cualquier elección de $t_i \geq 0$, $i = 1, \dots, n$ y $n \geq 1$. Consecuentemente, sus trayectorias son diferenciables en ninguna parte.

Una indicación adicional de la irregularidad de las trayectorias Brownianas esta dada por la siguiente proposición.

Proposición 2.2. *Las trayectorias Brownianas no tienen variación acotada sobre cualquier intervalo finito $[0, T]$. Esto significa que*

$$v(B(\omega)) = \sup_{\tau} \sum_{i=1}^n |B_{t_i}(\omega) - B_{t_{i-1}}(\omega)| = \infty \quad \text{c.s.},$$

donde el supremo es tomado sobre todas las posibles particiones $\tau : 0 = t_0 < \dots < t_n = T$ de $[0, T]$, $\omega \in \Omega$ y c.s. denota que converge casi seguramente o con probabilidad 1.

Demostración. Por conveniencia se supone que $T = 1$. Adicionalmente se supone que $v(B(\omega)) < \infty$ para un $\omega \in \Omega$ dado. Sea (τ_n) una sucesión de particiones $\tau_n : 0 = t_0 < t_1 < \dots < t_{n-1} < t_n = 1$, tales que malla $(\tau_n) \rightarrow 0$ (véase ecuación (3.2)). Teniendo en cuenta que $\Delta_i B = B_{t_i} - B_{t_{i-1}}$, se tiene la siguiente cadena de desigualdades:

$$\begin{aligned} Q_n(\omega) &= \sum_{i=1}^n (\Delta_i B(\omega))^2 \\ &\leq \max_{i=1, \dots, n} |\Delta_i B(\omega)| \sum_{i=1}^n |\Delta_i B(\omega)| \\ &\leq \max_{i=1, \dots, n} |\Delta_i B(\omega)| v(B(\omega)). \end{aligned} \quad (2.1)$$

Ya que B tiene trayectorias continuas con probabilidad 1, se puede suponer que $B_t(\omega)$ es una función continua en t . Es incluso uniformemente continua en $[0, 1]$, lo cual, en combinación con que malla $(\tau_n) \rightarrow 0$, implica que

$$\max_{i=1, \dots, n} |\Delta_i B(\omega)| \rightarrow 0.$$

Por lo tanto el lado derecho de (2.1) converge a cero, implicando que

$$Q_n(\omega) \rightarrow 0. \quad (2.2)$$

Por otro lado se tiene que Q_n converge en probabilidad a 1 ($Q_n \xrightarrow{P} 1$), por lo tanto para una sucesión adecuada (n_k) se obtiene que $Q_{n_k} \xrightarrow{\text{c.s.}} 1$ (Q_{n_k} converge casi seguramente a 1). Así (2.2) solamente es posible que se cumpla sobre el conjunto vacío, y por tanto

$$P(\{\omega : v(B(\omega)) = \infty\}) = 1.$$

□

A este punto, la variación no acotada y la no diferenciabilidad de las trayectorias Brownianas son razones de peso para que los métodos de integración clásica fallen cuando los aplicamos a estas trayectorias y para la introducción del cálculo estocástico.

2.2. Procesos derivados del movimiento Browniano.

Varios de los procesos estocásticos gaussianos y no gaussianos de relevancia práctica pueden ser derivados del movimiento Browniano. Presentaremos algunos ejemplos de procesos derivados del movimiento Browniano que utilizaremos posteriormente. Como antes, $B = (B_t, t \in [0, \infty))$ denota al movimiento Browniano.

Ejemplo 2.1. (Puente Browniano.)

Considere el proceso

$$X_t = B_t - tB_1, \quad 0 \leq t \leq 1.$$

Obviamente,

$$X_0 = B_0 - 0B_1 = 0 \quad \text{y} \quad X_1 = B_1 - 1B_1 = 0.$$

Por esta simple razón, el proceso X lleva el nombre de *puente Browniano o movimiento Browniano sujeto*. Este es un proceso gaussiano, por tanto, utilizando la definición [1.2] podemos calcular las funciones de esperanza y covarianza

$$\begin{aligned} \mu_X(t) &= E(X_t) \\ &= E(B_t - tB_1) \\ &= E(B_t) - tE(B_1) \\ &= \mu_B(t) - t\mu_B(1) = 0, \quad s, t \in [0, 1]. \end{aligned}$$

$$\begin{aligned} \text{cov}_X(t, s) &= E(X_t X_s) \\ &= E[(B_t - tB_1)(B_s - sB_1)] \\ &= E(B_t B_s) - tE(B_1 B_s) - sE(B_1 B_t) + stE(B_1^2) \\ &= \text{cov}_B(t, s) - t\text{cov}_B(1, s) - s\text{cov}_B(1, t) + st\text{cov}_B(1, 1) \\ &= \text{mín}(t, s) - st, \quad s, t \in [0, 1]. \end{aligned}$$

Ya que X es gaussiano, el puente Browniano es caracterizado por estas dos funciones. El puente Browniano aparece como el proceso límite de funciones de distribución empíricas normalizadas de una muestra iid de variables aleatorias uniformes $U(0, 1)$. ■

Ejemplo 2.2. (Movimiento Browniano geométrico.)

El proceso sugerido por Black, Scholes y Merton está dado por

$$X_t = e^{\mu t + \sigma B_t}, \quad t \geq 0,$$

el cual es un proceso no gaussiano.

Con el propósito de utilizar estos resultados posteriormente, calcularemos las funciones de esperanza y covarianza del movimiento Browniano geométrico. Para calcular la función de esperanza consideremos primeramente una variable aleatoria $Z \sim N(0, 1)$ y calculemos lo siguiente:

$$\begin{aligned} E(e^{\lambda Z}) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\lambda z} e^{-\frac{z^2}{2}} dz \\ &= e^{\frac{\lambda^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(z-\lambda)^2}{2}} dz \\ &= e^{\frac{\lambda^2}{2}}, \quad \lambda \in \mathbb{R}. \end{aligned}$$

Aquí se utilizó el hecho de que $(2\pi)^{-1/2} \exp\{-(z-\lambda)^2/2\}$ es la función de densidad de probabilidad de una variable aleatoria $N(\lambda, 1)$. Del hecho de que $E(\exp\{\lambda Z\}) = \exp\{\lambda^2/2\}$ y de la auto-similitud del movimiento Browniano se sigue inmediatamente que:

$$\begin{aligned} \mu_X(t) &= e^{\mu t} E(e^{\sigma B_t}) \\ &= e^{\mu t} E(e^{\sigma t^{1/2} B_1}) \\ &= e^{(\mu+0.5\sigma^2)t}. \end{aligned}$$

Para $s \leq t$, $B_t - B_s$ y B_s son independientes y $B_t - B_s \stackrel{d}{=} B_{t-s}$. Por lo tanto,

$$\begin{aligned} \text{cov}_X(t, s) &= E(X_t X_s) - E(X_t)E(X_s) \\ &= e^{\mu(t+s)} E(e^{\sigma(B_t+B_s)}) - e^{(\mu+0.5\sigma^2)(t+s)} \\ &= e^{\mu(t+s)} E(e^{\sigma[(B_t-B_s)+2B_s]}) - e^{(\mu+0.5\sigma^2)(t+s)} \\ &= e^{\mu(t+s)} E(e^{\sigma(B_t-B_s)}) E(e^{2\sigma B_s}) - e^{(\mu+0.5\sigma^2)(t+s)} \\ &= e^{(\mu+0.5\sigma^2)(t+s)} (e^{\sigma^2 s} - 1). \end{aligned}$$

En particular, el movimiento Browniano geométrico tiene función de varianza:

$$\sigma_X^2(t) = e^{(2\mu+\sigma^2)t} (e^{\sigma^2 t} - 1).$$

■

2.3. Movimiento Browniano de una partícula libre, ecuación de difusión.

2.3.1. Deducción de la ecuación de difusión, consideración macroscópica.

La heurística y la forma más simple de deducir la ecuación que describe el comportamiento de las partículas en un medio es la siguiente. Considere un número grande de partículas las cuales realizan el movimiento Browniano a lo largo de algún eje (por simplicidad, consideramos al principio, el movimiento en una dimensión) y las cuales no interactúan entre sí. Sea $u(t, x)dx$ que denote el número de partículas en un pequeño intervalo dx alrededor de la posición x , al tiempo t (es decir, la densidad de las partículas) y $j(t, x)$ denote el flujo de partículas, es decir, el número neto de partículas Brownianas que pasan por el punto x en la dirección de incremento de los valores de x por unidad de tiempo. Es conocido como un hecho experimental que el flujo de partículas es proporcional al gradiente de su densidad:

$$j(t, x) = -a^2 \frac{\partial u}{\partial x}(t, x). \quad (2.3)$$

Esta relación incluso sirve como la definición de la constante de difusión a^2 . Si las partículas no son creadas ni destruidas, la densidad y el flujo cumplen la ecuación de continuidad

$$\frac{\partial u}{\partial t}(t, x) = -\frac{\partial j}{\partial x}(t, x),$$

la cual, debido a (2.3), puede ser escrita de la siguiente forma:

$$\frac{\partial u}{\partial t}(t, x) = a^2 \frac{\partial^2 u}{\partial x^2}(t, x). \quad (2.4)$$

La ecuación (2.4) es llamada la *ecuación de difusión*.

2.3.2. Solución de la ecuación de difusión.

Consideremos nuevamente la ecuación de difusión (2.4) con las siguientes condiciones iniciales y de frontera.

$$\begin{aligned} CI : \quad u(0, x) &= \psi(x); \\ CF : \quad u(t, 0) &= u(t, L) = 0. \end{aligned} \quad (2.5)$$

Por el método de separación de variables proponemos la siguiente solución

$$u(t, x) = T(t)\phi(x),$$

sustituyendo esta solución en la ecuación (2.4), dividiendo entre $u(t, x)$ e introduciendo la constante de separación de variables λ^2 , obtenemos las siguientes ecuaciones diferenciales

$$\begin{aligned} T'(t) &= -\lambda^2 a^2 t, \\ \phi''(x) &= -\lambda^2 \phi(x). \end{aligned}$$

cuyas soluciones son

$$\begin{aligned} T(t) &= C e^{-\lambda^2 a^2 t}, \\ \phi(x) &= A \cos(\lambda x) + B \operatorname{sen}(\lambda x). \end{aligned}$$

Por lo tanto, obtenemos que la solución de (2.4) es:

$$u(t, x) = \left(C e^{-\lambda^2 a^2 t} \right) (A \cos(\lambda x) + B \operatorname{sen}(\lambda x)).$$

Luego, aplicando el principio de superposición generalizado, obtenemos

$$u(t, x) = \int_{-\infty}^{\infty} e^{-\lambda^2 a^2 t} (A(\lambda) \cos(\lambda x) + B(\lambda) \operatorname{sen}(\lambda x)) d\lambda, \quad (2.6)$$

donde $A(\lambda) = CA$ y $B(\lambda) = CB$.

Utilizando las condiciones iniciales (2.5), tenemos que

$$u(0, x) = \int_{-\infty}^{\infty} (A(\lambda) \cos(\lambda x) + B(\lambda) \operatorname{sen}(\lambda x)) d\lambda = \psi(x). \quad (2.7)$$

Por otro lado consideremos la siguiente integral de Fourier

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} d\lambda \int_{-\infty}^{\infty} \psi(x_0) \cos(\lambda(x_0 - x)) dx_0. \quad (2.8)$$

Desarrollando (2.8) y haciendo

$$\begin{aligned} A(\lambda) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi(x_0) \cos(\lambda x_0) dx_0, \\ B(\lambda) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi(x_0) \operatorname{sen}(\lambda x_0) dx_0, \end{aligned}$$

tenemos que las ecuaciones (2.7) y (2.8) son equivalentes.

Utilizando (2.8) se puede reescribir (2.6) como sigue:

$$\begin{aligned} u(t, x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} d\lambda \int_{-\infty}^{\infty} e^{-\lambda^2 a^2 t} \psi(x_0) \cos(\lambda(x_0 - x)) dx_0 \\ &= \frac{1}{\pi} \int_{-\infty}^{\infty} \psi(x_0) dx_0 \int_0^{\infty} e^{-\lambda^2 a^2 t} \cos(\lambda(x_0 - x)) d\lambda. \end{aligned} \quad (2.9)$$

Tomando $z = a\lambda\sqrt{t}$ y $\lambda(x - x_0) = \mu z$ podemos transformar la segunda integral en (2.9) como sigue:

$$\begin{aligned} \int_0^{\infty} e^{-\lambda^2 a^2 t} \cos(\lambda(x_0 - x)) d\lambda &= \frac{1}{a\sqrt{t}} \int_0^{\infty} e^{-z^2} \cos(\mu z) dz \\ &= \frac{1}{a\sqrt{t}} J(\mu), \end{aligned} \quad (2.10)$$

donde $J(\mu) = \int_0^{\infty} \exp\{-z^2\} \cos(\mu z) dz$. Derivando $J(\mu)$ con respecto de μ y utilizando integración por partes, tenemos que se cumple la siguiente ecuación diferencial:

$$J'(\mu) = -\frac{\mu}{2} J(\mu). \quad (2.11)$$

Resolviendo la ecuación diferencial (2.11), se tiene que su solución es:

$$J(\mu) = \frac{\sqrt{\pi}}{2} e^{-\frac{\mu^2}{2}}. \quad (2.12)$$

Así, sustituyendo (2.12) en (2.10), y reescribiendo (2.9) obtenemos finalmente la solución de la ecuación de difusión

$$u(t, x) = \frac{1}{2a\sqrt{\pi t}} \int_{-\infty}^{\infty} \psi(x_0) e^{-\frac{(x_0 - x)^2}{4a^2 t}} dx_0, \quad (2.13)$$

la cual es llamada la *solución fundamental de la ecuación de difusión*.

Capítulo 3

Cálculo Estocástico de Itô.

3.1. Integración estocástica.

3.1.1. Las integrales de Riemann y Riemann-Stieltjes.

La integral ordinaria de Riemann.

Supóngase, por simplicidad, que f es una función de valores reales definida en $[0, 1]$, pero incluso podemos considerar cualquier intervalo $[a, b]$.

Consideremos una partición del intervalo $[0, 1]$:

$$\tau_n : 0 = t_0 < t_1 < \dots < t_{n-1} < t_n = 1,$$

y definamos

$$\Delta_i = t_i - t_{i-1}, \quad i = 1, \dots, n.$$

Una partición intermedia σ_n de τ_n esta dada por los valores y_i que satisfagan que $t_{i-1} \leq y_i \leq t_i$ para $i = 1, \dots, n$. Para las particiones dadas σ_n y τ_n se puede definir la suma de Riemann

$$S_n = S_n(\tau_n, \sigma_n) = \sum_{i=1}^n f(y_i) (t_i - t_{i-1}) = \sum_{i=1}^n f(y_i) \Delta_i. \quad (3.1)$$

Así, una suma de Riemann no es nada mas que una medida ponderada de los valores $f(y_i)$, donde los pesos son las correspondientes longitudes Δ_i de los intervalos $[t_{i-1}, t_i]$. Sabemos también que S_n es una aproximación del área entre la gráfica de la función f y el eje t , suponiendo que f sólo asume valores no negativos.

Ahora dada una malla de la partición τ_n que tiende a cero, es decir,

$$\text{malla}(\tau_n) := \max_{i=1, \dots, n} \Delta_i = \max_{i=1, \dots, n} (t_i - t_{i-1}) \rightarrow 0. \quad (3.2)$$

Si procedemos de esta manera, los puntos $t_i = t_i^{(n)}$ claramente tienen que depender en n , pero eliminamos esta dependencia en nuestra notación.

Definición 3.1. Si el límite

$$S = \lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(y_i) \Delta_i,$$

existe a medida que $\text{malla}(\tau_n) \rightarrow 0$ y S sea independiente de la elección de las particiones τ_n y sus particiones intermedias σ_n , entonces S es llamada la **integral ordinaria o de Riemann de f sobre $[0, 1]$** .

Escribimos

$$S = \int_0^1 f(t) dt.$$

La integral de Riemann es tomada como un modelo para la definición de cualquier tipo de integral. El nuevo tipo de integral debe tener tantas propiedades en común con la integral de Riemann como sea posible. Tales propiedades son dadas a continuación.

Para funciones integrables de Riemann f , f_1 y f_2 sobre $[0, 1]$ las siguientes propiedades se cumplen

- La integral de Riemann es lineal, es decir, para cualquiera constantes c_1 y c_2 :

$$\int_0^1 [c_1 f_1(t) + c_2 f_2(t)] dt = c_1 \int_0^1 f_1(t) dt + c_2 \int_0^1 f_2(t) dt.$$

- La integral de Riemann es lineal en intervalos adyacentes

$$\int_0^1 f(t) dt = \int_0^a f(t) dt + \int_a^1 f(t) dt \quad \text{para } 0 \leq a \leq 1.$$

- Se puede definir la integral indefinida de Riemann como función del límite superior

$$\int_0^s f(t) dt = \int_0^1 f(t) I_{[0,s]}(t) dt, \quad 0 \leq s \leq 1.$$

La integral Riemann-Stieltjes.

En la teoría de la probabilidad es usual el denotar la esperanza de una variable aleatoria X por

$$E(X) = \int_{-\infty}^{\infty} t dF_X(t),$$

donde F_X denota la función de distribución de probabilidad acumulada de X . Esta notación se refiere al hecho que $E(X)$ esta definida como una integral de

Riemann-Stieltjes o como una integral de Lebesgue-Stieltjes. Esto significa, en términos generales, que

$$\int_{-\infty}^{\infty} t dF_X(t) \approx \sum_i y_i [F_X(t_i) - F_X(t_{i-1})],$$

para una partición (t_i) de \mathbb{R} y una correspondiente partición intermedia (y_i) . Recordando también que se puede definir la integral $\int_0^1 f(t) dg(t)$ como $\int_0^1 f(t) g'(t) dt$, dado que la derivada $g'(t)$ exista.

Ambas, $\int_{-\infty}^{\infty} t dF_X(t)$ y $\int_0^1 f(t) dg(t)$, son ejemplos de como se puede aproximar el problema de integración de una función f con respecto a otra función g . El objetivo de este capítulo es el de sugerir algunos métodos, en particular, se quiere obtener integrales del tipo $\int_0^1 f(t) dB(\omega)$, donde f sea una función o un proceso estocástico en $[0, 1]$ y $B_t(\omega)$ sea una trayectoria Browniana. Existe una dificultad mayor en definir tal integral ya que las trayectorias $B_t(\omega)$ no tienen derivada, sin embargo, se puede encontrar algún tipo de integral de trayectoria la cual algunas veces permita evaluar la integral $\int_0^1 f(t) dB(\omega)$. Este tipo de integral es llamada la integral de Riemann-Stieltjes.

En lo que sigue se dará una definición precisa de la integral de Riemann-Stieltjes. Como para integral de Riemann, se considera una partición del intervalo $[0, 1]$:

$$\tau_n = t_0 < t_1 < \dots < t_{n-1} < t_n = 1,$$

y una partición intermedia σ_n de τ_n :

$$\sigma_n : t_{i-1} \leq y_i \leq t_i \text{ para } i = 1, \dots, n.$$

Sean f y g dos funciones reales en $[0, 1]$ y definimos

$$\Delta_i g = g(t_i) - g(t_{i-1}), \quad i = 1, \dots, n.$$

La *suma de Riemann-Stieltjes* correspondiente a τ_n y σ_n esta dada por

$$S_n = S_n(\tau_n, \sigma_n) = \sum_{i=1}^n f(y_i) \Delta_i g = \sum_{i=1}^n f(y_i) [g(t_i) - g(t_{i-1})].$$

Esta tiene alguna similitud con la suma de Riemann (3.1), donde en ese caso, $g(t) = t$. Así, una suma de Riemann-Stieltjes es obtenida ponderando los valores $f(y_i)$ con los incrementos $\Delta_i g$ de g en los intervalos $[t_{i-1}, t_i]$.

Definición 3.2. Si el límite

$$S = \lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(y_i) \Delta_i g,$$

existe a medida que $\text{malla}(\tau_n) \rightarrow 0$ y S sea independiente de la elección de las particiones τ_n y sus particiones intermedias σ_n , entonces S es llamada la **integral de Riemann-Stieltjes de f con respecto a g sobre $[0, 1]$** .

Escribimos

$$S = \int_0^1 f(t) dg(t).$$

Sin entrar en detalles, daremos una condición suficiente para la integrabilidad de Riemann-Stieltjes la cual es cercana a la condición de necesidad. Primero daremos una definición.

Definición 3.3. La función real h sobre $[0, 1]$ se dice que tiene **p -variación acotada** para algún $p > 0$ si

$$\sup_{\tau} \sum_{i=1}^n |h(t_i) - h(t_{i-1})|^p < \infty,$$

donde el supremo es tomado sobre todas las particiones τ de $[0, 1]$.

Nota 3.1. h tiene variación acotada si $p = 1$.

Observación 3.1. La integral de Riemann-Stieltjes $\int_0^1 f(t) dg(t)$ existe si las siguientes condiciones son satisfechas:

- i) Las funciones f y g no tienen discontinuidades en el mismo punto $t \in [0, 1]$.
- ii) La función f tiene p -variación acotada y la función g tiene q -variación acotada para algunos $p > 0$ y $q > 0$ tales que $p^{-1} + q^{-1} > 1$.

3.2. La integral de Itô.

3.2.1. Un caso especial.

Considérese la suma de Riemann-Stieltjes

$$S_n = \sum_{i=1}^n B_{t_{i-1}} \Delta_i B,$$

donde

$$\tau_n : 0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = t,$$

es una partición de $[0, t]$ y, para cualquier función f en $[0, t]$

$$\Delta f : \Delta_i f = f(t_i) - f(t_{i-1}), \quad i = 1, \dots, n,$$

son los incrementos correspondientes de f y

$$\Delta_i = t_i - t_{i-1}, \quad i = 1, \dots, n.$$

Así, la suma de Riemann-Stieltjes S_n correspondiente a la partición τ_n y la partición intermedia (y_i) con $y_i = t_{i-1}$, es decir, y_i es el punto final izquierdo del intervalo $[t_{i-1}, t_i]$. Esta elección es típica para la definición de la integral estocástica de Itô.

Nótese que S_n puede ser escrito en la siguiente forma:

$$S_n = \frac{1}{2}B_t^2 - \frac{1}{2}\sum_{i=1}^n (\Delta_i B)^2 =: \frac{1}{2}B_t^2 - \frac{1}{2}Q_n(t).$$

Dado que el movimiento Browniano tiene incrementos independientes y estacionarios (véase definición [2.1]), entonces tenemos que

$$E(\Delta_i B \Delta_j B) = \begin{cases} 0 & \text{si } i \neq j, \\ \text{var}(\Delta_i B) = t_i - t_{i-1} = \Delta_i & \text{si } i = j. \end{cases}$$

Luego se tiene que

$$E(Q_n(t)) = \sum_{i=1}^n E(\Delta_i B)^2 = \sum_{i=1}^n \Delta_i = t,$$

y nuevamente por los incrementos independientes del movimiento Browniano y ya que se cumple que $\text{var}(X) = E(X^2) - E(X)^2$, entonces

$$\text{var}(Q_n(t)) = \sum_{i=1}^n \text{var}([\Delta_i B]^2) = \sum_{i=1}^n [E(\Delta_i B)^4 - \Delta_i^2].$$

Por otro lado, de la definición [2.1], tenemos que la variable aleatoria $B_1 \sim N(0, 1)$, entonces $E(B_1^4) = 3$. Por lo tanto

$$E(\Delta_i B)^4 = E(B_{t_i - t_{i-1}}^4) = E\left[(\Delta_i)^{\frac{1}{2}} B_1\right]^4 = 3\Delta_i^2,$$

lo cual implica

$$\text{var}(Q_n(t)) = 2 \sum_{i=1}^n \Delta_i^2.$$

Así, si

$$\text{malla}(\tau_n) = \max_{i=1, \dots, n} \Delta_i \rightarrow 0,$$

obtenemos que

$$\text{var}(Q_n(t)) \leq 2 \text{malla}(\tau_n) \sum_{i=1}^n \Delta_i = 2t \text{malla}(\tau_n) \rightarrow 0.$$

Dado que $\text{var}(Q_n(t)) = E[(Q_n(t) - t)^2]$ se ha mostrado que:

$Q_n(t)$ converge a t en la media cuadrada, por lo tanto, en probabilidad. La función límite $f(t) = t$ es característica solamente para el movimiento Browniano. Esta es llamada la variación cuadrática del movimiento Browniano en $[0, t]$.

Ya que $S_n = 0.5 [B_t^2 - Q_n(t)]$ converge en la media cuadrada a $0.5 (B_t^2 - t)$, se puede tomar este límite como el valor de la integral, $\int_0^t B_s dB_s$, es decir

$$\int_0^t B_s dB_s = 0.5 (B_t^2 - t).$$

Nota 3.2. La fórmula $\int_0^t B_s dB_s = 0.5 (B_t^2 - t)$ sugiere que la regla clásica de la cadena de integración no se cumple para la integración estocástica de Itô.

Por otra parte, el incremento $\Delta_i B = B_{t_i} - B_{t_{i-1}}$ en el intervalo $[t_{i-1}, t_i]$ satisface que $E(\Delta_i B) = 0$ y $E[(\Delta_i B)^2] = \Delta_i = t_i - t_{i-1}$. El límite cuadrático medio de $Q_n(t)$ es t . Estas propiedades sugieren que $(\Delta_i B)^2$ es del orden Δ_i .

Definición 3.4. En términos de diferenciales, escribimos

$$(dB_t)^2 = (B_{t+dt} - B_t)^2 = dt, \quad (3.3)$$

y en términos de integrales, se tiene

$$\int_0^t (dB_s)^2 = \int_0^t ds = t.$$

El lado derecho es la variación cuadrática del movimiento Browniano en $[0, t]$.

3.2.2. Integración estocástica de Itô para procesos simples.

Comenzaremos el estudio de la integral estocástica de Itô para una clase de procesos cuyas trayectorias sólo toman un número finito de valores. Como es usual, $B = (B_t \geq 0)$ denota el movimiento Browniano, y

$$\mathcal{F}_t = \sigma(B_s, s \leq t), \quad t \geq 0,$$

es la correspondiente filtración natural. Esto es, de la definición [A.4] del apéndice A, se tiene que un proceso estocástico $X = (X_t, t \geq 0)$ es adaptado a un movimiento Browniano si X es adaptado a $(\mathcal{F}_t, t \geq 0)$. Esto significa que, para cada t , X_t es una función del pasado y presente del movimiento Browniano.

Primeramente introduciremos una clase apropiada de procesos integrables de Itô.

Definición 3.5. *El proceso estocástico $C = (C_t, t \in [0, T])$ se dice que es **simple** si satisface las siguientes propiedades:*

i) *Existe una partición*

$$\tau_n : 0 = t_0 < t_1 < \dots < t_{n-1} < t_n = T,$$

ii) *y una sucesión $(Z_i, i = 0, \dots, n)$ de variables aleatorias tal que:*

$$C_t = \begin{cases} Z_n & \text{si } t = T, \\ Z_i & \text{si } t_{i-1} \leq t < t_i, \quad i = 1, \dots, n. \end{cases}$$

La sucesión (Z_i) es adaptada a $(\mathcal{F}_{t_{i-1}}, i = 1, \dots, n)$, es decir, Z_i es una función del movimiento Browniano a partir del tiempo t_{i-1} , y satisface que $E(Z_i^2) < \infty$ para toda i .

Definición 3.6. *La **integral estocástica de Itô para un proceso simple** C en $[0, T]$, esta dada por*

$$\int_0^T C_s dB_s := \sum_{i=1}^n C_{t_{i-1}} (B_{t_i} - B_{t_{i-1}}) = \sum_{i=1}^n Z_i \Delta_i B.$$

*Adicionalmente, la **integral estocástica de Itô para un proceso simple** C en $[0, t]$, $t_{k-1} \leq t \leq t_k$, $k = 1, \dots, n$, esta dada por*

$$\int_0^t C_s dB_s := \int_0^T C_s I_{[0,t]}(s) dB_s = \sum_{i=1}^{k-1} Z_i \Delta_i B + Z_k (B_t - B_{t_{k-1}}), \quad (3.4)$$

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

Así, el valor de la integral estocástica de Itô, $\int_0^t C_s dB_s$, es la suma Riemann-Stieltjes de la trayectoria de C , evaluada en los puntos finales izquierdos de los intervalos $[t_{i-1}, t_i]$, con respecto al movimiento Browniano. Si $t < t_n$, el punto t puede ser formalmente interpretado como el último punto de la partición de $[0, t]$.

Observación 3.2. *El proceso estocástico $I_t(C) = \int_0^t C_s dB_s$, $t \in [0, T]$ es una martingala con respecto a la filtración natural Browniana $(\mathcal{F}_t, t \in [0, T])$ (véase definición [A.6]), es decir, se verifican las siguientes propiedades.*

- i) $E|I_t(C)| < \infty$ para todo $t \in [0, T]$.
- ii) $I(C)$ es adaptado a (\mathcal{F}_t) .
- iii) $E(I_t(C) | \mathcal{F}_s) = I_s(C)$ para $s < t$.

Por otro lado se puede ver que $E(I_t(C)) = 0$, ya que Z_i y $\Delta_i B$ en la definición [3.6] son independientes, y $E(Z_i \Delta_i B) = E(Z_i)E(\Delta_i B) = 0$. Así, se obtiene lo siguiente:

Observación 3.3. *La integral estocástica de Itô tiene esperanza cero.*

Alternativamente, se puede discutir la observación [3.3] como sigue: ya que $I(C)$ es una martingala, esta tiene una función de esperanza constante. Más aún, por la definición, $I_0(C) = 0$ y así $E(I_t(C)) = E(I_0(C)) = 0$.

Otra propiedad de la integral estocástica de Itô para procesos simples resultará crucial para la definición de la integral estocástica de Itô general.

Teorema 3.1. *La integral estocástica de Itô satisface la propiedad de isometría:*

$$E \left(\int_0^t C_s dB_s \right)^2 = \int_0^t E(C_s^2) ds, \quad t \in [0, T]. \quad (3.5)$$

Demostración. Para la comodidad de la demostración, suponemos que $t = t_k$ para algún k . En efecto, si $t_{k-1} < t < t_k$ obsérvese que $0 = t_0 < \dots < t_{k-1} < t'_k := t$ es una partición de $[0, t]$, así que se puede considerar formalmente a t como un punto de la partición. Tenemos, para $W_i = Z_i \Delta_i B$,

$$E[I_t(C)]^2 = \sum_{i=1}^k \sum_{j=1}^k E(W_i W_j). \quad (3.6)$$

Se cumple que, para $i > j$ las variables aleatorias W_i y W_j son no correlacionadas, además hay que tener en cuenta que W_j y Z_i son funciones del movimiento

Browniano al tiempo t_{i-1} , por lo tanto son independientes de $\Delta_i B$; de esta manera se concluye que

$$E(W_i W_j) = E(W_j Z_i) E(\Delta_i B) = 0.$$

Así, los términos $E(W_i W_j)$ en (3.6) desaparecen para $i \neq j$, y por tanto se obtiene,

$$E[I_t(C)]^2 = \sum_{i=1}^k E[(Z_i \Delta_i B)^2] = \sum_{i=1}^k E(Z_i^2) E[(\Delta_i B)^2] = \sum_{i=1}^k E(Z_i^2) (t_i - t_{i-1}).$$

El lado derecho no es más que la integral de Riemann, $\int_0^t f(s) ds$, de la función escalón $f(s) = E(C_s^2)$, la cual coincide con $E(Z_i^2)$ para $t_{i-1} \leq t < t_i$. Así hemos probado la propiedad de isometría de la integral estocástica de Itô. \square

La integral estocástica de Itô tiene varias propiedades en común con las integrales de Riemann y de Riemann-Stieltjes.

Propiedad 3.1. *La integral estocástica de Itô,*

1. *es lineal:*

Para constantes c_1, c_2 y procesos simples $C^{(1)}$ y $C^{(2)}$ en $[0, T]$,

$$\int_0^t [c_1 C_s^{(1)} + c_2 C_s^{(2)}] dB_s = c_1 \int_0^t C_s^{(1)} dB_s + c_2 \int_0^t C_s^{(2)} dB_s.$$

2. *y es lineal en intervalos adyacentes:*

Para $0 \leq t \leq T$,

$$\int_0^T C_S dB_s = \int_0^t C_S dB_s + \int_t^T C_S dB_s.$$

Finalmente, establecemos la siguiente propiedad:

Propiedad 3.2. *El proceso $I(C)$ tiene trayectorias continuas.*

Esto se sigue de la definición de $I(C)$: la relación

$$I(C) = I_{t_{i-1}}(C) + Z_i(B_t - B_{t_{i-1}}), \quad t_{i-1} \leq t \leq t_i,$$

se sigue cumpliendo y las trayectorias del movimiento Browniano son continuas.

3.2.3. La integral estocástica general de Itô.

En la sección 3.2.2 se ha introducido la integral estocástica para procesos simples C , es decir, para procesos estocásticos cuyas trayectorias son funciones escalón. La integral estocástica de Itô, $\int_0^t C_s dB_s$, es entonces simplemente la suma de Riemann-Stieltjes de C , evaluada en los puntos finales de la izquierda del intervalo $[t_{i-1}, t_i]$, con respecto al movimiento Browniano. En la sección 3.1.1 se mostró que, en general no se puede definir la integral estocástica de Itô como un límite de sumas de Riemann-Stieltjes. En la sección 3.2.1 se sugirió que se puede definir la integral estocástica de Itô como el límite cuadrado medio de sumas de Riemann-Stieltjes adecuadas. La idea trabaja bajo condiciones bastante generales, y es el objetivo de esta sección el defender este enfoque. En cierto punto necesitaremos algunas herramientas de la teoría de espacios de Hilbert. También dependeremos de algunos argumentos heurísticos. En lo que sigue, los procesos $C = (C_t, t \in [0, T])$ sirven como los integrandos de la integral estocástica de Itô. Suponemos que las siguientes condiciones se satisfacen.

Nota 3.3. Suposiciones de los procesos integrantes C :

- C es adaptado al movimiento Browniano en $[0, T]$, es decir, C_t es una función de B_s , $s \leq t$.
- La integral $\int_0^T E(C_s^2) ds$ es finita.

Las suposiciones en la nota [3.3] son trivialmente satisfechas por los procesos simples. Otra clase de integrandos admisibles consiste de las funciones determinísticas $c(t)$ en $[0, t]$ con $\int_0^T c^2(t) dt < \infty$, esto incluye las funciones continuas sobre $[0, T]$.

En lo que sigue se dará una aproximación heurística a la integral estocástica de Itô. Primero recordemos como procedimos para la definición de la integral, $\int_0^t B_s dB_s$, en la sección 3.2.1:

- Para un t fijo y una partición dada (t_i) de $[0, t]$ se introdujo la suma de Riemann-Stieltjes $\sum_{i=1}^n B_{t_{i-1}}(B_{t_i} - B_{t_{i-1}})$.
- Nótese que: en la sección 3.2.2 se definió exactamente estas sumas de Riemann-Stieltjes como las integrales $I(C^{(n)}) = \int_0^t C_s^{(n)} dB_s$ de las funciones simples $C^{(n)}$ la cual asume el valor $B_{t_{i-1}}$ sobre el intervalo $[t_{i-1}, t_i]$.
- Entonces se consideró el límite cuadrático medio de las sumas de Riemann-Stieltjes $I(C^{(n)})$ y se obtuvo el valor $0.5(B_t^2 - t)$.

Esta aproximación puede ser hecha para que funcione con procesos C que satisfagan las suposiciones de la nota [3.3]. Comenzaremos con una observación interesante:

Observación 3.4. Sea C un proceso que satisfaga las suposiciones en la nota [3.3]. Entonces se puede encontrar una sucesión $(C^{(n)})$ de procesos simples tales que

$$\int_0^T E[C_s - C_s^{(n)}]^2 ds \rightarrow 0.$$

Así los procesos simples $C^{(n)}$ convergen en sentido cuadrático medio al proceso integrante C . Ya que $C^{(n)}$ es simple podemos evaluar las integrales estocásticas de Itô, $I_t(C^{(n)}) = \int_0^t C_s^{(n)} dB_s$, para cada n y t .

El siguiente paso es mostrar que la sucesión $(I(C^{(n)}))$ de integrales estocásticas de Itô convergen en sentido cuadrático medio a un proceso límite único. En efecto, se puede mostrar la existencia de un proceso $I(C)$ sobre $[0, T]$, tal que

$$E \sup_{0 \leq t \leq T} [I_t(C) - I_t(C^{(n)})]^2 \rightarrow 0.$$

Definición 3.7. El límite cuadrático medio $I(C)$ es llamado la **integral estocástica de Itô de C** . Es denotado por

$$I_t(C) = \int_0^t C_s dB_s, \quad t \in [0, T].$$

Para un proceso simple C , la integral estocástica de Itô tiene la representación de suma de Riemann-Stieltjes (3.4).

Después de esta definición general de la integral estocástica de Itô se ha perdido la intuición porque ya no somos capaces de escribir la integral, $\int_0^t C_s dB_s$, en términos simples del movimiento Browniano. En casos particulares se podrá obtener una fórmula explícita para las integrales estocásticas de Itô, pero esto requiere el conocimiento del lema de Itô. Para nuestra intuición general, y para propósitos prácticos, la siguiente regla es útil:

Nota 3.4. Las integrales estocásticas de Itô $I_t(C) = \int_0^t C_s dB_s$, $t \in [0, T]$, constituyen un proceso estocástico. Para una partición dada

$$\tau_n : 0 = t_0 < t_1 < \dots < t_{n-1} < t_n = T,$$

y $t \in [t_{k-1}, t_k]$, la variable aleatoria $I_t(C)$ es cercana a la suma de Riemann-Stieltjes

$$\sum_{i=1}^{k-1} C_{t_{i-1}}(B_{t_i} - B_{t_{i-1}}) + C_{t_{k-1}}(B_t - B_{t_{k-1}}),$$

y esta aproximación es cercana (en el sentido cuadrático medio) al valor de $I_t(C)$, para la partición más densa τ_n en $[0, T]$.

Ahora veamos algunas propiedades de la integral estocástica general de Itô. La integral estocástica general de Itô hereda las propiedades de la integral estocástica de Itô para procesos simples. En general, no desarrollaremos las pruebas de estas propiedades.

Propiedad 3.3. *El proceso estocástico $I_t(C) = \int_0^t C_s dB_s$, $t \in [0, T]$, es una martingala con respecto a la filtración natural Browniana $(\mathcal{F}_t, t \in [0, T])$ (véase definición [A.6]).*

Esto resulta de la elección particular de la aproximación de la sumas de Riemann-Stieltjes de C , las cuales son evaluadas en los puntos finales izquierdos de los intervalos $[t_{i-1}, t_i]$.

Propiedad 3.4. *La integral estocástica de Itô tiene esperanza cero.*

Para la prueba de la existencia de la integral estocástica de Itô, la propiedad de isometría (3.5) para procesos simples es esencial. Esto es válido también para el caso general.

Propiedad 3.5. *La integral estocástica de Itô satisface la propiedad de isometría:*

$$E \left(\int_0^t C_s dB_s \right)^2 = \int_0^t E(C_s^2) ds, \quad t \in [0, T].$$

La integral estocástica de Itô incluso tiene algunas propiedades en común con las integrales de Riemann y de Riemann-Stieltjes.

Propiedad 3.6. *La integral estocástica de Itô,*

1. *es lineal:*

Para constantes c_1, c_2 y procesos $C^{(1)}$ y $C^{(2)}$ en $[0, T]$, que satisfacen las suposiciones en la nota [3.3],

$$\int_0^t [c_1 C_s^{(1)} + c_2 C_s^{(2)}] dB_s = c_1 \int_0^t C_s^{(1)} dB_s + c_2 \int_0^t C_s^{(2)} dB_s.$$

2. *y es lineal en intervalos adyacentes:*

Para $0 \leq t \leq T$,

$$\int_0^T C_s dB_s = \int_0^t C_s dB_s + \int_t^T C_s dB_s.$$

Finalmente, establecemos la siguiente propiedad:

Propiedad 3.7. *Los procesos $I(C)$ tienen trayectorias continuas.*

3.3. El lema de Itô.

3.3.1. La regla de la cadena clásica de diferenciación.

El lema de Itô es el análogo estocástico de la regla de la cadena clásica de diferenciación. Recordando que, para una función diferenciable $b(s)$,

$$\frac{1}{2} \frac{db^2(s)}{ds} = b(s) \frac{db(s)}{ds}.$$

Esta relación implica que, con $b(0) = 0$,

$$\frac{1}{2} \int_0^t \frac{db^2(s)}{ds} ds = \frac{1}{2} b^2(t) = \int_0^t b(s) \frac{db(s)}{ds} ds = \int_0^t b(s) db(s). \quad (3.7)$$

No se puede simplemente reemplazar la función $b(s)$ con una trayectoria $B_s(\omega)$ del movimiento Browniano en (3.7), ya que tal trayectoria es diferenciable en ninguna parte. Usando algunos recursos elementales, descubrimos en la sección 3.2.1 que la integral estocástica de Itô, $\int_0^t B_s dB_s$, tienen valor, $0.5(B_t^2 - t)$, lo que contrasta con (3.7). Nótese que el valor de la integral, $0.5(B_t^2 - t)$, es el valor $0.5(B_t^2)$, el cual se espera obtener de la regla de la cadena clásica de diferenciación, corregido por $-0.5t$. Esto sugiere que se debe encontrar un término de corrección a la regla de la cadena clásica. Otra apelación a la discusión en la sección 3.2.1 nos dice que estos términos vienen del límite cuadrático medio del funcional cuadrático, $\sum_{i=1}^n (B_{t_i} - B_{t_{i-1}})^2$.

Con el objetivo de entender la regla de la cadena estocástica, primero se recordara la regla de la cadena determinística. Por simplicidad, escribimos $h'(t)$, $h''(t)$, etc., para las derivadas ordinarias de la función h de t .

Definición 3.8. Sean f y g funciones diferenciables. Entonces la regla clásica de la cadena de diferenciación nos dice:

$$[f(g(s))]' = f'(g(s))g'(s). \quad (3.8)$$

Dado que estamos interesados en integrales, reescribimos (3.8) en forma integral, es decir, integrando a ambos lados de (3.8) en $[0, t]$.

Definición 3.9. La regla de la cadena en forma integral esta dada por:

$$f(g(t)) - f(g(0)) = \int_0^t f'(g(s))g'(s)ds = \int_0^t f'(g(s))dg(s).$$

Damos un argumento heurístico para la validez de la regla de la cadena, el cual tomaremos como motivación para el lema de Itô. En el lenguaje de diferenciales, (3.8) se convierte en $df(g) = f'dg$. Esta diferencial puede ser interpretado como el término de primer orden en una expansión de Taylor:

$$f(g(t) + dg(t)) - f(g(t)) = f'(g(t))dg(t) + \frac{1}{2}f''(g(t))[dg(t)]^2 + \dots \quad (3.9)$$

Aquí $dg(t) = g(t + dt) - g(t)$ es el incremento de g sobre $[t, t + dt]$. Bajo condiciones adecuadas, los términos de segundo y orden mayor son despreciables para dt pequeños.

3.3.2. Una versión simple del lema de Itô.

Ahora supóngase que f es una función dos veces diferenciable, pero reemplazamos $g(t)$ en (3.9) con una trayectoria $B_t(\omega)$ del movimiento Browniano. Escribimos $dB_t = B_{t+dt} - B_t$ para el incremento de B en $[t, t + dt]$. Usando los mismos argumentos anteriores, obtenemos que

$$f(B_t + dB_t) - f(B_t) = f'(B_t)dB_t + \frac{1}{2}f''(B_t)(dB_t)^2 + \dots \quad (3.10)$$

En la ecuación (3.3) establecimos que la diferencial cuadrada $(dB_t)^2$ puede ser interpretada como dt . Así:

En contraste al caso determinístico, la contribución del término de segundo orden en la expansión de Taylor (3.10) no es despreciable.

Este hecho es la razón para la desviación de la regla clásica de la cadena.

Integrando ambos lados de (3.10) en un sentido formal y despreciando términos de orden mayor que 3 del lado derecho, obtenemos para $s < t$:

$$\int_s^t df(B_x) = f(B_t) - f(B_s) \quad (3.11)$$

$$= \int_s^t f'(B_x)dB_x + \frac{1}{2} \int_s^t f''(B_x)dx. \quad (3.12)$$

Donde las integrales en (3.11) y (3.12) se pueden interpretar como sigue. La primera integral en (3.12) es la integral estocástica de Itô de $f'(B)$ y la segunda integral ha de ser interpretada como la integral de Riemann de $f''(B)$. La relación (3.11) define las cantidades, $\int_s^t df(B_s)$, la cual es análoga a una suma telescópica, y por lo tanto es natural el asignar el valor $f(B_t) - f(B_s)$ a esta. En lo que sigue, daremos siempre el valor $V_t - V_s$ a la integral simbólica $\int_s^t dV_x$, cualquiera que sea el proceso estocástico V .

Definición 3.10. Sea f una función dos veces continuamente diferenciable.

La fórmula

$$f(B_t) - f(B_s) = \int_s^t f'(B_x)dB_x + \frac{1}{2} \int_s^t f''(B_x)dx, \quad s < t,$$

es una forma simple del **lema de Itô** o de la **fórmula de Itô**.

3.3.3. Versiones extendidas del lema de Itô.

En esta sección extenderemos el lema de Itô, dado en (3.11)-(3.12), en varias formas.

Comenzaremos con un lema de Itô para el proceso estocástico $f(t, B_t)$. Suponga que $f(t, x)$ tiene derivadas parciales de al menos segundo orden. Una modificación del argumento de la expansión de Taylor usado en la sección 3.3.2 es incluso exitoso en este caso más general. Recordando que la expansión de Taylor de segundo orden produce

$$\begin{aligned} f(t + dt, B_{t+dt}) - f(t, B_t) = & \quad (3.13) \\ & f_1(t, B_t)dt + f_2(t, B_t)dB_t \\ & + \frac{1}{2} [f_{11}(t, B_t)(dt)^2 + 2f_{12}(t, B_t)dtdB_t + f_{22}(t, B_t)(dB_t)^2] + \dots \end{aligned}$$

Nota 3.5. De aquí y en lo que sigue, se usará la siguiente notación para las derivadas parciales de f :

$$\begin{aligned} f_i(t, x) &= \left. \frac{\partial}{\partial x_i} f(x_1)f(x_2) \right|_{x_1=t, x_2=x}, \quad i = 1, 2, \\ f_{ij}(t, x) &= \left. \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} f(x_1)f(x_2) \right|_{x_1=t, x_2=x}, \quad i, j = 1, 2. \end{aligned}$$

Como en el cálculo clásico, los términos de orden mayor en (3.13) son despreciables, y así también los términos con factores $dtdB_t$ y $(dt)^2$. Sin embargo, ya que interpretamos $(dB_t)^2$ como dt , el término con $(dB_t)^2$ no puede ser despreciado. Procediendo en la misma forma que en la sección 3.3.2, es decir, integrando formalmente el lado derecho e izquierdo en (3.13) y agrupando todos los términos con dt y dB_t separadamente, concluimos con la ecuación (3.14) dada en la siguiente definición.

Definición 3.11. Extensión I del lema de Itô.

Sea $f(t, x)$ una función cuyas derivadas parciales de segundo orden son continuas. Entonces

$$f(t, B_t) - f(s, B_s) = \int_s^t \left[f_1(x, B_x) + \frac{1}{2} f_{22}(x, B_x) \right] dx + \int_s^t f_2(x, B_x) dB_x, \quad s < t. \quad (3.14)$$

Ejemplo 3.1. (Movimiento Browniano geométrico.)

Consideremos una forma particular del movimiento Browniano geométrico

$$X_t = f(t, B_t) = e^{(c-0.5\sigma^2)t + \sigma B_t}, \quad (3.15)$$

donde c y $\sigma > 0$ son constantes. Nótese que

$$f(t, x) = e^{(c-0.5\sigma^2)t + \sigma x}, \quad f_1(t, x) = (c - 0.5\sigma^2)f(t, x),$$

$$f_2(t, x) = \sigma f(t, x), \quad f_{22}(t, x) = \sigma^2 f(t, x).$$

Una aplicación del lema de Itô (3.14) hace que el proceso X satisfaga la ecuación diferencial estocástica lineal

$$X_t - X_0 = c \int_0^t X_s ds + \sigma \int_0^t X_s dB_s. \quad (3.16)$$

■

Para un uso posterior se necesitará una versión más general del lema de Itô. Se consideran procesos de la forma $f(t, X_t)$, donde X está dado por:

$$X_t = X_0 + \int_0^t A_s^{(1)} ds + \int_0^t A_s^{(2)} dB_s \quad (3.17)$$

y ambos, $A^{(1)}$ y $A^{(2)}$, son adaptados al movimiento Browniano. Aquí se supone que las integrales en (3.17) están bien definidas en el sentido de Riemann e Itô, respectivamente.

Nota 3.6. Un proceso X , el cual tiene representación (3.17), es llamado un **proceso de Itô**. Se puede mostrar que los procesos $A^{(1)}$ y $A^{(2)}$ son únicamente determinados en el sentido que, si X tiene representación (3.17), donde los $A^{(i)}$'s son reemplazados con procesos adaptados $D^{(i)}$, entonces $A^{(i)}$ y $D^{(i)}$ necesariamente coinciden.

Recordemos que el movimiento Browniano geométrico (3.15) satisface (3.16). Por lo tanto, es un proceso de Itô con $A^{(1)} = cX$ y $A^{(2)} = \sigma X$.

Ahora, usando un argumento similar con una expansión de Taylor como se hizo anteriormente, se obtiene la ecuación (3.18) dada en la siguiente definición.

Definición 3.12. Extensión II del lema de Itô.

Sea X un proceso de Itô con representación (3.17) y $f(t, x)$ una función cuyas derivadas parciales de segundo orden son continuas. Entonces

$$f(t, X_t) - f(s, X_s) = \int_s^t \left[f_1(y, X_y) + A_y^{(1)} f_2(y, X_y) + \frac{1}{2} [A_y^{(2)}]^2 f_{22}(y, X_y) \right] dy + \int_s^t A_y^{(2)} f_2(y, X_y) dB_y, \quad s < t. \quad (3.18)$$

La fórmula (3.18) es frecuentemente dada en la siguiente forma:

$$f(t, X_t) - f(s, X_s) = \int_s^t \left[f_1(y, X_y) + \frac{1}{2} [A_y^{(2)}]^2 f_{22}(y, X_y) \right] dy + \int_s^t f_2(y, X_y) dX_y, \quad (3.19)$$

donde

$$dX_y = A_y^{(1)} dy + A_y^{(2)} dB_y.$$

La última identidad es una forma simbólica de escribir la representación de Itô (3.17). La integral con respecto a X en (3.19) ha de ser interpretada como sigue:

$$\int_s^t f_2(y, X_y) dX_y = \int_s^t A_y^{(1)} f_2(y, X_y) dy + \int_s^t [A_y^{(2)}]^2 f_2(y, X_y) dB_y, \quad (3.20)$$

Finalmente, consideramos el lema de Itô para un proceso estocástico de la forma $f(t, X_t^{(1)}, X_t^{(2)})$, donde ambos, $X_t^{(1)}$ y $X_t^{(2)}$, son procesos de Itô con respecto al mismo movimiento Browniano:

$$X_t^{(i)} = X_0^{(i)} + \int_0^t A_s^{(1,i)} ds + \int_0^t A_s^{(2,i)} dB_s, \quad i = 1, 2. \quad (3.21)$$

Un argumento similar de la expansión en serie de Taylor al utilizado con anterioridad produce la ecuación (3.22) dada en la siguiente definición.

Definición 3.13. Extensión III del lema de Itô.

Sean $X^{(1)}$ y $X^{(2)}$ dos procesos de Itô dados por (3.21) y $f(t, x_1, x_2)$ una función cuyas derivadas parciales de segundo orden son continuas. Entonces para $s < t$,

$$\begin{aligned}
 f(t, X_t^{(1)}, X_t^{(2)}) - f(s, X_s^{(1)}, X_s^{(2)}) = & \quad (3.22) \\
 & \int_s^t f_1(y, X_y^{(1)}, X_y^{(2)}) dy \\
 & + \sum_{i=2}^3 \int_s^t f_i(y, X_y^{(1)}, X_y^{(2)}) dX_y^{(i)} \\
 & + \frac{1}{2} \sum_{i=2}^3 \sum_{j=2}^3 \int_s^t f_{ij}(y, X_y^{(1)}, X_y^{(2)}) A_y^{(2,i)} A_y^{(2,j)} dy.
 \end{aligned}$$

Aquí $f_i(y, x_1, x_2)$, $f_{ij}(y, x_1, x_2)$ son las derivadas parciales de $f(y, x_1, x_2)$ con respecto a la variable i -ésima y las variables i -ésima y j -ésima, respectivamente.

Las integrales con respecto a $X_y^{(i)}$ han de ser interpretadas en la misma forma que en (3.20). La fórmula (3.22) puede ser extendida de manera natural a funciones $f(t, X_t^{(1)}, \dots, X_t^{(m)})$, donde las $X_t^{(i)}$'s son procesos de Itô con respecto al mismo movimiento Browniano.

Capítulo 4

Ecuaciones Diferenciales Estocásticas.

4.1. Ecuaciones diferenciales determinísticas.

La idea subyacente de una ecuación diferencial es simple: Damos una relación funcional

$$f(t, x(t), x'(t), x''(t), \dots) = 0, \quad 0 \leq t \leq T, \quad (4.1)$$

envolviendo al tiempo t , una función desconocida $x(t)$ y sus derivadas. El objetivo es encontrar la función $x(t)$ la cual satisface (4.1). Esta es llamada una solución de la ecuación diferencial (4.1). Un hecho importante es que esta solución es única para una condición inicial dada, por ejemplo, $x(0) = x_0$ para ecuaciones diferenciales de orden 1.

Las ecuaciones diferenciales más simples son aquellas de orden 1. Estas involucran solamente a t , $x(t)$ y su primera derivada $x'(t)$. Idealmente, tal ecuación diferencial esta dada de la forma

$$x'(t) = \frac{dx(t)}{dt} = a(t, x(t)), \quad x(0) = x_0, \quad (4.2)$$

para una función conocida $a(t, x)$. Es mas usual escribir (4.2) en una forma más intuitiva

$$dx(t) = a(t, x(t))dt, \quad x(0) = x_0. \quad (4.3)$$

Si se interpreta $x(t)$ como la posición de una partícula de una dimensión en el espacio al tiempo t , (4.3) describe el cambio de posición de la partícula en un pequeño intervalo de tiempo $[t, t + dt]$. La ecuación (4.3) entonces nos dice que $dx(t) = x(t + dt) - x(t)$ es proporcional al incremento de tiempo dt con factor $a(t, x(t))$. Alternativamente, (4.2) nos dice que la velocidad $x'(t)$ de la partícula es una función dada del tiempo t y la posición $x(t)$.

Ejemplo 4.1. (Algunas ecuaciones diferenciales simples.)

Suponga que la velocidad es solo función de t :

$$x'(t) = a(t).$$

Integrando ambos lados produce la solución

$$x(t) = x(0) + \int_0^t a(s)ds$$

Esta es la forma más simple de una ecuación diferencial, pero es incluso trivial dado que el lado derecho no depende de la función desconocida $x(t)$.

Ahora suponemos que la velocidad $x'(t)$ es proporcional a la posición $x(t)$:

$$x'(t) = cx(t)$$

para alguna constante c . Aquí la simple integración no sirve, pero se conoce la solución para esta ecuación diferencial: esta es la función exponencial $x(t) = x(0) \exp\{ct\}$. ■

Ejemplo 4.2. (Separación de variables.)

Suponga que el lado derecho de la ecuación (4.2) puede ser separado en el producto de dos funciones:

$$x'(t) = a_1(t)a_2(x(t)).$$

Simbólicamente, se reescribe esta ecuación diferencial como

$$\frac{dx}{a_2(x)} = a_1(t)dt. \quad (4.4)$$

Ahora integrando ambos lados de (4.4) obtenemos:

$$\int_{x(0)}^{x(t)} \frac{dx}{a_2(x)} = \int_0^t a_1(s)ds. \quad (4.5)$$

En el lado izquierdo obtenemos una función de $x(t)$, en el lado derecho tenemos una función de t . Afortunadamente, obtenemos una forma explícita de la función $x(t)$. Este acercamiento es justificado por la diferenciación en ambos lados de (4.5) como integrales del límites superiores. Entonces se obtiene

$$\frac{x'(t)}{a_2(x(t))} = a_1(t),$$

la cual es otra forma de escribir (4.4).

Aplicamos este método a la ecuación diferencial $x'(t) = cx(t)$ para algún $x(0) \neq 0$. Entonces

$$\int_{x(0)}^{x(t)} \frac{dx}{x} = c \int_0^t ds,$$

dando que $x(t) = x(0) \exp\{ct\}$. ■

Resumiendo, algunos aspectos importantes de las ecuaciones diferenciales ordinarias son:

- *Las soluciones de las ecuaciones diferenciales son funciones. Estas describen la evolución o la dinámica de procesos de la vida real sobre un periodo de tiempo dado.*
- *Con el objetivo de obtener una solución única, se tiene que conocer la condición inicial $x(0) = x_0$. Si esta solución única $x(t)$ comienza del punto x_0 en el presente $t = 0$, se dice, que la función $x(t)$ esta completamente determinada en el futuro, es decir, para $t > 0$.*
- *Las soluciones explícitas de las ecuaciones diferenciales son la excepción de la regla. En general, se tiene que recurrir a soluciones numéricas de las ecuaciones diferenciales.*
- *Integrando ambos lados de la ecuación diferencial (4.2), se obtiene una ecuación integral equivalente:*

$$x(t) = x(0) + \int_0^t a(s, x(s)) ds.$$

Aunque esta ecuación transformada no es en general de gran utilidad para encontrar la solución de (4.4), esta da una idea de cómo se puede definir una ecuación diferencial estocástica: como una ecuación integral estocástica.

4.2. Ecuaciones diferenciales estocásticas de Itô.

4.2.1. Definición de una ecuación diferencial estocástica.

Considerese la ecuación diferencial determinística

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

La forma mas fácil de introducir aleatoriedad en esta ecuación es aleatorizar la condición inicial. La solución $x(t)$ entonces se convierte en un proceso estocástico ($X_t, t \in [0, T]$):

$$dX_t = a(t, X_t)dt, \quad X_0(\omega) = Y(\omega).$$

Tal ecuación es llamada una *ecuación diferencial aleatoria*. Su solución no requiere de cálculo estocástico; se pueden utilizar los métodos clásicos y ajustar la solución a la correspondiente salida de la condición inicial. Las ecuaciones

diferenciales aleatorias pueden ser consideradas como ecuaciones diferenciales determinísticas con una condición inicial perturbada.

Para nuestros propósitos, la aleatoriedad en la ecuación diferencial es introducida a través de un término adicional de ruido aleatorio:

$$dX_t = a(t, X_t)dt + b(t, X_t)dB_t, \quad X_0(\omega) = Y(\omega) \quad (4.6)$$

Aquí, como es usual, $B = (B_t, t \geq 0)$ denota el movimiento Browniano, y $a(t, x)$ y $b(t, x)$ son funciones determinísticas. La solución X , si esta existe, es entonces un proceso estocástico. La aleatoriedad de $X = (X_t, t \in [0, T])$ resulta, de un lado, de la condición inicial, y por otro lado, del ruido generado por el movimiento Browniano.

Una interpretación simple de (4.6) nos dice que el cambio $dX_t = X_{t+dt} - X_t$ es causado por un cambio dt del tiempo, con un factor $a(t, X_t)$, en combinación con un cambio $dB_t = B_{t+dt} - B_t$ del movimiento Browniano, con un factor $b(t, X_t)$. Ya que el movimiento Browniano no tiene trayectorias diferenciables, no es claro como se puede interpretar (4.6). Ahora que tenemos el cálculo de Itô, se puede proponer lo siguiente.

Definición 4.1. *Se interpreta (4.6) como la ecuación integral estocástica*

$$X_t = X_0 + \int_0^t a(s, X_s)ds + \int_0^t b(s, X_s)dB_s, \quad 0 \leq t \leq T \quad (4.7)$$

donde la primera integral del lado derecho es una integral de Riemann, y la segunda es una ecuación integral de Itô.

La ecuación (4.7) es llamada una **ecuación diferencial estocástica de Itô**.

El llamar a la ecuación (4.7) una ecuación diferencial no es intuitivo. Su nombre es originado a partir de la ecuación simbólica (4.6), pero esta última es absurda a menos que se diga que son las diferenciales. Seguiremos la costumbre general y llamaremos a (4.7) una ecuación diferencial estocástica de Itô.

Definición 4.2. *El movimiento Browniano B es llamado el **proceso conductor** de la ecuación estocástica de Itô.*

Es posible reemplazar el movimiento Browniano por otro proceso conductor, pero esto requiere el definir integrales estocásticas más generales.

No es claro a priori como las integrales en (4.7) estén bien definidas. Se tiene que asegurar que $b(s, X_s)$ es adaptado al movimiento Browniano y que, $\int_0^T E[b(s, X_s)]^2 ds$, sea finita, ya que estas son condiciones cruciales para la definición de una integral estocástica de Itô. Una limitante para poder verificar estas

condiciones es que no se conoce la solución X . Pronto se mostrará que existen condiciones simples para la existencia y unicidad de las soluciones para las ecuaciones integrales estocásticas de Itô.

Primeramente trataremos de ver qué es una solución de la ecuación diferencial estocástica de Itô (4.7). Se pueden encontrar dos tipos de soluciones a una ecuación diferencial estocástica. Estas son llamadas soluciones fuertes y soluciones débiles.

Definición 4.3. Una **solución fuerte** de la ecuación diferencial estocástica de Itô (4.7) es un proceso estocástico $X = (X_t, t \in [0, T])$ el cual satisface las siguientes condiciones:

- i) X es adaptado al movimiento Browniano (véase apéndice A), es decir, al tiempo t es una función de $B_s, s \leq t$.
- ii) Las integrales en (4.7) están bien definidas como integrales de Riemann e integrales estocásticas de Itô, respectivamente.
- iii) X es una función de la trayectoria Browniana subyacente y de las funciones coeficiente $a(t, x)$ y $b(t, x)$.

Así una solución fuerte de (4.7) está basada en la trayectoria del movimiento Browniano subyacente. Si cambiáramos el movimiento Browniano por otro movimiento Browniano, obtendríamos otra solución fuerte la cual estaría dada por la misma relación funcional, pero con este nuevo movimiento Browniano en esta.

Por otro lado, para las soluciones débiles el comportamiento de la trayectoria no es esencial, sólo se está interesado en la distribución de X . La condición inicial X_0 y las funciones coeficiente $a(t, x)$ y $b(t, x)$ son dadas, y tenemos que encontrar el movimiento Browniano tal que (4.7) se cumpla.

Las soluciones débiles de X son suficientes con el objetivo de determinar las características distribucionales de X , tales como la esperanza, funciones de varianza y covarianza del proceso. En este caso, no tenemos que conocer las trayectorias de X .

Definición 4.4. Una solución fuerte o débil de X de la ecuación diferencial estocástica de Itô (4.7) es llamada una **difusión**. En particular, tomando $a(t, x) = 0$ y $b(t, x) = 1$ en (4.7), se verifica que el movimiento Browniano es un proceso de difusión.

Nota 4.1. En lo que sigue, sólo se considerarán soluciones fuertes de las ecuaciones diferenciales estocásticas de Itô.

Primero daremos las condiciones suficientes para la existencia y unicidad de tales soluciones.

Teorema 4.1. *Supóngase que la condición inicial X_0 tiene segundo momento finito: $E(X_0^2) < \infty$, y es independiente de $(B_t, t \geq 0)$.*

Adicionalmente supóngase que, para todo $t \in [0, T]$ y $x, y \in \mathbb{R}$, las funciones coeficiente $a(t, x)$ y $b(t, x)$ satisfacen las siguientes condiciones:

1. *Son continuas.*
2. *Satisfacen la condición de Lipschitz con respecto a la segunda variable:*

$$|a(t, x) - a(t, y)| + |b(t, x) - b(t, y)| \leq K|x - y|.$$

Entonces la ecuación diferencial estocástica de Itô (4.7) tiene una única solución fuerte X sobre $[0, T]$.

4.2.2. Resolución de ecuaciones diferenciales estocásticas de Itô mediante el lema de Itô.

En esta sección se resolverán algunas ecuaciones diferenciales estocásticas de Itô elementales mediante el lema de Itô.

Ejemplo 4.3. (El movimiento Browniano geométrico como solución de una ecuación diferencial estocástica lineal de Itô con ruido multiplicativo.)

Consideremos la ecuación diferencial estocástica lineal de Itô siguiente

$$X_t = X_0 + c \int_0^t X_s ds + \sigma \int_0^t X_s dB_s, \quad t \in [0, T], \quad (4.8)$$

con c y $\sigma > 0$ constantes dadas.

En la segunda integral de (4.8), la integral de Itô, el movimiento Browniano y el proceso X están conectados en un modo multiplicativo.

De (3.16) tenemos que:

El movimiento Browniano geométrico particular

$$X_t = X_0 e^{(c-0.5\sigma^2)t + \sigma B_t}, \quad t \in [0, T], \quad (4.9)$$

resuelve (4.8), y X es la única solución de (4.8).

Recordemos que se aplicó el lema de Itô para verificar que X satisface (4.8). Ahora estamos ante el problema inverso, usar el lema de Itô para encontrar la solución (4.9) a partir de (4.8).

Supongamos que $X_t = f(t, B_t)$ para alguna función suave $f(t, x)$ y a partir del lema de Itô (3.14):

$$X_t = X_0 + \int_0^t \left[f_1(s, B_s) + \frac{1}{2} f_{22}(s, B_s) \right] ds + \int_0^t f_2(s, B_s) dB_s. \quad (4.10)$$

El proceso X es un proceso de Itô, y por lo tanto podemos identificar los integrandos en las integrales de Riemann e Itô, respectivamente, de (4.8) y (4.10). Esto junto con la continuidad de las trayectorias del movimiento Browniano producen el siguiente sistema de dos ecuaciones diferenciales parciales para f :

$$cf(t, x) = f_1(t, x) + \frac{1}{2} f_{22}(t, x), \quad (4.11)$$

$$\sigma f(t, x) = f_2(t, x). \quad (4.12)$$

De (4.12) obtenemos

$$\sigma^2 f(t, x) = f_{22}(t, x).$$

Así, las dos ecuaciones diferenciales (4.11) y (4.12) pueden ser simplificadas:

$$(c - 0.5\sigma^2)f(t, x) = f_1(t, x), \quad \sigma f(t, x) = f_2(t, x). \quad (4.13)$$

Escribimos a $f(t, x)$ como el producto de dos funciones

$$f(t, x) = g(t)h(x).$$

Entonces (4.13) se transforma en

$$(c - 0.5\sigma^2)g(t) = g'(t), \quad \sigma h(x) = h'(x).$$

Resolviendo por separación de variables ambas ecuaciones se obtienen las soluciones

$$g(t) = g(0)e^{(c-0.5\sigma^2)t}, \quad h(x) = h(0)e^{\sigma x}.$$

Así, se obtiene

$$f(t, x) = g(0)h(0)e^{(c-0.5\sigma^2)t+\sigma x},$$

teniendo en cuenta que

$$X_0 = f(0, B_0) = f(0, 0) = g(0)h(0);$$

esto da finalmente

$$X_t = f(t, B_t) = X_0 e^{(c-0.5\sigma^2)t+\sigma B_t}, \quad t \in [0, T].$$

De esta manera la solución de una ecuación diferencial estocástica de Itô puede a veces ser derivada como solución de una ecuación diferencial parcial (determinística). ■

Ejemplo 4.4. (El proceso de Ornstein-Uhlenbeck.)

Consideremos otra ecuación diferencial estocástica lineal:

$$X_t = X_0 + c \int_0^t X_s ds + \sigma \int_0^t dB_s, \quad t \in [0, T]. \quad (4.14)$$

La ecuación (4.14) es usualmente conocida como la *ecuación de Langevin*.

En contraste con (4.8), el movimiento Browniano y el proceso X no están directamente conectados en la parte de la integral de Itô de (4.14).

Escribimos (4.14) en la forma intuitiva

$$dX_t = cX_t dt + \sigma dB_t,$$

y formalmente fijamos $dt = 1$. Entonces tenemos

$$X_{t+1} - X_t = cX_t + \sigma(B_{t+1} - B_t)$$

ó

$$X_{t+1} = \phi X_t + Z_t,$$

donde $\phi = c + 1$ es una constante y las variables aleatorias $Z_t = \sigma(B_{t+1} - B_t)$ constituyen una sucesión iid de variables aleatorias $N(0, \sigma^2)$.

Para resolver (4.14) es conveniente usar la siguiente transformación de X :

$$Y_t = e^{-ct} X_t$$

con la cual, ambos procesos X y Y satisfacen la misma condición inicial $X_0 = Y_0$.

Aplicando el lema de Itô (3.18) con

$$f(t, x) = e^{-ct} x, \quad f_1(t, x) = -cf(t, x), \quad f_2(t, x) = e^{-ct}, \quad f_{22} = 0,$$

y

$$A^{(1)} = cX \quad \text{y} \quad A^{(2)} = \sigma,$$

obtenemos

$$\begin{aligned} Y_t - Y_0 &= \int_0^t \left[f_1(s, X_s) + cX_s f_2(s, X_s) + \frac{1}{2} \sigma^2 f_{22}(s, X_s) \right] ds \\ &\quad + \int_0^t [\sigma f_2(s, X_s)] dB_s \\ &= \int_0^t [-cY_s + cY_s + 0] ds + \int_0^t [\sigma e^{-cs}] dB_s \\ &= \int_0^t [\sigma e^{-cs}] dB_s. \end{aligned}$$

De esta manera se concluye que el proceso

$$X_t = e^{ct} X_0 + \sigma e^{ct} \int_0^t e^{-cs} dB_s$$

es solución de la ecuación diferencial estocástica de Langevin (4.14). Para una condición inicial constante X_0 , este proceso es llamado un proceso de *Ornstein-Uhlenbeck*. ■

4.3. La ecuación diferencial general lineal.

Definición 4.5. *La ecuación diferencial estocástica general lineal.*

$$X_t = X_0 + \int_0^t [c_1(s)X_s + c_2(s)]ds + \int_0^t [\sigma_1(s)X_s + \sigma_2(s)]dB_s, \quad t \in [0, T] \quad (4.15)$$

Las funciones coeficiente (determinísticas) c_i y σ_i son continuas, por tanto acotadas sobre $[0, T]$, y así no es difícil de ver que las condiciones de existencia y unicidad, dadas en la observación [4.1], garantizan que tiene una única solución fuerte.

Esta ecuación es importante para muchas aplicaciones. Esta es particularmente atractiva porque tiene una solución explícita en términos de las funciones coeficiente y de la trayectoria Browniana subyacente. En lo que sigue, derivaremos esta solución mediante múltiples usos de diferentes variantes del lema de Itô.

4.3.1. Ecuaciones lineales con ruido aditivo.

Estableciendo $\sigma_1(t) = 0$ en (4.15), obtenemos

Definición 4.6. *La ecuación lineal con ruido aditivo.*

$$X_t = X_0 + \int_0^t [c_1(s)X_s + c_2(s)]ds + \int_0^t \sigma_2(s)dB_s, \quad t \in [0, T]. \quad (4.16)$$

En términos de diferenciales esto puede ser escrito como

$$dX_t = [c_1(t)X_t + c_2(t)]dt + \sigma_2(t)dB_t, \quad t \in [0, T].$$

El proceso X no está directamente envuelto en la integral estocástica, por consiguiente (4.16) de ahí gana su nombre. La solución de (4.16) es particularmente simple.

Una forma de resolver la ecuación diferencial es el intuir su forma a partir de algunos ejemplos particulares. Esta no es quizá la aproximación más satisfactoria ya que ciertamente se necesita mucha experiencia para tal cometido. De cualquier forma, vamos a suponer que

$$X_t = [y(t)]^{-1}Y_t,$$

donde

$$y(t) = \exp \left\{ - \int_0^t c_1(s) ds \right\} \text{ y } Y_t = f(t, X_t),$$

para alguna función suave $f(t, x)$. Una aplicación del lema de Itô (3.18) produce

$$dY_t = d(y(t)X_t) = c_2(t)y(t)dt + \sigma_2(t)y(t)dB_t.$$

Integrando ambos lados y notando que $y(0) = 1$, por lo tanto $X_0 = Y_0$, obtenemos:

La solución de la ecuación lineal con ruido aditivo:

$$X_t = [y(t)]^{-1} \left(X_0 + \int_0^t c_2(s)y(s)ds + \int_0^t \sigma_2(s)y(s)dB_s \right).$$

Si X_0 es una constante, esto define un proceso Gaussiano.

Ejemplo 4.5. (La ecuación de Langevin.)

En el ejemplo 4.4 ya hemos considerado la ecuación de Langevin. Esta es la ecuación (4.16) con $c_1(t) = c$, $c_2(t) = 0$ y $\sigma_2(t) = \sigma$. La solución está dada por

$$X_t = e^{ct}X_0 + \sigma e^{ct} \int_0^t e^{-cs}dB_s, \quad t \in [0, T].$$

Se mencionó en el ejemplo 4.4 que X es llamado un proceso de Ornstein-Uhlenbeck considerando X_0 constante. ■

4.3.2. Ecuaciones homogéneas con ruido multiplicativo.

Otro caso especial de la ecuación diferencial estocástica lineal general (4.15) esta dado por:

Definición 4.7. La ecuación lineal homogénea:

$$X_t = X_0 + \int_0^t c_1(s)X_s ds + \int_0^t \sigma_1(s)X_s dB_s, \quad t \in [0, T], \quad (4.17)$$

o, en el lenguaje de diferenciales:

$$dX_t = c_1(t)X_t dt + \sigma_1(t)X_t dB_t, \quad t \in [0, T].$$

Ya que X_t aparece como factor de incrementos del movimiento Browniano, (4.17) es llamada una ecuación diferencial estocástica con ruido multiplicativo. Es llamada homogénea porque $c_2(t) = \sigma_2(t) = 0$ para toda t .

Puesto que podemos dividir ambos lados de (4.17) por X_t entonces podemos suponer que $X_0 = 1$. (El caso $X_0 = 0$ corresponde al caso trivial $X_t = 0$ para toda t el cual no es de interés). Ya que esperamos una forma exponencial de la solución, suponemos que $X_t > 0$ para toda t . Esto nos permite considerar a $Y_t = \ln X_t = f(X_t)$ y aplicando el lema de Itô (3.18) con

$$f(t, x) = \ln x, \quad f_1(t, x) = 0, \quad f_2(t, x) = x^{-1}, \quad f_{22}(t, x) = -x^{-2}.$$

Obtenemos

$$dY_t = [c_1(t) - 0.5\sigma_1^2(t)]dt + \sigma_1(t)dB_t. \quad (4.18)$$

Primero integramos ambos lados de la ecuación (4.18), después tomando exponenciales y finalmente corrigiendo para el valor inicial si $X_0 \neq 1$; se obtiene:

La solución de la ecuación homogénea

$$X_t = X_0 \exp \left\{ \int_0^t [c_1(s) - 0.5\sigma_1^2(s)]ds + \int_0^t \sigma_1(s)dB_s \right\} \quad t \in [0, T]. \quad (4.19)$$

Ejemplo 4.6. (Movimiento Browniano geométrico.)

Un ejemplo de ecuación diferencial estocástica homogénea con ruido multiplicativo fue tratada en el ejemplo 4.3: con $c_1(t) = c$, $\sigma_1(t) = \sigma$ para c y σ constantes. Se encontró que el movimiento Browniano geométrico

$$X_t = X_0 e^{(c-0.5\sigma^2)t + \sigma B_t}, \quad t \in [0, T],$$

es la única solución fuerte en este caso. ■

4.3.3. El caso general.

Ahora se resolverá la ecuación diferencial estocástica lineal general (4.15). Esto se logra acoplando (4.15) en un sistema de dos ecuaciones diferenciales estocásticas. Recordando la solución Y de la ecuación diferencial estocástica homogénea (4.17) con $Y_0 = 1$ la cual esta dada por la fórmula (4.19) (con X reemplazado por Y). Considere los dos procesos

$$X_t^{(1)} = Y_t^{-1} \text{ y } X_t^{(2)} = X_t.$$

Aplicando el lema de Itô (3.18) a Y_t^{-1} (con $f(t, x) = x^{-1}$). Después de algunos cálculos obtenemos

$$dX_t^{(1)} = [-c_1(t) + \sigma_1^2(t)]X_t^{(1)}dt - \sigma_1(t)X_t^{(1)}dB_t,$$

apelando a la fórmula de integración por partes, produce

$$d(X_t^{(1)}X_t^{(2)}) = [c_2(t) - \sigma_1(t)\sigma_2(t)]X_t^{(1)}dt + \sigma_2(t)X_t^{(1)}dB_t.$$

Luego integrando ambos lados, recordando que $Y_0 = 1$ y teniendo en mente las formas particulares de $X_t^{(1)}$ y $X_t^{(2)}$, obtenemos finalmente:

La solución de la ecuación lineal general (4.15):

$$X_t = Y_t \left(X_0 + \int_0^t [c_2(s) - \sigma_1(s)\sigma_2(s)]Y_s^{-1}ds + \int_0^t \sigma_2(s)Y_s^{-1}dB_s \right), \quad t \in [0, T].$$

Nota 4.2. *Aquí, el proceso Y es la solución (4.19) de la ecuación homogénea (4.17) (donde X tiene que ser reemplazado con Y) con la condición inicial $Y_0 = 1$.*

De hecho, derivamos la solución de la correspondiente ecuación diferencial determinística.

$$dx(t) = [c_1(t)x(t) + c_2(t)]dt, \quad t \in [0, t].$$

Simplemente establecemos $\sigma_1(t) = \sigma_2(t) = 0$ para toda t .

Capítulo 5

Aplicaciones del Cálculo Estocástico a Matemáticas Financieras.

5.1. Conceptos y notación de matemáticas financieras.

5.1.1. Conceptos de finanzas.

Suponemos que el precio X_t de una posesión peligrosa (llamada acción) en el tiempo t esta dado por el movimiento Browniano geométrico de la forma

$$X_t = f(t, B_t) = X_0 e^{(c-0.5\sigma^2)t + \sigma B_t}, \quad (5.1)$$

donde, como es usual, $B = (B_t, t \geq 0)$ es el movimiento Browniano y X_0 se supone independiente de B . La motivación para esta suposición sobre X viene del hecho que X es la única solución fuerte de la ecuación diferencial estocástica lineal

$$X_t = X_0 + c \int_0^t X_s ds + \sigma \int_0^t X_s dB_s. \quad (5.2)$$

La ecuación (5.2) se puede escribir formalmente como

$$dX_t = cX_t dt + \sigma X_t dB_t.$$

Si se interpreta esta ecuación de una forma sencilla, tenemos en $[t, t + dt]$

$$X_{t+dt} - X_t = cX_t dt + \sigma X_t dB_t.$$

Equivalentemente

$$\frac{X_{t+dt} - X_t}{X_t} = cdt + \sigma dB_t.$$

La cantidad del lado izquierdo es el *retorno relativo* de la posesión en el periodo de tiempo $[t, t + dt]$. Esto nos dice que existe una tendencia lineal $c dt$, la cual es perturbada por un término de ruido estocástico σdB_t . La constante $c > 0$ es la así llamada *tasa de retorno media*, y $\sigma > 0$ es la *volatilidad*. Un vistazo a la fórmula (5.1) nos dice que, para σ grande, más grandes son las fluctuaciones de X_t . Por lo tanto, σ es una medida de riesgo de la posesión.

Es cierto que el modelo (5.2) es razonable, aunque, una primera aproximación burda a un proceso de precio real. Si se olvida por el momento el término que contiene a σ , es decir, suponemos que $\sigma = 0$, entonces (5.2) es una ecuación diferencial determinística la cual tiene una solución bien conocida: $X_t = X_0 \exp\{ct\}$. De esta manera, si $\sigma > 0$ esperamos obtener una función exponencial perturbada aleatoriamente, y esta es el movimiento Browniano geométrico (5.1).

Ahora supóngase que no se tiene ninguna posesión peligrosa como una cuenta bancaria. En la teoría financiera, esto es llamado un *bono*. Suponemos que una inversión de β_0 en el bono produce un monto de

$$\beta_t = \beta_0 e^{rt},$$

al tiempo t . Por lo tanto, el capital inicial de β_0 ha sido agravado constantemente a una tasa de interés constante $r > 0$. Esto es una idealización ya que la tasa de interés cambia con el tiempo. Notese que β satisface la ecuación integral determinística

$$\beta_t = \beta_0 + r \int_0^t \beta_s ds. \quad (5.3)$$

En general, se desea mantener cierto monto de posesiones: a_t en acciones y b_t en bonos, esto constituyen un *portafolio*. Suponemos que a_t y b_t son procesos estocásticos adaptados al movimiento Browniano y llamamos al par

$$(a_t, b_t), \quad t \in [0, T],$$

estrategia de mercado. Claramente se desea escoger una estrategia donde no haya pérdidas. Cómo elegir (a_t, b_t) de un modo razonable, será discutido posteriormente. Nótese que la *riqueza* V_t (o el valor del portafolio) al tiempo t esta dado por

$$V_t = a_t X_t + b_t \beta_t.$$

Se permite que ambos, a_t y b_t , tomen cualquier valor positivo o negativo. Un valor negativo de a_t significa una venta breve de la acción, es decir, se vende la acción al tiempo t . Un valor negativo de b_t significa que se pide prestado a la tasa de interés segura r del bono. En realidad, se debería pagar un costo de transacción por las operaciones sobre las acciones y venta, pero despreciamos este costo por simplicidad. Más aún, no suponemos que a_t y b_t , sean acotados.

Así, en principio, se debería tener un monto (cantidad) potencialmente infinito de capital, y se debería tener deudas inconsolidadas (no-acotadas) en la cuenta también. Claramente esto es una simplificación, lo cual hace nuestros problemas matemáticos más fáciles. Finalmente, suponemos que no se gasta dinero en otros propósitos, es decir, no se hace el portafolio mas pequeño por el consumo.

Suponemos que la estrategia de portafolio (a_t, b_t) está autofinanciada. Esto significa que el incremento de la riqueza V_t resulta sólo de los cambios de los precios de las posesiones X_t y β_t . Formulemos la condición de autofinanciamiento en términos de diferenciales

$$dV_t = d(a_t X_t + b_t \beta_t) = a_t dX_t + b_t d\beta_t,$$

el cual interpretamos en el sentido de Itô como la relación

$$V_t - V_0 = \int_0^t d(a_s X_s + b_s \beta_s) = \int_0^t a_s dX_s + \int_0^t b_s d\beta_s.$$

Las integrales en el lado derecho claramente tienen sentido si reemplazamos dX_s con $cX_s ds + \sigma X_s dB_s$, véase (5.2) y $d\beta_s$ con $r\beta_s ds$, véase (5.3). Por lo tanto, el valor V_t del portafolio al tiempo t es precisamente igual a la inversión inicial V_0 más el capital ganado de las acciones y el bono al tiempo t .

5.1.2. Opciones de compra y venta europeas.

Ahora supóngase que se adquiere un boleto, llamado una *opción*, al tiempo $t = 0$ el cual da la facultad para comprar una parte de la acción hasta o al tiempo T , el *tiempo de maduración* o *tiempo de expiración* de la opción. Si se puede ejercer esta opción al precio fijo K llamado *precio de ejercicio* de la opción, solo al tiempo de maduración T , esta es llamada *la opción de compra europea* (call option). Si se puede ejercer hasta o al tiempo T , esta es llamada una *opción de compra americana*. Hay muchas más opciones en el mundo real de las finanzas, sin embargo, éste estudio se enfocará en las opciones de compra y venta europeas.

El tenedor de una opción de compra no está obligado a ejercer esta. Por lo tanto, si al tiempo T el precio X_T es menor que K , el tenedor del boleto sería tonto de ejercer esta (se podría comprar una acción por $\$X_T$ en el mercado) y así el boleto expira como un contrato sin valor. Si el precio X_T excede K , vale la pena el ejercer la llamada, es decir, se puede comprar la acción al precio K , entonces gira y se vende esta al precio X_T para un beneficio neto $X_T - K$.

En resumen, el comprador de la opción de compra europea es facultado a una retribución de

$$(X_T - K)^+ = \max(0, X_T - K) = \begin{cases} X_T - K & \text{si } X_T > K, \\ 0 & \text{si } X_T \leq K. \end{cases}$$

Una opción de venta europea (put option) es una opción para vender una acción al precio dado K en o hasta una fecha particular de maduración T . Una opción de venta europea es ejercida solamente al tiempo de maduración, una opción de venta americana puede ser ejercida hasta o al tiempo T . El comprador de una opción de venta europea hace una ganancia

$$(K - X_T)^+ = \begin{cases} K - X_T & \text{si } X_T < K, \\ 0 & \text{si } X_T \geq K. \end{cases}$$

En nuestras consideraciones teóricas nos restringiremos a las opciones de compra y venta europeas, como se había mencionado anteriormente. Esto tiene una razón simple: en este caso se puede derivar explícitamente las soluciones y compactar la formulación para el problema de valuación de opciones.

Nota 5.1. *De ahora en adelante, una opción es una opción europea.*

Por otra parte, es interesante notar que la situación puede ser imaginada como un juego donde la recompensa es el pago de la opción y el tenedor de la opción paga los honorarios (el precio de la opción) por jugar el juego.

5.2. Análisis de la ecuación de Black-Scholes.

5.2.1. Una formulación matemática del problema de valuación de opciones.

Supongamos que queremos encontrar la estrategia de autofinanciamiento (a_t, b_t) y un valor asociado al proceso V_t tal que

$$V_t = a_t X_t + b_t \beta_t = u(T - t, X_t) \quad t \in [0, T],$$

para alguna función determinística suave $u(t, x)$. Claramente esto es una restricción: se supone que el valor de V_t del portafolio depende en forma suave de t y X_t . Es nuestro objetivo encontrar esta función $u(t, x)$. Ya que el valor V_t del portafolio al tiempo de maduración T será $(X_T - K)^+$ obtenemos la condición terminal

$$V_t = u(0, X_T) = (X_T - K)^+. \quad (5.4)$$

En la literatura financiera, el proceso de construir una estrategia de autofinanciamiento tal que (5.4) se cumpla es llamada *equilibrar la denuncia en contra de la contingencia* $(X_T - K)^+$.

Se intentará aplicar el lema de Itô al valor del proceso $V_t = u(T - t, X_t)$.
Escribimos $f(t, x) = u(T - t, x)$ y notese que

$$f_1(t, x) = -u_1(T - t, x), \quad f_2(t, x) = u_2(T - t, x), \quad f_{22}(t, x) = u_{22}(T - t, x).$$

Incluso es necesario recordar que X satisface la ecuación integral de Itô

$$X_t = X_0 + c \int_0^t X_s ds + \sigma \int_0^t X_s dB_s.$$

Ahora, la aplicación del lema de Itô (3.18) con $A^{(1)} = cX$ y $A^{(2)} = \sigma X$ produce

$$\begin{aligned} V_t - V_0 &= f(t, X_t) - f(0, X_0) \\ &= \int_0^t [f_1(s, X_s) + cX_s f_2(s, X_s) + \frac{1}{2} \sigma^2 X_s^2 f_{22}(s, X_s)] ds \\ &\quad + \int_0^t [\sigma X_s f_2(s, X_s)] dB_s \\ &= \int_0^t [-u_1(T - s, X_s) + cX_s u_2(T - s, X_s) + \frac{1}{2} \sigma^2 X_s^2 u_{22}(T - s, X_s)] ds \\ &\quad + \int_0^t [\sigma X_s u_2(T - s, X_s)] dB_s. \end{aligned} \tag{5.5}$$

Por otra parte, (a_t, b_t) es el autofinanciamiento

$$V_t - V_0 = \int_0^t a_s dX_s + \int_0^t b_s d\beta_s. \tag{5.6}$$

Ya que $\beta_t = \beta_0 e^{rt}$,

$$d\beta_t = r\beta_0 e^{rt} dt = r\beta_t dt. \tag{5.7}$$

Más aún, $V_t = a_t X_t + b_t \beta_t$, por lo tanto

$$b_t = \frac{V_t - a_t X_t}{\beta_t}. \tag{5.8}$$

Combinando (5.6)-(5.8), obtenemos la expresión:

$$\begin{aligned} V_t - V_0 &= \int_0^t a_s dX_s + \int_0^t \frac{V_s - a_s X_s}{\beta_s} r \beta_s ds \\ &= \int_0^t a_s dX_s + \int_0^t r(V_s - a_s X_s) ds \\ &= \int_0^t c a_s X_s ds + \int_0^t \sigma a_s X_s dB_s + \int_0^t r(V_s - a_s X_s) ds \\ &= \int_0^t [(c - r)a_s X_s + rV_s] ds + \int_0^t [\sigma a_s X_s] dB_s. \end{aligned} \tag{5.9}$$

Ahora comparamos las ecuaciones (5.5) y (5.9). De acuerdo con la nota [3.6] tenemos que las funciones coeficiente de los procesos de Itô deben coincidir. Así, se puede formalmente identificar los integrandos de las integrales de Riemann e Itô, respectivamente, en (5.5) y (5.9) como:

$$\begin{aligned} a_t &= u_2(T-t, X_t), \\ (c-r)a_t X_t + ru(T-t, X_t) &= (c-r)u_2(T-t, X_t)X_t + ru(T-t, X_t) \\ &= -u_1(T-t, X_t) + cX_t u_2(T-t, X_t) \\ &\quad + \frac{1}{2}\sigma^2 X_t^2 u_{22}(T-t, X_t). \end{aligned}$$

Ya que X_t puede tomar cualquier valor positivo, podemos escribir la última identidad como una ecuación diferencial parcial:

$$ru(T-t, x) = -u_1(T-t, x) + \frac{1}{2}\sigma^2 x^2 u_{22}(T-t, x) + rxu_2(T-t, x) \quad x > 0.$$

Haciendo el cambio de variable $\tau = T-t$ y regresando a la variable t , es decir, tomando $u(\tau, x) = V(t, x)$, entonces tenemos que la última ecuación diferencial parcial se transforma en:

$$rV(t, x) = V_1(t, x) + \frac{1}{2}\sigma^2 x^2 V_{22}(t, x) + rxV_2(t, x) \quad x > 0, t \in [0, T]. \quad (5.10)$$

la cual es llamada la *ecuación de Black-Scholes* (véase apéndice D para una reseña histórica).

Recordando la condición terminal (5.4) y teniendo en cuenta el cambio de variable, requerimos incluso que

$$V_T = V(T, X_T) = (X_T - K)^+.$$

Esto produce la condición terminal determinística

$$V(T, x) = (x - K)^+, \quad x > 0.$$

5.3. Planteamiento de problemas de valor a la frontera con la ecuación de Black-Scholes.

5.3.1. Problema de valor a la frontera para una opción de compra.

Sea el problema de valoración de la opción de compra $C(t, S)$, análogo al dado por la ecuación (5.10),

$$rC(t, S) = C_1(t, S) + \frac{1}{2}\sigma^2 S^2 C_{22}(t, S) + rSC_2(t, S) \quad (5.11)$$

con las siguientes condiciones de frontera

$$\begin{aligned} C(T, S) &= \text{máx}(S - K, 0), \\ C(t, S) &= S - Ke^{-r(T-t)} \quad \text{cuando } S \rightarrow \infty, \\ C(t, 0) &= 0 \quad \forall t, \end{aligned}$$

donde $S > 0$ es el valor del subyacente, σ la volatilidad fija, r el tipo de interés libre constante, $t \in [0, T]$ y T es el tiempo de vencimiento o maduración de la opción.

Con el objetivo de resolver la ecuación diferencial parcial, se transforma la ecuación (5.11) en una ecuación de difusión, para la cual ya hemos encontrado su solución en la sección 2.3.2. Por lo tanto, introducimos el siguiente cambio de variable

$$\tau = \frac{1}{2}\sigma^2(T - t), \quad (5.12)$$

$$x = \ln\left(\frac{S}{K}\right), \quad (5.13)$$

$$C(t, S) = KV(\tau, x). \quad (5.14)$$

Primero calculemos las siguientes derivadas parciales

$$\begin{aligned} V_1(\tau, x) &= \frac{1}{K} \frac{\partial C}{\partial \tau} = \frac{1}{K} \left[\frac{\partial C}{\partial t} \frac{\partial t}{\partial \tau} + \frac{\partial C}{\partial S} \frac{\partial S}{\partial \tau} \right] \\ &= \frac{1}{K} \left[\frac{\partial C}{\partial t} \left(-\frac{2}{\sigma^2} \right) \right] = -\frac{2}{K\sigma^2} C_1(t, S). \end{aligned}$$

De aquí tenemos que,

$$C_1(t, S) = -\frac{K\sigma^2}{2} V_1(\tau, x). \quad (5.15)$$

Ahora, calculemos $V_2(\tau, x)$

$$\begin{aligned} V_2(\tau, x) &= \frac{1}{K} \frac{\partial C}{\partial x} = \frac{1}{K} \left[\frac{\partial C}{\partial t} \frac{\partial t}{\partial x} + \frac{\partial C}{\partial S} \frac{\partial S}{\partial x} \right] \\ &= \frac{1}{K} \left[\frac{\partial C}{\partial S} (Ke^x) \right] = e^x C_2(t, S). \end{aligned}$$

Con lo cual se obtiene,

$$C_2(t, S) = \frac{K}{S} V_2(\tau, x). \quad (5.16)$$

Por último, calculemos $V_{22}(\tau, x)$

$$\begin{aligned}
 V_{22}(\tau, x) &= \frac{1}{K} \frac{\partial^2 C}{\partial x^2} = \frac{\partial}{\partial x} \left[\frac{1}{K} \frac{\partial C}{\partial x} \right] \\
 &= \frac{\partial}{\partial x} \left[e^x \frac{\partial C}{\partial S} \right] = \frac{\partial C}{\partial S} \frac{\partial}{\partial x} (e^x) + e^x \frac{\partial}{\partial x} \left(\frac{\partial C}{\partial S} \right) \\
 &= e^x \frac{\partial C}{\partial S} + e^x \left[\frac{\partial^2 C}{\partial t \partial S} \frac{\partial t}{\partial x} + \frac{\partial^2 C}{\partial S^2} \frac{\partial S}{\partial x} \right] \\
 &= V_2(\tau, x) + e^x \left[\frac{\partial^2 C}{\partial S^2} (K e^x) \right] = V_2(\tau, x) + K e^{2x} C_{22}(t, S).
 \end{aligned}$$

Obteniendo así,

$$C_{22}(t, S) = \frac{K}{S^2} [V_{22}(\tau, x) - V_2(\tau, x)]. \quad (5.17)$$

Sustituyendo las ecuaciones (5.14)-(5.17) en (5.11), obtenemos la siguiente ecuación diferencial parcial

$$-V_1(\tau, x) + V_{22}(\tau, x) + \left[\frac{2r}{\sigma^2} - 1 \right] V_2(\tau, x) = \frac{2r}{\sigma^2} V(\tau, x), \quad (5.18)$$

con la condición inicial

$$\begin{aligned}
 V(x, 0) &= \frac{1}{K} C(T, S) = \frac{1}{K} \max(S - K, 0) \\
 &= \frac{1}{K} \max(K e^x - K, 0) = \frac{1}{K} \max(K(e^x - 1), 0) \\
 &= \max(e^x - 1, 0).
 \end{aligned}$$

Ahora utilizamos la siguiente transformación

$$U(\tau, x) = e^{mx+l\tau} V(\tau, x), \quad (5.19)$$

y calculamos las siguientes derivadas parciales

$$\begin{aligned}
 V_1(\tau, x) &= \frac{\partial}{\partial \tau} (e^{-mx-l\tau} U(\tau, x)) \\
 &= e^{-mx-l\tau} U_1(\tau, x) - l e^{-mx-l\tau} U(\tau, x) \\
 &= e^{-mx-l\tau} [U_1(\tau, x) - lU(\tau, x)].
 \end{aligned}$$

Por lo tanto,

$$V_1(\tau, x) = e^{-mx-l\tau} [U_1(\tau, x) - lU(\tau, x)]. \quad (5.20)$$

$$\begin{aligned}
V_2(\tau, x) &= \frac{\partial}{\partial x} (e^{-mx-l\tau} U(\tau, x)) \\
&= e^{-mx-l\tau} U_2(\tau, x) - m e^{-mx-l\tau} U(\tau, x) \\
&= e^{-mx-l\tau} [U_2(\tau, x) - mU(\tau, x)].
\end{aligned}$$

Luego,

$$V_2(\tau, x) = e^{-mx-l\tau} [U_2(\tau, x) - mU(\tau, x)]. \quad (5.21)$$

$$\begin{aligned}
V_{22}(\tau, x) &= \frac{\partial^2}{\partial x^2} (e^{-mx-l\tau} U(\tau, x)) \\
&= \frac{\partial}{\partial x} (e^{-mx-l\tau} [U_2(\tau, x) - mU(\tau, x)]) \\
&= e^{-mx-l\tau} U_{22}(\tau, x) - 2m e^{-mx-l\tau} U_2(\tau, x) + m^2 e^{-mx-l\tau} U(\tau, x) \\
&= e^{-mx-l\tau} [U_{22}(\tau, x) - 2mU_2(\tau, x) + m^2U(\tau, x)].
\end{aligned}$$

Finalmente,

$$V_{22}(\tau, x) = e^{-mx-l\tau} [U_{22}(\tau, x) - 2mU_2(\tau, x) + m^2U(\tau, x)]. \quad (5.22)$$

Reescribiendo (5.18) en términos de las ecuaciones (5.20)-(5.22), dividiendo por $\exp\{-mx - l\tau\}$ y simplificando, obtenemos la siguiente ecuación diferencial parcial

$$\begin{aligned}
-U_1(\tau, x) + U_{22}(\tau, x) + \left[\frac{2r}{\sigma^2} - 1 - 2m \right] U_2(\tau, x) \\
+ \left[l + m^2 - m \left(\frac{2r}{\sigma^2} - 1 \right) - \frac{2r}{\sigma^2} \right] U(\tau, x) = 0. \quad (5.23)
\end{aligned}$$

Luego para obtener la ecuación de difusión, tenemos que encontrar valores de m y l tales que los coeficientes de $U_2(\tau, x)$ y $U(\tau, x)$, en (5.23), sean iguales a cero.

Del coeficiente de $U_2(\tau, x)$, se tiene

$$\begin{aligned}
0 &= \frac{2r}{\sigma^2} - 1 - 2m \\
2m &= \frac{2r}{\sigma^2} - 1 \\
m &= \frac{1}{2} \left(\frac{2r}{\sigma^2} - 1 \right).
\end{aligned}$$

y del coeficiente de $U(\tau, x)$ y sustituyendo el valor de m , se sigue

$$\begin{aligned} 0 &= l + m^2 - m \left(\frac{2r}{\sigma^2} - 1 \right) - \frac{2r}{\sigma^2} \\ l &= -m^2 + m \left(\frac{2r}{\sigma^2} - 1 \right) + \frac{2r}{\sigma^2} \\ l &= -m^2 + 2m^2 + 2m + 1 \\ l &= m^2 + 2m + 1 \\ l &= (1 + m)^2. \end{aligned}$$

De esta manera, con los valores de m y l , la ecuación (5.23) se transforma en la ecuación de difusión

$$U_1(\tau, x) = U_{22}(\tau, x), \quad (5.24)$$

la cual tiene coeficiente de difusión igual a 1, y que, usualmente se escribe como

$$\frac{\partial U}{\partial \tau} = \frac{\partial^2 U}{\partial x^2},$$

con condición inicial

$$\begin{aligned} U(0, x) &= U_0(x) = e^{mx} \text{máx}(e^x - 1, 0) \\ &= \text{máx}(e^{mx}(e^x - 1), 0) \\ &= \text{máx}(e^{mx+x} - e^{mx}, 0) \\ &= \text{máx}(e^{(m+1)x} - e^{mx}, 0) \\ &= \text{máx}(e^{\frac{1}{2}(\alpha+1)x} - e^{\frac{1}{2}(\alpha-1)x}, 0) \end{aligned} \quad (5.25)$$

donde $\alpha = 2r/\sigma^2$.

Ahora utilizamos la solución de la ecuación de difusión (2.13) obtenida en la sección 2.3.2, tomando la constante de difusión $a^2 = 1$ y $U_0(x) = \psi(x)$ en (2.13), así tenemos que

$$U(\tau, x) = \frac{1}{2\sqrt{\pi\tau}} \int_{-\infty}^{\infty} U_0(x_0) e^{-\frac{(x_0-x)^2}{4\tau}} dx_0, \quad (5.26)$$

es la solución de la ecuación de difusión (5.24).

Hacemos el cambio de variable $y = (x_0 - x)/\sqrt{2\tau}$, de esta manera la ecuación (5.26) se convierte en:

$$U(\tau, x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} U_0(\sqrt{2\tau}y + x) e^{-\frac{y^2}{2}} dy.$$

Sustituyendo la condición inicial U_0 dada por la ecuación (5.25) en la última ecuación, esta se transforma en:

$$\begin{aligned} U(\tau, x) &= \frac{1}{\sqrt{2\pi}} \int_{-\frac{x}{\sqrt{2\tau}}}^{\infty} \left(e^{\frac{1}{2}(\alpha+1)(\sqrt{2\tau}y+x)} - e^{\frac{1}{2}(\alpha-1)(\sqrt{2\tau}y+x)} \right) e^{-\frac{y^2}{2}} dy \\ &= \frac{1}{\sqrt{2\pi}} \left[\int_{-\frac{x}{\sqrt{2\tau}}}^{\infty} e^{\frac{1}{2}(\alpha+1)(\sqrt{2\tau}y+x)} e^{-\frac{y^2}{2}} dy - \int_{-\frac{x}{\sqrt{2\tau}}}^{\infty} e^{\frac{1}{2}(\alpha-1)(\sqrt{2\tau}y+x)} e^{-\frac{y^2}{2}} dy \right] \\ &= I_1 - I_2, \end{aligned}$$

donde

$$I_1 = \frac{1}{\sqrt{2\pi}} \int_{-\frac{x}{\sqrt{2\tau}}}^{\infty} e^{\frac{1}{2}(\alpha+1)(\sqrt{2\tau}y+x)} e^{-\frac{y^2}{2}} dy, \quad (5.27)$$

$$I_2 = \frac{1}{\sqrt{2\pi}} \int_{-\frac{x}{\sqrt{2\tau}}}^{\infty} e^{\frac{1}{2}(\alpha-1)(\sqrt{2\tau}y+x)} e^{-\frac{y^2}{2}} dy. \quad (5.28)$$

Reescribiendo (5.27) tenemos que:

$$I_1 = \frac{1}{\sqrt{2\pi}} e^{\frac{1}{2}(\alpha+1)x + \frac{1}{4}(\alpha+1)^2\tau} \int_{-\frac{x}{\sqrt{2\tau}}}^{\infty} e^{-\frac{1}{2}(y - \frac{1}{2}(\alpha+1)\sqrt{2\tau})^2} dy.$$

Utilizando el cambio de variable $z = y - \frac{1}{2}(\alpha+1)\sqrt{2\tau}$ tenemos que la última igualdad se transforma en

$$I_1 = \frac{1}{\sqrt{2\pi}} e^{\frac{1}{2}(\alpha+1)x + \frac{1}{4}(\alpha+1)^2\tau} \int_{-d_1}^{\infty} e^{-\frac{z^2}{2}} dz.$$

Teniendo en cuenta que $\frac{1}{\sqrt{2\pi}} \int_{-d_1}^{\infty} \exp\{-z^2/2\} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_1} \exp\{-z^2/2\} dz$, ya que es la función de distribución de probabilidad acumulada de una normal estándar, entonces tenemos finalmente que:

$$I_1 = e^{\frac{1}{2}(\frac{2r}{\sigma^2}+1)x + \frac{1}{4}(\frac{2r}{\sigma^2}+1)^2\tau} \Phi(d_1),$$

donde

$$d_1 = \frac{x + (\frac{2r}{\sigma^2} + 1)\tau}{\sqrt{2\tau}}. \quad (5.29)$$

Análogamente transformamos (5.28) con lo que se tiene que

$$I_2 = e^{\frac{1}{2}(\frac{2r}{\sigma^2}-1)x + \frac{1}{4}(\frac{2r}{\sigma^2}-1)^2\tau} \Phi(d_2),$$

con

$$d_2 = \frac{x + \left(\frac{2r}{\sigma^2} - 1\right)\tau}{\sqrt{2\tau}}. \quad (5.30)$$

De esta manera, la solución de (5.24) es:

$$U(\tau, x) = e^{\frac{1}{2}\left(\frac{2r}{\sigma^2}+1\right)x + \frac{1}{4}\left(\frac{2r}{\sigma^2}+1\right)^2\tau} \Phi(d_1) - e^{\frac{1}{2}\left(\frac{2r}{\sigma^2}-1\right)x + \frac{1}{4}\left(\frac{2r}{\sigma^2}-1\right)^2\tau} \Phi(d_2). \quad (5.31)$$

Despejando $V(\tau, x)$ y sustituyendo (5.31) junto con los valores de m y l en (5.19), se tiene que

$$V(\tau, x) = e^x \Phi(d_1) - e^{-\frac{2r}{\sigma^2}\tau} \Phi(d_2). \quad (5.32)$$

Por último, utilizando las ecuaciones (5.12)-(5.14) y (5.32), se tiene que la solución de la ecuación diferencial parcial (5.11) esta dada por:

$$C(t, S) = S\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2), \quad (5.33)$$

con

$$d_1 = \frac{\ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{(T-t)}} \quad \text{y} \quad d_2 = \frac{\ln\left(\frac{S}{K}\right) + \left(r - \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{(T-t)}}.$$

La ecuación (5.33) es conocida como la *fórmula de Black-Scholes para valoración de opciones de compra*.

5.3.2. Problema de valor a la frontera para una opción de venta.

Ahora consideremos el problema de valoración de la opción de venta $P(t, S)$, análogo al dado por la ecuación (5.10),

$$rP(t, S) = P_1(t, S) + \frac{1}{2}\sigma^2 S^2 P_{22}(t, S) + rSP_2(t, S) \quad (5.34)$$

con las siguientes condiciones de frontera

$$\begin{aligned} P(T, S) &= \text{máx}(K - S, 0), \\ P(t, S) &= 0 \quad \text{cuando } S \rightarrow \infty, \\ P(t, S) &= Ke^{-r(T-t)} - S \quad \text{cuando } S \rightarrow 0, \end{aligned}$$

donde $S > 0$ es el valor del subyacente, σ la volatilidad fija, r el tipo de interés libre constante, $t \in [0, T]$ y T es el tiempo de vencimiento o maduración de la opción.

Con el objetivo de resolver la ecuación diferencial parcial se transforma la ecuación (5.34) en una ecuación de difusión, para la cual ya hemos encontrado su solución en la sección 2.3.2. Para hacer esto, introducimos el siguiente cambio de variable

$$\tau = \frac{1}{2}\sigma^2(T - t), \quad (5.35)$$

$$x = \ln\left(\frac{S}{K}\right), \quad (5.36)$$

$$P(t, S) = KV(\tau, x). \quad (5.37)$$

En forma análoga a como se obtuvieron las ecuaciones (5.15)-(5.17), tenemos que:

$$P_1(t, S) = -\frac{K\sigma^2}{2}V_1(\tau, x). \quad (5.38)$$

$$P_2(t, S) = \frac{K}{S}V_2(\tau, x). \quad (5.39)$$

$$P_{22}(t, S) = \frac{K}{S^2}[V_{22}(\tau, x) - V_2(\tau, x)]. \quad (5.40)$$

Sustituyendo las ecuaciones (5.37)-(5.40) en (5.34), obtenemos la ecuación diferencial parcial dada en (5.18)

$$-V_1(\tau, x) + V_{22}(\tau, x) + \left[\frac{2r}{\sigma^2} - 1\right]V_2(\tau, x) = \frac{2r}{\sigma^2}V(\tau, x),$$

sólo que ahora la condición inicial esta dada por

$$\begin{aligned} V(x, 0) &= \frac{1}{K}P(T, S) = \frac{1}{K}\max(K - S, 0) \\ &= \frac{1}{K}\max(K - Ke^x, 0) = \frac{1}{K}\max(K(1 - e^x), 0) \\ &= \max(1 - e^x, 0). \end{aligned}$$

De la misma manera como se procedió en la sección 5.3.1, utilizamos la transformación dada por la ecuación (5.19), $U(\tau, x) = \exp\{mx + l\tau\}V(\tau, x)$, por lo que, la ecuación (5.34) se convierte finalmente en la ecuación de difusión dada por (5.24)

$$U_1(\tau, x) = U_{22}(\tau, x).$$

con la diferencia que este problema tiene la condición inicial

$$\begin{aligned}
U(0, x) &= U_0(x) = e^{mx} \text{máx}(1 - e^x, 0) \\
&= \text{máx}(e^{mx}(1 - e^x), 0) \\
&= \text{máx}(e^{mx} - e^{m+x}, 0) \\
&= \text{máx}(e^{mx} - e^{(m+1)x}, 0) \\
&= \text{máx}(e^{\frac{1}{2}(\alpha-1)x} - e^{\frac{1}{2}(\alpha+1)x}, 0),
\end{aligned} \tag{5.41}$$

con $\alpha = 2r/\sigma^2$.

Ahora utilizamos la solución de la ecuación de difusión (2.13) obtenida en la sección 2.3.2, tomando la constante de difusión $a^2 = 1$ y $U_0(x) = \psi(x)$ en (2.13), así, tenemos que

$$U(\tau, x) = \frac{1}{2\sqrt{\pi\tau}} \int_{-\infty}^{\infty} U_0(x_0) e^{-\frac{(x_0-x)^2}{4\tau}} dx_0, \tag{5.42}$$

es la solución de la ecuación de difusión asociada a este problema.

Hacemos el cambio de variable $y = (x_0 - x)/\sqrt{2\tau}$, de esta manera la ecuación (5.42) se convierte en:

$$U(\tau, x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} U_0(\sqrt{2\tau}y + x) e^{-\frac{y^2}{2}} dy.$$

Sustituyendo la condición inicial U_0 dada por la ecuación (5.41) en la última ecuación, esta se transforma en:

$$\begin{aligned}
U(\tau, x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\frac{x}{\sqrt{2\tau}}} \left(e^{\frac{1}{2}(\alpha-1)(\sqrt{2\tau}y+x)} - e^{\frac{1}{2}(\alpha+1)(\sqrt{2\tau}y+x)} \right) e^{-\frac{y^2}{2}} dy \\
&= \frac{1}{\sqrt{2\pi}} \left[\int_{-\infty}^{-\frac{x}{\sqrt{2\tau}}} e^{\frac{1}{2}(\alpha-1)(\sqrt{2\tau}y+x)} e^{-\frac{y^2}{2}} dy - \int_{-\infty}^{-\frac{x}{\sqrt{2\tau}}} e^{\frac{1}{2}(\alpha+1)(\sqrt{2\tau}y+x)} e^{-\frac{y^2}{2}} dy \right] \\
&= I_1 - I_2,
\end{aligned}$$

donde

$$I_1 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\frac{x}{\sqrt{2\tau}}} e^{\frac{1}{2}(\alpha-1)(\sqrt{2\tau}y+x)} e^{-\frac{y^2}{2}} dy, \tag{5.43}$$

$$I_2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\frac{x}{\sqrt{2\tau}}} e^{\frac{1}{2}(\alpha+1)(\sqrt{2\tau}y+x)} e^{-\frac{y^2}{2}} dy. \tag{5.44}$$

Reescribiendo (5.43) tenemos que:

$$I_1 = \frac{1}{\sqrt{2\pi}} e^{\frac{1}{2}(\alpha-1)x + \frac{1}{4}(\alpha-1)^2\tau} \int_{-\infty}^{-\frac{x}{\sqrt{2\tau}}} e^{-\frac{1}{2}(y - \frac{1}{2}(\alpha-1)\sqrt{2\tau})^2} dy.$$

Utilizando el cambio de variable $z = y - \frac{1}{2}(\alpha - 1)\sqrt{2\tau}$ tenemos que la última igualdad se transforma en

$$I_1 = \frac{1}{\sqrt{2\pi}} e^{\frac{1}{2}(\alpha-1)x + \frac{1}{4}(\alpha-1)^2\tau} \int_{-\infty}^{-d_2} e^{-\frac{z^2}{2}} dz.$$

Teniendo en cuenta que $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-d_2} \exp\{-z^2/2\} dz$ es la función de distribución de probabilidad acumulada de una normal estándar, entonces tenemos finalmente que:

$$I_1 = e^{\frac{1}{2}(\frac{2r}{\sigma^2}-1)x + \frac{1}{4}(\frac{2r}{\sigma^2}-1)^2\tau} \Phi(-d_2),$$

donde d_2 se define como en (5.30).

Análogamente transformamos (5.44) con lo que se tiene que

$$I_2 = e^{\frac{1}{2}(\frac{2r}{\sigma^2}+1)x + \frac{1}{4}(\frac{2r}{\sigma^2}+1)^2\tau} \Phi(-d_1),$$

con d_1 definido como se hizo en (5.29).

De esta manera reescribimos la solución (5.42) como:

$$U(\tau, x) = e^{\frac{1}{2}(\frac{2r}{\sigma^2}-1)x + \frac{1}{4}(\frac{2r}{\sigma^2}-1)^2\tau} \Phi(-d_2) - e^{\frac{1}{2}(\frac{2r}{\sigma^2}+1)x + \frac{1}{4}(\frac{2r}{\sigma^2}+1)^2\tau} \Phi(-d_1). \quad (5.45)$$

Despejando $V(\tau, x)$ de la transformación dada por (5.19), utilizando los valores de m y l obtenidos en la sección 5.3.1 y sustituyendo (5.45), se tiene que

$$V(\tau, x) = e^{-\frac{2r}{\sigma^2}\tau} \Phi(-d_2) - e^x \Phi(-d_1). \quad (5.46)$$

Por último, utilizando las ecuaciones (5.35)-(5.37) y (5.46) se tiene que la solución de la ecuación diferencial parcial (5.34) esta dada por:

$$P(t, S) = K e^{-r(T-t)} \Phi(-d_2) - S \Phi(-d_1), \quad (5.47)$$

con

$$d_1 = \frac{\ln\left(\frac{S}{K}\right) + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{(T-t)}} \quad \text{y} \quad d_2 = \frac{\ln\left(\frac{S}{K}\right) + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{(T-t)}}.$$

La ecuación (5.47) es conocida como la *fórmula de Black-Scholes para valuación de opciones de venta*.

Capítulo 6

Métodos en Diferencias Finitas

Para Solucionar Problemas Matemáticos en Matemáticas Financieras.

6.1. Ecuaciones diferenciales parabólicas.

La ecuación diferencial parabólica que estudiaremos es la de calor o difusión,

$$u_1(t, x) = \alpha^2 u_{22}(t, x), \quad 0 < x < l, \quad t > 0, \quad (6.1)$$

sujeta a las condiciones

$$u(t, 0) = u(t, l) = 0, \quad t > 0,$$

y

$$u(0, x) = f(x), \quad 0 \leq x \leq l.$$

El método que usamos para aproximar la solución de este problema contiene diferencias finitas.

Primero seleccionamos un entero $m > 0$ y sea $h = l/m$. Después seleccionamos un tamaño de paso de tiempo k . Los puntos de red para este caso son (t_j, x_i) , donde $t_j = jk$, para $j = 0, 1, \dots$ y $x_i = ih$ para $i = 0, 1, \dots, m$.

El método de diferencias finitas se obtiene al usar la serie de Taylor en t para formar el cociente de diferencias

$$\frac{\partial u}{\partial t}(t_j, x_i) = \frac{u(t_j + k, x_i) - u(t_j, x_i)}{k} - \frac{k}{2} \frac{\partial^2 u}{\partial t^2}(\mu_j, x_i), \quad (6.2)$$

para alguna $\mu_j \in (t_j, t_{j+1})$, y la serie de Taylor en x para formar el cociente de diferencias

$$\frac{\partial^2 u}{\partial x^2}(t_j, x_i) = \frac{u(t_j, x_i + h) - 2u(t_j, x_i) + u(t_j, x_i - h)}{h^2} - \frac{h^2}{12} \frac{\partial^4 u}{\partial x^4}(t_j, \xi_i), \quad (6.3)$$

donde $\xi_i \in (x_{i-1}, x_{i+1})$.

La ecuación diferencial parcial parabólica (6.1) implica que en los puntos de red interiores (t_j, x_i) para toda $i = 1, 2, \dots, m-1$ y $j = 1, 2, \dots$, tendremos

$$\frac{\partial u}{\partial t}(t_j, x_i) - \alpha^2 \frac{\partial^2 u}{\partial x^2}(t_j, x_i) = 0,$$

así que el método que utiliza los cocientes de diferencias (6.2) y (6.3) es

$$\frac{w_{j+1,i} - w_{ji}}{k} - \alpha^2 \frac{w_{j,i+1} - 2w_{ji} + w_{j,i-1}}{h^2} = 0, \quad (6.4)$$

donde w_{ji} aproxima a $u(t_j, x_i)$.

El error local de truncamiento para esta ecuación en diferencias es

$$\tau_{ji} = \frac{k}{2} \frac{\partial^2 u}{\partial t^2}(\mu_j, x_i) - \alpha^2 \frac{h^2}{12} \frac{\partial^4 u}{\partial x^4}(t_j, \xi_i). \quad (6.5)$$

Al resolver la ecuación (6.4) para $w_{j+1,i}$, obtenemos

$$w_{j+1,i} = \left(1 - \frac{2\alpha^2 k}{h^2}\right) w_{ji} + \alpha^2 \frac{k}{h^2} (w_{j,i+1} + w_{j,i-1}), \quad (6.6)$$

para toda $i = 1, 2, \dots, m-1$ y $j = 1, 2, \dots$. Dado que la condición inicial $u(0, x) = f(x)$, para toda $0 \leq x \leq l$, implica que $w_{0,i} = f(x_i)$, para toda $i = 0, 1, \dots, m$, entonces se pueden utilizar estos valores en la ecuación (6.6) para calcular el valor de $w_{1,i}$ para toda $i = 1, 2, \dots, m-1$. Las condiciones de frontera $u(t, 0) = 0$ y $u(t, l) = 0$ implican que $w_{1,0} = w_{1,m} = 0$, y por tanto, podemos determinar todos los elementos de la forma $w_{1,i}$. Si volvemos a aplicar el procedimiento una vez conocidas todas las aproximaciones $w_{1,i}$ podemos obtener en forma semejante los valores $w_{2,i}, w_{3,i}, \dots$

La naturaleza explícita del método de diferencias implica que la matriz de $(m-1) \times (m-1)$ asociada a este sistema puede escribirse en forma tridiagonal

$$A = \begin{bmatrix} (1-2\lambda) & \lambda & 0 & \cdots & 0 \\ \lambda & (1-2\lambda) & \lambda & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \lambda \\ 0 & \cdots & 0 & \lambda & (1-2\lambda) \end{bmatrix}$$

donde $\lambda = \alpha^2(k/h^2)$. Si se utiliza

$$\mathbf{w}^{(0)} = (f(x_1), f(x_2), \dots, f(x_{m-1}))^t,$$

y

$$\mathbf{w}^{(j)} = (w_{j,1}, w_{j,2}, \dots, w_{j,m-1})^t, \text{ para toda } j = 1, 2, \dots,$$

entonces la solución aproximada esta dada por

$$\mathbf{w}^{(j)} = A\mathbf{w}^{(j-1)}, \text{ para toda } j = 1, 2, \dots,$$

por tanto, $\mathbf{w}^{(j)}$ se obtiene para $\mathbf{w}^{(j-1)}$ por una matriz simple de multiplicación. A esto se le conoce con el nombre de *método de diferencias progresivas*. Si la solución a la ecuación diferencial parcial tiene cuatro derivadas parciales continuas en x y dos en t , entonces la ecuación (6.5) implica que el método es de orden $O(k+h^2)$.

Ahora, si al representar los datos iniciales se comete un error $\mathbf{e}^{(0)} = (e_1^{(0)}, e_2^{(0)}, \dots, e_{m-1}^{(0)})^t$, al tomar

$$\mathbf{w}^{(0)} = (f(x_1), f(x_2), \dots, f(x_{m-1}))^t,$$

(o si en cualquier paso la elección del paso inicial se realiza simplemente por razones de comodidad) un error de $A\mathbf{e}^{(0)}$ se propaga en $\mathbf{w}^{(1)}$, ya que

$$\mathbf{w}^{(1)} = A(\mathbf{w}^{(0)} + \mathbf{e}^{(0)}) = A\mathbf{w}^{(0)} + A\mathbf{e}^{(0)}.$$

Este proceso continúa. En el n -ésimo paso de tiempo, el error de $\mathbf{w}^{(n)}$ debido a $\mathbf{e}^{(0)}$ es $A^n\mathbf{e}^{(0)}$.

En consecuencia, el método es estable si, y sólo si con cualquier error inicial $\mathbf{e}^{(0)}$ tenemos $\|A^n\mathbf{e}^{(0)}\| \leq \|\mathbf{e}^{(0)}\|$ para toda n . Esto significa que $\|A^n\| \leq 1$, condición que requiere que el radio espectral $\rho(A^n) = (\rho(A))^n \leq 1$. Así, el método de diferencias progresivas será estable sólo si $\rho(A) \leq 1$.

Se puede demostrar que los valores característicos de A son

$$\mu_i = 1 - 4\lambda \left(\text{sen} \left(\frac{i\pi}{2m} \right) \right)^2, \text{ para cada } i = 1, 2, \dots, m-1.$$

Por tanto, la condición de estabilidad se reduce a determinar si

$$\rho(A) = \max_{1 \leq i \leq m-1} \left| 1 - 4\lambda \left(\text{sen} \frac{i\pi}{2m} \right)^2 \right| \leq 1,$$

lo cual se simplifica y se transforma en

$$0 \leq \lambda \left(\text{sen} \left(\frac{i\pi}{2m} \right) \right)^2 \leq \frac{1}{2}, \text{ para cada } i = 1, 2, \dots, m-1.$$

Se requiere que esta condición de desigualdad se conserve cuando $h \rightarrow 0$, o, en forma equivalente, cuando $m \rightarrow \infty$; por ello el hecho de que:

$$\lim_{m \rightarrow \infty} \left[\operatorname{sen} \left(\frac{(m-1)\pi}{2m} \right) \right]^2 = 1,$$

significa que la estabilidad ocurrirá sólo si $0 \leq \lambda \leq \frac{1}{2}$. Puesto que $\lambda = \alpha^2(k/h^2)$, esta desigualdad requiere elegir h y k , de modo que

$$\alpha^2 \frac{k}{h^2} \leq \frac{1}{2}.$$

Así, se llama *condicionalmente estable* al método de diferencias progresivas, y hacemos énfasis en que converge a la solución de la ecuación con la rapidez de convergencia $O(k+h^2)$, a condición de que $\alpha^2(k/h^2) \leq 1/2$ y se cumplan las condiciones requeridas de continuidad.

Para obtener un método *incondicionalmente estable*, se considera un método de diferencias implícitas que se obtiene al usar el cociente de diferencias regresivas para $\frac{\partial u}{\partial t}(t_j, x_i)$, en la forma

$$\frac{\partial u}{\partial t}(t_j, x_i) = \frac{u(t_j, x_i) - u(t_{j-1}, x_i)}{k} + \frac{k}{2} \frac{\partial^2 u}{\partial t^2}(\mu_j, x_i),$$

donde μ_j está en (t_{j-1}, t_j) . Al sustituir esta ecuación junto con la ecuación (6.3) para $u_{22}(t, x)$, en la ecuación diferencial parcial, se obtiene

$$\begin{aligned} \frac{u(t_j, x_i) - u(t_{j-1}, x_i)}{k} - \alpha^2 \frac{u(t_j, x_{i+1}) - 2u(t_j, x_i) + u(t_j, x_{i-1}))}{h^2} \\ = -\frac{k}{2} \frac{\partial^2 u}{\partial t^2}(\mu_j, x_i) - \alpha^2 \frac{h^2}{12} \frac{\partial^4 u}{\partial x^4}(t_j, \xi_i), \end{aligned}$$

para alguna $\xi_i \in (x_{i-1}, x_{i+1})$. El *método de diferencias regresivas* que resulta es

$$\frac{w_{ji} - w_{j-1,i}}{k} - \alpha^2 \frac{w_{j,i+1} - 2w_{ji} + w_{j,i-1}}{h^2} = 0, \quad (6.7)$$

para toda $i = 1, 2, \dots, m-1$, y $j = 1, 2, \dots$

El método de diferencias regresivas incluye, en un paso típico, los puntos de red

$$(t_j, x_i), (t_{j-1}, x_i), (t_j, x_{i-1}) \text{ y } (t_j, x_{i+1}).$$

Las condiciones de frontera relacionadas con el problema suministran información en los puntos circundantes de la red; por ello, no es posible utilizar procedimientos explícitos para resolver la ecuación (6.7). En el método de diferencias progresivas se utilizaron las aproximaciones en

$$(t_j, x_{i-1}), (t_j, x_i), (t_{j+1}, x_i) \text{ y } (t_j, x_{i+1}),$$

de modo que se dispuso de un método explícito para calcular las aproximaciones, que tenía como base la información proveniente de las condiciones iniciales y de frontera.

Si una vez más denotamos con λ la cantidad $\alpha^2(k/h^2)$, el método de diferencias regresivas se convierte en

$$(1 + 2\lambda)w_{ji} - \lambda w_{j,i+1} - \lambda w_{j,i-1} = w_{j-1,i},$$

para toda $i = 1, 2, \dots, m-1$ y $j = 1, 2, \dots$. Aplicando el hecho de que $w_{0,i} = f(x_i)$ para toda $i = 1, 2, \dots, m-1$ y $w_{j,m} = w_{j,0} = 0$ para toda $j = 1, 2, \dots$, este método de diferencias tiene la representación matricial:

$$\begin{bmatrix} (1+2\lambda) & -\lambda & 0 & \cdots & 0 \\ -\lambda & (1+2\lambda) & -\lambda & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -\lambda \\ 0 & \cdots & 0 & -\lambda & (1+2\lambda) \end{bmatrix} \begin{bmatrix} w_{j,1} \\ w_{j,2} \\ \vdots \\ w_{j,m-1} \end{bmatrix} = \begin{bmatrix} w_{j-1,1} \\ w_{j-1,2} \\ \vdots \\ w_{j-1,m-1} \end{bmatrix} \quad (6.8)$$

o $A\mathbf{w}^{(j)} = \mathbf{w}^{(j-1)}$ para toda $j = 1, 2, \dots$.

Así, debemos resolver un sistema lineal para obtener $\mathbf{w}^{(j)}$ a partir de $\mathbf{w}^{(j-1)}$. Dado que $\lambda > 0$, la matriz A es definida positiva y estrictamente dominante en forma diagonal, además de ser tridiagonal. Para resolver el sistema, se puede emplear la *factorización de Crout* para sistemas tridiagonales.

El método de diferencias regresivas no plantea los problemas de estabilidad del método de diferencias progresivas. Esto lo comprobamos al analizar los valores característicos de la matriz. En el método de diferencias regresivas los valores característicos son

$$\mu_i = 1 + 4\lambda \left[\text{sen} \left(\frac{i\pi}{2m} \right) \right]^2, \quad \text{para cada } i = 1, 2, \dots, m-1,$$

y como $\lambda > 0$, tendremos que $\mu_i > 1$ para toda $i = 1, 2, \dots, m-1$. Esto implica, que existe A^{-1} porque cero no es un valor característico de A . Un error $\mathbf{e}^{(0)}$ en los datos iniciales genera un error $(A^{-1})^n \mathbf{e}^{(0)}$ en el n -ésimo paso. Y como los valores característicos de A^{-1} son los recíprocos de los valores característicos de A , el radio espectral de A^{-1} está acotado superiormente por 1 y el método es estable, independientemente de la elección de $\lambda = \alpha^2(k/h^2)$. De esta manera se llama *incondicionalmente estable* al método de diferencias regresivas. El error de truncamiento de esta técnica es del orden de $O(k + h^2)$, siempre y cuando la solución de la ecuación diferencial satisfaga las condiciones normales

de diferenciabilidad. En este caso, el método converge a la solución de la ecuación diferencial parcial con la misma rapidez.

La debilidad del método de diferencias regresivas radica en el hecho de que el error local de truncamiento tiene una parte con orden $O(k)$, la cual requiere hacer mucho más pequeños los intervalos de tiempo, que los de espacio. Sin duda, sería conveniente tener un procedimiento cuyo error local de truncamiento fuese de $O(k^2 + h^2)$. El primer paso en esta dirección consiste en emplear una ecuación de diferencias que tenga un error de $O(k^2)$ para $u_1(t, x)$ en vez de las que hemos usado antes, cuyo error fue de $O(k)$. Esto podemos hacerlo utilizando la serie de Taylor en t para la función $u(t, x)$ en el punto (t_j, x_i) y evaluando después en (t_{j+1}, x_i) y en (t_{j-1}, x_i) para obtener la fórmula de diferencias centrales

$$\frac{\partial u}{\partial t}(t_j, x_i) = \frac{u(t_{j+1}, x_i) - u(t_{j-1}, x_i)}{2k} + \frac{k^2}{6} \frac{\partial^3 u}{\partial t^3}(\mu_j, x_i),$$

donde $\mu_j \in (t_{j-1}, t_{j+1})$. El método de diferencias que resulta al sustituir esto y el cociente común de diferencias para $(u_{22}(t, x))$, ecuación (6.3), en la ecuación diferencial, recibe el nombre de *método de Richardson* y está dado por

$$\frac{w_{j+1,i} - w_{j-1,i}}{2k} - \alpha^2 \frac{w_{j,i+1} - 2w_{j,i} + w_{j,i-1}}{h^2} = 0. \quad (6.9)$$

El método de Richardson tiene un error local de truncamiento del orden $O(k^2 + h^2)$, pero lamentablemente también presenta serios problemas de estabilidad.

Un método más eficaz se deriva al promediar el método de diferencias progresivas en el j -ésimo paso en t ,

$$\frac{w_{j+1,i} - w_{j,i}}{k} - \alpha^2 \frac{w_{j,i+1} - 2w_{j,i} + w_{j,i-1}}{h^2} = 0,$$

que tiene error local de truncamiento

$$\tau_F = \frac{k}{2} \frac{\partial^2 u}{\partial t^2}(\mu_j, x_i) + O(h^2),$$

y el método en el $(j + 1)$ -ésimo paso en t ,

$$\frac{w_{j+1,i} - w_{j,i}}{k} - \alpha^2 \frac{w_{j+1,i+1} - 2w_{j+1,i} + w_{j+1,i-1}}{h^2} = 0,$$

que tiene error local de truncamiento

$$\tau_B = -\frac{k}{2} \frac{\partial^2 u}{\partial t^2}(\hat{\mu}_j, x_i) + O(h^2).$$

Si suponemos que

$$\frac{\partial^2 u}{\partial t^2}(\hat{\mu}_j, x_i) \approx \frac{\partial^2 u}{\partial t^2}(\mu_j, x_i)$$

entonces el método de diferencias promediadas,

$$\frac{w_{j+1,i} - w_{j,i}}{k} - \frac{\alpha^2}{2} \left[\frac{w_{j,i+1} - 2w_{j,i} + w_{j,i-1}}{h^2} - \frac{w_{j+1,i+1} - 2w_{j+1,i} + w_{j+1,i-1}}{h^2} \right] = 0,$$

tiene un error local de truncamiento del orden $O(k^2 + h^2)$, siempre y cuando se cumplan las condiciones normales de diferenciabilidad. A esto se le llama *método de Crank-Nicolson* y está representado en la forma matricial

$$A\mathbf{w}^{(j+1)} = B\mathbf{w}^{(j)}, \quad \text{para cada } j = 1, 2, \dots \quad (6.10)$$

donde

$$\lambda = \alpha^2 \frac{k}{h^2}, \quad \mathbf{w}^{(j)} = (w_{j,1}, w_{j,2}, \dots, w_{j,m-1})^t,$$

y las matrices A y B estan dadas por:

$$A = \begin{bmatrix} (1 + \lambda) & -\frac{\lambda}{2} & 0 & \cdots & 0 \\ -\frac{\lambda}{2} & (1 + \lambda) & -\frac{\lambda}{2} & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -\frac{\lambda}{2} \\ 0 & \cdots & 0 & -\frac{\lambda}{2} & (1 + \lambda) \end{bmatrix}$$

y

$$B = \begin{bmatrix} (1 - \lambda) & \frac{\lambda}{2} & 0 & \cdots & 0 \\ \frac{\lambda}{2} & (1 - \lambda) & \frac{\lambda}{2} & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \frac{\lambda}{2} \\ 0 & \cdots & 0 & \frac{\lambda}{2} & (1 - \lambda) \end{bmatrix}.$$

Por tanto, A es no singular, porque es una matriz definida positiva, estrictamente dominante en forma diagonal y tridiagonal. Se puede usar la factorización de Crout para obtener $\mathbf{w}^{(j+1)}$ a partir de $\mathbf{w}^{(j)}$ para toda $j = 1, 2, \dots$. Por último, falta decir que el método de Crank-Nicolson es incondicionalmente estable y tiene orden de convergencia $O(k^2 + h^2)$.

Ejemplo 6.1. Considérese la ecuación de calor

$$u_1(t, x) - u_{22}(t, x) = 0, \quad 0 < x < l, \quad 0 \leq t,$$

con las condiciones de frontera

$$u(t, 0) = u(t, l) = 0, \quad 0 < t,$$

y la condición inicial

$$u(0, x) = \text{sen}(\pi x), \quad 0 \leq x \leq 1.$$

La solución a este problema es

$$u(t, x) = e^{-\pi^2 t} \text{sen}(\pi x).$$

La solución en $t = 0.5$ será aproximada mediante el método de Crank-Nicolson, por lo cual, tomamos el número de particiones del intervalo de x , $m = 10$ y el número de particiones del intervalo de t , $N = 50$. Fijamos el intervalo de t hasta 0.5, de esta manera, tenemos que los tamaños de paso para x y t están dados por $h = 0.1$ y $k = 0.01$ respectivamente, por tanto, $\lambda = 1$. En el Cuadro 6.1 tenemos la comparación de las aproximaciones, $w_{50,i}$, obtenidas por el método de Crank-Nicolson, las cuales son muy buenas con respecto a los valores reales de la solución $u(0.5, x_i)$, debido a que el error absoluto cometido al aproximar la solución es muy pequeño.

x_i	$w_{50,i}$	$u(0.5, x_i)$	$ w_{50,i} - u(0.5, x_i) $	error relativo
0.0	0	0		
0.1	0.00230512	0.00222241	8.271×10^{-5}	3.722×10^{-2}
0.2	0.00438461	0.00422728	1.573×10^{-4}	3.722×10^{-2}
0.3	0.00603489	0.00581836	2.165×10^{-4}	3.722×10^{-2}
0.4	0.00709444	0.00683989	2.546×10^{-5}	3.722×10^{-2}
0.5	0.00745954	0.00719188	2.677×10^{-4}	3.722×10^{-2}
0.6	0.00709444	0.00683989	2.546×10^{-4}	3.722×10^{-2}
0.7	0.00603489	0.00581836	2.165×10^{-4}	3.722×10^{-2}
0.8	0.00438461	0.00422728	1.573×10^{-4}	3.722×10^{-2}
0.9	0.00230512	0.00222241	8.271×10^{-5}	3.722×10^{-2}
1.0	0	0		

Cuadro 6.1: Tabla de comparación.

La Figura 6.1, muestra la aproximación de la solución y la solución real para los puntos en los cuales se utilizó el método de Crank-Nicolson. Las gráficas se muestran por separado debido a que por la exactitud de la aproximación no se aprecia la diferencia cuando se ponen en una misma gráfica. ■

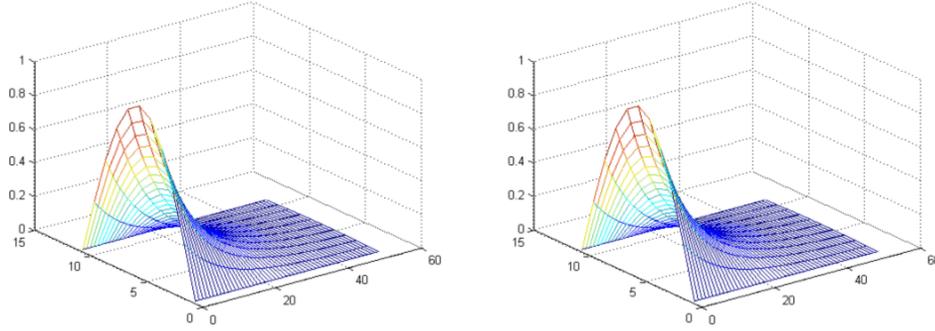


Figura 6.1: Comparación gráfica de la solución real (gráfica de la derecha) y su aproximación por el método de Crank-Nicolson (gráfica de la izquierda).

6.2. Aplicación de la fórmula de Black-Scholes y de métodos de diferencias finitas en problemas de valuación de opciones.

Consideremos el problema de calcular el valor de la opción de compra dado por la ecuación (5.11),

$$rC(t, S) = C_1(t, S) + \frac{1}{2}\sigma^2 S^2 C_{22}(t, S) + rSC_2(t, S),$$

donde S es el valor del subyacente, $\sigma > 0$ la volatilidad fija, r el tipo de interés libre constante y T es el tiempo de vencimiento o maduración de la opción.

En la sección 5.3.1 ya hemos obtenido la solución de la ecuación (5.11), la cual está dada por la fórmula de Black-Scholes (5.33). Con la fórmula de Black-Scholes podemos obtener el valor de la opción de compra para cualquier valor de $t \in [0, T]$ y $S > 0$. Ahora, utilizaremos un método de diferencias finitas, el método de Crank-Nicolson, para obtener el valor de la opción de compra.

Para poder utilizar el método de Crank-Nicolson expuesto en la sección 6.1, es necesario convertir la ecuación (5.11) en una ecuación de difusión, lo que ya se hizo en la sección 5.3.1. Es decir, consideramos la ecuación diferencial parabólica dada por la ecuación (5.24), y análoga a la ecuación (6.1) con $\alpha = 1$,

$$u_1(t, x) = u_{22}(t, x)$$

sujeta a las condiciones de frontera

$$u(t, 0) = 0,$$

$$u(t, l) = e^{\frac{1}{2}(\gamma+1)l + \frac{1}{4}(\gamma+1)^2 t} - e^{\frac{1}{2}(\gamma-1)l + \frac{1}{4}(\gamma-1)^2 t},$$

y la condición inicial dada por la ecuación (5.25)

$$u(0, x) = \text{máx}(e^{\frac{1}{2}(\gamma+1)x} - e^{\frac{1}{2}(\gamma-1)x}, 0),$$

con $\gamma = \frac{2r}{\sigma^2}$.

Dado que estamos utilizando un método numérico, es necesario fijar intervalos en los cuales se pueda hacer la aproximación de la solución de la ecuación (5.11), y como la variable S no está acotada superiormente es necesario tomar un intervalo adecuado para esta. Considerando el cambio de variable hecho en la sección 5.3.1 para τ y x dados por las ecuaciones (5.12) y (5.13) respectivamente, así, tenemos que los intervalos para aproximar la función quedan como, $t \in [0, \frac{\sigma^2}{2}T]$ y $x \in [\ln(S_0/2K), \ln(2S_0/K)]$, donde S_0 es el valor inicial del subyacente.

En adición a esto, es necesario hacer otra observación. Tomando en cuenta las condiciones de frontera, tenemos que $u(l, t) \neq 0$ entonces $w_{j,m} \neq 0$ para $j = 1, 2, \dots, N$, por lo tanto, el método se tiene que cambiar un poco para tomar en cuenta estos puntos. De esta manera el método en forma matricial dado por la ecuación (6.10), se transforma en:

$$A\mathbf{w}^{(j+1)} = B\mathbf{w}^{(j)} + \mathbf{b} - \mathbf{a}, \quad \text{para cada } j = 1, 2, \dots, N, \quad (6.11)$$

donde $A, B, \mathbf{w}^{(j)}, \lambda$ se definen como antes y los vectores \mathbf{a} y \mathbf{b} de tamaño $(m-1) \times 1$, están definidos por:

$$\mathbf{a} = \begin{bmatrix} 0 \\ \vdots \\ -\frac{\lambda}{2}w_{j+1,m} \end{bmatrix} \quad \text{y} \quad \mathbf{b} = \begin{bmatrix} 0 \\ \vdots \\ \frac{\lambda}{2}w_{j,m} \end{bmatrix}.$$

Con el objetivo de comparar que tan buena es la aproximación que se obtiene utilizando el método de Crank-Nicolson, utilizaremos los siguientes datos. El valor inicial del subyacente es $S_0 = 21$, el precio de ejercicio de la opción $K = 20$, la volatilidad $\sigma = 0.317$, el tipo de interés libre $r = 0.05$ y el tiempo de vencimiento o maduración de la opción $T = 0.25$ (3 meses).

Para hacer esto, se dividió el intervalo de la variable S con $m = 30$ particiones; análogamente para t , se dividió el intervalo con $N = 80$ particiones, de este modo, los tamaños de paso están dados por $h = 0.0462$ y $k = 1.5701 \times 10^{-4}$, consiguiendo así que λ sea igual a 0.0735, aunque no importa mucho debido a que el método es incondicionalmente estable.

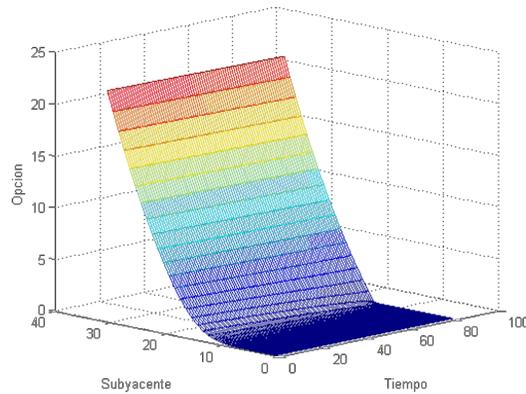


Figura 6.2: Gráfica de la valuación de la opción de compra mediante la fórmula de Black-Scholes.

Evaluando en la fórmula de Black-Scholes los mismos puntos para los que se hicieron las aproximaciones de la solución, se obtiene la gráfica que se muestra en la Figura 6.2.

La comparación gráfica es mostrada en la Figura 6.3, para ello hubo de utilizar los cambios de variable dados por las ecuaciones (5.14) y (5.19), ya que, lo que se aproxima es la solución de la ecuación de difusión.

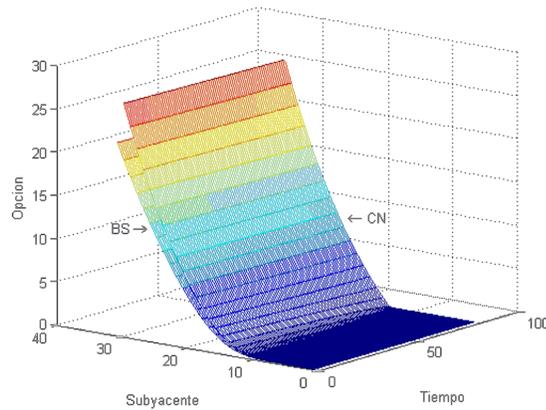


Figura 6.3: Comparación gráfica de la valuación de la opción de compra mediante la fórmula de Black-Scholes y la aproximación obtenida por el método de Crank-Nicolson.

En la Figura 6.3, se nota que la gráfica de la aproximación hecha está ligeramente por arriba de la gráfica obtenida utilizando la fórmula de Black-Scholes,

esto se debe en parte a los valores tomados para S y la condición de frontera $u(t, l)$.

Otro punto de comparación entre la fórmula de Black-Scholes y la aproximación hecha mediante el método de Crank-Nicolson es el valor inicial o el precio racional al tiempo $t = 0$ que se debe pagar por la adquisición de la opción de compra. Este valor se obtiene evaluando la fórmula de Black-Scholes con $t = 0$ y S_0 , es decir, $C_0 = C(0, S_0) = S\Phi(d_1) - Ke^{-rT}\Phi(d_2)$, obteniendo para nuestros valores particulares que, $C_0 = 2.0061$. El valor obtenido mediante la aproximación fue $C_0 = 2.0316$, obteniendo un error absoluto de 2.55×10^{-2} y el error relativo de 1.27×10^{-2} , siendo este no muy considerable y mostrando que el problema de valuación de una opción de compra puede ser muy bien aproximado utilizando métodos de diferencias finitas.

Capítulo 7

Nuevas Soluciones de la Ecuación de Black-Scholes.

7.1. Simetría de la ecuación de Black-Scholes y leyes de conservación.

La idea básica del enfoque de Sukhomlin consiste en la búsqueda y en el estudio de ciertas relaciones entre el valor de la opción $V(t, x)$ y la rapidez de su cambio relacionado con el precio. El hecho de la existencia de las relaciones que se conservan durante todo el proceso modelado, lleva a información práctica muy importante. Por ejemplo, esto permite evaluar la volatilidad, o sugiere la formulación apropiada de los problemas de frontera.

La simetría de la ecuación de difusión es bien conocida. Dado que la ecuación de Black-Scholes se puede reducir a una ecuación de difusión (véase sección 5.3.1), su simetría es similar. Sin hacer los cálculos pasamos directamente a los resultados.

El conjunto de los operadores de simetría de la ecuación (5.10) diferenciales lineales de primer orden y que conmutan con el operador A , definido como

$$A \equiv \frac{\partial}{\partial t} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2}{\partial x^2} + rx \frac{\partial}{\partial x} - r,$$

constituye un espacio vectorial con la base:

$$B_1 = x \frac{\partial}{\partial x} - \beta, \quad (7.1)$$

$$B_2 = \sigma^2 tx \frac{\partial}{\partial x} - \ln x - \sigma^2 t \beta. \quad (7.2)$$

y también un operador de simetría trivial de orden cero: $B_3 = 1$. Obsérvese que:

$$B_2 = \sigma^2 t B_1 - \ln x, \quad (7.3)$$

y que se verifican las condiciones de conmutación:

$$[B_2, B_1] = B_3, \quad [B_2, B_3] = [B_1, B_3] = 0. \quad (7.4)$$

La simetría mencionada es intrínseca de la ecuación de Black-Scholes (5.10). Es fácil verificar que el operador (7.1) define totalmente al operador A:

$$A = \sigma^2 \left\{ \frac{1}{\sigma^2} \frac{\partial}{\partial t} + \frac{1}{2} B_1^2 - \frac{1}{2} (\beta - 1)^2 \right\}. \quad (7.5)$$

Visto que $[A, B_i] = 0$ ($i = 1, 2, 3$), es obvio que toda función de los operadores B_1, B_2, B_3 conmutará con A y en consecuencia, será un operador de simetría de la ecuación (5.10). Este hecho permitirá en las secciones 7.3 y 7.4, clasificar los operadores de simetría y construir las nuevas soluciones exactas de la ecuación de Black-Scholes usando el método de separación de variables en el enfoque de V. Shapovalov.

Se hallan las soluciones de la ecuación (5.10) que son las funciones propias del operador de simetría (7.1), es decir se resuelve el sistema:

$$AV(t, x) = 0, \quad (7.6)$$

$$BV(t, x) = \lambda V(t, x). \quad (7.7)$$

con $B = B_1$ en este caso. Después de algunos cálculos se encuentra una familia de soluciones parametrizadas por λ :

$$V_\lambda(t, x) = Cx^{\lambda+\beta} \exp \left\{ \frac{\sigma^2}{2} [(\beta - 1)^2 - \lambda^2] t \right\}, \quad C = \text{const.} \quad (7.8)$$

Esta familia de soluciones corresponde a la ley de conservación que se expresa por la relación (7.7): elasticidad-precio del valor de la opción que se conserva en este caso:

$$E_p \equiv \frac{x}{V} \frac{\partial V}{\partial x} = (\lambda + \beta) = \text{const.} \quad (7.9)$$

Este hecho significa que al elegir la solución (7.8), se elige la relación (7.9) entre el valor V y la rapidez de su variación. La existencia de esta ley de conservación permite experimentalmente evaluar la potencia del precio en la función (7.8): $\lambda + \beta$.

El segundo operador de simetría de primer orden B_2 define otra ley de conservación: la relación (7.7) informa que la misma elasticidad ahora no se conserva, sino depende linealmente del logaritmo del precio y es inversamente proporcional al tiempo. En consecuencia se conserva la expresión:

$$\sigma^2 t \left[\frac{x}{V} \frac{\partial V}{\partial x} - \beta \right] - \ln x = \lambda = \text{const.} \quad (7.10)$$

Esta igualdad se construye usando la fórmula (7.3). También es posible interpretar esta ley como conservación del valor inicial del precio x_0 : para cuando

el tiempo $t \rightarrow 0$ el operador B_2 se reduce a la expresión: $-(\ln x_0)$, entonces $x_0 = \exp\{-\lambda\}$. Ahora la relación (7.10) se puede escribir:

$$E_p \equiv \frac{x}{V} \frac{\partial V}{\partial x} = \frac{1}{\sigma^2 t} \ln \left(\frac{x}{x_0} \right) + \beta. \quad (7.11)$$

La resolución del sistema de tipo (7.6)-(7.7) con el operador B_2 , ecuación (7.2), lleva al siguiente resultado:

$$V_\lambda(t, x) = \frac{C}{\sqrt{t}} \exp \left\{ \frac{[\ln x + \sigma^2 \beta t + \lambda]^2}{2\sigma^2 t} \right\}, \quad C = \text{const.} \quad (7.12)$$

Si $r = \frac{\sigma^2}{2}$ entonces la solución (7.12) queda:

$$V_\lambda(t, x) = \frac{C}{\sqrt{t}} \left(\frac{x}{x_0} \right)^{\frac{1}{2\sigma^2 t} \ln \left(\frac{x}{x_0} \right)}.$$

Se puede constatar que la función (7.12) es cóncava con la condición:

$$\ln \left(\frac{x}{x_0} \right) - r t^2 + 1 < \left(\frac{\sigma^2 t}{2} - 1 \right),$$

lo que se verifica a partir del momento de tiempo $t = \frac{4}{\sigma^2}$.

7.2. Grupo de equivalencia de la ecuación de Black-Scholes.

Sea la ecuación de Black-Scholes (5.10). Estudiamos las transformaciones de variables:

$$\tau = \tau(t), \quad \xi = \xi(t, x), \quad V(t, x) = a(t, x)Q(\tau, \xi), \quad (7.13)$$

tales que la ecuación (5.10) no cambia su estructura, salvo que aparece un factor multiplicativo a la izquierda:

$$f(t, x) \left\{ Q_1(\tau, \xi) + \frac{1}{2} \sigma_0^2 \xi^2 Q_{22}(\tau, \xi) + r_0 \xi Q_2(\tau, \xi) - r_0 Q(\tau, \xi) \right\} = 0, \quad (7.14)$$

donde Q es la nueva incógnita.

Proposición 7.1. *Sea la ecuación de Black-Scholes (5.10). La transformación siguiente:*

$$\tau = \frac{\alpha^2 \sigma^2}{\sigma_0^2} (t - t_0), \quad \xi = \frac{x}{x_0}, \quad (7.15)$$

$$V(t, x) = Q(\tau, \xi) \exp \left\{ (\beta - \alpha\beta_0) \ln x + \frac{\sigma^2}{2} [(\beta - 1)^2 - \alpha^2(\beta_0 - 1)^2] t \right\}, \quad (7.16)$$

convierte la ecuación (5.10) en la siguiente ecuación

$$f(t, x) \left\{ Q_1(\tau, \xi) + \frac{\sigma_0^2}{2\alpha^2} \xi^2 Q_{22}(\tau, \xi) + \frac{\sigma_0^2}{\alpha^2} \left(\frac{1}{2} - \alpha\beta_0 \right) Q_2(\tau, \xi) - r_0 Q(\tau, \xi) \right\} = 0 \quad (7.17)$$

con el factor exterior:

$$f(t, x) = \frac{\alpha^2 \sigma^2}{\sigma_0^2} \exp \left\{ (\beta - \alpha\beta_0) \ln x + \frac{\sigma^2}{2} [(\beta - 1)^2 - \alpha^2(\beta_0 - 1)^2] t \right\},$$

para cualesquiera valores de las constantes: α , t_0 , x_0 ($\alpha \neq 0$, $x_0 \neq 0$). Aquí las constantes β y β_0 están definidas por $\beta \equiv \frac{1}{2} - \frac{r}{\sigma^2}$ y $\beta_0 \equiv \frac{1}{2} - \frac{r_0}{\sigma_0^2}$. Si tomamos $\alpha = 1$, entonces la ecuación (7.17) se convierte en una ecuación tipo (7.14).

Demostración. De (7.15) y (7.16) se tiene que:

$$t = \frac{\sigma_0^2}{\alpha^2 \sigma^2} \tau + t_0, \quad x = x_0 \xi, \quad Q(\tau, \xi) = V(t, x) e^{-\delta},$$

con

$$\delta = (\beta - \alpha\beta_0) \ln x + \frac{\sigma^2}{2} [(\beta - 1)^2 - \alpha^2(\beta_0 - 1)^2] t.$$

Dado este cambio de variable, obtenemos las siguientes ecuaciones diferenciales parciales

$$Q_1(\tau, \xi) = e^{-\delta} \frac{\sigma_0^2}{\alpha^2 \sigma^2} \left\{ V_1(t, x) - \frac{\sigma^2}{2} [(\beta - 1)^2 - \alpha^2(\beta_0 - 1)^2] V(t, x) \right\}; \quad (7.18)$$

$$Q_2(\tau, \xi) = e^{-\delta} \left\{ x_0 V_2(t, x) - \frac{(\beta - \alpha\beta_0)}{\xi} V(t, x) \right\}; \quad (7.19)$$

$$Q_{22}(\tau, \xi) = e^{-\delta} \left\{ x_0^2 V_{22}(t, x) - 2(\beta - \alpha\beta_0) \frac{x_0}{\xi} V_2(t, x) + [(\beta - \alpha\beta_0) + 1] \frac{(\beta - \alpha\beta_0)}{\xi^2} V(t, x) \right\}. \quad (7.20)$$

Luego, de (7.18)-(7.20), se tiene

$$V_1(t, x) = \frac{\alpha^2 \sigma^2}{\sigma_0^2} e^{\delta} \left\{ Q_1(\tau, \xi) + \frac{\sigma_0^2}{2\alpha^2} [(\beta - 1)^2 - \alpha^2(\beta_0 - 1)^2] Q(\tau, \xi) \right\}; \quad (7.21)$$

$$V_2(t, x) = \frac{\alpha^2 \sigma^2 e^\delta}{\sigma_0^2 x_0} \left\{ \frac{\sigma_0^2}{\alpha^2 \sigma^2} Q_2(\tau, \xi) + \frac{\sigma_0^2 (\beta - \alpha \beta_0)}{\alpha^2 \sigma^2 \xi} Q(\tau, \xi) \right\}; \quad (7.22)$$

$$V_{22}(\tau, \xi) = \frac{\alpha^2 \sigma^2 e^\delta}{\sigma_0^2 x_0^2} \left\{ \frac{\sigma_0^2}{\alpha^2 \sigma^2} Q_{22}(\tau, \xi) + \frac{2\sigma_0^2}{\alpha^2 \sigma^2} (\beta - \alpha \beta_0) \frac{1}{\xi} Q_2(\tau, \xi) \right. \\ \left. + \frac{\sigma_0^2}{\alpha^2 \sigma^2} [(\beta - \alpha \beta_0) - 1] \frac{(\beta - \alpha \beta_0)}{\xi^2} Q(\tau, \xi) \right\}. \quad (7.23)$$

Sustituyendo (7.21)-(7.23) en la ecuación de Black-Scholes (5.10), agrupando términos y simplificando, obtenemos finalmente la ecuación (7.17):

$$\frac{\alpha^2 \sigma^2}{\sigma_0^2} e^\delta \left\{ Q_1(\tau, \xi) + \frac{\sigma_0^2}{2\alpha^2} \xi^2 Q_{22}(\tau, \xi) + \frac{\sigma_0^2}{\alpha^2} \left(\frac{1}{2} - \alpha \beta_0 \right) \xi Q_2(\tau, \xi) - r_0 Q(\tau, \xi) \right\} = 0.$$

□

Las transformaciones (7.15) y (7.16) constituyen un grupo continuo con tres parámetros: la constante α corresponde a un cambio en la escala del tiempo con la variación simultánea de la coordenada x ; la constante t_0 representa la libertad de la elección del origen del tiempo y la constante x_0 esta ligada con la libertad de elección del precio inicial (notamos la imposición: $x_0 \neq 0$).

Llamamos este grupo galileano G por analogía con el grupo correspondiente en la teoría de difusión. El grupo galileano G crea una relación de equivalencia sobre el conjunto de las soluciones de la ecuación (5.10). Además de este grupo, existe una transformación especial que tampoco cambia la estructura de la ecuación (5.10).

Proposición 7.2. *Sea la ecuación de Black-Scholes (5.10). La transformación siguiente:*

$$\tau = -\frac{1}{\sigma_0^2 \sigma^2} \frac{1}{t}, \quad \xi = \exp \left\{ -\frac{\ln x}{\sigma^2 t} \right\}, \quad (7.24)$$

$$V(t, x) = \frac{Q(\tau, \xi)}{\sqrt{t}} \exp \left\{ \frac{(\ln x + \beta_0)^2}{2\sigma^2 t} + \beta \ln x + \frac{\sigma^2 (\beta - 1)^2}{2} t + \frac{r_0}{\sigma_0^2 \sigma^2} \frac{1}{t} \right\} \quad (7.25)$$

convierte la ecuación (5.10) en la ecuación tipo (7.14) con el factor exterior:

$$f(t, x) = \frac{1}{\sigma_0^2 \sigma^2 t^2 \sqrt{t}} \exp \left\{ \frac{(\ln x + \beta_0)^2}{2\sigma^2 t} + \beta \ln x + \frac{\sigma^2 (\beta - 1)^2}{2} t + \frac{r_0}{\sigma_0^2 \sigma^2} \frac{1}{t} \right\}.$$

Demostración. De (7.24) y (7.25) tenemos que:

$$t = -\frac{1}{\sigma_0^2 \sigma^2} \frac{1}{\tau}, \quad x = e^{-\sigma^2 t \ln \xi}, \quad Q(\tau, \xi) = V(t, x) \sqrt{t} e^{-\delta},$$

con

$$\delta = \frac{(\ln x + \beta_0)^2}{2\sigma^2 t} + \beta \ln x + \frac{\sigma^2(\beta - 1)^2}{2} t + \frac{r_0}{\sigma_0^2 \sigma^2} \frac{1}{t}.$$

De esta manera se obtienen las siguientes ecuaciones diferenciales parciales

$$Q_1(\tau, \xi) = \frac{\sqrt{t} e^{-\delta}}{\sigma_0^2 \sigma^2 \tau^2} \left\{ V_1(t, x) + \sigma^2 \ln \xi x V_2(t, x) - \left[\frac{1}{2t} + \frac{(\ln x + \beta_0)^2}{2\sigma^2 t^2} + \frac{\ln x + \beta_0}{t} \ln \xi + \sigma^2 \beta \ln \xi - \frac{\sigma^2(\beta - 1)^2}{2} + \frac{r_0}{\sigma_0^2 \sigma^2 t^2} \right] V(t, x) \right\}; \quad (7.26)$$

$$Q_2(\tau, \xi) = \frac{-\sigma^2 t \sqrt{t} e^{-\delta}}{\xi} \left\{ x V_2(t, x) - \left[\frac{\ln x + \beta_0}{\sigma^2 t} + \beta \right] V(t, x) \right\}; \quad (7.27)$$

$$Q_{22}(\tau, \xi) = \frac{\sigma^4 t^2 \sqrt{t} e^{-\delta}}{\xi^2} \left\{ x^2 V_{22}(t, x) + \left[1 + \frac{1}{\sigma^2 t} - 2 \left(\frac{\ln x + \beta_0}{\sigma^2 t} + \beta \right) \right] x V_2(t, x) + \left[\left(\frac{\ln x + \beta_0}{\sigma^2 t} + \beta \right)^2 - \left(\frac{\ln x + \beta_0}{\sigma^4 t^2} + \frac{\beta}{\sigma^2 t} \right) - \frac{1}{\sigma^2 t} \right] V(t, x) \right\}. \quad (7.28)$$

Así, de (7.26)-(7.28), se tiene

$$V_1(t, x) = \frac{e^\delta}{\sqrt{t}} \left\{ \frac{1}{\sigma_0^2 \sigma^2 t^2} Q_1(\tau, \xi) - \frac{\ln \xi}{t} \xi Q_2(\tau, \xi) - \left[\frac{1}{2t} + \frac{(\ln x + \beta_0)^2}{2\sigma^2 t^2} - \frac{\sigma^2(\beta - 1)^2}{2} + \frac{r_0}{\sigma_0^2 \sigma^2 t^2} \right] Q(\tau, \xi) \right\}; \quad (7.29)$$

$$x V_2(t, x) = \frac{e^\delta}{\sqrt{t}} \left\{ -\frac{1}{\sigma^2 t} \xi Q_2(\tau, \xi) + \left[\frac{\ln x + \beta_0}{\sigma^2 t} + \beta \right] Q(\tau, \xi) \right\}; \quad (7.30)$$

$$x^2 V_{22}(t, x) = \frac{e^\delta}{\sqrt{t}} \left\{ \frac{1}{\sigma^4 t^2} \xi^2 Q_{22}(\tau, \xi) + \frac{1}{\sigma^2 t} \left[1 + \frac{1}{\sigma^2 t} - 2 \left(\frac{\ln x + \beta_0}{\sigma^2 t} + \beta \right) \right] \xi Q_2(\tau, \xi) - \left[\left(\frac{\ln x + \beta_0}{\sigma^2 t} + \beta \right) - \left(\frac{\ln x + \beta_0}{\sigma^2 t} + \beta \right)^2 - \frac{1}{\sigma^2 t} \right] Q(\tau, \xi) \right\}. \quad (7.31)$$

Sustituyendo (7.29)-(7.31) en la ecuación de Black-Scholes (5.10), agrupando términos y simplificando se tiene finalmente:

$$\frac{e^\delta}{\sigma_0^2 \sigma^2 t^2 \sqrt{t}} \left\{ Q_1(\tau, \xi) + \frac{1}{2} \sigma_0^2 \xi^2 Q_{22}(\tau, \xi) + r_0 \xi Q_2(\tau, \xi) - r_0 Q(\tau, \xi) \right\} = 0. \quad (7.32)$$

□

En realidad las constantes en la fórmula (7.24) se pueden eliminar por una transformación del grupo galileano G : por ejemplo, usando las transformaciones (7.15)-(7.16), la ecuación (5.10) se puede reducir a otra de la misma estructura con $\sigma_0 = 1$ y la constante r_0 puede ser igual a cero o una constante cualquiera. En este caso la transformación (7.24) se presenta por:

$$x^I = \exp \left\{ -\frac{\ln x}{t} \right\}, \quad t^I = -\frac{1}{t}. \quad (7.33)$$

La denotamos ν . Estudiamos ahora las potencias de ν :

ν^2 corresponde a $x^{II} = \exp \left\{ -\frac{\ln x^I}{t^I} \right\} = \frac{1}{x}$, $t^{II} = -\frac{1}{t^I} = t$.

ν^3 corresponde a $x^{III} = \exp \left\{ \frac{\ln x}{t} \right\}$, $t^{III} = -\frac{1}{t}$.

ν^4 corresponde a $x^{IV} = x$, $t^{IV} = t$: transformación idéntica i .

Se concluye que el conjunto $\{\nu, \nu^2, \nu^3, \nu^4 = i\}$ constituye un grupo discreto que denotamos N .

Teorema 7.1. *Sea la ecuación de Black-Scholes (5.10) con las constantes σ ($\sigma \neq 0$), r dadas. El grupo de equivalencia más amplio que admite la ecuación (5.10) es $\Gamma = G \otimes N$ donde G es el grupo galileano y N es el grupo discreto definidos arriba por las fórmulas (7.14)-(7.15) y (7.24)-(7.25).*

7.3. Clasificación de los operadores de simetría de la ecuación de Black-Scholes hasta el tercer orden.

Teorema 7.2. *Sea la ecuación de Black-Scholes (5.10) con el operador A en la forma (7.5). Todos los operadores lineales diferenciales de hasta tercer orden que conmutan con A entran en una de las siguientes 5 clases de equivalencia y en 8 agrupaciones de clases de equivalencia:*

1. *Todos los operadores de primer orden constituyen una sola clase de equivalencia cuyo representante más sencillo es B_1 (ó B_2) definido por la fórmula (7.1).*
2. *Todos los operadores de segundo orden constituyen cinco clases de equivalencia cuyos representantes más sencillos son:*

a) $B_2^2 + B_1^2$.

b) $B_2^2 - B_1^2$.

- c) $B_1B_2 + B_2B_1$.
- d) $B_1^2 + B_2$.
- e) B_1^2 .

3. Todos los operadores de tercer orden constituyen ocho agrupaciones de clases de equivalencia parametrizadas por tres o cuatro constantes arbitrarias ω , μ , ν , η y cuyos representantes más sencillos son:

- a) $B_2^3 + (B_2B_1^2 + B_1^2B_2) + \omega B_1^3 + \mu B_1^2 + \nu B_2 + \eta B_1$.
- b) $B_2^3 - (B_2B_1^2 + B_1^2B_2) + \omega B_1^3 + \mu B_1^2 + \nu B_2 + \eta B_1$.
- c) $B_2^3 + B_1^3 + \mu B_1^2 + \nu B_2 + \eta B_1$.
- d) $B_2^3 - B_1^3 + \mu B_1^2 + \nu B_2 + \eta B_1$.
- e) $B_2^3 + \mu B_1^2 + \nu B_2 + \eta B_1$.
- f) $(B_2^2B_1 + B_1B_2^2) + B_1^3 + \mu B_1^2 + \nu B_2 + \eta B_1$.
- g) $(B_2^2B_1 + B_1B_2^2) - B_1^3 + \mu B_1^2 + \nu B_2 + \eta B_1$.
- h) $(B_2^2B_1 + B_1B_2^2) + \mu B_1^2 + \nu B_2 + \eta B_1$.

También, cada operador de simetría corresponde a un sistema privilegiado de coordenadas que permite separar las variables en la ecuación (5.10) y resolverla exactamente.

7.4. Nuevas soluciones de la ecuación de Black-Scholes.

1. Sea la clase de equivalencia definida por el operador (2a) del teorema [7.2]: $B = B_2^2 + B_1^2 = (\sigma^4t^2 + 1)B_1^2 - 2\sigma^2t(\ln x)B_1 + \ln^2 x - \sigma^2t$.

Se puede verificar que la función siguiente cumple el sistema (7.6)-(7.7) y en consecuencia es la solución de Black-Scholes:

$$V_\lambda(t, x) = \frac{\Psi(\xi)}{(\sigma^4t^2 + 1)^{\frac{1}{4}}} \exp \left\{ \frac{\sigma^2t}{2(\sigma^4t^2 + 1)} \ln^2 x + \beta \ln x - \frac{\lambda}{2} \arctg(\sigma^2t) + \frac{\sigma^2}{2} (\beta - 1)^2t \right\}, \quad (7.34)$$

$$\xi \equiv (\sigma^4t^2 + 1)^{-\frac{1}{2}} \ln x.$$

La función $\Psi(\xi)$ verifica la ecuación diferencial ordinaria siguiente:

$$\Psi'' + \xi^2\Psi = \lambda\Psi,$$

cuyas soluciones son las funciones del cilindro parabólico. Las soluciones correspondientes a la familia de soluciones (7.34) describen en la mecánica cuántica los estados llamados coherentes.

2. Sea la clase definida por el operador (2b) del teorema [7.2]:

$$B = B_2^2 - B_1^2 = (\sigma^4 t^2 - 1)B_1^2 - 2\sigma^2 t(\ln x)B_1 + \ln^2 x - \sigma^2 t.$$

En este caso la solución de la ecuación de Black-Scholes se presenta en la forma:

$$V_\lambda(t, x) = \frac{\Psi(\xi)}{|\sigma^4 t^2 - 1|} \exp \left\{ \frac{\sigma^2 t}{2(\sigma^4 t^2 - 1)} \ln^2 x + \beta \ln x + \frac{\lambda}{2} \operatorname{arcth}(\sigma^2 t) + \frac{\sigma^2}{2} (\beta - 1)^2 t \right\}, \quad (7.35)$$

$$\xi \equiv |\sigma^4 t^2 - 1|^{-\frac{1}{2}} \ln x.$$

La función $\Psi(\xi)$ verifica la ecuación diferencial ordinaria de tipo de oscilador armónico en la mecánica:

$$\Psi'' - \xi^2 \Psi = \lambda \Psi.$$

Las soluciones que se anulan al infinito se expresan por los polinomios de Tchebyshev-Hermite $H_n(\xi)$; el espectro de los valores propios del operador mencionado es discreto:

$$\Psi_n(\xi) = C_n H_n(\xi) \exp \left\{ -\frac{\xi^2}{2} \right\}, \quad \lambda_n = -(2n + 1), \quad n = 0, 1, 2, \dots,$$

con las funciones $H_n(\xi)$ que verifican la ecuación de Hermite:

$$H_n'' - 2\xi H_n' + 2n H_n = 0.$$

Se sabe que las funciones $\Psi_n(\xi)$ representan una base ortonormal en L^2 . Las soluciones (7.35) se definen sobre un intervalo finito del tiempo y se pueden fácilmente adaptar al modelo de Black-Scholes.

3. Sea la clase de equivalencia definida por el operador (2c) del teorema [7.2]:

$$B = B_1 B_2 + B_2 B_1 = 2\sigma^2 t B_1^2 - 2(\ln x) B_1 - 1.$$

La solución de la ecuación de Black-Scholes se presenta como:

$$V_\lambda(t, x) = \Psi(\xi) \exp \left\{ \frac{1}{4\sigma^2 t} \ln^2 x + \beta \ln x - \frac{\lambda + 1}{4} \ln(\sigma^2 t) + \frac{\sigma^2}{2} (\beta - 1)^2 t \right\}, \quad (7.36)$$

$$\xi \equiv (\sigma^2 t)^{-\frac{1}{2}} \ln x,$$

$$\Psi'' - \frac{\xi^2}{4} \Psi = \frac{\lambda}{2} \Psi.$$

Las funciones $\Psi(\xi)$ de la última ecuación, como en el caso anterior, se expresan por los polinomios de Tchebyshev-Hermite H_n ; el espectro de los valores propios del operador mencionado es también discreto:

$$\Psi_n(\xi) = C_n H_n \left(\frac{\xi}{\sqrt{2}} \right) \exp \left\{ -\frac{\xi^2}{4} \right\}, \quad \lambda_n = -2n - 1, \quad n = 0, 1, 2, \dots$$

Finalmente la solución (7.36) en este caso se presenta como:

$$V_\lambda(t, x) = C_n t^{\frac{n}{2}} H_n \left(\frac{\xi}{\sqrt{2}} \right) \exp \left\{ \beta \ln x + \frac{\sigma^2}{2} (\beta - 1)^2 t \right\}. \quad (7.37)$$

Esta solución tiene una propiedad interesante: ella verifica una de las condiciones complementarias del modelo de Black-Scholes (si $\beta > 0$):

$$\begin{aligned} x \rightarrow 0^+ &\Rightarrow V_n \sim x^\beta (\ln x)^n \\ &\Rightarrow \lim_{x \rightarrow 0^+} (V_n) = \lim_{x \rightarrow 0^+} (\text{const } x^\beta) = 0. \end{aligned}$$

Por otro lado, se constató que si $x \rightarrow \infty \Rightarrow V_n \sim x^\beta (\ln x)^n$. Esto permite aproximar bastante bien la otra condición de frontera si β es pequeña y construir una serie de tipo exponencial usando las soluciones (7.37). Más exactamente para $C_n = \sigma^n$:

$$V^\circ(t, x) \equiv \sum_{n=0}^{\infty} \frac{1}{n!} V_n(t, x) = x^{1+\beta} \exp \left\{ \frac{\sigma^2}{2} (\beta - 1)^2 t \right\}. \quad (7.38)$$

Esta función es el caso particular de la solución (7.8) para $\lambda = 1$. En la aproximación $\beta \approx 0$, $\beta \neq 0$, la solución de la ecuación de Black-Scholes que verifica las condiciones complementarias, será la combinación lineal de (7.38) y de (7.37) para $n = 0$:

$$V(t, x) = x^\beta \exp \left\{ \frac{\sigma^2}{2} (\beta - 1)^2 (t - T) \right\} \left[x e^{-\frac{\sigma^2}{2}(t-T)} - K \right]. \quad (7.39)$$

Precisamente la aproximación es en la presencia del factor $x^\beta \approx 1$ si $\beta \approx 0$. Por ejemplo, para $r = 0.05$; $\sigma = 0.317$ implica $\beta = 0.002$.

4. Sea la clase de equivalencia definida por el operador (2d) del teorema [7.2]: $B = B_1^2 + B_2$. La solución de la ecuación de Black-Scholes en este caso tiene la forma siguiente:

$$V_\lambda(t, x) = \Psi(\xi) \exp \left\{ \left(\beta - \frac{\sigma^2 t}{2} \right) \ln x - \frac{\sigma^6 t^3}{12} + \frac{\sigma^2}{2} [(\beta - 1)^2 - \lambda] t \right\},$$

$$\xi \equiv \ln x + \frac{\sigma^4 t^2}{4}.$$

La función $\Psi(\xi)$ verifica la ecuación de Airy: $\Psi'' - \xi\Psi = \lambda\Psi$.

5. La última clase de equivalencia de los operadores de segundo orden tiene como operador más sencillo a (2e) del teorema [7.2]: $B = B_1^2$. Las soluciones de la ecuación de Black-Scholes con $B = B_1^2$ para cualesquiera valores del parámetro λ son parecidas a las funciones dadas en (7.8), pero aquí dependen de dos constantes arbitrarias C, D (porque el operador de simetría en este caso es de segundo orden):

$$V_\lambda(t, x) = (C e^{\sqrt{\lambda} \ln x} + D e^{-\sqrt{\lambda} \ln x}) \exp \left\{ \beta \ln x + \frac{\sigma^2}{2} [(\beta - 1)^2 - \lambda] t \right\}.$$

También la nueva solución de la ecuación de Black-Scholes que pertenece a la clase de equivalencia definida por la solución (7.37) es:

$$V'_n(t, x) = C'_n t^{-\frac{n+1}{2}} H_n(\rho) \exp \left\{ \frac{\ln^2 x}{2\sigma^2 t} + \beta \ln x + \frac{\sigma^2}{2} (\beta - 1)^2 t \right\},$$

$$\rho \equiv \frac{\ln x}{\sqrt{-2\sigma^2 t}}.$$

La unidad imaginaria $\sqrt{-1}$ en la expresión ρ no influye en el resultado porque los polinomios de número par $H_{2k}(\rho)$ contienen únicamente los términos con potencias pares de ρ y por lo tanto son reales. En los polinomios $H_{2k+1}(\rho)$ todos los términos de potencia impar y la unidad imaginaria se pueden sacar como factor común, el que queda incluido en la constante C'_{2k+1} , que será en este caso también imaginaria.

La solución clásica de la ecuación de Black-Scholes es la función siguiente:

$$V(t, x) = x\Phi(\theta) - K e^{-r(T-t)}\Phi(\omega) \quad (7.40)$$

donde, como se mencionó antes, Φ representa la función de distribución de probabilidad acumulada de una normal estándar y con:

$$\theta \equiv \frac{\ln x - (1 - \beta)\sigma^2 t + T\sigma^2(1 - \beta) - \ln K}{\sigma\sqrt{T - t}}$$

$$\omega \equiv \frac{\ln x + \beta\sigma^2 t - (T\sigma^2\beta + \ln K)}{\sigma\sqrt{T - t}}$$

la cual es una forma análoga a la ecuación (5.33) que se obtuvo al final de la sección 5.3.1.

En realidad cada término de la fórmula (7.40) representa una solución de la ecuación de Black-Scholes. Dichos términos tienen la simetría superior: el operador de simetría correspondiente es un operador integral. Observando que las funciones x , $\exp\{-r(T - t)\}$ son las soluciones particulares de la ecuación de Black-Scholes (5.10), se puede constatar que cada uno de los términos de la solución (7.40) representa un producto de dicha solución particular por la función $\Phi(z)$ que en consecuencia verifica la ecuación diferencial ordinaria: $\Phi'' + z\Phi' = 0$.

La nueva solución de la ecuación de Black-Scholes, que es equivalente a (7.40), relacionada con el grupo discreto N es la función siguiente:

$$V'(t, x) = V_1 - V_2 \quad (7.41)$$

donde cada término es también una solución de la ecuación de Black-Scholes:

$$V_1 \equiv \frac{\Phi(\nu(t, x))}{\sqrt{\sigma^2 t}} \exp \left\{ \frac{\ln^2 x}{2\sigma^2 t} + \left(\beta + \frac{1 - b}{\sigma^2 t} \right) \ln x + \frac{(b - 1)^2}{2\sigma^2 t} + \left(r + \frac{\sigma^2 \beta^2}{2} \right) t \right\},$$

$$\nu(t, x) \equiv \frac{\ln x + (1 - b) + [T\sigma^2(1 - b) - \ln K]\sigma^2 t}{\sqrt{\sigma^2 t(Ts^2\sigma^2 t + 1)}},$$

$$V_2 \equiv \frac{K e^{-\rho T} \exp\left(\frac{2b-1}{2\sigma^2 t}\right)}{\sqrt{\sigma^2 t}} \Phi(\mu(t, x))$$

$$\cdot \exp \left\{ \frac{\ln^2 x}{2\sigma^2 t} + \left(\beta - \frac{b}{\sigma^2 t} \right) \ln x + \frac{(b - 1)^2}{2\sigma^2 t} + \left(r + \frac{\sigma^2 \beta^2}{2} \right) t \right\},$$

$$\mu(t, x) \equiv \frac{\ln x - b - [T\sigma^2 b + \ln K]\sigma^2 t}{\sqrt{\sigma^2 t(Ts^2\sigma^2 t + 1)}}.$$

Se concluye que la función (7.41), $V'(t, x) = V_1 - V_2$, es una solución de la ecuación de Black-Scholes (5.10) para cualesquiera valores de las constantes K , T , s , ρ ($K > 0$, $T > 0$, $s > 0$, $b \equiv \frac{1}{2} - \frac{\rho}{s^2}$).

7.5. Comparación de las soluciones obtenidas por Sukhomlin y la fórmula de Black-Scholes.

Vamos a comparar la fórmula de Black-Scholes para una opción de compra (véase sección 5.3.1) y la fórmula obtenida por Sukhomlin dada por la ecuación (7.39). Para esto tomaremos nuevamente los datos utilizados en la sección 6.2, ya que, para los valores de r y σ considerados anteriormente $\beta = 0.002$, cumpliendo así la condición de que $\beta \approx 0$.

Primeramente simulamos el movimiento Browniano geométrico Figura 7.1. Se utilizan los valores obtenidos del subyacente S_t para evaluar en ambas fórmulas.

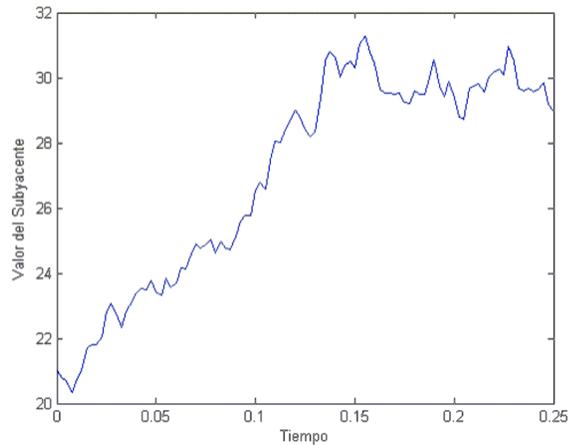


Figura 7.1: Simulación del movimiento Browniano geométrico.

La Figura 7.2 muestra la comparación entre ambas fórmulas, con lo cual se ve que la fórmula obtenida por Sukhomlin es una buena aproximación a la fórmula de Black-Scholes para valores grandes del subyacente y también conforme pasa el tiempo.

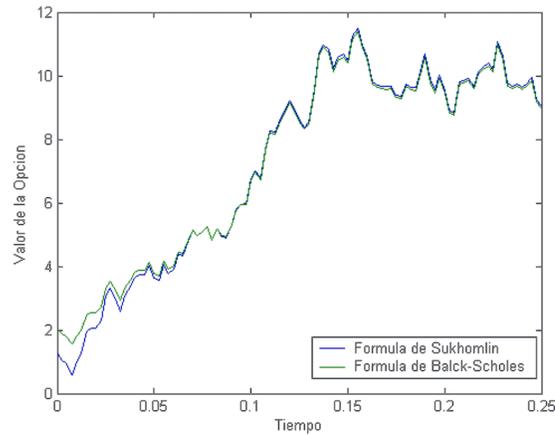


Figura 7.2: Gráfica de comparación entre la fórmula de Sukhomlin y la fórmula de Balck-Scholes mediante la simulación de una trayectoria del movimiento Browniano geométrico.

Por otro lado, la Figura 7.3 muestra la comparación en un mallado, obteniendo nuevamente que la fórmula de Sukhomlin es una buena aproximación de la fórmula de Black-Scholes, aunque para valores del subyacente menores que el precio de ejercicio K de la opción de compra esta toma valores negativos lo que no ocurre para la fórmula de Black-Scholes.

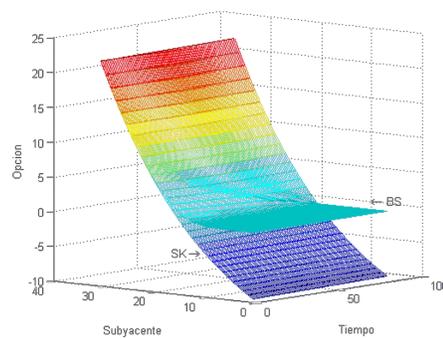


Figura 7.3: Gráfica de comparación entre la fórmula de Sukhomlin y Balck-Scholes sobre un mallado.

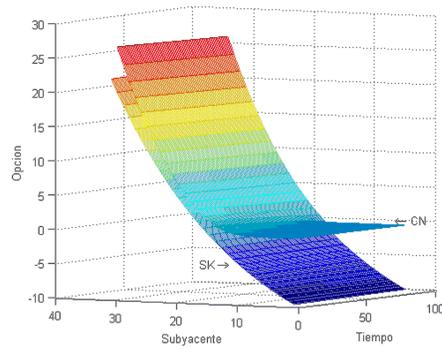


Figura 7.4: Gráfica de comparación entre la fórmula de Sukhomlin y la solución numérica mediante el método de Crank-Nicolson.

Por último la Figura 7.4 muestra la comparación de la fórmula de Sukhomlin y la solución numérica obtenida mediante el método de Crank-Nicolson en la sección 6.2.

Conclusiones.

Las matemáticas financieras y el cálculo de Itô abren una nueva perspectiva para plantear nuevos problemas matemáticos utilizando ecuaciones en derivadas parciales y métodos numéricos, ya que muchas ecuaciones en esta área no tienen solución analítica.

Dentro de los métodos más utilizados para realizar modelación computacional en el área de matemáticas financieras están los métodos de diferencias finitas, para los cuales, es necesario realizar modificaciones, ya que las condiciones de frontera e iniciales no son continuas.

En la tesis se utilizó el método de Crank-Nicolson, debido a las ventajas que este presenta y que fueron expuestas en el capítulo 6, para modelar la valuación de opciones. En la sección 6.2 se hizo la modelación de valuación de opciones de compra europeas, ya que, para este caso, se encontró la solución analítica dada por la ecuación de Black-Scholes, por lo tanto se pudo hacer una comparación entre los cálculos obtenidos por la modelación con el método de Crank-Nicolson y los que se obtuvieron usando la fórmula de Black-Scholes. De esta manera, para trabajos futuros, se podría utilizar el método de Crank-Nicolson con el objetivo de modelar la valuación de otro tipo de opciones, para las cuales los problemas de valor a la frontera asociados a estas no tengan solución analítica y sólo se dependa de métodos numéricos.

Las relaciones obtenidas por Sukhomlin abren un nuevo camino para resolver ecuaciones en derivadas parciales en el área de matemáticas financieras. Una de las soluciones es bastante general ya que toma valores negativos como lo muestra la Figura 7.3, pero esta solución está restringida a la aproximación $x^\beta \approx 1$ en la ecuación (7.39), lo cual implica que $\beta \approx 0$. Esto tiene como consecuencia que dicha solución no se pueda generalizar para valores de $\beta \gg 0$, donde este valor corresponde a muchas de las aplicaciones que se realizan en matemáticas financieras. Es necesario modificar la aproximación de x^β para que la ecuación se pueda aplicar para cualquier valor de una opción de compra.

Las otras soluciones de Sukhomlin no se compararon debido a que el mismo Sukhomlin [13], no prueba que dichas soluciones cumplen con las condiciones iniciales y de frontera, ya que el enfoque que utiliza es local.

Por último, los resultados obtenidos en los capítulos 6 y 7 fueron obtenidos mediante la implementación en MATLAB de los algoritmos que se dan en el apéndice C, específicamente, el algoritmo C.1 fue utilizado en la sección 6.1, el algoritmo C.2 fue utilizado en la sección 6.2 y el algoritmo C.3 fue utilizado en la sección 7.5. Cabe señalar que los algoritmos C.2 y C.3 fueron desarrollados de acuerdo a las necesidades de la tesis, además, todos fueron implementados en MATLAB debido a las facilidades que este software presentó al momento de programar estos algoritmos.

Apéndice.

A. Martingalas.

Definición A.1. La colección $(\mathcal{F}_t, t \geq 0)$ de σ -álgebras sobre Ω es llamada una **filtración** si

$$\mathcal{F}_s \subset \mathcal{F}_t \text{ para toda } 0 \leq s \leq t.$$

Así, una filtración es un flujo creciente de información.

Si $(\mathcal{F}_n, n = 0, 1, \dots)$ es una sucesión de σ -álgebras sobre Ω y $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ para toda n , llamamos también a (\mathcal{F}_n) una filtración.

Definición A.2. El proceso estocástico $Y = (Y_t, t \geq 0)$ se dice que es **adaptado a la filtración** $(\mathcal{F}_t, t \geq 0)$ si

$$\sigma(Y_t) \subset \mathcal{F}_t \text{ para toda } t \geq 0.$$

El proceso estocástico Y siempre es adaptado a la **filtración natural generada por Y** :

$$\mathcal{F}_s = \sigma(Y_t, s \leq t).$$

Definición A.3. Si $Y = (Y_n, n = 0, 1, \dots)$ es un proceso de tiempo discreto se define la adaptación de forma análoga: para una filtración $(\mathcal{F}_n, n = 0, 1, \dots)$ se requiere que $\sigma(Y_n) \subset \mathcal{F}_n$.

Definición A.4. Si el proceso estocástico Y es adaptado a la filtración natural Browniana $(\mathcal{F}_t, t \geq 0)$, diremos que Y es **adaptado al movimiento Browniano**. Esto significa que Y_t es una función de $B_s, s \leq t$.

Definición A.5. El proceso estocástico $X = (X_t, t \geq 0)$ es llamado una **martingala de tiempo continuo con respecto a la filtración** $(\mathcal{F}_t, t \geq 0)$, y lo denotamos por $(X, (\mathcal{F}_t))$, si

i) $E|X_t| < \infty$ para toda $t \geq 0$.

ii) X es adaptado a (\mathcal{F}_t) .

iii) $E(X_t | \mathcal{F}_s) = X_s$ para toda $0 \leq s \leq t$, es decir, X_s es la mejor predicción de X_t dada \mathcal{F}_s .

Definición A.6. El proceso estocástico $X = (X_n, n = 0, 1, \dots)$ es llamado una **martingala de tiempo discreto con respecto a la filtración** $(\mathcal{F}_n, n = 0, 1, \dots)$, y lo denotamos por $(X, (\mathcal{F}_n))$, si

i) $E|X_n| < \infty$ para toda $n = 0, 1, \dots$

ii) X es adaptado a (\mathcal{F}_n) .

iii) $E(X_{n+1}|\mathcal{F}_n) = X_n$ para toda $n = 0, 1, \dots$, es decir, X_n es la mejor predicción de X_{n+1} dada \mathcal{F}_n .

B. Modos de convergencia.

Definición B.1. La sucesión (A_n) **converge en distribución o converge débilmente a la variable aleatoria** A ($A_n \xrightarrow{d} A$) si para todas las funciones continuas acotadas f , se cumple que

$$Ef(A_n) \rightarrow Ef(A), \quad n \rightarrow \infty.$$

Definición B.2. La sucesión (A_n) **converge en probabilidad a la variable aleatoria** A ($A_n \xrightarrow{P} A$) si para todo $\epsilon > 0$, se cumple que

$$P(|A_n - A| > \epsilon) \rightarrow 0, \quad n \rightarrow \infty.$$

Definición B.3. La sucesión (A_n) **converge casi seguramente (c.s.) o con probabilidad 1 a la variable aleatoria** A ($A_n \xrightarrow{c.s.} A$) si el conjunto de ω 's con

$$A_n(\omega) \rightarrow A(\omega), \quad n \rightarrow \infty.$$

tiene probabilidad 1, es decir,

$$P(A_n \rightarrow A) = P(\omega : A_n(\omega) \rightarrow A(\omega)) = 1.$$

C. Algoritmos.

Algoritmo C.1. Método de Crank-Nicolson.

Para aproximar la solución de la ecuación diferencial parabólica

$$u_1(t, x) - \alpha^2 u_{22}(t, x) = 0, \quad 0 < x < l, \quad 0 < t < T,$$

sujeta a las condiciones de frontera

$$u(t, 0) = u(t, l) = 0, \quad 0 < t < T,$$

y las condiciones iniciales

$$u(x, 0) = f(x), \quad 0 \leq x \leq l.$$

ENTRADA: extremo l ; tiempo máximo T ; constante α ; enteros $m \geq 3$, $N \geq 1$.

SALIDA: aproximaciones $w_{j,i}$ a $u(t_j, x_i)$ para toda $i = 1, \dots, m - 1$ y $j = 1, \dots, N$.

Paso 1 Tomar $h = l/m$; $k = T/N$; $\lambda = \alpha^2 k/h^2$; $w_m = 0$.

Paso 2 Para $i = 1, \dots, m - 1$, tomar $w_i = f(ih)$. (Valores iniciales.)

(Los pasos 3-11 resuelven el sistema tridiagonal utilizando factorización de Crout.)

Paso 3 Tomar $l_1 = 1 + \lambda$; $u_1 = -\lambda/(2l_1)$.

Paso 4 Para $i = 1, \dots, m - 2$, tomar $l_i = 1 + \lambda + \lambda u_{i-1}/2$; $u_i = -\lambda/(2l_i)$.

Paso 5 Tomar $l_{m-1} = 1 + \lambda + \lambda u_{m-2}/2$.

Paso 6 Para $j = 1, \dots, N$ hacer los pasos 7-11.

Paso 7 Tomar $t = jk$; (t_j actual)

$$z_1 = \frac{(1 - \lambda)w_1 + \frac{\lambda}{2}w_2}{l_1}.$$

Paso 8 Para $i = 2, \dots, m - 1$, tomar

$$z_i = \frac{(1 - \lambda)w_i + \frac{\lambda}{2}(w_{i+1} + w_{i-1} + z_{i-1})}{l_i}.$$

Paso 9 Tomar $w_{m-1} = z_{m-1}$.

Paso 10 Para $i = m - 2, \dots, 1$, tomar $w_i = z_i - u_i w_{i+1}$.

Paso 11 SALIDA (t); (Nota: $t = t_j$). Para $i = 1, \dots, m - 1$, tomar $x = ih$; SALIDA (x, w_i). (Nota: $w_i = w_{i,j}$.)

Paso 12 PARAR. (Procedimiento terminado.)

Algoritmo C.2. Método de Crank-Nicolson modificado.

Para aproximar la solución de la ecuación diferencial parabólica

$$u_1(t, x) - u_{22}(t, x) = 0, \quad x_{min} < x < x_{max}, \quad 0 < t \leq T_1,$$

sujeta a las condiciones de frontera

$$u(t, 0) = 0,$$

$$u(t, x_{max}) = g(t) = e^{\frac{1}{2}(\gamma+1)x_{max} + \frac{1}{4}(\gamma+1)^2 t} - e^{\frac{1}{2}(\gamma-1)x_{max} + \frac{1}{4}(\gamma-1)^2 t}, \quad 0 \leq t \leq T_1$$

y la condición inicial dada por la ecuación (5.25)

$$u(0, x) = f(x) = \max(e^{\frac{1}{2}(\gamma+1)x} - e^{\frac{1}{2}(\gamma-1)x}, 0), \quad x_{min} \leq x \leq x_{max}$$

con $\gamma = \frac{2r}{\sigma^2}$.

ENTRADA: extremos x_{min} y x_{max} ; tiempo máximo T_1 ; enteros $N_x \geq 3$, $N_t \geq 1$.

SALIDA: aproximaciones $w_{j,i}$ a $u(t_j, x_i)$ para toda $i = 1, \dots, N_x - 1$ y $j = 1, \dots, N_t$.

Paso 1 Tomar $L = x_{max} - x_{min}$; $dx = L/N_x$; $dt = L/N_t$; $\lambda = dt/dx^2$.

Paso 2 Para $i = 1, \dots, N_x$, tomar $w_{0,i} = f(x_{min} + idx)$. (Valores iniciales.)

Paso 3 Para $j = 0, \dots, N_t$, tomar $w_{j,0} = 0$ y $w_{j,x_{max}} = g(jdt)$. (Valores de frontera.)

Paso 4 Tomar $diagpB = 1 - \lambda$; $diagB = \lambda/2$; $diagpA = 1 + \lambda$ y $diagA = -\lambda/2$.

Paso 5 Tomar $A_{1,1} = diagpA$; $A_{1,2} = diagA$; $A_{N_x-1, N_x-2} = diagA$ y $A_{N_x-1, N_x-1} = diagpA$.

Paso 6 Tomar $B_{1,1} = diagpB$; $B_{1,2} = diagB$; $B_{N_x-1, N_x-2} = diagB$ y $B_{N_x-1, N_x-1} = diagpB$.

Paso 7 Para $k = 3, \dots, N_x - 1$, tomar $A_{1,k} = 0$; $A_{N_x-1, N_x-k} = 0$ y $B_{1,k} = 0$; $B_{N_x-1, N_x-k} = 0$.

Paso 8 Para $i = 2, \dots, N_x - 1$ hacer los pasos 9-13. (Construcción de las matrices A y B .)

Paso 9 Para $k = 1, \dots, N_x - 1$ hacer los pasos 10-13.

Paso 10 Si $i - k = 1$ o $k - i = 1$, tomar $A_{i,k} = \text{diag}A$ y $B_{i,k} = \text{diag}B$.

Paso 11 De lo contrario, hacer los pasos 12-13.

Paso 12 Si $i = k$, tomar $A_{i,k} = \text{diag}pA$ y $B_{i,k} = \text{diag}pB$.

Paso 13 De lo contrario, tomar $A_{i,k} = 0$ y $B_{i,k} = 0$.

Paso 14 Para $j = 1, \dots, N_t$, tomar

$$\mathbf{w}^{(j)} = A^{-1}(B\mathbf{w}^{(j-1)} + \mathbf{b} - \mathbf{a}).$$

(Nota: $\mathbf{a} = (0, \dots, (-\lambda/2)w_{j,N_x})$ y $\mathbf{b} = (0, \dots, (\lambda/2)w_{j-1,N_x})$.)

Paso 15 Para $i = 1, \dots, N_x - 1$, tomar $x = x_i = x_{min} + idx$ y para $j = 1, \dots, N_t$, tomar $t = t_j = jdt$.

SALIDA $(x, t, \mathbf{w}^{(j)})$. (Nota: $\mathbf{w}^{(j)} = (w_{j,1}, w_{j,2}, \dots, w_{j,N_x-1})^t$.)

Paso 16 PARAR. (Procedimiento terminado.)

Algoritmo C.3. Simulación de trayectorias del movimiento Browniano geométrico utilizando la propiedad de incrementos independientes del movimiento Browniano.

ENTRADA: tasa de interés r ; volatilidad σ ; valor inicial X_0 ; tiempo máximo T y $N > 0$ número de particiones.

SALIDA: t el tiempo para $t \in [0, T]$ y X_t el valor del movimiento Browniano geométrico al tiempo t

Paso 1 Tomar $c = \ln r - 0.5\sigma^2$ y $\mu = c - 0.5\sigma^2$.

Paso 2 Tomar $B_0 = 0$; $t_0 = 0$ y $dB = T/N$.

Paso 3 Para $j = 1, \dots, N$, tomar $t_j = jdB$ y $B_j = \sqrt{t_j - t_{j-1}} \cdot \text{randn} + B_{j-1}$.
(Nota: randn es un número aleatorio generado según una distribución $N(0, 1)$.)

Paso 4 Tomar $X_{t_0} = X_0$.

Paso 5 Para $j = 1, \dots, N$, tomar $X_{t_j} = X_{t_0} \exp\{\mu * t_j + \sigma B_j\}$.
SALIDA (t, X_t) . (Nota: $t = t_j$ y $X_t = X_{t_j}$.)

Paso 6 PARAR. (Procedimiento terminado.)

D. Reseña Histórica.

Uno de los instrumentos más usados por los financieros son las opciones de compra o venta a futuro, lo cual al ser a futuro acarrea una cierta incertidumbre. Un modelo matemático para poner precio a las opciones fue desarrollado en 1973 por Fisher Black y Myron Scholes. Ellos usaron, para modelar los cambios de precio, las matemáticas que describen los movimientos de las moléculas de un gas en un recipiente cerrado, conocido como el movimiento Browniano. Con este modelo Black y Scholes comparan las alzas y bajas de precios de las opciones con las moléculas de gas chocando rozándose una con otra. Este modelo ha sido muy valioso en el área de las matemáticas como de las altas finanzas.

Las bases teóricas de estos estudios las puso Kiyoshi Itô en 1942 cuando desarrollo y probó lo que actualmente se conoce como el lema de Itô.

El mundo económico tuvo un cambio importante cuando en 1970 los Estados Unidos de Norteamérica abandonaron el patrón oro para que las monedas entraran en lo que se conoce como un sistema de flotación. Como resultado existe la posibilidad de que los precios de las monedas, las tasas de interés, los bonos, los futuros, las opciones, etc., puedan fluctuar de manera aterradora en un instante. Hemos indicado ya que en esa misma época se usan las descripciones de los fenómenos aleatorios para explicar las fluctuaciones de los productos financieros. El lema de Itô se aplica a esta economía cambiante y los matemáticos se interesan cada vez más en el estudio de los fenómenos aleatorios.

En 1997 el premio Nobel de economía se otorgo a Robert C. Merton y a Myron S. Scholes en reconocimiento a su contribución en el área de ingeniería financiera. Merton adoptó el lema de Itô para introducir una nueva fórmula de fluctuación de precios.

Así, en el mundo de la economía y las finanzas el uso de las matemáticas es cada vez mayor, en particular en campos como la investigación financiera o la administración de riesgo.

Bibliografía

- [1] Burden L. Richard, Faires Douglas J.; *Análisis Numérico*. Thomson Learning, Séptima Edición 2002.
- [2] Chaichian M.; *Path Integrals in Physics Volume I: Stochastic Processes and Quantum Mechanics*. Institute of Physics Publishing Bristol and Philadelphia 2001.
- [3] Debnath Lokenath; *Nonlinear Partial Differential Equations for Scientist and Engineers*. Birkhäuser, Boston 1997.
- [4] Evans G. Blackledge and Yardley P.; *Numerical Methods for Partial Differential Equations*. Springer-Verlag 2000.
- [5] Fernández Viviana; *Teoría de Opciones: una Síntesis*. Revista de Análisis Económico, Vol. 14, No.2, Pág. 87-116 (Noviembre de 1999).
- [6] Ganzha G. Victor, Vorozhtsov V. Evgenii; *Numerical Solutions for Partial Differential Equations*. CRC PRESS.
- [7] Goodman J., Moon K. S. ; *Stochastic and Partial Differential Equations with Adapted Numerics*.
- [8] Haberman, Richard; *Ecuaciones en Derivadas Parciales con Series de Fourier y Problemas de Contorno*. 3ª edición; Prentice Hall 2003.
- [9] Higham J. Desmond; *An Introduction to Financial Option Valuation: Mathematics, Stochastic and Computation*. Cambridge University Press 2004.
- [10] Koshlyakov N. S., Smirnov M. M., Gliner E. B.; *Differential Equations of Mathematical Physics*. North-Holland Publishing Company, Amsterdam; 1964.
- [11] Mikosch Thomas; *Elementary Stochastic Calculus with Finance in View*. Advanced Series on Statistical Science and Applied Probability Vol. 6. 2003.

-
- [12] Mordecki Ernesto; *Modelos Matemáticos en Finanzas: Valuación de Opciones*. Centro de Matemática, Facultad de Ciencias, Montevideo, Uruguay.
- [13] Morton K. W., Mayers D. F.; *Numerical Solutions Of Partial Differential Equations*. Cambridge University Press, Segunda Edición 2005.
- [14] Sukhomlin Nikolay; *Simetría y Nuevas Soluciones de la Ecuación de Black-Scholes*. Boletín de la Asociación Matemática Venezolana, Vol. XI, No. 2 (2004) Pág.175-189.
- [15] Villalón G Julio, Barbeito Mtz. Josefina; *Métodos Modernos de Valoración de los Instrumentos Financieros*. Universidad de Valladolid y Universidad de A. Coruña.