



Universidad Juárez Autónoma de Tabasco
División Académica de Ciencias Básicas



**Aplicaciones del Cálculo Estocástico en la
Biología**

TESIS

Que para obtener el grado de:

Maestro en Ciencias Matemáticas

Presenta:

Dorilian García Cerino

Director de tesis:

Dr. Aroldo Pérez Pérez

Cunduacán, Tabasco, México

Diciembre 2017



**UNIVERSIDAD JUÁREZ
AUTÓNOMA DE TABASCO**

"ESTUDIO EN LA DUDA. ACCIÓN EN LA FE"



División
Académica
de Ciencias
Básicas

DIRECCIÓN

7 de diciembre de 2017

Lic. Dorilian García Cerino

Pasante de la Maestría en Ciencias Matemáticas

Presente

Por medio del presente y de la manera más cordial, me dirijo a Usted para hacer de su conocimiento que proceda a la impresión del trabajo titulado "**Aplicaciones del Cálculo Estocástico en la Biología**", en virtud de que reúne los requisitos para el EXAMEN PROFESIONAL para obtener el grado de Maestro en Ciencias Matemáticas.

Sin otro particular, reciba un cordial saludo.

Atentamente,

Dr. Gerardo Delgadillo Piñón
Director



DIVISIÓN ACADÉMICA DE
CIENCIAS BÁSICAS

C.c.p.- Archivo
Dr'GDP/Dr'JGPS/emt

Miembro CUMEX desde 2008



Km. 1, Carretera Cunduacán - Jalpa de Méndez, A.P. 24, C.P. 86690, Cunduacán, Tabasco.
Tel/Fax (914)3360928, (993)3581500 Ext. 6701 E-mail: direccion.dacb@ujat.mx

CARTA DE AUTORIZACIÓN

El que suscribe, autoriza por medio del presente escrito a la Universidad Juárez Autónoma de Tabasco para que utilice tanto física como digitalmente la tesis de grado denominada "**Aplicaciones del Cálculo Estocástico en la Biología**", de la cual soy autor y titular de los Derechos de Autor.

La finalidad del uso por parte de la Universidad Juárez Autónoma de Tabasco de la tesis antes mencionada, será única y exclusivamente para difusión, educación y sin fines de lucro; autorización que se hace de manera enunciativa mas no limitativa para subirla a la Red Abierta de Bibliotecas Digitales (RABID) y a cualquier otra red académica con las que la Universidad tenga relación institucional.

Por lo antes manifestado, libero a la Universidad Juárez Autónoma de Tabasco de cualquier reclamación legal que pudiera ejercer respecto al uso y manipulación de la tesis mencionada y para los fines estipulados en este documento.

Se firma la presente autorización en la ciudad de Villahermosa, Tabasco a los 8 días del mes de diciembre del año 2017.

AUTORIZÓ



Dorilian García Cerino
152A21001

Dedicatoria

A mi Familia

Universidad Juárez Autónoma de Tabasco.
México.

Agradecimientos

A nuestro Creador, por brindarme todo lo necesario para llevar acabo este trabajo y poder finalizarlo.

A mi Hermosa Mamá, la Sra. Dora María Cerino Hernández, por su confianza, por sus valiosos consejos y principalmente por su apoyo incondicional.

A mi Querido Hermano, el Lic. Raúl García Cerino, por motivarme y guiarme en el camino del buen saber.

A mi Maestro, el Dr. Aroldo Pérez Pérez, por su paciencia, por sus valiosas aportaciones para la finalización de este trabajo, y principalmente, por ese entusiasmo y gran empeño que ha puesto en mi formación como profesionista.

A cada uno de los Revisores de esta tesis por sus sugerencias para la mejora de la misma.

En general, a la DACB-UJAT.

Al CONACYT por darnos la oportunidad de obtener estudios de posgrado.

Índice general

Dedicatoria	I
Agradecimientos	II
Introducción	VII
1. Preliminares del cálculo avanzado	1
1.1. Fórmula de Taylor	1
1.2. Condiciones de Hölder y de crecimiento	2
1.3. Continuidad por la izquierda y por la derecha	4
1.4. Variación de una función	5
1.5. Integral de Riemann-Stieltjes	7
1.6. Espacios de medida	9
1.7. Funciones medibles	10
1.8. Integral de Lebesgue	11
2. Conceptos de la teoría de la probabilidad	13
2.1. Espacios de probabilidad filtrados	13
2.2. Probabilidad condicional e independencia de σ -álgebras	14
2.3. Variables aleatorias e independencia de variables aleatorias	15
2.4. Filtraciones generadas por variables aleatorias	17
2.5. Esperanza de una variable aleatoria	17
2.6. Esperanza condicional	19
2.7. Función generadora de momentos y función característica	20
2.8. Distribuciones gaussiana, gamma y exponencial	21
2.9. Convergencia de variables aleatorias	23

3. Conceptos de la teoría de los procesos estocásticos	24
3.1. Procesos estocásticos a tiempo continuo	24
3.2. Tiempos de paro y procesos estocásticos previsibles	26
3.3. Martingalas, submartingalas y supermartingalas	28
3.4. Procesos uniformemente integrables	29
3.5. Algunos resultados sobre la convergencia de martingalas	30
3.6. Martingalas locales	31
3.7. Procesos de Markov y procesos de Markov fuertes	33
3.8. Proceso de Wiener	34
4. Cálculo estocástico para el proceso de Wiener	37
4.1. La integral de Itô	37
4.2. Algunas propiedades del proceso integral de Itô	44
4.3. Procesos de Itô	47
4.4. Ecuaciones diferenciales estocásticas	49
4.5. Tiempos de salida de un intervalo	51
5. Aplicaciones en la biología	55
5.1. Difusión de ramificación de Feller	55
5.2. Difusión de Wright-Fisher	59
Conclusión	62
Bibliografía	64

Introducción

Cuando Newton y Leibniz inventaron el cálculo, una enorme gama de problemas de las ciencias naturales, sociales y biológicas quedaron bajo el dominio de la teoría de funciones de una variable real. Las principales componentes de este invento fueron el uso de la diferenciación para describir razones de cambio, el uso de la integración para pasar al límite en la aproximación por sumas y el teorema fundamental del cálculo, el cual relaciona ambos conceptos y hace más accesible los cálculos de estos límites de sumas (integrales). Esta teoría dio lugar al concepto de ecuaciones diferenciales ordinarias, y la aplicación de estas ecuaciones para modelar fenómenos del mundo real revela el enorme potencial del cálculo.

El cálculo estocástico surge de la necesidad de dar sentido a ecuaciones diferenciales ordinarias que, con la finalidad de obtener modelos matemáticos más realistas del fenómeno en cuestión, involucran procesos estocásticos. A manera de ilustración considere el siguiente modelo sencillo de crecimiento de población

$$dX_t = \alpha(t)X_t dt, \quad X_0 = x_0,$$

donde X_t es el tamaño de la población al tiempo t y $\alpha(t)$ es la tasa relativa de crecimiento al tiempo t . Podría ocurrir que dicha tasa de crecimiento no esté completamente determinada debido a su dependencia de algunos fenómenos ambientales aleatorios, de manera que se tendría

$$\alpha(t) = \mu(t) + R_t,$$

donde la función $\mu(t)$ es determinista y se desconoce el comportamiento exacto del “ruido”, R_t . Formalmente el proceso R_t es definido como la derivada de un proceso estocástico W_t , es decir, $R_t = dW_t/dt$, y así, formalmente, se obtiene la ecuación diferencial estocástica

$$dX_t = \mu(t)X_t dt + X_t dW_t.$$

Sin embargo, la mayoría de los procesos estocásticos importantes, tales como el proceso de Wiener (sección 3.8), no son diferenciables. Por tal motivo, el cálculo estocástico toma el camino opuesto al del cálculo ordinario; primero se define la integral estocástica

$$Y_t = \int_0^t X_s dW_s$$

de cierto proceso estocástico $\{X_t\}_{t \geq 0}$ con respecto a cierto proceso estocástico $\{W_t\}_{t \geq 0}$ y luego se le da sentido a la diferencial estocástica a través del “teorema” fundamental del cálculo. Este “teorema” es en realidad una definición en el cálculo estocástico, debido a que la diferencial no tiene más significado que el que se le asigna mediante una integral estocástica, esto es, si $\{X_t\}_{t \geq 0}$ es un proceso estocástico que satisface la ecuación integral estocástica

$$X_t = X_0 + \int_0^t \mu(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s,$$

donde $\mu(t, x)$ y $\sigma(t, x)$ son funciones dadas, se escribe

$$dX_t = \mu(t, X_t) dt + \sigma(t, X_t) dW_t. \quad (1)$$

Existen muchos textos que consideran el caso donde $\{W_t\}_{t \geq 0}$ es una semimartingala (para su definición véase por ejemplo, [18] pág. 211), sin embargo, en este trabajo de tesis se considera sólo el caso en que $\{W_t\}_{t \geq 0}$ es un proceso de Wiener (capítulo 4).

Las ecuaciones diferenciales estocásticas (1) son el análogo probabilista de las ecuaciones diferenciales deterministas, y como se ejemplificó arriba, aparecen cuando se presenta aleatoriedad en los coeficientes de las ecuaciones deterministas. Hoy en día, la teoría de las ecuaciones diferenciales estocásticas constituye un tema bastante desarrollado de la teoría de los procesos estocásticos, con aplicaciones muy importantes en varios campos del conocimiento tales como la economía, ingeniería, física, medicina y biología. En particular, en la biología el cálculo estocástico se usa para modelar los efectos de la variabilidad estocástica en la reproducción y evolución genética de poblaciones (véase por ejemplo, [6] y [18]).

En este trabajo se considerarán dos ejemplos de soluciones (difusiones) de ciertas ecuaciones diferenciales estocásticas como modelos básicos para el crecimiento de poblaciones y evolución genética. Se abordará en primer lugar la difusión de ramificación de Feller, modelo para el crecimiento de cierto tipo de poblaciones, que como se verá (sección 5.1) tiene un crecimiento medio exponencial; y se encontrará además su probabilidad de extinción. Este tipo de procesos, aparte de ser de interés por sí mismos, surgen también como límite de sucesiones de procesos de ramificación simples cuando el tamaño de la población crece a infinito y la escala de tiempo se comprime de una manera específica (véase por ejemplo, [3] pág. 262). En segundo lugar, se abordará la difusión de Wright-Fisher, modelo útil en la aproximación de frecuencias alélicas en genética de poblaciones. En este modelo se considera una población de tamaño $2N$ en donde cada individuo tiene dos copias de cada gen, y se asume además que un rasgo es determinado por un solo gen en un sitio (locus) particular y que este gen tiene únicamente dos alelos diferentes, A_1 y A_2 . Con estas suposiciones y en ausencia de la mutación y de la selección, se encontrará la probabilidad de que la fijación ocurra en A_1 y el tiempo esperado para que esto suceda.

Este trabajo está estructurado en cinco capítulos. En los dos primeros se presentarán algunos resultados básicos del cálculo avanzado y de probabilidad empleados en el desarrollo futuro del mismo.

En el capítulo tercero se introducirán algunos conceptos básicos de la teoría de los procesos estocásticos a tiempo continuo; a saber, se abordarán los conceptos de tiempos de paro, martingalas, martingalas locales, procesos de Markov y finalmente se dará el concepto del proceso de Wiener, el cual es uno de los procesos más importantes en el cálculo estocástico. Se expondrán también algunos resultados concernientes a estos conceptos con sus respectivas demostraciones.

Se abordará, en el cuarto capítulo, una construcción de la integral estocástica, concepto introducido a mediados de 1940 por el matemático Japonés K. Itô y llamada actualmente integral estocástica de Itô. Se dará la prueba de algunas propiedades que posee el proceso integral de Itô; se verá el concepto de ecuación diferencial estocástica y se presentarán algunos teoremas de existencia y unicidad. Se finalizará este capítulo demostrando dos resultados concernientes a ecuaciones diferenciales estocásticas con coeficientes homogéneos en el tiempo (proposición 4.8 y proposición 4.7), los cuales serán de gran utilidad para el desarrollo de las aplicaciones consideradas en este trabajo.

Finalmente, en el quinto capítulo, se desarrollarán dos aplicaciones del cálculo estocástico en la biología; a saber, la difusión de ramificación de Feller y la difusión de Wright-Fisher.

Universidad Autónoma de Tabasco.
México.

Capítulo 1

Preliminares del cálculo avanzado

1.1. Fórmula de Taylor

Recuerde que si $f : [a, b] \rightarrow \mathbb{R}$ es una función diferenciable en $x_0 \in (a, b)$, entonces f es aproximadamente una función lineal en las proximidades de x_0 ; a saber, $f(x) \approx f(x_0) + f'(x_0)(x - x_0)$ cuando x está suficientemente próximo a x_0 . La fórmula de Taylor dice que, en general, f puede aproximarse por medio de un polinomio de grado n si es de clase C^n en $[a, b]$, es decir, f tiene derivadas continuas hasta de orden n en $[a, b]$. Más aún, esta fórmula proporciona una expresión útil para estimar el error cometido en esta aproximación. Las demostraciones de los teoremas que se exponen a continuación pueden ser consultadas, por ejemplo, en [1] páginas 113 y 361, respectivamente.

Teorema 1.1. Sea $f : [a, b] \rightarrow \mathbb{R}$ una función de clase C^n en $[a, b]$ y con derivada $f^{(n+1)}$ en (a, b) . Sea $x_0 \in [a, b]$. Entonces para cada $x \in [a, b]$, $x \neq x_0$, existe x_1 en el intervalo abierto con extremos x_0 y x tal que

$$f(x) = \sum_{i=0}^n \frac{f^{(i)}(x_0)}{i!} (x - x_0)^i + \frac{f^{(n+1)}(x_1)}{(n+1)!} (x - x_0)^{n+1}.$$

Un caso particular de la fórmula de Taylor multidimensional que es de gran importancia en el cálculo estocástico es dada en el siguiente resultado.

Teorema 1.2. Sea U un subconjunto abierto de $\mathbb{R}^n$ y $f : U \rightarrow \mathbb{R}$ una función de clase C^2 en U , esto es, f tiene derivadas parciales continuas hasta de segundo orden en U . Si el segmento de línea que une $\mathbf{x} = (x_1, \dots, x_n)$ con $\mathbf{x} + \mathbf{h}$, $\mathbf{h} = (h_1, \dots, h_n)$, está en U , entonces

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(\mathbf{x}) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{z}),$$

donde $\mathbf{z}$ es un punto en el segmento de línea que une $\mathbf{x}$ con $\mathbf{x} + \mathbf{h}$.

1.2. Condiciones de Hölder y de crecimiento

La condición de Hölder es requerida como una condición sobre los coeficientes en los resultados concernientes a la existencia y unicidad de soluciones de ecuaciones diferenciales ordinarias y estocásticas.

Definición 1.1. Una función $f : \mathbb{R} \rightarrow \mathbb{R}$ satisface la **condición de Hölder** de orden $\alpha \in (0, 1]$ si existe $K > 0$ tal que para toda $x, y \in \mathbb{R}$

$$|f(x) - f(y)| \leq K|x - y|^\alpha.$$

Cuando f satisface la condición de Hölder de orden $\alpha = 1$, se dice que f satisface la **condición de Lipschitz** o que es una **función de Lipschitz**.

Por ejemplo, la función $f : [0, \infty) \rightarrow \mathbb{R}$ definida por $f(x) = \sqrt{x}$ satisface la condición de Hölder de orden $\alpha = 1/2$. Para ver esto, sean $x, y \in [0, \infty)$ y note que si $y = 0$, entonces $|\sqrt{x} - \sqrt{y}| \leq K\sqrt{|x - y|}$ se cumple para toda $x \geq 0$ eligiendo $K = 1$. Ahora si $y > 0$, dividiendo por $\sqrt{y}$, es suficiente con demostrar que el cociente $\frac{|\sqrt{\frac{x}{y}} - 1|}{\sqrt{|\frac{x}{y} - 1|}}$ es acotado para toda $x \geq 0$; o equivalentemente, que para toda $z \geq 0$ la función

$$g(z) = \frac{|\sqrt{z} - 1|}{\sqrt{|z - 1|}} = \begin{cases} \frac{1 - \sqrt{z}}{\sqrt{1 - z}} & \text{si } z < 1 \\ \frac{\sqrt{z} - 1}{\sqrt{z - 1}} & \text{si } z > 1 \end{cases}$$

es acotada. Observe que $g(z)$ es continua en toda $z \neq 1$, y debido a la regla de L' Hopital $\lim_{z \rightarrow 1} g(z) = 0$ pues

$$\lim_{z \uparrow 1} g(z) = \lim_{z \uparrow 1} \frac{(-1/2)z^{-1/2}}{(-1/2)(1 - z)^{-1/2}} = \lim_{z \uparrow 1} \frac{(1 - z)^{1/2}}{z^{1/2}} = 0$$

y

$$\lim_{z \downarrow 1} g(z) = \lim_{z \downarrow 1} \frac{(1/2)z^{-1/2}}{(1/2)(z - 1)^{-1/2}} = \lim_{z \downarrow 1} \frac{(z - 1)^{1/2}}{z^{1/2}} = 0.$$

Además $\lim_{z \rightarrow \infty} g(z) = 1$ ya que

$$\lim_{z \rightarrow \infty} g(z) = \lim_{z \rightarrow \infty} \frac{\sqrt{z} - 1}{\sqrt{z - 1}} = \lim_{z \rightarrow \infty} \frac{1 - (1/\sqrt{z})}{\sqrt{1 - 1/z}} = 1.$$

Por lo tanto $g(z)$ es acotada, y en conclusión $f(x) = \sqrt{x}$ satisface la **condición de Hölder de orden $\alpha = 1/2$** .

Definición 1.2. Una función $f : [0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}$ es **localmente Lipschitz en x uniformemente en t** si para cada $N > 0$ existe $K(N) > 0$ tal que para toda x, y con $|x|, |y| \leq N$ y para toda $t \geq 0$

$$|f(t, x) - f(t, y)| \leq K|x - y|.$$

Igual que en el caso de la condición de Hölder, la condición de crecimiento lineal se presenta como una condición sobre los coeficientes en los resultados sobre la existencia y unicidad de soluciones de ecuaciones diferenciales.

Definición 1.3. Una función $f : \mathbb{R} \rightarrow \mathbb{R}$ satisface la **condición de crecimiento lineal** si existe $K > 0$ tal que para toda $x \in \mathbb{R}$

$$|f(x)| \leq K(|x| + 1).$$

Teorema 1.3. (**Desigualdad de Gronwall**) Sean $f, g : [0, T] \rightarrow [0, \infty)$ funciones y $\alpha > 0$ tales que

$$f(t) \leq g(t) + \alpha \int_0^t f(s) ds \quad (1.1)$$

para toda $t \in [0, T]$. Entonces para $t \in [0, T]$

$$f(t) \leq g(t) + \alpha \int_0^t g(s) e^{\alpha(t-s)} ds. \quad (1.2)$$

Demostración. Debido a la desigualdad (1.1) se tiene que

$$\frac{d(e^{-\alpha t} \int_0^t f(s) ds)}{dt} = e^{-\alpha t} f(t) - \alpha e^{-\alpha t} \int_0^t f(s) ds \leq g(t) e^{-\alpha t}.$$

Luego, integrando de 0 a t y después multiplicando por $e^{\alpha t}$ se obtiene que

$$\int_0^t f(s) ds \leq \int_0^t g(s) e^{\alpha(t-s)} ds.$$

De este modo, empleando nuevamente (1.1) se llega a que

$$f(t) \leq g(t) + \alpha \int_0^t f(s) ds \leq g(t) + \alpha \int_0^t g(s) e^{\alpha(t-s)} ds.$$

Por lo tanto la desigualdad (1.2) se satisface. $\blacktriangle$

1.3. Continuidad por la izquierda y por la derecha

Recuerde que una función $f : [a, b] \rightarrow \mathbb{R}$ es continua por la izquierda (derecha) en $x_0 \in (a, b]$ ($x_0 \in [a, b)$) si el límite por la izquierda (derecha) de f en x_0 es $f(x_0)$, esto es,

$$\lim_{x \uparrow x_0} f(x) = f(x_0) \quad (\lim_{x \downarrow x_0} f(x) = f(x_0)).$$

La versión continua por la izquierda (por la derecha) de f es definida en $(a, b]$ (en $[a, b)$) por

$$f(x-) := \lim_{y \uparrow x} f(y) \quad (f(x+) := \lim_{y \downarrow x} f(y)).$$

Definición 1.4. Sea $f : [a, b] \rightarrow \mathbb{R}$ una función. Un punto $x \in (a, b)$ es llamado una **discontinuidad de primer tipo** o **discontinuidad de salto** de f si $f(x+)$ y $f(x-)$ existen y no son iguales. El salto de f en x es definido como la diferencia, $f(x+) - f(x-)$. Cualquier otra discontinuidad se dice ser de **segundo tipo**.

Notación. En algunas ocasiones se usará $\Delta f(x)$ para denotar el tamaño del salto de f en x , esto es, $\Delta f(x) := f(x+) - f(x-)$, y en otras, será usado para denotar el incremento de f en $[x, x + \Delta]$, es decir, $\Delta f(x) = f(x + \Delta) - f(x)$. La interpretación de $\Delta f(x)$ será clara del contexto.

La demostración de la siguiente proposición puede verse, por ejemplo, en [27] pág. 107.

Proposición 1.1. Sea $f : [a, b] \rightarrow \mathbb{R}$ una función. El conjunto de discontinuidades de salto de f en $[a, b]$ es a lo más contable.

El resultado que sigue dice que la derivada finita de una función no puede tener discontinuidades de salto.

Proposición 1.2. Sea $f : (a, b) \rightarrow \mathbb{R}$ una función. Si f es diferenciable con derivada f' finita en (a, b) , entonces, en cualquier punto, f' es continua o tiene una discontinuidad de segundo tipo.

Demostración. Sea $x_0 \in (a, b)$. Suponga que f' no tiene una discontinuidad de segundo tipo en x_0 . Así $f'(x_0-) = \lim_{x \uparrow x_0} f'(x)$ y $f'(x_0+) = \lim_{x \downarrow x_0} f'(x)$ existen. Luego, como $f'(x_0)$ existe, por el teorema del valor medio se tiene que

$$f'(x_0) = \lim_{x \uparrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \uparrow x_0} f'(y) = f'(x_0-),$$

lo cual prueba que f' es continua por la izquierda en x_0 . Análogamente se ve que f' es continua por la derecha en x_0 , y así f' es continua en x_0 .

Una clase de funciones de gran importancia en el cálculo estocástico son las llamadas funciones regulares.

Definición 1.5. Sea $f : [a, b] \rightarrow \mathbb{R}$ una función.

- a) Se dice que f es una **función regular** si tiene límites por la izquierda y por la derecha en cada $x \in (a, b)$ y si los límites por la derecha en a y por la izquierda en b existen.
- b) Si f es una función regular que es continua por la izquierda en cada $x \in (a, b]$, se dice que f es una **función càglàd**.
- c) Si f es una función regular que es continua por la derecha en cada $x \in [a, b)$, se dice que f es una **función càdlàg**.

Note que si f es una función regular y tiene discontinuidades, entonces por definición, éstas son de salto. Así, por la proposición 1.1, el conjunto formado por estas discontinuidades es a lo más contable.

1.4. Variación de una función

Recuerde que una partición $\mathcal{P}$ de $[a, b]$ es un conjunto finito de puntos $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ con $a = x_0 < x_1 < \dots < x_n = b$. Al número $|\mathcal{P}| = \max_{1 \leq i \leq n} (x_i - x_{i-1})$ se le llama la **mall**a de la partición $\mathcal{P}$.

La **variación** de una función $f : [a, b] \rightarrow \mathbb{R}$ sobre una partición $\mathcal{P} = \{x_0, \dots, x_n\}$ de $[a, b]$ es el número

$$V(f, \mathcal{P}) := \sum_{i=1}^n |f(x_i) - f(x_{i-1})|,$$

y la **variación** de f sobre $[a, b]$ se define como

$$V(f, [a, b]) := \sup V(f, \mathcal{P}),$$

donde el supremo es tomado sobre todas las particiones $\mathcal{P}$ de $[a, b]$. Si $V(f, [a, b])$ es finito, se dice que f es una **función de variación acotada** sobre $[a, b]$. Observe que si $\mathcal{P}_1$ es un refinamiento de una partición $\mathcal{P}$, es decir, $\mathcal{P} \subseteq \mathcal{P}_1$, entonces $V(f, \mathcal{P}) \leq V(f, \mathcal{P}_1)$. Por lo tanto, se tiene que

$$V(f, [a, b]) = \lim_{|\mathcal{P}| \rightarrow 0} V(f, \mathcal{P}).$$

Definición 1.6. La **función variación** de una función $f : [0, \infty) \rightarrow \mathbb{R}$ es definida en $[0, \infty)$ por $V_f(t) = V(f, [0, t])$. La función f es de **variación finita** sobre $[0, \infty)$ si $V_f(t) < \infty$ para cada $t \in [0, \infty)$, y es de **variación acotada** sobre $[0, \infty)$ si $\sup_{t \in [0, \infty)} V_f(t) < \infty$.

Proposición 1.3. Sea f una función continua en $[0, \infty)$. Si f' existe y es acotada en $(0, \infty)$, entonces f es de variación finita.

Demostración. Sea $\mathcal{P} = \{t_0, t_1, \dots, t_n\}$ una partición de $[0, t]$. Usando el teorema del valor medio, se tiene que para cada $i = 1, \dots, n$, existe $s_i \in (t_{i-1}, t_i)$ tal que $f(t_i) - f(t_{i-1}) = f'(s_i)(t_i - t_{i-1})$. Luego, tomando $M > 0$ tal que $|f'(s)| \leq M$ para toda $s \in (0, \infty)$ se obtiene que

$$V(f, \mathcal{P}) = \sum_{i=1}^n |f'(s_i)| (t_i - t_{i-1}) \leq M \sum_{i=1}^n (t_i - t_{i-1}) = tM < \infty,$$

de lo cual se sigue que $V_f(t) < \infty$ para cada $t \in [0, \infty)$. $\blacktriangle$

Teorema 1.4. (Descomposición de Jordan) Una función $f : [0, \infty) \rightarrow \mathbb{R}$ es de variación finita si, y solo si, puede ser expresada como la diferencia de dos funciones no decrecientes.

Es sencillo ver que V_f y $V_f - f$ son funciones no decrecientes. De modo que, una descomposición en el teorema 1.4 está dada por $f = V_f - (V_f - f)$. Para una demostración completa véase por ejemplo, [1] pág. 132.

Como consecuencia de la proposición 1.1 y el teorema 1.4 se tiene el siguiente resultado.

Proposición 1.4. Sea $f : [0, \infty) \rightarrow \mathbb{R}$ una función de variación finita. El conjunto formado por las discontinuidades de f es a lo más contable. Más aún, todas las discontinuidades son de salto.

Definición 1.7. Sean $f, g : [0, \infty) \rightarrow \mathbb{R}$ funciones.

a) La **variación cuadrática** de f sobre $[0, t]$ es definida como el límite (cuando exista)

$$[f, f](t) = \lim_{|\mathcal{P}| \rightarrow 0} \sum_{i=1}^n [f(t_i) - f(t_{i-1})]^2,$$

1 donde el límite es tomado sobre las particiones $\mathcal{P} = \{t_0, t_1, \dots, t_n\}$ de $[0, t]$ con $|\mathcal{P}| \rightarrow 0$.

b) La **covariación cuadrática** de f y g sobre $[0, t]$ es definida mediante la **identidad de polarización**

$$[f, g](t) = \frac{1}{2} \{ [f + g, f + g](t) - [f, f](t) - [g, g](t) \}.$$

Note que la variación cuadrática de f , que será denotada simplemente por $[f]$, es una función no decreciente en $[0, \infty)$, y así por el teorema 1.4, ésta es de variación finita sobre $[0, \infty)$. En [18] pág. 8 se demuestra que si f es de variación finita y continua, entonces $[f] = 0$. Ligeras modificaciones se pueden realizar en esta prueba para obtener el siguiente resultado.

Proposición 1.5. Sean $f, g : [0, \infty) \rightarrow \mathbb{R}$ funciones. Si f es continua y g es de variación finita, entonces $[f, g](t) = 0$.

1.5. Integral de Riemann-Stieltjes

Sean $f, F : [a, b] \rightarrow \mathbb{R}$ funciones acotadas. Para una partición $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ de $[a, b]$ y una elección de puntos $y_i \in [x_{i-1}, x_i]$, $i = 1, 2, \dots, n$, se define una **suma de Riemann-Stieltjes** de f respecto de F por

$$S(f, F, \mathcal{P}) := \sum_{i=1}^n f(y_i) [F(x_i) - F(x_{i-1})].$$

Definición 1.8. Se dice que f es **Riemann-Stieltjes integrable** respecto de F en $(a, b]$ si existe un número real, denotado por $\int_a^b f dF$ o por $\int_a^b f(x) dF(x)$, con la propiedad de que, para cada $\epsilon > 0$, existe $\delta > 0$ tal que $|S(f, F, \mathcal{P}) - \int_a^b f dF| < \epsilon$ para cualquier partición $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ de $[a, b]$ con $|\mathcal{P}| < \delta$ y para cualquier elección de puntos $y_i \in [x_{i-1}, x_i]$, $i = 1, 2, \dots, n$. El número $\int_a^b f dF$ es llamado la **integral de Riemann-Stieltjes** de f respecto de F en $(a, b]$, y en este caso se escribe

$$\int_a^b f dF = \lim_{|\mathcal{P}| \rightarrow 0} S(f, F, \mathcal{P}).$$

Observe que en el caso particular en que $F(x) = x$ se tiene la definición de que una función sea Riemann integrable.

La siguiente proposición brinda una condición suficiente para reducir una integral de Riemann-Stieltjes a una integral de Riemann. Su prueba puede verse, por ejemplo, en [7] pág. 232.

Proposición 1.6. Sean $f, F : [a, b] \rightarrow \mathbb{R}$ funciones acotadas. Si F' existe y es Riemann integrable en $[a, b]$, entonces f es Riemann-Stieltjes integrable respecto de F en $(a, b]$ si, y solo si, fF' es Riemann integrable en $[a, b]$. En cuyo caso,

$$\int_a^b f(x) dF(x) = \int_a^b f(x) F'(x) dx.$$

Considere una función $F : [0, t] \rightarrow \mathbb{R}$ de variación acotada. Se ha mencionado que $F = V_F - (V_F - F)$, donde las funciones V_F y $V_F - F$ son no decrecientes. Usando el teorema 14.16 de [7] se puede deducir que si $f : [0, t] \rightarrow \mathbb{R}$ es una función acotada y

$$V(t) = \int_0^t |f| |dF| := \int_0^t |f| dV_F < \infty,$$

entonces f es Riemann-Stieltjes integrable respecto de F en $(0, t]$ y se tiene que

$$\int_0^t f dF = \int_0^t f dV_F - \int_0^t f d(V_F - F).$$

Cuando la función f es Riemann- Stieltjes integrable respecto a la función de variación acotada F , a $V(t)$ se le llama la **variación de la integral**.

Puede demostrarse (véase por ejemplo, [1] pág. 159) que la integral de Riemann-Stieltjes de una función continua respecto de una función de variación acotada existe. Desafortunadamente, ésta no necesariamente existe cuando se integra respecto de una función de variación infinita. Esto es mostrado por el siguiente teorema, cuyas ideas generales de su prueba son dadas, por ejemplo, en [25] pág. 145.

Teorema 1.5. Sean $F : [a, b] \rightarrow \mathbb{R}$ una función y $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ una partición de $[a, b]$. Si $\lim_{|\mathcal{P}| \rightarrow 0} \sum_{i=1}^n f(x_{i-1})[F(x_i) - F(x_{i-1})]$ existe para cualquier función continua $f : [a, b] \rightarrow \mathbb{R}$, entonces F es de variación acotada.

Las demostraciones de las fórmulas dadas en los teoremas que se presentan a continuación pueden ser consultadas, por ejemplo, en [25] pág. 6.

Teorema 1.6. (Integración por partes) Sean $f, F : [a, b] \rightarrow \mathbb{R}$ funciones de variación acotada y continuas por la derecha. Se tiene que

$$\begin{aligned} f(b)F(b) - f(a)F(a) &= \int_a^b f(x-)dF(x) + \int_a^b F(x-)df(x) + \sum_{a < x \leq b} \Delta f(x)\Delta F(x) \\ &= \int_a^b f(x-)dF(x) + \int_a^b F(x)df(x). \end{aligned} \quad (1.3)$$

Note que por la proposición 1.4 la suma en (1.3) está bien definida. Observe además que si f es continua, esta igualdad se reduce a la fórmula usual de integración por partes

$$f(b)F(b) - f(a)F(a) = \int_a^b f(x)dF(x) + \int_a^b F(x)df(x).$$

Teorema 1.7. (Cambio de variables) Sean $f, F : [0, t] \rightarrow \mathbb{R}$ funciones tales que f es diferenciable y F es de variación acotada. Si f' y F son continuas, entonces

$$f[F(t)] - f[F(0)] = \int_0^t f'[F(s)]dF(s) = \int_{F(0)}^{F(t)} f'(u)du.$$

Con la hipótesis más débil de que F sea de variación acotada, pero continua por la derecha, se tiene que

$$f[F(t)] - f[F(0)] = \int_0^t f'[F(s-)]dF(s) + \sum_{0 < s \leq t} \{f[F(s)] - f[F(s-)] - f'[F(s-)]\Delta F(s)\}.$$

1.6. Espacios de medida

Sea Ω un conjunto no vacío. Una σ -álgebra $\mathcal{F}$ sobre Ω es una colección de subconjuntos de Ω que satisface

- a) El conjunto vacío, $\emptyset$, pertenece a $\mathcal{F}$.
- b) Si $A \in \mathcal{F}$, entonces $\Omega \setminus A \in \mathcal{F}$.
- c) Si $A_1, A_2, \dots \in \mathcal{F}$, entonces $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Al par $(\Omega, \mathcal{F})$ se le llama **espacio medible** y a los elementos $A \in \mathcal{F}$ **conjuntos medibles**. Observe que de las propiedades a), b) y c) se obtiene

- d) $\Omega = \Omega \setminus \emptyset \in \mathcal{F}$.
- e) Si $A_1, \dots, A_n \in \mathcal{F}$, entonces $\bigcup_{i=1}^n A_i = A_1 \cup \dots \cup A_n \cup \emptyset \cup \emptyset \cup \dots \in \mathcal{F}$.
- f) Si $A_1, A_2, \dots \in \mathcal{F}$, entonces $\bigcap_{n=1}^{\infty} A_n = \Omega \setminus (\bigcup_{n=1}^{\infty} \Omega \setminus A_n) \in \mathcal{F}$, y por lo tanto, $A_1, \dots, A_n \in \mathcal{F}$ implica $\bigcap_{i=1}^n A_i = A_1 \cap \dots \cap A_n \cap \Omega \cap \Omega \cap \dots \in \mathcal{F}$.
- g) Si $A, B \in \mathcal{F}$, entonces $A \setminus B = A \cap (\Omega \setminus B) \in \mathcal{F}$.

Es claro que la intersección de una colección arbitraria de σ -álgebras es una σ -álgebra.

Definición 1.9. Sea $\mathcal{A}$ una colección no vacía de subconjuntos de Ω . La intersección de todas las σ -álgebras que contienen a $\mathcal{A}$ es llamada la σ -álgebra generada por $\mathcal{A}$ o la **mínima σ -álgebra** que contiene a $\mathcal{A}$ y es denotada por $\sigma(\mathcal{A})$.

Por ejemplo, si $\Omega = \mathbb{R}$ y $\mathcal{A}$ es la colección de todos los intervalos de $\mathbb{R}$, entonces $\sigma(\mathcal{A})$ es la σ -álgebra de Borel de $\mathbb{R}$ y será denotada por $\mathcal{B}(\mathbb{R})$. Si $A \in \mathcal{B}(\mathbb{R})$, se representará por $\mathcal{B}(A)$ a la σ -álgebra $\{B \in \mathcal{B}(\mathbb{R}) : B \subseteq A\} = \{A \cap B : B \in \mathcal{B}(\mathbb{R})\}$.

Sean $(\Omega_j, \mathcal{F}_j)$, $j = 1, 2$, espacios medibles. Un **rectángulo medible** en $\Omega = \Omega_1 \times \Omega_2$ es un conjunto $A = A_1 \times A_2$, donde $A_j \in \mathcal{F}_j$, $j = 1, 2$. La σ -álgebra **producto**, $\mathcal{F}_1 \times \mathcal{F}_2$, es la σ -álgebra generada por todos los rectángulos medibles.

Definición 1.10. Sea $\mathcal{F}$ una σ -álgebra sobre Ω . Una función $\mu : \mathcal{F} \rightarrow [0, \infty]$ es una **medida** sobre $\mathcal{F}$ si es contablemente aditiva, es decir, si $A_1, A_2, \dots \in \mathcal{F}$ y $A_i \cap A_j = \emptyset$ para toda $i \neq j$, entonces

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

A la terna $(\Omega, \mathcal{F}, \mu)$ se le llama **espacio de medida**.

Por ejemplo, la **medida de Lebesgue** en $\mathbb{R}$, es la única medida sobre $\mathcal{B}(\mathbb{R})$ que a cualquier intervalo le asigna su longitud.

1.7. Funciones medibles

Definición 1.11. Sea $(\Omega, \mathcal{F})$ un espacio medible. Una función $X : \Omega \rightarrow \mathbb{R}$ es $\mathcal{F}$ -medible (o medible) si para toda $B \in \mathcal{B}(\mathbb{R})$,

$$\{X \in B\} := \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}.$$

Por ejemplo, la **función indicadora** de un conjunto $A \subseteq \Omega$, $1_A : \Omega \rightarrow \mathbb{R}$, es $\mathcal{F}$ -medible si, y solo si, $A \in \mathcal{F}$. Recuerde que $1_A(\omega) = 1$ si $\omega \in A$ y $1_A(\omega) = 0$ si $\omega \notin A$.

No es complicado demostrar (véase por ejemplo, [2] pág. 40) que la suma, la diferencia, el producto, el cociente y la composición de dos funciones medibles es medible.

Proposición 1.7. Una función $X : \Omega \rightarrow \mathbb{R}$ es $\mathcal{F}$ -medible si, y solo si, $\{X \in (a, \infty)\} \in \mathcal{F}$ para cualquier $a \in \mathbb{R}$.

Demostración. Note que la necesidad se sigue directamente de la definición 1.11. Ahora, para probar la suficiencia, suponga que $\{X \in (a, \infty)\} \in \mathcal{F}$ para cualquier $a \in \mathbb{R}$. Se puede verificar fácilmente que la colección $\mathcal{A} = \{A \subseteq \mathbb{R} : \{X \in A\} \in \mathcal{F}\}$ es una σ -álgebra sobre $\mathbb{R}$. Así, por la hipótesis, $(a, \infty) \in \mathcal{A}$ para cualquier $a \in \mathbb{R}$, y por ende, $\mathcal{A}$ contiene a todos los intervalos. Por lo tanto, debido a que $\mathcal{B}(\mathbb{R})$ es la σ -álgebra más pequeña que contiene a todos los intervalos se concluye que $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{A}$. ▲

Usando la proposición 1.7 se obtiene que si $X_n : \Omega \rightarrow \mathbb{R}$ es una sucesión de funciones medibles, entonces las funciones $\max_{n \leq k} X_n$ y $\sup_n X_n$ son medibles; pues para toda $a \in \mathbb{R}$ los conjuntos $\{\max_{n \leq k} X_n \in (a, \infty)\} = \bigcup_{n=1}^k \{\omega \in \Omega : X_n(\omega) > a\}$ y $\{\sup_n X_n \in (a, \infty)\} = \bigcup_{n=1}^{\infty} \{\omega \in \Omega : X_n(\omega) > a\}$ son medibles, debido a que cada X_n es medible. En particular, las funciones $\min_{n \leq k} X_n = -\max_{n \leq k} (-X_n)$ e $\inf_n X_n = -\sup_n (-X_n)$ son medibles. Por consiguiente, también las funciones $\limsup_{n \rightarrow \infty} X_n := \inf_n \{\sup_{m \geq n} X_m\}$ y $\liminf_{n \rightarrow \infty} X_n := \sup_n \{\inf_{m \geq n} X_m\}$ son medibles. Consecuentemente, si X_n converge puntualmente, por decir, a X , entonces X es medible; pues en este caso $\limsup_{n \rightarrow \infty} X_n = \liminf_{n \rightarrow \infty} X_n = X$.

Notación. Si X_n es una sucesión de funciones que converge puntualmente a X , se escribirá esta situación simplemente como $X_n \rightarrow X$, y si además X_n es no decreciente, se escribirá $X_n \uparrow X$ en lugar de $X_n \rightarrow X$. Finalmente, si X_n es no creciente y $X_n \rightarrow X$, se escribirá $X_n \downarrow X$.

Dada una función $X : \Omega \rightarrow \mathbb{R}$ se definen su **parte positiva** y su **parte negativa** por

$$X^+ := \max\{X, 0\} \quad \text{y} \quad X^- := \max\{-X, 0\},$$

respectivamente. Es claro que estas funciones no negativas satisfacen que $X = X^+ - X^-$. Observe además que X es medible si, y solo si, X^+ y X^- son medibles, y a su vez, de esto se deduce que si X es medible, entonces $|X| = X^+ + X^-$ también lo es.

1.8. Integral de Lebesgue

En esta sección se abordará la definición de la integral de Lebesgue de una función medible. Para ello, se considerará un espacio de medida $(\Omega, \mathcal{F}, \mu)$ fijo.

Recuerde que una función $X : \Omega \rightarrow \mathbb{R}$ es **simple** si es medible y toma un número finito de valores. Claramente, si X toma valores, por decir, $x_1, x_2, \dots, x_n$, entonces $X = \sum_{i=1}^n x_i 1_{A_i}$, donde $A_i = \{X \in \{x_i\}\} \in \mathcal{F}$ y 1_{A_i} es la función indicadora de A_i , $i = 1, 2, \dots, n$.

Definición 1.12. La **integral de Lebesgue** de la función simple X es definida como $\int_{\Omega} X d\mu = \sum_{i=1}^n x_i \mu(A_i)$.

Puede demostrarse que cualquier función medible no negativa puede aproximarse de forma creciente por funciones simples no negativas; esto puede verse, por ejemplo, en [2] pág. 40.

Definición 1.13. Sea $X : \Omega \rightarrow \mathbb{R}$ una función medible no negativa. La integral de Lebesgue de X es definida por $\int_{\Omega} X d\mu = \sup \left\{ \int_{\Omega} Y d\mu : Y \text{ es simple y } 0 \leq Y \leq X \right\}$. (Observe que este supremo siempre existe y puede ser ∞ .)

Considere ahora una función medible $X : \Omega \rightarrow \mathbb{R}$ arbitraria, no necesariamente positiva. Se vio ya que $X = X^+ - X^-$, donde las funciones no negativas X^+ y X^- son las partes positiva y negativa de X , respectivamente.

Definición 1.14. La función X es μ -**integrable** (o **integrable**) si

$$\int_{\Omega} X^+ d\mu < \infty \quad \text{y} \quad \int_{\Omega} X^- d\mu < \infty.$$

Cuando esto ocurre se define la **integral de Lebesgue** de X por

$$\int_{\Omega} X d\mu = \int_{\Omega} X^+ d\mu - \int_{\Omega} X^- d\mu.$$

Si $\int_{\Omega} X^+ d\mu = \int_{\Omega} X^- d\mu = \infty$, entonces $\int_{\Omega} X d\mu$ no está definida, en caso contrario se dice que la integral de X **existe**. La integral de X sobre un conjunto $A \in \mathcal{F}$ se define como $\int_A X d\mu = \int_{\Omega} 1_A X d\mu$.

Notación. La integral de Lebesgue sobre un conjunto $B \in \mathcal{B}(\mathbb{R})$ de una función $f : \mathbb{R} \rightarrow \mathbb{R}$, $\mathcal{B}(\mathbb{R})$ -medible, respecto a la medida de Lebesgue será escrita como $\int_B f(x) dx$.

Tres importantes resultados bien conocidos en la teoría de integración de Lebesgue son los siguientes. La prueba de cada uno de estos puede verse, por ejemplo, en [2] págs. 49-51.

Teorema 1.8. Sea $X, Y : \Omega \rightarrow \mathbb{R}$ funciones medibles y $X_n : \Omega \rightarrow \mathbb{R}$ una sucesión de funciones medibles. Se tiene que

(**Teorema de convergencia monótona**) Si $Y \leq X_n \leq X_{n+1}$ para toda n con Y integrable y $X_n \uparrow X$, entonces

$$\lim_{n \rightarrow \infty} \int_{\Omega} X_n d\mu = \int_{\Omega} X d\mu.$$

(**Lema de Fatou**) Si $X_n \geq Y$ para toda n con $\int_{\Omega} Y d\mu > -\infty$, entonces

$$\liminf_{n \rightarrow \infty} \int_{\Omega} X_n d\mu \geq \int_{\Omega} \liminf_{n \rightarrow \infty} X_n d\mu.$$

(**Teorema de convergencia dominada**) Si $|X_n| \leq Y$ para toda n con Y integrable y $X_n \rightarrow X$, entonces X es integrable y

$$\lim_{n \rightarrow \infty} \int_{\Omega} X_n d\mu = \int_{\Omega} X d\mu.$$

Las demostraciones de los siguientes teoremas pueden encontrarse, por ejemplo, en [2] págs. 102-108.

Teorema 1.9. Sea $(\Omega_1, \mathcal{F}_1, \mu_1)$ un espacio de medida con μ_1 σ -finita sobre $\mathcal{F}_1$, esto es, $\Omega_1 = \bigcup_{n=1}^{\infty} A_n$ donde $A_n \in \mathcal{F}_1$ y $\mu_1(A_n) < \infty$ para toda n . Sea $(\Omega_2, \mathcal{F}_2)$ un espacio medible. Suponga que para cada $\omega_1 \in \Omega_1$, $\mu(\omega_1, \cdot)$ es una medida sobre $\mathcal{F}_2$ y que para cada $B \in \mathcal{F}_2$, $\mu(\cdot, B)$ es $\mathcal{F}_1$ -medible. Suponga también que las medidas $\mu(\omega_1, \cdot)$ son uniformemente σ -finitas, es decir, $\Omega_2 = \bigcup_{n=1}^{\infty} B_n$ donde para algunas constantes $0 < K_n < \infty$ se cumple que $\mu(\omega_1, B_n) \leq K_n$ para toda $\omega_1 \in \Omega_1$. Entonces

(**Teorema de la medida producto**) Existe una medida única μ , σ -finita, sobre $\mathcal{F}_1 \times \mathcal{F}_2$ tal que $\mu(A \times B) = \int_A \mu(\cdot, B) d\mu_1$ para cualesquiera $A \in \mathcal{F}_1$ y $B \in \mathcal{F}_2$; a saber,

$$\mu(F) = \int_{\Omega_1} \mu(\cdot, F(\omega_1)) d\mu_1, \quad F \in \mathcal{F}_1 \times \mathcal{F}_2,$$

donde $F(\omega_1) = \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in F\}$.

(**Teorema de Fubini**) Sea $X : \Omega_1 \times \Omega_2 \rightarrow \mathbb{R}$ una función $\mathcal{F}_1 \times \mathcal{F}_2$ -medible.

1) Si X es no negativa, entonces $\int_{\Omega_2} X(\omega_1, \cdot) d\mu(\omega_1, \cdot)$ existe y es una función $\mathcal{F}_1$ -medible. Además

$$\int_{\Omega_1 \times \Omega_2} X d\mu = \int_{\Omega_1} \left(\int_{\Omega_2} X(\omega_1, \cdot) d\mu(\omega_1, \cdot) \right) d\mu_1. \quad (1.4)$$

2) Si $\int_{\Omega_1 \times \Omega_2} X d\mu$ existe (resp. es finita), entonces $\int_{\Omega_2} X(\omega_1, \cdot) d\mu(\omega_1, \cdot)$ existe (resp. es finita) para μ_1 -casi cualquier ω_1 y define una función $\mathcal{F}_1$ -medible si es tomada como cualquier función $\mathcal{F}_1$ -medible sobre el conjunto de μ_1 -medida 0. Además, en este caso también se cumple la igualdad (1.4).

Capítulo 2

Conceptos de la teoría de la probabilidad

2.1. Espacios de probabilidad filtrados

Un espacio de medida $(\Omega, \mathcal{F}, P)$ es llamado **espacio de probabilidad** si $P(\Omega) = 1$, y en este caso, a P se le llama **medida de probabilidad** y a los conjuntos $A \in \mathcal{F}$ se les llama **eventos**. Observe, usando la propiedad contablemente aditiva de la medida de probabilidad P , que

- a) $P(\emptyset) = 0$.
- b) Si $A_1, \dots, A_n \in \mathcal{F}$ y $A_i \cap A_j = \emptyset$ para toda $i \neq j$, entonces $P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$, esto es, P es finitamente aditiva. Consecuentemente,
- c) Si $A, B \in \mathcal{F}$ y $A \subseteq B$, entonces $P(A) \leq P(B)$.
- d) Si $A, B \in \mathcal{F}$, entonces $P(A \setminus B) = P(A) - P(A \cap B)$. En particular, $P(\Omega \setminus B) = 1 - P(B)$.

Las propiedades de continuidad de una medida de probabilidad son demostradas en la siguiente proposición.

Proposición 2.1. *Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad. Se tiene que*

- 1) Si $A_1, A_2, \dots \in \mathcal{F}$ y $A_n \subseteq A_{n+1}$ para toda n , entonces $P(\bigcup_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} P(A_n)$.
- 2) Si $A_1, A_2, \dots \in \mathcal{F}$ y $A_{n+1} \subseteq A_n$ para toda n , entonces $P(\bigcap_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} P(A_n)$.

Demostración. 1) Para cada $n = 1, 2, \dots$, sea $B_n = A_n \setminus A_{n-1}$, donde $A_0 = \emptyset$. Es claro que $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n$ y $B_i \cap B_j = \emptyset$ para toda $i \neq j$. De este modo, por la propiedad

contablemente aditiva de P , b) y el hecho de que $A_n = \bigcup_{i=1}^n B_i$ se obtiene que

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = P\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} P(B_n) = \lim_{n \rightarrow \infty} \sum_{i=1}^n P(B_i) = \lim_{n \rightarrow \infty} P\left(\bigcup_{i=1}^n B_i\right) = \lim_{n \rightarrow \infty} P(A_n).$$

2) Se sigue de las leyes de De Morgan, del inciso 1) de esta proposición y d). $\blacktriangle$

Definición 2.1. Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad.

a) $(\Omega, \mathcal{F}, P)$ es **completo** si para cada $A \in \mathcal{F}$ con $P(A) = 0$ se tiene que $B \in \mathcal{F}$ para todo $B \subseteq A$.

b) Una **filtración** sobre $(\Omega, \mathcal{F})$ es una colección de sub- σ -álgebras $\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$ de $\mathcal{F}$ con $\mathcal{F}_s \subseteq \mathcal{F}_t$ para $s \leq t$.

c) Si $\mathbb{F}$ es una filtración, a $(\Omega, \mathcal{F}, \mathbb{F}, P)$ se le llama **espacio de probabilidad filtrado**.

2.2. Probabilidad condicional e independencia de σ -álgebras

En esta sección y en las secciones posteriores de este capítulo se considerará un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ fijo.

Definición 2.2. Sean $A, B \in \mathcal{F}$ con $P(B) > 0$. La **probabilidad condicional** de A dado B es definida por

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \quad (2.1)$$

Note que para cada $B \in \mathcal{F}$ fijo con $P(B) > 0$, la probabilidad condicional $P(\cdot|B)$ es una medida de probabilidad sobre $\mathcal{F}$. Por otro lado, para cada $A \in \mathcal{F}$ fijo, la probabilidad condicional $P(A|\cdot)$ satisface la llamada **fórmula de probabilidad total**.

Teorema 2.1. (Fórmula de probabilidad total) Sea $A \in \mathcal{F}$ y sea $B_1, B_2, \dots \in \mathcal{F}$ una partición contable de Ω , es decir, $\Omega = \bigcup_n B_n$ y $B_i \cap B_j = \emptyset$ para toda $i \neq j$. Si $P(B_n) > 0$ para toda n , entonces

$$P(A) = \sum_n P(A|B_n)P(B_n).$$

Demostración. Note que $A = A \cap \Omega = A \cap (\bigcup_n B_n) = \bigcup_n (A \cap B_n)$ y $(A \cap B_i) \cap (A \cap B_j) = A \cap (B_i \cap B_j) = \emptyset$ para toda $i \neq j$. Por lo tanto, el resultado se sigue de (2.1) y del hecho de que $P(A) = P(\bigcup_n (A \cap B_n)) = \sum_n P(A \cap B_n)$. $\blacktriangle$

Usando el teorema 2.1 se deduce fácilmente el siguiente resultado.

Teorema 2.2. (Regla ⁴ de Bayes) Sea $B_1, B_2, \dots \in \mathcal{F}$ una partición contable de Ω tal que $P(B_n) > 0$ para toda n . Entonces para cualquier $A \in \mathcal{F}$ con $P(A) > 0$ se tiene que

$$P(B_m|A) = \frac{P(A|B_m)P(B_m)}{\sum_n P(A|B_n)P(B_n)}.$$

Intuitivamente, parece natural decir que un evento A es independiente de otro evento B si el saber que ha ocurrido B no influye en la probabilidad de ocurrencia de A ; esto es, $P(A|B) = P(A)$. Por (2.1), esta relación implica la igualdad $P(A \cap B) = P(A)P(B)$, la cual es usualmente tomada como la definición de independencia de dos eventos. La ventaja práctica de ésta es que no es necesario pedir que $P(B) > 0$.

Definición 2.3. a) Los eventos $A_1, \dots, A_n$ son **independientes** si para todo subconjunto $\{\alpha_1, \dots, \alpha_k\}$ de $\{1, \dots, n\}$, $k = 1, \dots, n$, se satisface que

$$P\left(\bigcap_{i=1}^k A_{\alpha_i}\right) = \prod_{i=1}^k P(A_{\alpha_i}).$$

b) Una colección arbitraria de eventos es **independiente** si cualquier sub-colección finita es independiente.

Definición 2.4. a) Las sub- σ -álgebras $\mathcal{F}_1, \dots, \mathcal{F}_n$ de $\mathcal{F}$ son **independientes** si cualesquiera n eventos $A_1 \in \mathcal{F}_1, \dots, A_n \in \mathcal{F}_n$ son independientes.

b) Una colección arbitraria de sub- σ -álgebras es **independiente** si cualquier sub-colección finita es independiente.

2.3. Variables aleatorias e independencia de variables aleatorias

Definición 2.5. Una **variable aleatoria** $X : \Omega \rightarrow \mathbb{R}$ sobre $(\Omega, \mathcal{F}, P)$ es una función $\mathcal{F}$ -medible.

²² La **distribución de probabilidad** de una variable aleatoria X es la medida de probabilidad sobre $\mathcal{B}(\mathbb{R})$ definida por

$$P_X(B) := P(X \in B).$$

La función $F_X : \mathbb{R} \rightarrow [0, 1]$ dada por

¹⁸
$$F_X(x) := P(X \leq x)$$

es llamada la **función de distribución** de X . Se sigue de las propiedades de la medida de probabilidad P que F_X es no decreciente, continua por la derecha, $\lim_{x \rightarrow \infty} F_X(x) = 1$ y $\lim_{x \rightarrow -\infty} F_X(x) = 0$. Debido a la monotonía, F_X es de variación finita.

Definición 2.6. Sea X una variable aleatoria sobre $(\Omega, \mathcal{F}, P)$.

- a) Si existe una función $f_X : \mathbb{R} \rightarrow [0, \infty)$, $\mathcal{B}(\mathbb{R})$ -medible, tal que para cualquier $B \in \mathcal{B}(\mathbb{R})$, $P_X(B) = \int_B f_X(x) dx$ se dice que X es **absolutamente continua** y f_X es llamada la **densidad** de X .
- b) Si existe una colección contable $x_1, x_2, \dots$ de números reales con $x_i \neq x_j$ para toda $i \neq j$ tal que para cualquier $B \in \mathcal{B}(\mathbb{R})$, $P_X(B) = \sum_{x_i \in B} P(X = x_i)$ se dice que X es **discreta** con valores $x_1, x_2, \dots$. La función $f_X : \mathbb{R} \rightarrow [0, 1]$ dada por $f_X(x) := P(X = x)$ se llama **función de probabilidad** de X .

Observe que si una variable aleatoria X es absolutamente continua y tiene densidad f_X continua en x , entonces $F_X(x) = \int_{-\infty}^x f_X(t) dt$ y $F'_X(x) = f_X(x)$.

Definición 2.7. Un **vector aleatorio** n -dimensional sobre $(\Omega, \mathcal{F}, P)$ es una n -ada $\mathbf{X} = (X_1, \dots, X_n)$ de variables aleatorias.

La función de distribución de un vector aleatorio $\mathbf{X} = (X_1, \dots, X_n)$ es la función $F_{\mathbf{X}} : \mathbb{R}^n \rightarrow [0, 1]$ definida como

$$F_{\mathbf{X}}(x_1, \dots, x_n) := P(X_1 \leq x_1, \dots, X_n \leq x_n).$$

Esta función es también conocida como la **función de distribución conjunta** de las variables aleatorias $X_1, \dots, X_n$.

Definición 2.8. Las variables aleatorias $X_1, \dots, X_n$ sobre $(\Omega, \mathcal{F}, P)$ son **independientes** si para cualesquiera $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$,

$$P(X_1 \in B_1, \dots, X_n \in B_n) = P(X_1 \in B_1) \cdots P(X_n \in B_n).$$

Note que si $X_1, \dots, X_n$ son variables aleatorias independientes y $f_1, \dots, f_n : \mathbb{R} \rightarrow \mathbb{R}$ son funciones $\mathcal{B}(\mathbb{R})$ -medibles, las variables aleatorias $f_1(X_1), \dots, f_n(X_n)$ son independientes.

La demostración del siguiente resultado puede verse, por ejemplo, en [2] pág. 179.

Proposición 2.2. Sean $X_1, \dots, X_n$ variables aleatorias sobre $(\Omega, \mathcal{F}, P)$ y $\mathbf{X} = (X_1, \dots, X_n)$. Entonces $X_1, \dots, X_n$ son independientes si, y solo si,

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

2.4. Filtraciones generadas por variables aleatorias

Definición 2.9. a) La σ -álgebra generada por una variable aleatoria X es

$$\sigma(X) = \{\{X \in B\} : B \in \mathcal{B}(\mathbb{R})\}.$$

b) La σ -álgebra generada por una colección $\{X_t : t \in I \subseteq \mathbb{R}\}$ de variables aleatorias es definida por

$$\sigma(X_t : t \in I) = \sigma\left(\bigcup_{t \in I} \sigma(X_t)\right).$$

Proposición 2.3. Sean $X_1, \dots, X_n$ variables aleatorias. Entonces $X_1, \dots, X_n$ son independientes si, y solo si, $\sigma(X_1), \dots, \sigma(X_n)$ son independientes.

Demostración. Sean $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$. La suficiencia se sigue del hecho de que la independencia de $\sigma(X_1), \dots, \sigma(X_n)$ implica, en particular, que $P(X_1 \in B_1, \dots, X_n \in B_n) = P(X_1 \in B_1) \cdots P(X_n \in B_n)$. Ahora, para probar la necesidad, observe en primer lugar que la independencia de $X_1, \dots, X_n$ implica la independencia de $X_1, \dots, X_i$ para $i < n$; pues $P(X_1 \in B_1, \dots, X_i \in B_i) = P(X_1 \in B_1, \dots, X_i \in B_i, X_{i+1} \in \mathbb{R}, \dots, X_n \in \mathbb{R}) = P(X_1 \in B_1) \cdots P(X_n \in B_n)$. De este modo, para demostrar que $\sigma(X_1), \dots, \sigma(X_n)$ son independientes basta con mostrar que $P(X_1 \in B_1, \dots, X_n \in B_n) = P(X_1 \in B_1) \cdots P(X_n \in B_n)$, lo cual es cierto debido a que las variables aleatorias $X_1, \dots, X_n$ son independientes. $\blacktriangle$

³ Se hablará algunas veces de independencia entre una variable aleatoria y una σ -álgebra. Se dice que una variable aleatoria X es independiente de una sub- σ -álgebra $\mathcal{G}$ de $\mathcal{F}$ si $\sigma(X)$ y $\mathcal{G}$ son independientes.

Definición 2.10. La filtración $\{\mathcal{F}_t\}_{t \geq 0}$ con $\mathcal{F}_t = \sigma(X_s : s \leq t)$ es conocida como la **filtración generada** por la colección $\{X_t : t \geq 0\}$ de variables aleatorias.

2.5. Esperanza de una variable aleatoria

Definición 2.11. Sea X una variable aleatoria sobre $(\Omega, \mathcal{F}, P)$. La **esperanza** de X es definida por

$$EX = \int_{\Omega} X dP$$

siempre que la integral exista.

Sean X y Y variables aleatorias. Las siguientes son algunas propiedades que posee la esperanza (véase por ejemplo, [2] págs. 42, 46, 48 y 49).

- 1) EX es finita si, y solo si, $E|X|$ es finita.
- 2) Si EX y EY son finitas y $a, b \in \mathbb{R}$, entonces $E(aX + bY) = aEX + bEY$.
- 3) Si $X \leq Y$ y EX y EY son finitas, entonces $EX \leq EY$.
- 4) Si $|X| \leq Y$ con EY finita, entonces EX es finita.
- 5) Si $X \geq 0$, entonces $EX = 0$ si, y solo si, $P(X = 0) = 1$.

Proposición 2.4. Sean X una variable aleatoria y $g : \mathbb{R} \rightarrow \mathbb{R}$ una función $\mathcal{B}(\mathbb{R})$ -medible tales que $E[g(X)]$ es finita. Entonces

$$E[g(X)] = \int_{\mathbb{R}} g(x) dP_X(x) := \int_{\mathbb{R}} g(x) dF_X(x). \quad (2.2)$$

Demostración. Note que (2.2) es válida para funciones indicadoras, ya que si g es una función indicadora, $g = 1_B$ donde $B \in \mathcal{B}(\mathbb{R})$, entonces

$$E[1_B(X)] = E(1_{\{X \in B\}}) = P(X \in B) = P_X(B) = \int_{\mathbb{R}} 1_B dP_X.$$

Ahora, si g es una función simple, $g = \sum_{i=1}^n a_i 1_{B_i}$, donde $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$ forman una partición de $\mathbb{R}$, entonces

$$E[g(X)] = E\left[\sum_{i=1}^n a_i 1_{B_i}(X)\right] = \sum_{i=1}^n a_i E[1_{B_i}(X)] = \sum_{i=1}^n a_i \int_{\mathbb{R}} 1_{B_i} dP_X = \int_{\mathbb{R}} g dP_X,$$

y así, de nuevo, (2.2) se satisface. De esta manera, si g es no negativa y g_n es una sucesión de funciones simples tales que $0 \leq g_n \leq g_{n+1}$ y $g_n \uparrow g$, entonces por el teorema de convergencia monótona se tiene que

$$E[g(X)] = E\left[\lim_{n \rightarrow \infty} g_n(X)\right] = \lim_{n \rightarrow \infty} E[g_n(X)] = \lim_{n \rightarrow \infty} \int_{\mathbb{R}} g_n dP_X = \int_{\mathbb{R}} g dP_X.$$

Finalmente, si g es una función de Borel, entonces $g = g^+ - g^-$ y se obtiene que

$$E[g(X)] = E[g^+(X)] - E[g^-(X)] = \int_{\mathbb{R}} g^+ dP_X - \int_{\mathbb{R}} g^- dP_X = \int_{\mathbb{R}} g dP_X. \quad \blacktriangle$$

Proposición 2.5. (Desigualdad de Chebyshev) Sean $0 < r < \infty$ y $0 < \delta < \infty$. Si X es una variable aleatoria no negativa, entonces

$$P(X \geq \delta) \leq \frac{E(X^r)}{\delta^r}.$$

Demostración. $E(X^r) \geq E(X^r 1_{\{X \geq \delta\}}) \geq E(\delta^r 1_{\{X \geq \delta\}}) = \delta^r P(X \geq \delta)$. ▲

La prueba del siguiente resultado puede ser consultada, por ejemplo, en [10] pág. 46.

Proposición 2.6. Sean $X_1, \dots, X_n$ variables aleatorias independientes. Si $X_i \geq 0$ para toda i o si EX_i es finita para toda i , entonces $E(X_1 \cdots X_n)$ existe y

$$E(X_1 \cdots X_n) = EX_1 \cdots EX_n.$$

Sean X y Y variables aleatorias. Si éstas tienen esperanzas finitas y XY tiene también esperanza finita, se define la **covarianza** de X y Y por

$$\text{Cov}(X, Y) := E[(X - EX)(Y - EY)] = E(XY) - (EX)(EY).$$

La **varianza** de X es la covarianza consigo misma, esto es, $\text{Var}(X) := \text{Cov}(X, X)$. La desigualdad de Cauchy-Schwarz $(E[XY])^2 \leq E(X^2)E(Y^2)$ garantiza que la covarianza existe para variables aleatorias cuadrado integrables. Note de la proposición 2.6 que si X y Y son independientes, entonces son **incorrelacionadas**, es decir, $\text{Cov}(X, Y) = 0$.

2.6. Esperanza condicional

Definición 2.12. Sean X una variable aleatoria con esperanza finita y $\mathcal{G}$ una sub- σ -álgebra de $\mathcal{F}$. La **esperanza condicional** de X dada $\mathcal{G}$, $E(X|\mathcal{G})$, es una variable aleatoria $\mathcal{G}$ -medible tal que para cualquier $A \in \mathcal{G}$,

$$\int_A E(X|\mathcal{G}) dP = \int_A X dP.$$

Si X y Y son variables aleatorias con $EX < \infty$, la **esperanza condicional de X dada Y** , $E(X|Y)$, es definida como $E(X|\sigma(Y))$. La existencia y la unicidad casi segura de la esperanza condicional son garantizadas por el teorema de Radon-Nikodym (véase por ejemplo, [10] pág. 412). Que una condición se satisfaga casi seguramente, lo cual se abreviará como **c.s.**, significa que existe un evento A con $P(A) = 0$ y la condición se satisface en $\Omega \setminus A$.

La esperanza condicional satisface un número importante de propiedades, algunas son dadas a continuación. Debido a que la esperanza condicional es única casi seguramente, todas las igualdades y desigualdades que se presentan enseguida deben entenderse en el sentido casi seguro. Las respectivas pruebas pueden ser consultadas, por ejemplo, en [5] págs. 29-31.

Proposición 2.7. Sean X y Y variables aleatorias con $EX < \infty$ y sea $\mathcal{G}$ una sub- σ -álgebra de $\mathcal{F}$. Se tiene que

- 1) Si $EY < \infty$ y $a, b \in \mathbb{R}$, entonces $E(aX + bY|\mathcal{G}) = aE(X|\mathcal{G}) + bE(Y|\mathcal{G})$.

- 2) Si $X \geq 0$, entonces $E(X|\mathcal{G}) \geq 0$.
- 3) Si Y es $\mathcal{G}$ -medible y $E(XY) < \infty$, entonces $E(XY|\mathcal{G}) = YE(X|\mathcal{G})$.
- 4) Si X es independiente de $\mathcal{G}$, entonces $E(X|\mathcal{G}) = EX$.
- 5) Si $\mathcal{H}$ es una sub- σ -álgebra de $\mathcal{G}$, entonces $E[E(X|\mathcal{G})|\mathcal{H}] = E(X|\mathcal{H})$.
- 6) Si X es $\mathcal{G}$ -medible, entonces $E(X|\mathcal{G}) = X$.
- 7) $E[E(X|\mathcal{G})] = EX$.

En [2] pá. 253 se demuestra que si $g : \mathbb{R} \mapsto \mathbb{R}$ es una **función convexa**, es decir, para cualesquiera $x, y \in \mathbb{R}$ y $\lambda \in [0, 1]$, $g[\lambda x + (1 - \lambda)y] \leq \lambda g(x) + (1 - \lambda)g(y)$, entonces existen sucesiones a_n y b_n de números reales tales que $g(x) = \sup\{a_n x + b_n : n \in \mathbb{N}\}$ para toda $x \in \mathbb{R}$.

Teorema 2.3. (Desigualdad de Jensen) Sean $g : \mathbb{R} \rightarrow \mathbb{R}$ una función convexa y X una variable aleatoria con $EX < \infty$ tales que $E[g(X)] < \infty$. Entonces para cualquier sub- σ -álgebra $\mathcal{G}$ de $\mathcal{F}$ se tiene que

$$E[g(X)|\mathcal{G}] \geq g(E[X|\mathcal{G}]) \quad \text{c.s.}$$

Demostración. Sea $\mathcal{G}$ una sub- σ -álgebra de $\mathcal{F}$. Como g es una función convexa, existen sucesiones a_n y b_n de números reales tales que $g(x) = \sup\{a_n x + b_n : n \in \mathbb{N}\}$ para toda $x \in \mathbb{R}$. Entonces, para un n fijo, $g(X(\omega)) \geq a_n X(\omega) + b_n$ para toda $\omega \in \Omega$; esto es, $g(X(\omega)) - (a_n X(\omega) + b_n) \geq 0$ para toda $\omega \in \Omega$. Luego, por 1) y 2) de la proposición 2.7 se obtiene que

$$E[g(X)|\mathcal{G}](\omega) \geq E[a_n X + b_n|\mathcal{G}](\omega) = a_n E[X|\mathcal{G}](\omega) + b_n$$

para toda $\omega \in \Omega \setminus A_n$, donde $P(A_n) = 0$. De este modo, debido a que $P(A) = 0$, donde $A = \bigcup_n A_n$, se sigue que para toda $n = 1, 2, \dots$

$$E[g(X)|\mathcal{G}](\omega) \geq a_n E[X|\mathcal{G}](\omega) + b_n \quad \text{para toda } \omega \in \Omega \setminus A.$$

Así, tomando el supremo en el lado derecho de esta última desigualdad se concluye que $E[g(X)|\mathcal{G}] \geq g(E[X|\mathcal{G}])$ c.s. $\blacktriangle$

2.7. Función generadora de momentos y función característica

Sean X una variable aleatoria y $r > 0$. Si $E(X^r) < \infty$, es llamado el **r -ésimo momento** de X . Si existe $\delta > 0$ tal que para toda $u \in (-\delta, \delta)$, $E(e^{uX}) < \infty$, se define la **función**

generadora de momentos de X por

$$m_X(u) := E(e^{uX}), \quad u \in (-\delta, \delta).$$

Debido a que $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ y haciendo uso del teorema de convergencia dominada se tiene que

$$m_X(u) = E\left(\sum_{n=0}^{\infty} \frac{u^n X^n}{n!}\right) = \sum_{n=0}^{\infty} \frac{u^n}{n!} E(X^n).$$

De esta manera, diferenciando término a término, n veces, en esta última suma y evaluando en $u = 0$ se obtiene $E(X^n)$. La existencia de la función generadora de momentos de X en una vecindad de 0 implica la existencia de todos los momentos de X . Si algún momento de X no existe, entonces su función generadora de momentos no existe. La función generadora de momentos determina unívocamente a la distribución, como lo muestra el siguiente resultado. Su prueba puede verse, por ejemplo, en [4] pág. 408.

Proposición 2.8. Sean X y Y variables aleatorias. Si $m_X(u)$ y $m_Y(u)$ existen y $m_X(u) = m_Y(u)$ para toda $u \in (-\delta, \delta)$, para algún $\delta > 0$, entonces $F_X(u) = F_Y(u)$ para toda $u \in \mathbb{R}$.

La **función característica** de una variable aleatoria X es definida por

$$h_X(u) := E(e^{iuX}) = E(\cos uX) + iE(\sin uX),$$

donde $i = \sqrt{-1}$. Análogamente a la función generadora de momentos, la función característica determina unívocamente a la distribución, ya que se puede demostrar (véase por ejemplo, [2] pág. 295) que si P_1 y P_2 son medidas de probabilidad sobre $\mathcal{B}(\mathbb{R})$ y $\int_{\mathbb{R}} e^{iux} dP_1(x) = \int_{\mathbb{R}} e^{iux} dP_2(x)$ para toda $u \in \mathbb{R}$, entonces $P_1 \equiv P_2$. La ventaja de la función característica sobre la función generadora de momentos es que existe para cualquier variable aleatoria, puesto que las funciones $\cos x$ y $\sin x$ son acotadas en $\mathbb{R}$, mientras que la función e^x no es acotada, y así, la función generadora de momentos pudiera no existir.

2.8. Distribuciones gaussiana, gamma y exponencial

Una variable aleatoria X tiene una **distribución gaussiana** con parámetros $\mu \in \mathbb{R}$ y $\sigma^2 > 0$, denotado por $X \sim N(\mu, \sigma^2)$, si tiene densidad dada por

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

Usando que $\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}$ y $\int_{\mathbb{R}} x^2 e^{-x^2} dx = \sqrt{\pi}/2$ resulta sencillo ver que $EX = \mu$ y $Var(X) = \sigma^2$. Note además que $Z = (X - \mu)/\sigma$ tiene una **distribución normal estándar**,

esto es, $Z \sim N(0, 1)$; pues haciendo el cambio de variables $y = (x - \mu)/\sigma$ se obtiene que

$$P(Z \leq z) = P(X \leq z\sigma + \mu) = \int_{-\infty}^{z\sigma + \mu} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy.$$

No es difícil ver que la función generadora de momentos de X es

$$m_X(u) = E(e^{uX}) = \int_{\mathbb{R}} e^{ux} f_X(x) dx = e^{u\mu + \frac{u^2\sigma^2}{2}}.$$

Una importante propiedad de la familia gaussiana es que una combinación lineal de variables gaussianas independientes resulta en una variable gaussiana. Esto es establecido en el siguiente resultado, cuya prueba puede verse, por ejemplo, en [8] pág. 156.

Proposición 2.9. Sean $X_1, X_2, \dots, X_n$ variables aleatorias independientes. Si $X_i \sim N(\mu_i, \sigma_i^2)$ para toda $i = 1, 2, \dots, n$, entonces $a_1X_1 + a_2X_2 + \dots + a_nX_n \sim N(a_1\mu_1 + a_2\mu_2 + \dots + a_n\mu_n, a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \dots + a_n^2\sigma_n^2)$.

Se dice que una variable aleatoria X tiene una **distribución gamma** con parámetros $\alpha > 0$ y $\beta > 0$, escrito como $X \sim \text{Gamma}(\alpha, \beta)$, si posee densidad dada por

$$f_X(x) = \frac{x^{\alpha-1} e^{-\frac{x}{\beta}}}{\beta^\alpha \Gamma(\alpha)} 1_{(0, \infty)}(x),$$

donde $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$, es la **función gamma**. Integrando por partes se puede ver que $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$, y consecuentemente, usando el hecho de que $\Gamma(1) = 1$, se tiene por inducción que $\Gamma(n + 1) = n!$ para $n = 1, 2, \dots$. También, haciendo uso de la propiedad $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$ se obtiene fácilmente que $EX = \alpha\beta$ y $\text{Var}(X) = \alpha\beta^2$. La función generadora de momentos de X es

$$m_X(u) = E(e^{uX}) = \int_{\mathbb{R}} e^{ux} f_X(x) dx = \left(\frac{1}{1 - u\beta} \right)^\alpha \quad \text{si } u < \frac{1}{\beta}.$$

La distribución gamma con $\alpha = 1$ es mejor conocida como la **distribución exponencial**. En este caso, se dice que una variable aleatoria X tiene distribución exponencial con parámetro $\beta > 0$, denotado por $X \sim \text{Exp}(\beta)$, si tiene densidad dada por

$$f_X(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}} 1_{(0, \infty)}(x).$$

En este caso particular se tiene que $EX = \beta$, $\text{Var}(X) = \beta^2$ y $m_X(u) = 1/(1 - u\beta)$ si $u < 1/\beta$. Además, es sencillo verificar que esta distribución tiene la propiedad "sin memoria", esto es, $P(X > a + b | X > a) = P(X > b)$ para $a > 0$ y $b > 0$.

2.9. Convergencia de variables aleatorias

Las demostraciones de los resultados mencionados en esta sección pueden ser consultadas, por ejemplo, en las secciones 2.4, 2.5, 2.8, 7.1 y 7.2 de [2].

Definición 2.13. Sean $X, X_1, X_2, X_3, \dots$ variables aleatorias.

a) X_n **converge casi seguramente** a X , escrito como $X_n \rightarrow X$ c.s., si

$$P(\omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)) = 1.$$

b) X_n **converge en probabilidad** a X , denotado por $X_n \xrightarrow{P} X$, si para cualquier $\epsilon > 0$,

$$P(|X_n - X| > \epsilon) \rightarrow 0 \quad \text{cuando } n \rightarrow \infty.$$

La convergencia casi segura implica la convergencia en probabilidad y la convergencia en probabilidad implica la convergencia casi segura de alguna subsucesión, esto es, si X_n converge en probabilidad a X , entonces existe una subsucesión X_{k_n} de X_n que converge casi seguramente a X .

Definición 2.14. Sean $X, X_1, X_2, X_3, \dots$ variables aleatorias. X_n **converge en distribución** a X , escrito como $X_n \xrightarrow{d} X$, si $F_{X_n}(x)$ converge a $F_X(x)$ en cualquier punto de continuidad de F_X .

Se conoce que X_n converge en distribución a X si, y solo si, sus funciones características (o funciones generadoras de momentos) convergen a la función característica (a la función generadora de momentos) de X , es decir, si $h_{X_n}(u)$ (o $m_{X_n}(u)$) convergen a $h_X(u)$ (a $m_X(u)$). La convergencia en distribución es también equivalente a que $E[g(X_n)]$ converja a $E[g(X)]$ para todas las funciones $g : \mathbb{R} \rightarrow \mathbb{R}$ continuas y acotadas. Es conocido también que la convergencia en probabilidad es más fuerte que la convergencia en distribución, y que la convergencia en distribución a una constante es la misma que la convergencia en probabilidad a la misma constante.

Definición 2.15. Sean $X, X_1, X_2, X_3, \dots$ variables aleatorias. X_n **converge a X en L^r** , $r \geq 1$, denotado por $X_n \xrightarrow{r} X$, si $E(|X_n - X|^r) \rightarrow 0$ cuando $n \rightarrow \infty$.

Recuerde que $L^r = L^r(\Omega, \mathcal{F}, P)$ es el conjunto de todas las variables aleatorias X definidas sobre $(\Omega, \mathcal{F}, P)$ tales que $E|X|^r < \infty$.

Se sabe que la convergencia en L^r es más fuerte que la convergencia en probabilidad, es decir, si $X_n \xrightarrow{r} X$, entonces $X_n \xrightarrow{P} X$.

Capítulo 3

Conceptos de la teoría de los procesos estocásticos

3.1. Procesos estocásticos a tiempo continuo

Un **proceso estocástico a tiempo continuo** es una colección $X = \{X_t\}_{t \geq 0}$ de variables aleatorias definidas sobre un espacio de probabilidad $(\Omega, \mathcal{F}, P)$, las cuales toman valores en un espacio medible, llamado **espacio de estados**. En este trabajo, el espacio de estados considerado es $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. El índice $t \in [0, \infty)$ es interpretado como el tiempo. Para cada $\omega \in \Omega$ fijo, la función $t \mapsto X_t(\omega)$, que mapea $[0, \infty)$ en $\mathbb{R}$, es una **trayectoria** o **realización** del proceso X . El proceso estocástico X se dice ser **continuo por la derecha** si sus trayectorias son continuas por la derecha. Similarmente se define que un proceso estocástico sea **continuo por la izquierda**, **càdlàg**, **càglàd** y de **variación finita**.

Definición 3.1. Sean $X = \{X_t\}_{t \geq 0}$ y $Y = \{Y_t\}_{t \geq 0}$ procesos estocásticos sobre $(\Omega, \mathcal{F}, P)$.

a) X es una **modificación** de Y si, para cada $t \geq 0$, $P(X_t = Y_t) = 1$.

b) X y Y son **indistinguibles** si $P(X_t = Y_t \text{ para toda } t \geq 0) = 1$.

Evidentemente si dos procesos son indistinguibles, entonces uno es una modificación del otro. El recíproco de esta afirmación no es válida en general (véase por ejemplo, [17] pág. 2). No obstante, se cumple un recíproco parcial.

Proposición 3.1. Sean $X = \{X_t\}_{t \geq 0}$ y $Y = \{Y_t\}_{t \geq 0}$ procesos estocásticos sobre $(\Omega, \mathcal{F}, P)$. Si X es una modificación de Y y ambos son continuos por la derecha c.s., entonces X y Y son indistinguibles.

Demostración. Sean A y B los eventos nulos donde las trayectorias de X y Y , respectivamente, no son continuas por la derecha. Sea $C = \bigcup_{t \in Q} \{X_t \neq Y_t\}$, donde $Q = [0, \infty) \cap \mathbb{Q}$. Entonces, dado que X es una modificación de Y , $P(C) = 0$. Sea $D = A \cup B \cup C$. Luego,

$P(D) = 0$. Además, sobre $\Omega \setminus D$, $X_t = Y_t$ para toda $t \in Q$. Si $t \notin Q$, por la densidad de Q existe $t_n \in Q$ tal que $t_n \downarrow t$. Por lo tanto, debido a que tanto X como Y son continuos por la derecha sobre $\Omega \setminus D$ se concluye que $X_t = \lim_{t_n \downarrow t} X_{t_n} = \lim_{t_n \downarrow t} Y_{t_n} = Y_t$. $\blacktriangle$

Corolario 3.1. Sean X y Y procesos estocásticos càdlàg. Si X es una modificación de Y , entonces X y Y son indistinguibles.

Sea $(\Omega, \mathcal{F}, \mathbb{F}, P)$ un espacio de probabilidad filtrado. A partir de ahora se hace la convención de que el espacio de probabilidad $(\Omega, \mathcal{F}, P)$ es completo y que la filtración $\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$ es completa, esto es, $\mathcal{F}_0$ contiene a todos los eventos nulos de $\mathcal{F}$. Se acuerda además que $\mathbb{F}$ es continua por la derecha, es decir, $\mathcal{F}_t = \bigcap_{s > t} \mathcal{F}_s$ para toda $t \geq 0$.

En el contexto de los procesos estocásticos, la filtración $\mathbb{F}$ especifica cómo la información es revelada en el tiempo. La propiedad de que $\mathbb{F}$ es creciente corresponde al hecho de que la información no es olvidada. Así, para cada tiempo t , la σ -álgebra $\mathcal{F}_t$ representa la información acumulada al tiempo t .

Definición 3.2. Un proceso estocástico $X = \{X_t\}_{t \geq 0}$ es **adaptado** a la filtración $\{\mathcal{F}_t\}_{t \geq 0}$ si, para cada $t \geq 0$, X_t es $\mathcal{F}_t$ -medible.

La filtración más pequeña respecto a la cual un proceso $X = \{X_t\}_{t \geq 0}$ es adaptado es la generada por sí mismo, $\{\mathcal{F}_t^X\}_{t \geq 0}$, dada por $\mathcal{F}_t^X = \sigma(X_s : s \leq t)$. Esta filtración es conocida como la **filtración canónica** de X y es interpretada como la información generada por el proceso X .

Note que un proceso estocástico $X = \{X_t\}_{t \geq 0}$ puede ser visto como una función $X : [0, \infty) \times \Omega \rightarrow \mathbb{R}$, dada por $(t, \omega) \mapsto X_t(\omega)$, tal que toda t -sección sea $\mathcal{F}$ -medible. El proceso X es **medible** si $X : [0, \infty) \times \Omega \rightarrow \mathbb{R}$ es $\mathcal{B}([0, \infty)) \times \mathcal{F}$ -medible. En este caso el teorema de Fubini garantiza que

1. Las trayectorias de X son $\mathcal{B}([0, \infty))$ -medibles c.s.
2. Si $M(t) = EX_t$ existe para toda $t \geq 0$, entonces M es $\mathcal{B}([0, \infty))$ -medible.
3. Si I es un subintervalo de $[0, \infty)$ tal que $\int_I E|X_t| dt < \infty$, entonces las trayectorias de X son integrables c.s. y

$$\int_I EX_t dt = E \int_I X_t dt.$$

Definición 3.3. Un proceso estocástico $X = \{X_t\}_{t \geq 0}$ es **progresivamente medible** respecto a la filtración $\{\mathcal{F}_t\}_{t \geq 0}$ si, para cada $t \geq 0$, la función $(s, \omega) \mapsto X_s(\omega)$, que mapea $[0, t] \times \Omega$ en $\mathbb{R}$, es $\mathcal{B}([0, t]) \times \mathcal{F}_t$ -medible.

Proposición 3.2. Sea $X = \{X_t\}_{t \geq 0}$ un proceso estocástico adaptado a la filtración $\{\mathcal{F}_t\}_{t \geq 0}$. Si X es continuo por la derecha o continuo por la izquierda, entonces X es progresivamente medible respecto a $\{\mathcal{F}_t\}_{t \geq 0}$.

Demostración. Se dará un bosquejo de la demostración para el caso en que X es continuo por la derecha. Para $t > 0$, $n \geq 1$, $k = 0, 1, \dots, 2^n - 1$ y $0 \leq s \leq t$, defina

$$X_0^n(\omega) = X_0(\omega) \quad \text{y} \quad X_s^n(\omega) = X_{(k+1)t/2^n}(\omega) \quad \text{si} \quad \frac{kt}{2^n} < s \leq \frac{(k+1)t}{2^n}.$$

Puede demostrarse que para cada n , la función $(s, \omega) \mapsto X_s^n(\omega)$, que mapea $[0, t] \times \Omega$ en $\mathbb{R}$, es $\mathcal{B}([0, t]) \times \mathcal{F}_t$ -medible. Usando el hecho de que $\frac{k+1}{2^n}t \downarrow s$ cuando $n \rightarrow \infty$ y la continuidad por la derecha, se obtiene que $\lim_{n \rightarrow \infty} X_s^n(\omega) = X_s(\omega)$ para todo $(s, \omega) \in [0, t] \times \Omega$. Por lo tanto, la aplicación $(s, \omega) \mapsto X_s(\omega)$ es también $\mathcal{B}([0, t]) \times \mathcal{F}_t$ -medible. $\blacktriangle$

Definición 3.4. Sean $X = \{X_t\}_{t \geq 0}$ y $Y = \{Y_t\}_{t \geq 0}$ procesos estocásticos.

a) La **variación cuadrática** de X sobre $[0, t]$ es definida como el límite en probabilidad (cuando exista)

$$[X, X]_t = \lim \sum_{i=1}^n (X_{t_i} - X_{t_{i-1}})^2,$$

donde el límite es tomado sobre las particiones $\mathcal{P} = \{t_0, t_1, \dots, t_n\}$ de $[0, t]$ con $|\mathcal{P}| \rightarrow 0$.

b) La **covariación cuadrática** de X y Y sobre $[0, t]$ es definida por

$$[X, Y]_t = \frac{1}{2} ([X + Y, X + Y]_t - [X, X]_t - [Y, Y]_t).$$

3.2. Tiempos de paro y procesos estocásticos previsibles

Definición 3.5. Una variable aleatoria $\tau : \Omega \rightarrow [0, \infty]$ es un **tiempo de paro** respecto a la filtración $\{\mathcal{F}_t\}_{t \geq 0}$ si, para cada $t \geq 0$, $\{\tau \leq t\} \in \mathcal{F}_t$.

La continuidad por la derecha de la filtración garantiza el siguiente resultado.

Proposición 3.3. Una variable aleatoria τ es un tiempo de paro respecto a $\{\mathcal{F}_t\}_{t \geq 0}$ si, y solo si, $\{\tau < t\} \in \mathcal{F}_t$ para cada $t \geq 0$.

Demostración. La necesidad se sigue de la igualdad $\{\tau < t\} = \bigcup_{n=1}^{\infty} \{\tau \leq t - 1/n\} \in \mathcal{F}_t$, y la suficiencia de la igualdad $\{\tau \leq t\} = \bigcap_{n=1}^{\infty} \{\tau < t + 1/n\} \in \bigcap_{n=1}^{\infty} \mathcal{F}_{t+1/n} = \mathcal{F}_t$, por la continuidad por la derecha de la filtración. $\blacktriangle$

Dados un proceso estocástico $X = \{X_t\}_{t \geq 0}$ y un conjunto $A \in \mathcal{B}(\mathbb{R})$, se define

$$\tau_A(\omega) := \inf\{t \geq 0 : X_t(\omega) \in A\}$$

como el **primer tiempo de llegada** de X a A , donde se conviene que $\inf \emptyset = \infty$.

Proposición 3.4. Sea $X = \{X_t\}_{t \geq 0}$ un proceso càdlàg adaptado a $\{\mathcal{F}_t\}_{t \geq 0}$ y sea $A \in \mathcal{B}(\mathbb{R})$.

- 1) Si A es abierto, τ_A es un tiempo de paro.
- 2) Si A es cerrado, $\tau(\omega) = \inf\{t > 0 : X_t(\omega) \in A \text{ o } X_{t-}(\omega) \in A\}$ es un tiempo de paro.

Demostración. 1) Note que $\{\tau_A < t\} = \{X_s \in A \text{ para algún } s < t\} = \bigcup_{s \in [0, t)} \{X_s \in A\}$. Como X es continuo por la derecha y A es abierto, entonces para cualquier $s \in [0, t)$ con $X_s \in A$ existe $s' \in [s, t) \cap \mathbb{Q}$ tal que $X_{s'} \in A$. Así, $\{\tau_A < t\} = \bigcup_{s \in [0, t)} \{X_s \in A\} = \bigcup_{s \in [0, t) \cap \mathbb{Q}} \{X_s \in A\} \in \mathcal{F}_t$ y el resultado se sigue de la proposición 3.3.

2) La demostración de este inciso puede ser consultada, por ejemplo, en [24] pág. 5. $\blacktriangle$

Definición 3.6. Sea τ un tiempo de paro respecto a la filtración $\{\mathcal{F}_t\}_{t \geq 0}$.

- a) Los eventos observados antes o en el tiempo τ son descritos por la σ -álgebra $\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ para toda } t \geq 0\}$.
- b) Los eventos observados antes de τ son dados por la σ -álgebra $\mathcal{F}_{\tau-}$, la cual es generada por la colección $\{A \cap \{t < \tau\} : A \in \mathcal{F}_t, t > 0\} \cup \mathcal{F}_0$.

Dado que $[(A \cap \{t < \tau\}) \cup B] \cap \{\tau \leq s\} = (A \cap \{t < \tau\} \cap \{\tau \leq s\}) \cup (B \cap \{\tau \leq s\}) \in \mathcal{F}_s$ para cualesquiera $t, s > 0$, $A \in \mathcal{F}_t$ y $B \in \mathcal{F}_0$, se deduce que $\mathcal{F}_{\tau-} \subset \mathcal{F}_\tau$. Como también $\{t < \tau\} \in \mathcal{F}_{\tau-}$ para toda $t \geq 0$, se obtiene que τ es $\mathcal{F}_{\tau-}$ -medible.

Definición 3.7. Sean τ un tiempo de paro y $X = \{X_t\}_{t \geq 0}$ un proceso estocástico.

- a) Se define la variable aleatoria X_τ sobre el evento $\{\tau < \infty\}$ como $X_\tau(\omega) = X_{\tau(\omega)}(\omega)$.
- b) El **proceso parado** en τ , $X^\tau = \{X_{t \wedge \tau}\}_{t \geq 0}$, es definido por $X_{t \wedge \tau}(\omega) = X_{t \wedge \tau(\omega)}(\omega)$.

Observe que el proceso parado está bien definido puesto que $t \wedge \tau := \min\{t, \tau\} < \infty$.

Denote por $\mathbb{L}$ a la colección de los procesos estocásticos, $X : [0, \infty) \times \Omega \rightarrow \mathbb{R}$, continuos por la izquierda y adaptados a $\{\mathcal{F}_t\}_{t \geq 0}$. La σ -álgebra **previsible** sobre $[0, \infty) \times \Omega$, que será denotada por $\mathfrak{P}$, es la σ -álgebra generada por los procesos en $\mathbb{L}$, esto es,

$$\mathfrak{P} = \sigma(\{[s, t) \times A : s < t \text{ y } A \in \mathcal{F}_{t-}\}),$$

donde $\mathcal{F}_{t-} = \sigma(\bigcup_{s < t} \mathcal{F}_s)$ para $t > 0$ y $\mathcal{F}_{0-} = \mathcal{F}_0$. La σ -álgebra $\mathcal{F}_{t-}$ representa la información disponible antes de $t > 0$.

Definición 3.8. Un proceso estocástico $X : [0, \infty) \times \Omega \rightarrow \mathbb{R}$ es **previsible** si es $\mathfrak{P}$ -medible.

Obviamente cualquier proceso continuo por la izquierda y adaptado es previsible, en particular, cualquier proceso continuo y adaptado también lo es.

Proposición 3.5. Sean τ un tiempo de paro y $X = \{X_t\}_{t \geq 0}$ un proceso previsible. Entonces

- 1) La variable aleatoria $X_\tau 1_{\{\tau < \infty\}}$ es $\mathcal{F}_{\tau-}$ -medible.
- 2) El proceso parado X^τ es previsible.

La demostración de la proposición 3.5 puede verse, por ejemplo, en [21] pág. 13.

3.3. Martingalas, submartingalas y supermartingalas

Definición 3.9. Sea $X = \{X_t\}_{t \geq 0}$ un proceso estocástico adaptado a una filtración $\{\mathcal{F}_t\}_{t \geq 0}$ con $E|X_t| < \infty$ para toda $t \geq 0$.

- a) X es una **martingala** si, para cualesquiera $0 \leq s < t$, $E(X_t | \mathcal{F}_s) = X_s$ c.s.
- b) X es una **submartingala** si, para cualesquiera $0 \leq s < t$, $E(X_t | \mathcal{F}_s) \geq X_s$ c.s.
- c) X es una **supermartingala** si, para cualesquiera $0 \leq s < t$, $E(X_t | \mathcal{F}_s) \leq X_s$ c.s.

Observe que si X es una martingala, entonces su función media, $t \mapsto EX_t$, es constante puesto que $EX_t = E[E(X_t | \mathcal{F}_s)] = EX_s$. Similarmente se ve que la función media de una submartingala es no decreciente y la función media de una supermartingala es no creciente. Además, es claro que si X es una supermartingala, entonces $-X$ es una submartingala.

La demostración del resultado siguiente hace uso de las hipótesis usuales y puede ser encontrada, por ejemplo, en [17] pág. 16.

Proposición 3.6. Una submartingala $\{X_t\}_{t \geq 0}$ tiene una modificación càdlàg si, y solo si, su función media, $t \mapsto EX_t$, es continua por la derecha. En particular, cualquier martingala tiene una modificación càdlàg.

Debido a que todas las martingalas tienen modificaciones càdlàg, se asumirá siempre que se está considerando esta clase de modificaciones.

Proposición 3.7. Sean $X = \{X_t\}_{t \geq 0}$ un proceso estocástico y g una función convexa tales que $E|g(X_t)| < \infty$ para toda $t \geq 0$.

- 1) Si X es una martingala, entonces $g(X) := \{g(X_t)\}_{t \geq 0}$ es una submartingala.
- 2) Si X es una submartingala y g es no decreciente, entonces $g(X)$ es una submartingala.

Demostración. 1) Note que esto se sigue de la propiedad de martingala y de la desigualdad de Jensen, puesto que si $0 \leq s < t$, entonces $g(X_s) = g(E[X_t|\mathcal{F}_s]) \leq E[g(X_t)|\mathcal{F}_s]$.

2) Análogamente a 1), usando la propiedad de submartingala y la desigualdad de Jensen se tiene que $g(X_s) \leq g(E[X_t|\mathcal{F}_s]) \leq E[g(X_t)|\mathcal{F}_s]$ si $0 \leq s < t$. $\blacktriangle$

Definición 3.10. Una martingala $\{X_t\}_{t \geq 0}$ es **cerrada** por una variable aleatoria Y si $E|Y| < \infty$ y $X_t = E(Y|\mathcal{F}_t)$ para cada $t \geq 0$.

Note que la definición 3.10 está bien hecha ya que efectivamente $\{E(Y|\mathcal{F}_t)\}_{t \geq 0}$ es una martingala, pues $E[E(Y|\mathcal{F}_t)|\mathcal{F}_s] = E(Y|\mathcal{F}_s)$ si $s < t$. Las martingalas de este tipo son conocidas como **martingalas de Doob-Lévy**.

El siguiente resultado dice que una martingala parada en un tiempo de paro es de nuevo una martingala. La prueba puede verse, por ejemplo, en [25] pág. 71.

Proposición 3.8. Si $X = \{X_t\}_{t \geq 0}$ es una martingala y τ es un tiempo de paro, ambos respecto a la filtración $\{\mathcal{F}_t\}_{t \geq 0}$, entonces el proceso parado $X^\tau = \{X_{t \wedge \tau}\}_{t \geq 0}$ es una martingala respecto a $\{\mathcal{F}_t\}_{t \geq 0}$. En particular, $EX_{t \wedge \tau} = EX_0$.

Corolario 3.2. Si $\{X_t\}_{t \geq 0}$ es una martingala y τ es un tiempo de paro acotado, entonces $EX_\tau = EX_0$.

3.4. Procesos uniformemente integrables

Definición 3.11. Un proceso estocástico $X = \{X_t\}_{t \geq 0}$ es **uniformemente integrable** si $E(|X_t|1_{\{|X_t|>n\}}) \rightarrow 0$ cuando $n \rightarrow \infty$ uniformemente en t , esto es,

$$\lim_{n \rightarrow \infty} \sup_{t \geq 0} E(|X_t|1_{\{|X_t|>n\}}) = 0.$$

Observe que si $\{X_t\}_{t \geq 0}$ es uniformemente integrable, entonces es **integrable**; esto es, $\sup_{t \geq 0} E|X_t| < \infty$. En efecto, note que $E|X_t| = E(|X_t|1_{\{|X_t|>n\}}) + E(|X_t|1_{\{|X_t| \leq n\}}) \leq E(|X_t|1_{\{|X_t|>n\}}) + n$ para toda t y para toda n . Por lo tanto, por la integrabilidad uniforme, $\sup_{t \geq 0} E|X_t| \leq \sup_{t \geq 0} E(|X_t|1_{\{|X_t|>n\}}) + n < \infty$ para alguna n suficientemente grande.

Proposición 3.9. Sean $X = \{X_t\}_{t \geq 0}$ un proceso y Y una variable aleatoria con $E|Y| < \infty$. Si X es dominado por Y , es decir, $|X_t| \leq Y$ para toda $t \geq 0$, entonces X es uniformemente integrable. Así, en particular, X es uniformemente integrable si $E(\sup_{t \geq 0} |X_t|) < \infty$.

Demostración. Note que $|X_t|1_{\{|X_t|>n\}} \leq |Y|1_{\{|Y|>n\}}$ para toda $t \geq 0$. Así, debido a que $|Y|1_{\{|Y|>n\}} \rightarrow 0$ cuando $n \rightarrow \infty$, $|Y|1_{\{|Y|>n\}} \leq |Y|$ y $E|Y| < \infty$, se tiene por el teorema

de convergencia dominada que $E(|X_t|1_{\{|X_t|>n\}}) \leq E(|Y|1_{\{|Y|>n\}}) \rightarrow 0$ cuando $n \rightarrow \infty$, y el resultado se sigue. ▲

La siguiente proposición, cuya demostración puede ser consultada, por ejemplo, en [11] pág. 494, da una condición suficiente y necesaria para que un proceso sea uniformemente integrable.

Proposición 3.10. *Un proceso estocástico $\{X_t\}_{t \geq 0}$ es uniformemente integrable si, y solo si, existe una función $g : [0, \infty) \rightarrow [0, \infty)$ creciente y convexa tal que $\lim_{x \rightarrow \infty} [g(x)/x] = \infty$ y $\sup_{t \geq 0} E g(|X_t|) < \infty$.*

Corolario 3.3. *Si $X = \{X_t\}_{t \geq 0}$ es un proceso estocástico **cuadrado integrable**, esto es, $\sup_{t \geq 0} E(X_t^2) < \infty$, entonces X es uniformemente integrable.*

3.5. Algunos resultados sobre la convergencia de martingalas

En esta sección se expondrán algunos resultados concernientes a convergencia y desigualdades de martingalas. El tratamiento riguroso del contenido de esta sección puede ser visto, por ejemplo, en [22] páginas 61 y 64.

Teorema 3.1. (Convergencia de martingalas) *Si $\{X_t\}_{t \geq 0}$ es una martingala integrable, es decir, $\sup_{t \geq 0} E|X_t| < \infty$, entonces $X_\infty = \lim_{t \rightarrow \infty} X_t$ existe c.s. y $E|X_\infty| < \infty$.*

Como consecuencia de este resultado se obtiene el siguiente corolario.

Corolario 3.4. *1) Las martingalas uniformemente integrables convergen casi seguramente. En particular, las martingalas cuadrado integrables convergen casi seguramente.*

2) Las martingalas positivas convergen casi seguramente.

Demostración. 1) Esto se sigue del hecho de que los procesos uniformemente integrables son integrables.

2) Si $\{X_t\}_{t \geq 0}$ es una martingala positiva, $\sup_{t \geq 0} E|X_t| = \sup_{t \geq 0} EX_t = EX_0 < \infty$. ▲

Proposición 3.11. *Si $X = \{X_t\}_{t \geq 0}$ es una martingala uniformemente integrable, entonces converge a una variable aleatoria X_∞ c.s. y en L^1 . Recíprocamente, si $X = \{X_t\}_{t \geq 0}$ es una martingala que converge en L^1 a una variable aleatoria X_∞ , entonces X es uniformemente integrable, y así, converge c.s. a X_∞ . En ambos casos, X es cerrada por X_∞ .*

Proposición 3.12. *Sean $\{X_t\}_{t \geq 0}$ una martingala y $X_t^* = \sup_{0 \leq s \leq t} |X_s|$. Se tiene que*

- 1) $P(X_t^* \geq \lambda) \leq \lambda^{-r} E(|X_t|^r)$ para toda $r \geq 1$ y para toda $\lambda > 0$.
- 2) $E[(X_t^*)^r] \leq \left(\frac{r}{r-1}\right)^r E(|X_t|^r)$ para toda $r > 1$.

Tomando $r = 2$ en 2) de la proposición 3.12 se obtiene que

$$E[(X_t^*)^2] \leq 4E(X_t^2),$$

la cual es la desigualdad más importante de las desigualdades de Doob para martingalas.

3.6. Martingalas locales

Definición 3.12. Un proceso estocástico $\{X_t\}_{t \geq 0}$ adaptado a una filtración $\{\mathcal{F}_t\}_{t \geq 0}$ es una **martingala local** si existe una sucesión creciente de tiempos de paro τ_n , llamada **sucesión localizante**, tal que $\tau_n \uparrow \infty$ cuando $n \rightarrow \infty$ y para cada n , $X^{\tau_n} = \{X_{t \wedge \tau_n}\}_{t \geq 0}$ es una martingala.

Se sigue de la proposición 3.8 y de la proposición 3.11, con $\tau_n = n$, que cualquier martingala es una martingala local.

Proposición 3.13. Si $X = \{X_t\}_{t \geq 0}$ es una martingala local tal que $|X_t| \leq Y$ para toda $t \geq 0$ con $EY < \infty$, entonces X es una martingala uniformemente integrable. En particular, esto se cumple si $E(\sup_{t \geq 0} |X_t|) < \infty$.

Demostración. Observe primero que $E|X_t| < \infty$ para toda $t \geq 0$, pues $E|X_t| \leq EY < \infty$. Ahora, sea τ_n una sucesión localizante tal que para cualquier n y para cualesquiera $0 \leq s < t$, $E(X_{t \wedge \tau_n} | \mathcal{F}_s) = X_{s \wedge \tau_n}$. Como para cada $t \geq 0$, $t \wedge \tau_n \uparrow t$ cuando $n \rightarrow \infty$ se tiene que $\lim_{n \rightarrow \infty} X_{t \wedge \tau_n} = X_t$. Usando esto y el hecho de que $|X_{t \wedge \tau_n}| \leq Y$ con $EY < \infty$ ²³ obtiene, por el teorema de convergencia dominada para esperanzas condicionales que $E(X_t | \mathcal{F}_s) = \lim_{n \rightarrow \infty} E(X_{t \wedge \tau_n} | \mathcal{F}_s) = \lim_{n \rightarrow \infty} X_{s \wedge \tau_n} = X_s$. Todo esto muestra que X es una martingala. La integrabilidad uniforme de X se sigue de la proposición 3.9 y con esto se finaliza la demostración. ▲

Definición 3.13. Un proceso estocástico $\{X_t\}_{t \geq 0}$ es de **clase Dirichlet** si la colección $\{X_\tau : \tau \text{ es un tiempo de paro finito}\}$ es uniformemente integrable.

Proposición 3.14. Si $X = \{X_t\}_{t \geq 0}$ es una martingala local de clase Dirichlet, entonces X es una martingala uniformemente integrable.

Demostración. Elija una sucesión localizante τ_n tal que para cualquier n y para cualesquiera $0 \leq s < t$, $E(X_{t \wedge \tau_n} | \mathcal{F}_s) = X_{s \wedge \tau_n}$. Debido a que X es de clase Dirichlet y como para cada n y para $t \geq 0$ fijo, $t \wedge \tau_n$ es un tiempo de paro finito, se tiene que $\{X_{t \wedge \tau_n}\}_{n \geq 1}$

es uniformemente integrable. Luego, como $X_{t \wedge \tau_n} \rightarrow X_t$, entonces $X_{t \wedge \tau_n} \rightarrow X_t$ en L^1 , es decir, $E|X_{t \wedge \tau_n} - X_t| \rightarrow 0$. De este modo, $E(X_{t \wedge \tau_n} | \mathcal{F}_s) \rightarrow E(X_t | \mathcal{F}_s)$ cuando $n \rightarrow \infty$ puesto que por propiedades de la esperanza condicional y por la desigualdad de Jensen,

$$\begin{aligned} E|E(X_{t \wedge \tau_n} | \mathcal{F}_s) - E(X_t | \mathcal{F}_s)| &= E|E(X_{t \wedge \tau_n} - X_t | \mathcal{F}_s)| \leq E[E(|X_{t \wedge \tau_n} - X_t| | \mathcal{F}_s)] \\ &= E|X_{t \wedge \tau_n} - X_t|. \end{aligned}$$

Así, se tiene que $E(X_t | \mathcal{F}_s) = \lim_{n \rightarrow \infty} E(X_{t \wedge \tau_n} | \mathcal{F}_s) = \lim_{n \rightarrow \infty} X_{s \wedge \tau_n} = X_s$, mostrando con ello que X posee la propiedad de martingala. La integrabilidad uniforme de X , y por lo tanto su integrabilidad, se obtiene usando el hecho de que X es de clase Dirichlet con $\tau = t$. $\blacktriangle$

Puede verse, por ejemplo, en [21] pág. 55 que si $X = \{X_t\}_{t \geq 0}$ es una martingala local, entonces su proceso de variación cuadrática $\{[X, X]_t\}_{t \geq 0}$ existe. Más aún, $\{X_t^2 - [X, X]_t\}_{t \geq 0}$ es una martingala local. Se dará un esbozo de la prueba de esta última conclusión bajo condiciones más fuertes.

Lema 3.1. Si $X = \{X_t\}_{t \geq 0}$ es una martingala con $E(X_t^2) < \infty$ para toda $t \geq 0$, entonces $\{X_t^2 - [X, X]_t\}_{t \geq 0}$ es una martingala.

Demostración. Note que $E[X_s(X_t - X_s) | \mathcal{F}_s] = X_s E(X_t - X_s | \mathcal{F}_s) = 0$ para cualesquiera $0 \leq s < t$. Luego, dado que $(X_t^2 - X_s^2) - (X_t - X_s)^2 = 2X_s(X_t - X_s)$, condicionando se obtiene que $E(X_t^2 - X_s^2 | \mathcal{F}_s) = E[(X_t - X_s)^2 | \mathcal{F}_s] = E[\sum_{i=1}^n (X_{t_i} - X_{t_{i-1}})^2 | \mathcal{F}_s]$, donde $\mathcal{P} = \{t_0, t_1, \dots, t_n\}$ es una partición de $[s, t]$. Por lo tanto, haciendo $|\mathcal{P}| \rightarrow 0$ se tiene que $E(X_t^2 - X_s^2 | \mathcal{F}_s) = E([X, X]_t - [X, X]_s | \mathcal{F}_s)$, de lo cual se sigue la propiedad de martingala de $\{X_t^2 - [X, X]_t\}_{t \geq 0}$. $\blacktriangle$

La demostración de la desigualdad que se expone a continuación puede ser consultada, por ejemplo, en [21] pág. 75.

Lema 3.2. (Desigualdad de Burkholder-Gundy) Sea $X = \{X_t\}_{0 \leq t \leq T \leq \infty}$ una martingala local y sea $X_t^* = \sup_{0 \leq s \leq t} |X_s|$. Si $1 < r < \infty$, entonces existen constantes k_r y K_r , ambas dependiendo únicamente de r , tales que

$$k_r E([X, X]_T^{r/2}) \leq E[(X_T^*)^r] \leq K_r E([X, X]_T^{r/2}).$$

Proposición 3.15. Sea $X = \{X_t\}_{t \geq 0}$ una martingala local. Se tiene que

- 1) Si $E[X, X]_t < \infty$ para toda $t \geq 0$, entonces X es una martingala con $E(X_t^2) = E[X, X]_t < \infty$ para toda $t \geq 0$.
- 2) Si $\sup_{t \geq 0} E[X, X]_t < \infty$, entonces X es una martingala cuadrado integrable.

Demostración. 1) Usando la desigualdad de Burkholder-Gundy con $r = 2$ se obtiene que $E[(\sup_{0 \leq t \leq T} |X_t|)^2] \leq K_r E[X, X]_T < \infty$ para alguna constante K_r y para cualquier $T < \infty$. Por consiguiente, $E(\sup_{0 \leq t \leq T} |X_t|) < \infty$ para cualquier $T < \infty$. Luego, por la proposición 3.13, X es una martingala. Así, si $E(X_t^2) < \infty$ para toda $t \geq 0$, entonces $\{X_t^2 - [X, X]_t\}_{t \geq 0}$ es una martingala y por lo tanto $E(X_t^2) = E[X, X]_t < \infty$ para toda $t \geq 0$.

2) Ahora, por 1) de esta proposición, X es una martingala con $E(X_t^2) = E[X, X]_t$ para toda $t \geq 0$. Debido a que ambos miembros de esta última igualdad son no decrecientes se tiene que $\sup_{t \geq 0} E(X_t^2) = \lim_{t \rightarrow \infty} E(X_t^2) = \lim_{t \rightarrow \infty} E[X, X]_t = \sup_{t \geq 0} E[X, X]_t < \infty$ y por lo tanto X es una martingala cuadrado integrable. ▲

3.7. Procesos de Markov y procesos de Markov fuertes

Definición 3.14. Sea $X = \{X_t\}_{t \geq 0}$ un proceso estocástico. Si para toda $y \in \mathbb{R}$ y para cualesquiera $0 \leq s \leq t$ se satisface

$$P(X_t \leq y | \mathcal{F}_s^X) = P(X_t \leq y | X_s) \quad \text{c.s.}, \quad (3.1)$$

se dice que X es un **proceso de Markov**.

La condición (3.1), conocida como la **propiedad de Markov**, dice intuitivamente que el comportamiento de X en el futuro depende únicamente de su estado presente y no de su comportamiento en el pasado.

La prueba del siguiente resultado puede ser consultada, por ejemplo, en [16] pág. 112.

Proposición 3.16. Un proceso $X = \{X_t\}_{t \geq 0}$ es un proceso de Markov si, y solo si, para cualesquiera $0 \leq s_1 \leq \dots \leq s_n \leq s \leq t$ y para cualquier función $f : \mathbb{R} \rightarrow \mathbb{R}$ acotada y $\mathcal{B}(\mathbb{R})$ -medible se tiene que

$$E[f(X_t) | X_{s_1}, \dots, X_{s_n}, X_s] = E[f(X_t) | X_s] \quad \text{c.s.}$$

Sea $X = \{X_t\}_{t \geq 0}$ un proceso de Markov. Debido a que se está considerando un espacio de probabilidad completo $(\Omega, \mathcal{F}, P)$, se tiene que existe (véase por ejemplo, [16] pág. 113) una función $P(s, x, t, B)$ definida para $0 \leq s < t$, $x \in \mathbb{R}$ y $B \in \mathcal{B}(\mathbb{R})$ tal que

1. $P(s, \cdot, t, B)$ es $\mathcal{B}(\mathbb{R})$ -medible.
2. $P(s, x, t, \cdot)$ es una medida de probabilidad sobre $\mathcal{B}(\mathbb{R})$.
3. $P(X_t \in B | \mathcal{F}_s^X) = P(s, X_s, t, B)$ c.s.

4. Para cualesquiera $0 \leq s < u < t$, se satisface la **ecuación de Chapman-Kolmogorov**

$$P(s, x, t, B) = \int_{\mathbb{R}} P(u, y, t, B) P(s, x, u, dy).$$

A $P(s, x, t, B)$ se le llama **probabilidad de transición** del proceso de Markov X .

Definición 3.15. Sea $X = \{X_t\}_{t \geq 0}$ un proceso estocástico. Si para cualquier tiempo de paro finito τ respecto a $\{\mathcal{F}_t^X\}_{t \geq 0}$ y para cualesquiera $y \in \mathbb{R}$ y $t \geq 0$ se cumple

$$P(X_{\tau+t} \leq y | \mathcal{F}_\tau) = P(X_{\tau+t} \leq y | X_\tau) \quad \text{c.s.}, \quad (3.2)$$

se dice que X es un **proceso de Markov fuerte**.

Análogamente a la propiedad de Markov, la condición (3.2), llamada **propiedad fuerte de Markov**, dice que dada la historia del proceso X hasta un tiempo de paro τ , su comportamiento en algún tiempo futuro $\tau + t$ es independiente de su pasado.

3.8. Proceso de Wiener

Definición 3.16. Un proceso estocástico $W = \{W_t\}_{t \geq 0}$ es un **proceso de Wiener o movimiento Browniano** si

a) $W_0 = 0$ c.s.

b) Tiene incrementos independientes; esto es, para cualesquiera $0 \leq t_1 < t_2 < \dots < t_n$, las variables aleatorias $W_{t_1}, W_{t_2} - W_{t_1}, \dots, W_{t_n} - W_{t_{n-1}}$ son independientes.

c) Posee incrementos Gaussianos, es decir, para cualesquiera $0 \leq s < t$, $W_t - W_s \sim N(0, t-s)$.

d) Tiene trayectorias continuas c.s.

Cuando W es un proceso adaptado a la filtración $\{\mathcal{F}_t\}_{t \geq 0}$ y satisface las propiedades a)-d) de la definición 3.16, se dice que W es un proceso de Wiener respecto a $\{\mathcal{F}_t\}_{t \geq 0}$. De este modo, si W es un proceso de Wiener, entonces es un proceso de Wiener respecto a su filtración canónica $\{\mathcal{F}_t^W\}_{t \geq 0}$, donde $\mathcal{F}_t^W = \sigma(W_s : s \leq t)$. En este caso, por b) de la definición 3.16 se tiene que para $s < t$, $W_t - W_s$ es independiente de $\mathcal{F}_s^W$.

La demostración de la existencia de un proceso de Wiener fue hecha por el matemático estadounidense Norbert Wiener en 1923, y ésta puede ser consultada, por ejemplo, en [17] págs. 49-71.

Proposición 3.17. Sea $W = \{W_t\}_{t \geq 0}$ un proceso de Wiener. Se tiene que

- 1) Para cualesquiera $s, t \geq 0$, $E(W_s W_t) = s \wedge t$.
- 2) Para cada $t \geq 0$, $[W, W]_t = t$.
- 3) W es una martingala.
- 4) $\{W_t^2 - t\}_{t \geq 0}$ es una martingala.
- 5) Para cualquier $u \in \mathbb{R}$, $\{e^{uW_t - u^2 t/2}\}_{t \geq 0}$ es una martingala.
- 6) Para cualquier $t \geq 0$, casi seguramente, las trayectorias de W son no diferenciables en t .

Demostración. 1) Se puede suponer, sin pérdida de generalidad, que $s < t$. Entonces, por los incisos a)-c) de la definición 3.16 se obtiene que

$$E(W_s W_t) = E[W_s(W_t - W_s + W_s)] = E[W_s(W_t - W_s)] + E(W_s^2) = E(W_s)E(W_t - W_s) + s = s.$$

2) Se dará la prueba para una sucesión de particiones $\mathcal{P}_n = \{t_0^n, t_1^n, \dots, t_n^n\}$ de $[0, t]$ con $\sum_{n=1}^{\infty} |\mathcal{P}_n| < \infty$. Un ejemplo de tales particiones es $\mathcal{P}_n = \{0, t/2^n, 2t/2^n, \dots, (2^n - 1)t/2^n, t\}$. Sea pues $X_n = \sum_{i=1}^n (W_{t_i^n} - W_{t_{i-1}^n})^2$ y note que

$$EX_n = \sum_{i=1}^n E[(W_{t_i^n} - W_{t_{i-1}^n})^2] = \sum_{i=1}^n (t_i^n - t_{i-1}^n) = t.$$

Ahora, debido a que $E(X^4) = 3\sigma^4$ si $X \sim N(0, \sigma^2)$ y usando el hecho de que W tiene incrementos independientes se tiene que

$$Var(X_n) = \sum_{i=1}^n Var[(W_{t_i^n} - W_{t_{i-1}^n})^2] = \sum_{i=1}^n 3(t_i^n - t_{i-1}^n)^2 \leq 3t|\mathcal{P}_n|.$$

Luego, por la hipótesis, $\sum_{n=1}^{\infty} E[(X_n - EX_n)^2] < \infty$ y consecuentemente, por el teorema de convergencia monótona, $E[\sum_{n=1}^{\infty} (X_n - EX_n)^2] < \infty$, y así, $\sum_{n=1}^{\infty} (X_n - EX_n)^2 < \infty$ c.s. Por lo tanto $X_n - EX_n \rightarrow 0$ c.s., es decir, $X_n \rightarrow t$ c.s.

Es posible mostrar (véase por ejemplo, [24] pág. 18) que $X_n \rightarrow t$ c.s. para cualquier sucesión de particiones $\mathcal{P}_n$ de $[0, t]$ tal que $\mathcal{P}_n \subseteq \mathcal{P}_{n+1}$ y $|\mathcal{P}_n| \rightarrow 0$ cuando $n \rightarrow \infty$.

3) Se sigue de la definición 3.16 que $W_t \sim N(0, t)$ para toda $t \geq 0$ y así, $EW_t = 0$. Ahora, sean $0 \leq s < t$. Dado que $W_t - W_s$ es independiente de $\mathcal{F}_s^W$ se obtiene que

$$E(W_t | \mathcal{F}_s^W) = E(W_s | \mathcal{F}_s^W) + E(W_t - W_s | \mathcal{F}_s^W) = W_s + E(W_t - W_s) = W_s.$$

4) Esto se sigue del lema 3.1 en conjunto con los incisos 2) y 3) de esta proposición.

5) Como $W_t \sim N(0, t)$, entonces su función generadora de momentos es $E(e^{uW_t}) = e^{u^2 t/2}$, lo cual implica que $E(e^{uW_t - u^2 t/2}) = 1$. Ahora, note que para $0 \leq s < t$, se tiene que

$$E(e^{uW_t} | \mathcal{F}_s^W) = e^{uW_s} E[e^{u(W_t - W_s)} | \mathcal{F}_s^W] = e^{uW_s} E[e^{u(W_t - W_s)}] = e^{uW_s} e^{u^2(t-s)/2}.$$

De este modo, la propiedad de martingala de $\{e^{uW_t - u^2 t/2}\}$ se obtiene multiplicando ambos lados esta última igualdad por $e^{-u^2 t/2}$.

6) Se tiene que

$$\frac{W_{t+\Delta} - W_t}{\Delta} = \frac{\sqrt{\Delta}Z}{\Delta} = \frac{Z}{\sqrt{\Delta}},$$

para una variable aleatoria normal estándar Z . De este modo, el cociente converge a ∞ en distribución, debido a que $P\left(\left|\frac{Z}{\sqrt{\Delta}}\right| > K\right) \rightarrow 1$ para cualquier constante K , cuando $\Delta \rightarrow 0$, lo que imposibilita la existencia de la derivada en t . $\blacktriangle$

Se sigue de d) de la definición 3.16, 2) de la proposición 3.17 y de la proposición 1.5 que casi cualquier trayectoria de un proceso de Wiener W tiene variación infinita sobre $[0, t]$ para toda $t > 0$. Esto tiene una importante consecuencia para la teoría de integración estocástica que se abordará en el siguiente capítulo, ya que una integral de la forma

$$\int_0^t f(s) dW_s$$

no puede ser definida, trayectoria por trayectoria (véase el teorema 1.5), como una integral de Riemann-Stieltjes si las trayectorias son de variación infinita.

Teorema 3.2. (Caracterización de Levy) *Un proceso $W = \{W_t\}_{t \geq 0}$ con $W_0 = 0$ es un proceso de Wiener si, y solo si, W es una martingala local continua con $[W, W]_t = t$ para toda $t \geq 0$.*

28 Demostración. La necesidad es una consecuencia inmediata de 2) y 3) de la proposición 3.17 y de d) de la definición 3.16. Ahora, se dará un esbozo de la prueba para la suficiencia. Para ello, suponga que W es una martingala local continua con $[W, W]_t = t$ para toda $t \geq 0$. Entonces, puede mostrarse que para cualquier u , $\{e^{uW_t - u^2 t/2}\}_{t \geq 0}$ es una martingala (véase por ejemplo, [18] pág. 227). Usando este resultado se tiene que $E(e^{uW_t - u^2 t/2} | \mathcal{F}_s^W) = e^{uW_s - u^2 s/2}$ para $0 \leq s < t$, y así, $E[e^{u(W_t - W_s)} | \mathcal{F}_s^W] = e^{u^2(t-s)/2}$. Esto implica que $W_t - W_s$ es independiente de $\mathcal{F}_s^W$ ya que $e^{u^2(t-s)/2}$ es determinista. También, tomando ²⁴ esperanza en ambos lados de esta última igualdad se tiene que $E[e^{u(W_t - W_s)}] = e^{u^2(t-s)/2}$, que es la función generadora de momentos de la distribución $N(0, t-s)$. Por lo tanto $W_t - W_s \sim N(0, t-s)$ y con esto se concluye la demostración. $\blacktriangle$

Capítulo 4

Cálculo estocástico para el proceso de Wiener

4.1. La integral de Itô

Considere un espacio de probabilidad completo $(\Omega, \mathcal{F}, P)$, y un proceso de Wiener $W = \{W_t\}_{t \geq 0}$ definido sobre éste. Sea $\mathcal{N}$ la colección de todos los eventos nulos de $\mathcal{F}$ y defina

$$\mathcal{F}_t := \sigma(\mathcal{F}_t^W \cup \mathcal{N}).$$

Puede demostrarse (véase por ejemplo, [29] pág. 24) que $\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$ es una filtración que satisface las hipótesis usuales, esto es, $\mathbb{F}$ es completa y continua por la derecha, y que además, W es también un proceso de Wiener respecto a $\mathbb{F}$. A lo largo de este capítulo se estará considerando al espacio de probabilidad filtrado $(\Omega, \mathcal{F}, \mathbb{F}, P)$ junto con el proceso de Wiener W .

Se denotará por L_T^2 al espacio de los procesos estocásticos $X = \{X_t\}_{0 \leq t \leq T}$ adaptados a $\mathbb{F}$, medibles ($X : [0, T] \times \Omega \rightarrow \mathbb{R}$ es $\mathcal{B}([0, T]) \times \mathcal{F}$ -medible) y tales que $E(\int_0^T X_t^2 dt) < \infty$.

Definición 4.1. Un proceso estocástico $X = \{X_t\}_{0 \leq t \leq T}$ en L_T^2 es un **proceso simple** si

$$X_0 = \xi_0 \quad y \quad X_t = \sum_{i=1}^n \xi_{i-1} 1_{(t_{i-1}, t_i]}(t) \quad \text{para } 0 < t \leq T, \quad (4.1)$$

donde $0 = t_0 < t_1 < \dots < t_n = T$ y $\xi_{i-1} \in L^2(\Omega, \mathcal{F}, P)$ con ξ_{i-1} $\mathcal{F}_{t_{i-1}}$ -medible, $i = 1, \dots, n$.

Definición 4.2. La **integral de Itô** del proceso simple X dado por (4.1) con respecto al proceso de Wiener W es definida como

$$\int_0^T X_t dW_t = \sum_{i=1}^n \xi_{i-1} (W_{t_i} - W_{t_{i-1}}). \quad (4.2)$$

Con el símbolo $\mathcal{S}$ se representará al espacio de todos los procesos simples. La integral de Itô definida sobre $\mathcal{S}$ satisface las siguientes propiedades.

Proposición 4.1. Sean $X, Y \in \mathcal{S}$ y $a \in \mathbb{R}$. Entonces

- 1) Linealidad: $\int_0^T (aX_t + Y_t) dW_t = a \int_0^T X_t dW_t + \int_0^T Y_t dW_t$.
- 2) Propiedad de media cero: $E \left(\int_0^T X_t dW_t \right) = 0$.
- 3) Propiedad de isometría: $E \left[\left(\int_0^T X_t dW_t \right)^2 \right] = E \left(\int_0^T X_t^2 dt \right)$.

Demostración. 1) Esta propiedad se sigue aplicando directamente la definición 4.2 y haciendo uso del hecho de que una combinación lineal de procesos simples es de nuevo un proceso simple.

2) Usando el hecho de que ξ_{i-1} es $\mathcal{F}_{t_{i-1}}$ -medible, $W_{t_i} - W_{t_{i-1}}$ es independiente de $\mathcal{F}_{t_{i-1}}$ y $E(W_{t_i} - W_{t_{i-1}}) = 0$ se deduce la propiedad de media cero, ya que

$$\begin{aligned} E \left[\sum_{i=1}^n \xi_{i-1} (W_{t_i} - W_{t_{i-1}}) \right] &= \sum_{i=1}^n E [\xi_{i-1} E(W_{t_i} - W_{t_{i-1}} | \mathcal{F}_{t_{i-1}})] \\ &= \sum_{i=1}^n E [\xi_{i-1} E(W_{t_i} - W_{t_{i-1}})] = 0. \end{aligned}$$

3) Observe que

$$\begin{aligned} E \left\{ \left[\sum_{i=1}^n \xi_{i-1} (W_{t_i} - W_{t_{i-1}}) \right]^2 \right\} &= \sum_{i=1}^n E [\xi_{i-1}^2 (W_{t_i} - W_{t_{i-1}})^2] \\ &\quad + 2 \sum_{i < j} E [\xi_{i-1} \xi_{j-1} (W_{t_i} - W_{t_{i-1}})(W_{t_j} - W_{t_{j-1}})]. \end{aligned}$$

Ahora, haciendo uso del hecho de que ξ_{i-1}^2 es $\mathcal{F}_{t_{i-1}}$ -medible, $(W_{t_i} - W_{t_{i-1}})^2$ es independiente de $\mathcal{F}_{t_{i-1}}$ y $E[(W_{t_i} - W_{t_{i-1}})^2] = t_i - t_{i-1}$, se obtiene que

$$\begin{aligned} \sum_{i=1}^n E [\xi_{i-1}^2 (W_{t_i} - W_{t_{i-1}})^2] &= \sum_{i=1}^n E \{ \xi_{i-1}^2 E[(W_{t_i} - W_{t_{i-1}})^2 | \mathcal{F}_{t_{i-1}}] \} \\ &= \sum_{i=1}^n E(\xi_{i-1}^2)(t_i - t_{i-1}) = \int_0^T E(X_t^2) dt, \end{aligned}$$

puesto que $X_t = \xi_{i-1}$ sobre $(t_{i-1}, t_i]$. Similarmente, condicionando y empleando el hecho de que, para $t < j$, ξ_{j-1} , ξ_{i-1} y $W_{t_i} - W_{t_{i-1}}$ son $\mathcal{F}_{t_{j-1}}$ -medibles se tiene que

$$\sum_{i < j} E [\xi_{i-1} \xi_{j-1} (W_{t_i} - W_{t_{i-1}})(W_{t_j} - W_{t_{j-1}})] = 0,$$

con lo cual se finaliza la demostración. $\blacktriangle$

El siguiente resultado es crucial para extender la integral de Itô a integrandos más generales en L_T^2 .

Lema 4.1. Si $X \in L_T^2$, entonces existe una sucesión X^n en $\mathcal{S}$ tal que

$$\lim_{n \rightarrow \infty} E \left(\int_0^T |X_t^n - X_t|^2 dt \right) = 0. \quad (4.3)$$

Demostración. Suponga primero que X es **continuo en media cuadrática**, esto es, $\lim_{s \rightarrow t} E(|X_s - X_t|^2) = 0$. En este caso es posible considerar la sucesión

$$X_t^n = X_0 1_{\{0\}}(t) + \sum_{i=1}^n X_{t_{i-1}} 1_{(t_{i-1}, t_i]}(t)$$

con $t_i = iT/n$, la cual pertenece claramente a $\mathcal{S}$. Por el teorema de Fubini y la continuidad en media cuadrática se tiene que

$$\begin{aligned} E \left(\int_0^T |X_t^n - X_t|^2 dt \right) &= \int_0^T E(|X_t^n - X_t|^2) dt = \sum_{i=1}^n \int_{t_{i-1}}^{t_i} E(|X_t^n - X_t|^2) dt \\ &= \sum_{i=1}^n \int_{t_{i-1}}^{t_i} E(|X_{t_{i-1}} - X_t|^2) dt \leq T \sup_{|s-t| \leq \frac{T}{n}} E(|X_s - X_t|^2) \rightarrow 0, \end{aligned}$$

cuando $n \rightarrow \infty$. Así en este caso la afirmación es válida.

Ahora, si X es arbitrario, por lo que ya se ha demostrado, es suficiente probar que existe una sucesión X^n de procesos en L_T^2 continuos en media cuadrática tal que (4.3) se cumpla. Para cada n defina la sucesión X^n por $X_t^n = X * n 1_{[0, 1/n]}(t)$, $0 \leq t \leq T$, donde el símbolo $*$ denota la convolución de funciones. Con la convención $X_t = 0$ para $t < 0$, se puede escribir

$$X_t^n = n \int_{t-\frac{1}{n}}^t X_s ds = n \left(\int_0^t X_s ds - \int_0^{t-\frac{1}{n}} X_s ds \right). \quad (4.4)$$

No es difícil mostrar que la sucesión X^n definida por (4.4) pertenece a L_T^2 . Note además, de (4.4), que X_t^n es absolutamente continua sobre $[0, T]$ y así, por el teorema de convergencia

dominada X^n es continua en media cuadrática. Haciendo uso del hecho de que la sucesión $n1_{[0,1/n]}$ forma un kernel positivo y por propiedades de la convolución (véase por ejemplo, la sección 2 de los capítulos VI y IX de [19]) se obtiene que

$$\lim_{n \rightarrow \infty} \int_0^T |X_t^n(\omega) - X_t(\omega)|^2 dt = 0 \quad \text{y} \quad \int_0^T |X_t^n(\omega)|^2 dt \leq \int_0^T |X_t(\omega)|^2 dt.$$

Por lo tanto el resultado se sigue aplicando el teorema de convergencia dominada. $\blacktriangle$

Sea $X \in L_T^2$. Por el lema 4.1 se puede elegir una sucesión X^n de procesos simples tal que el límite (4.3) se cumpla. La sucesión de variables aleatorias $\int_0^T X_t^n dW_t$ es de Cauchy en el espacio $L^2(\Omega, \mathcal{F}, P)$, pues por 3) de la proposición 4.1 (la propiedad de isometría) se tiene que

$$\begin{aligned} E \left(\left| \int_0^T X_t^n dW_t - \int_0^T X_t^m dW_t \right|^2 \right) &= E \left(\int_0^T |X_t^n - X_t^m|^2 dt \right) \leq 2E \left(\int_0^T |X_t^n - X_t|^2 dt \right) \\ &+ 2E \left(\int_0^T |X_t - X_t^m|^2 dt \right) \rightarrow 0, \end{aligned}$$

cuando $n, m \rightarrow \infty$, donde se ha usado la desigualdad $(a + b)^2 \leq 2(a^2 + b^2)$. Así, debido a que $L^2(\Omega, \mathcal{F}, P)$ es completo, se obtiene que la sucesión $\int_0^T X_t^n dW_t$ converge a una variable aleatoria. Usando nuevamente la propiedad de isometría puede demostrarse que este límite no depende de la sucesión aproximadora.

Definición 4.3. La *integral de Itô* de un proceso $X \in L_T^2$ es definida mediante el siguiente límite en $L^2(\Omega, \mathcal{F}, P)$

$$\int_0^T X_t dW_t = \lim_{n \rightarrow \infty} \int_0^T X_t^n dW_t,$$

donde X^n es una sucesión de procesos simples que satisface (4.3).

Algunas propiedades que satisface la integral de Itô definida sobre el espacio L_T^2 son las siguientes:

Proposición 4.2. Sean $X, Y \in L_T^2$ y $a \in \mathbb{R}$. Se tiene que

$$1) \text{ Linealidad: } \int_0^T (aX_t + Y_t) dW_t = a \int_0^T X_t dW_t + \int_0^T Y_t dW_t.$$

$$2) \text{ Propiedad de media cero: } E \left(\int_0^T X_t dW_t \right) = 0.$$

3) *Propiedad de isometría:* $E \left[\left(\int_0^T X_t dW_t \right)^2 \right] = E \left(\int_0^T X_t^2 dt \right).$

4) *Localidad:* $\int_0^T X_t dW_t = 0$ c.s. sobre el evento $A = \left\{ \int_0^T X_t^2 dt = 0 \right\}.$

Demostración. Las propiedades de linealidad, media cero e isometría se siguen con facilidad de la definición 4.3.

4) Esta propiedad se cumple ya que puede demostrarse que la sucesión aproximadora de procesos simples

$$X_t^{n,m} = \sum_{i=1}^n m \left[\int_{\frac{(i-1)T}{n}}^{\frac{iT}{n}} X_s ds \right] 1_{\left(\frac{(i-1)T}{n}, \frac{iT}{n}\right]}(t)$$

se anula sobre A . ▲

Para un proceso estocástico $X \in L_T^2$, se define el proceso estocástico $V = \{V_t\}_{0 \leq t \leq T}$ como

$$V_t = \int_0^t X_s^2 ds. \quad (4.5)$$

Observe que las trayectorias de V son no negativas y no decrecientes. Además, debido a que el proceso $X \in L_T^2$ es medible y adaptado a $\mathbb{F}$, entonces es progresivamente medible y por ende V es también adaptado a $\mathbb{F}$. De esta manera, si $K > 0$, el tiempo aleatorio $\tau_K = \inf\{t > 0 : V_t = K\}$ es un tiempo de paro respecto a $\mathbb{F}$ pues $\{t \leq \tau_K\} = \{V_t \leq K\} \in \mathcal{F}_t$.

Lema 4.2. Sean $X \in L_T^2$, $\delta > 0$ y $K > 0$. Entonces

$$P \left(\left| \int_0^T X_t dW_t \right| > \delta \right) \leq P \left(\int_0^T X_t^2 dt > K \right) + \frac{K}{\delta^2}. \quad (4.6)$$

Demostración. Considere el tiempo de paro τ_K y el proceso estocástico X^K dado por $X_t^K = X_t 1_{[0, \tau_K]}(t)$, $0 \leq t \leq T$. Note que

$$E \left(\int_0^T (X_t^K)^2 dt \right) = E \left(\int_0^{\tau_K} X_t^2 dt \right) = E(V_{\tau_K}) = K < \infty,$$

así usando el hecho de que $X_t^K = X_t 1_{\{t \leq \tau_K\}}(t)$ se obtiene que $X^K \in L_T^2$.

Ahora, observe que

$$X_t - X_t^K = X_t(1 - 1_{\{t \leq \tau_K\}}(t)) = X_t 1_{\{\tau_K < t\}}(t) = X_t 1_{\{K < V_t\}}(t),$$

y así, si $A_K = \{\int_0^T (X_t - X_t^K)^2 dt = 0\}$, entonces $\{\int_0^T X_t^2 dt \leq K\} \subset A_K$. Luego, por la propiedad de localidad se tiene que

$$\int_0^T X_t dW_t = \int_0^T X_t^K dW_t \quad \text{c.s.}$$

sobre $\{\int_0^T X_t^2 dt \leq K\}$. Por lo tanto, debido a la desigualdad de Chebyshev (proposición 2.5) y a la propiedad de isometría se obtiene que

$$\begin{aligned} P\left(\left|\int_0^T X_t dW_t\right| > \delta, \int_0^T X_t^2 dt \leq K\right) &= P\left(\left|\int_0^T X_t^K dW_t\right| > \delta, \int_0^T X_t^2 dt \leq K\right) \\ &\leq P\left(\left|\int_0^T X_t^K dW_t\right| > \delta\right) \\ &\leq \frac{1}{\delta^2} E\left[\left(\int_0^T X_t^K dW_t\right)^2\right] \\ &= \frac{1}{\delta^2} E\left(\int_0^T (X_t^K)^2 dt\right) \leq \frac{K}{\delta^2}, \end{aligned}$$

de lo cual, el resultado se sigue inmediatamente. $\blacktriangle$

Al espacio de los procesos estocásticos $X = \{X_t\}_{0 \leq t \leq T}$ adaptados a $\mathbb{F}$, $\mathcal{B}([0, T]) \times \mathcal{F}$ -medibles y tales que $P(\int_0^T X_t^2 dt < \infty) = 1$ se le representará por L_T .

Lema 4.3. Para cada $X \in L_T$ existe una sucesión X^n en L_T^2 tal que

$$\lim_{n \rightarrow \infty} \int_0^T |X_t^n - X_t|^2 dt = 0 \quad \text{c.s.}, \quad (4.7)$$

y por ende, en probabilidad.

Demostración. Considere, para cada n , el tiempo de paro $\tau_n = \inf\{t > 0 : V_t = n\}$ y la sucesión X^n definida por

$$X_t^n = X_t 1_{[0, \tau_n]}(t). \quad (4.8)$$

En la prueba del lema 4.2 se obtuvo que

$$X_t - X_t^n = X_t 1_{\{n < V_t\}}(t). \quad (4.9)$$

Fije $\omega \in \{\int_0^T X_t^2 dt < \infty\}$ y elija $N(\omega) \in \mathbb{N}$ de modo que $\int_0^T X_t^2 dt \leq N(\omega)$. Luego, por la ecuación (4.9), $X_t(\omega) = X_t^n(\omega)$ para toda $n > N(\omega)$. Por lo tanto (4.7) se cumple ya que $P(\int_0^T X_t^2 dt < \infty) = 1$. $\blacktriangle$

Lema 4.4. Para cada $X \in L_T$ existe una sucesión X^n en $\mathcal{S}$ tal que

$$\lim_{n \rightarrow \infty} \int_0^T |X_t^n - X_t|^2 dt = 0 \quad \text{en probabilidad.} \quad (4.10)$$

Demostración. Sea $\epsilon > 0$. Elija, por el lema 4.3, una sucesión Y^n en L_T^2 tal que

$$\lim_{n \rightarrow \infty} P \left(\int_0^T |Y_t^n - X_t|^2 dt > \frac{\epsilon}{4} \right) = 0.$$

Ahora, por el lema 4.1, puede elegirse, para cada Y^n , una sucesión X^n en $\mathcal{S}$ tal que

$$E \left(\int_0^T |X_t^n - Y_t^n|^2 dt \right) < \frac{1}{n}.$$

De este modo, haciendo uso de la desigualdad de Chebyshev se tiene que

$$\begin{aligned} P \left(\int_0^T |X_t^n - X_t|^2 dt > \epsilon \right) &\leq P \left(\int_0^T |X_t^n - Y_t^n|^2 dt > \frac{\epsilon}{4} \right) + P \left(\int_0^T |Y_t^n - X_t|^2 dt > \frac{\epsilon}{4} \right) \\ &< \frac{4}{n\epsilon} + P \left(\int_0^T |Y_t^n - X_t|^2 dt > \frac{\epsilon}{4} \right) \rightarrow 0, \end{aligned}$$

cuando $n \rightarrow \infty$. Esto finaliza la demostración. $\blacktriangle$

Definición 4.4. La *integral de Itô* de un proceso $X \in L_T$ es definida mediante el siguiente límite en probabilidad

$$\int_0^T X_t dW_t = \lim_{n \rightarrow \infty} \int_0^T X_t^n dW_t,$$

donde X^n es una sucesión en $\mathcal{S}$ que satisface (4.10).

La definición 4.4 está bien hecha, pues si $X \in L_T$, entonces por el lema 4.4 existe una sucesión X^n en $\mathcal{S}$ tal que el límite dado en (4.10) se satisface. Tomando $K = \delta^3/2$ en (4.6) se obtiene que

$$P \left(\left| \int_0^T (X_t^n - X_t^m) dW_t \right| > \delta \right) \leq P \left(\int_0^T |X_t^n - X_t^m|^2 dt > \frac{\delta^3}{2} \right) + \frac{\delta}{2}. \quad (4.11)$$

Observe que por (4.10) se tiene que

$$\begin{aligned} P \left(\int_0^T |X_t^n - X_t^m|^2 dt > \frac{\delta^3}{2} \right) &\leq P \left(\int_0^T |X_t^n - X_t|^2 dt > \frac{\delta^3}{8} \right) \\ &+ P \left(\int_0^T |X_t^m - X_t|^2 dt > \frac{\delta^3}{8} \right) \rightarrow 0, \end{aligned}$$

cuando $n, m \rightarrow \infty$. Así, existe $N \in \mathbb{N}$ tal que

$$P\left(\int_0^T |X_t^n - X_t^m|^2 dt > \frac{\delta^3}{2}\right) < \frac{\delta}{2} \quad \text{para toda } n, m \geq N. \quad (4.12)$$

Usando (4.11) y (4.12) se prueba que la sucesión $\int_0^T X_t^n dW_t$ es de Cauchy en probabilidad, y por consiguiente converge en probabilidad a una variable aleatoria $I(X)$. Este límite no depende de la sucesión aproximadora, pues si Y^n es otra sucesión en $\mathcal{S}$ que satisface el límite dado en (4.10), entonces

$$\begin{aligned} P\left(\left|\int_0^T X_t^n dW_t - \int_0^T Y_t^m dW_t\right| > \delta\right) &\leq P\left(\left|\int_0^T X_t^n dW_t - I(X)\right| > \frac{\delta}{2}\right) \\ &\quad + P\left(\left|\int_0^T Y_t^m dW_t - I(X)\right| > \frac{\delta}{2}\right). \end{aligned}$$

4.2. Algunas propiedades del proceso integral de Itô

Sea $X \in L_T$. Entonces, para cualquier $t \in [0, T]$, el proceso estocástico $X1_{[0,t]} := \{X_s 1_{[0,t]}(s)\}_{0 \leq s \leq T}$ también pertenece a L_T , y es posible entonces definir su integral de Itô como

$$\int_0^t X_s dW_s := \int_0^T X_s 1_{[0,t]}(s) dW_s.$$

De este modo se ha construido un nuevo proceso estocástico $Y = \{Y_t\}_{0 \leq t \leq T}$, usualmente llamado **proceso integral de Itô**, dado por

$$Y_t = \int_0^t X_s dW_s. \quad (4.13)$$

A continuación se exponen algunas propiedades que posee el proceso integral de Itô.

Proposición 4.3. Sean $X \in L_T$ y Y el proceso definido por (4.13). Se tiene que

- 1) Y es una martingala local respecto a $\mathbb{F}$.
- 2) Y tiene una modificación continua.
- 3) Para cada $t \in [0, T]$, la variación cuadrática de Y sobre $[0, t]$ es $[Y, Y]_t = \int_0^t X_s^2 ds$.

Demostración. 1) Suponga primero que $X \in \mathcal{S}$. Se probará que en este caso Y es una martingala. Para ello es suficiente con demostrar que para $s < t$,

$$E \left(\int_s^t X_u dW_u | \mathcal{F}_s \right) = 0, \quad (4.14)$$

pues de ser esto cierto, y usando el hecho de que Y está adaptado a $\mathbb{F}$, se tendría que

$$E(Y_t | \mathcal{F}_s) = E(Y_s | \mathcal{F}_s) + E \left(\int_s^t X_u dW_u | \mathcal{F}_s \right) = Y_s.$$

Note que (4.14) se cumple ya que si $X_t = \xi_0 1_{\{s\}}(t) + \sum_{i=1}^n \xi_{i-1} 1_{(t_{i-1}, t_i]}(t)$, donde $s = t_0 < t_1 < \dots < t_n = t$ y $\xi_{i-1} \in L^2(\Omega, \mathcal{F}, P)$ con ξ_{i-1} $\mathcal{F}_{t_{i-1}}$ -medible, $i = 1, \dots, n$, entonces

$$E \left(\int_s^t X_u dW_u | \mathcal{F}_s \right) = \sum_{i=1}^n E[\xi_{i-1} (W_{t_i} - W_{t_{i-1}}) | \mathcal{F}_s] = \sum_{i=1}^n E[\xi_{i-1} E(W_{t_i} - W_{t_{i-1}} | \mathcal{F}_{t_{i-1}}) | \mathcal{F}_s] = 0.$$

Ahora, sea $X \in L_T^2$. Se mostrará que Y es nuevamente una martingala. Elija una sucesión X^n en $\mathcal{S}$ tal que $\lim_{n \rightarrow \infty} E \left(\int_0^T |X_t^n - X_t|^2 dt \right) = 0$. Para cada n , considere el proceso

$$Y_t^n = \int_0^t X_u^n dW_u, \quad 0 \leq t \leq T.$$

Escribiendo, para $s < t$, $Y_t - Y_s = (Y_t - Y_t^n) + (Y_t^n - Y_s^n) + (Y_s^n - Y_s)$, usando la desigualdad de Jensen (teorema 2.3), la propiedad de martingala de cada Y^n y la propiedad de isometría, se tiene que

$$\begin{aligned} E \{ [(Y_t - Y_s | \mathcal{F}_s)]^2 \} &\leq 2E \{ [(Y_t - Y_t^n | \mathcal{F}_s)]^2 \} + 2E \{ [(Y_s^n - Y_s | \mathcal{F}_s)]^2 \} \\ &\leq 2E (|Y_t - Y_t^n|^2) + 2E (|Y_s^n - Y_s|^2) \\ &= 2E \left(\int_0^t |X_u^n - X_u|^2 du \right) + 2E \left(\int_0^s |X_u^n - X_u|^2 du \right) \rightarrow 0, \end{aligned}$$

cuando $n \rightarrow \infty$. Así $E \{ [(Y_t - Y_s | \mathcal{F}_s)]^2 \} = 0$, y por lo tanto $E(Y_t - Y_s | \mathcal{F}_s) = 0$.

Finalmente, suponga que $X \in L_T$. Considere la sucesión dada por (4.8) y note que

$$\begin{aligned} Y_{t \wedge \tau_n} &= \int_0^{t \wedge \tau_n} X_s dW_s = \int_0^T X_s 1_{[0, t \wedge \tau_n]}(s) dW_s = \int_0^T 1_{[0, t]}(s) X_s 1_{[0, \tau_n]}(s) dW_s \\ &= \int_0^T 1_{[0, t]}(s) X_s^n dW_s = \int_0^t X_s^n dW_s. \end{aligned}$$

De este modo, debido a que para cada n $X^n \in L_T^2$, y por lo que ya se ha demostrado, se tiene que el proceso parado Y^{τ_n} es una martingala respecto a $\mathbb{F}$. Como τ_n es una sucesión de tiempos de paro tal que $\tau_n \uparrow \infty$, se concluye que Y es una martingala local respecto a $\mathbb{F}$.

2) Se omite esta demostración; puede ser consultada, por ejemplo, en [20] pág. 72.

3) Sean $t \in [0, T]$ fijo y $\mathcal{P} = \{t_0, t_1, \dots, t_n\}$ una partición de $[0, t]$. Suponga en primer lugar que $X \in L_T^2$ y que el proceso V dado por (4.5) satisface que $V_t \leq K$ para alguna $K > 0$. En este caso $\{Y_s\}_{0 \leq s \leq t}$ es una martingala, y así resulta sencillo probar que $\{Y_s^2 - V_s\}_{0 \leq s \leq t}$ es también una martingala continua. Luego, $E[(Y_{t_i} - Y_{t_{i-1}})^2 - (V_{t_i} - V_{t_{i-1}}) | \mathcal{F}_{t_{i-1}}] = 0$, lo cual implica que

$$\begin{aligned} E \left(\left| \sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})^2 - V_t \right|^2 \right) &= E \left(\left| \sum_{i=1}^n [(Y_{t_i} - Y_{t_{i-1}})^2 - (V_{t_i} - V_{t_{i-1}})] \right|^2 \right) \\ &= \sum_{i=1}^n E \left[|(Y_{t_i} - Y_{t_{i-1}})^2 - (V_{t_i} - V_{t_{i-1}})|^2 \right] \\ &\leq 2 \sum_{i=1}^n E[(Y_{t_i} - Y_{t_{i-1}})^4] + 2 \sum_{i=1}^n E[(V_{t_i} - V_{t_{i-1}})^2]. \end{aligned}$$

Note que el proceso $Z = \{Z_s\}_{t_{i-1} \leq s \leq t_i}$ definido como $Z_s = Y_s - Y_{t_{i-1}}$ es una martingala continua y aplicando la desigualdad de Burkholder-Gundy se tiene que

$$E[(Y_{t_i} - Y_{t_{i-1}})^4] \leq E \left[\sup_{t_{i-1} \leq s \leq t_i} |Z_s|^4 \right] \leq K_2 E[(Z, Z_{t_i}^2)] = K_2 E[(V_{t_i} - V_{t_{i-1}})^2],$$

para alguna constante $K_2 > 0$. De este modo, tomando $C = 2(K_2 + 1)$ y usando el hecho de que $|V_{t_i} - V_{t_{i-1}}| = V_{t_i} - V_{t_{i-1}} \leq V_t \leq K$ se obtiene que

$$E \left(\left| \sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})^2 - V_t \right|^2 \right) \leq C \sum_{i=1}^n E[(V_{t_i} - V_{t_{i-1}})^2] \leq C K E \left(\sup_{|u-s| \leq |\mathcal{P}|} |V_u - V_s| \right) \rightarrow 0,$$

cuando $|\mathcal{P}| \rightarrow 0$ debido al teorema de convergencia dominada, ya que las trayectorias de V son uniformemente continuas sobre $[0, t]$ y $V_t \leq K$. Se ha mostrado así que

$$\lim_{|\mathcal{P}| \rightarrow 0} \sum_{i=1}^n |Y_{t_i} - Y_{t_{i-1}}|^2 = V_t := \int_0^t X_s^2 ds$$

en $L^2(\Omega, \mathcal{F}, P)$, y por consiguiente en probabilidad. Así el resultado se cumple sobre el evento $\{V_t \leq K\} = \{t \leq \tau_K\}$.

Ahora suponga que $X \in L_T$ y considere la sucesión localizante $\tau_n = \inf\{t > 0 : V_t = n\}$. En este caso $V_{t \wedge \tau_n} \leq n$ y por lo demostrado en la primera parte se tiene que

$$\lim_{|\mathcal{P}| \rightarrow 0} \sum_{i=1}^n |Y_{t_i \wedge \tau_n} - Y_{t_{i-1} \wedge \tau_n}|^2 = V_{t \wedge \tau_n}$$

en probabilidad. Luego, dado $\epsilon > 0$, existe $\delta > 0$ tal que si $|\mathcal{P}| < \delta$, entonces

$$P\left(\sum_{i=1}^n |Y_{t_i \wedge \tau_n} - Y_{t_{i-1} \wedge \tau_n}|^2 - V_{t \wedge \tau_n} > \epsilon\right) < \frac{\epsilon}{2}.$$

Como $\tau_n \uparrow \infty$, se puede elegir $n_0 \in \mathbb{N}$ tal que $P(\tau_{n_0} < t) < \epsilon/2$. Por lo tanto, si $|\mathcal{P}| < \delta$, entonces

$$\begin{aligned} P\left(\left|\sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})^2 - V_t\right| > \epsilon\right) &\leq P\left(\left|\sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})^2 - V_t\right| > \epsilon, t \leq \tau_{n_0}\right) + \frac{\epsilon}{2} \\ &\leq P\left(\left|\sum_{i=1}^n (Y_{t_i \wedge \tau_{n_0}} - Y_{t_{i-1} \wedge \tau_{n_0}})^2 - V_{t \wedge \tau_{n_0}}\right| > \epsilon, t \leq \tau_{n_0}\right) \\ &\quad + \frac{\epsilon}{2} < \epsilon, \end{aligned}$$

con lo cual se finaliza la demostración. $\blacktriangle$

4.3. Procesos de Itô

1 Con el símbolo L_T^0 se representará al espacio de los procesos estocásticos $X = \{X_t\}_{0 \leq t \leq T}$ adaptados a $\mathbb{F}$, $\mathcal{B}([0, T]) \times \mathcal{F}$ -medibles y tales que $P(\int_0^T |X_t| dt < \infty) = 1$.

Definición 4.5. Un proceso estocástico X es un **proceso de Itô** si, para cada $t \in [0, T]$,

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s, \quad (4.15)$$

donde $\mu \in L_T^0$, $\sigma \in L_T$ y X_0 es $\mathcal{F}_0$ -medible. En notación diferencial (4.15) se escribe como

$$dX_t = \mu_t dt + \sigma_t dW_t, \quad 0 \leq t \leq T. \quad (4.16)$$

Debido a que $\mu \in L_T^0$, el proceso de Itô X es adaptado a la filtración $\mathbb{F}$. Además, por (4.15) y 2) de la proposición 4.3 se sigue que X tiene una modificación continua.

Definición 4.6. Si X es un proceso de Itô que satisface (4.15) y Y es un proceso adaptado a $\mathbb{F}$ tal que $Y\mu \in L_T^0$ y $Y\sigma \in L_T$, entonces $\int_0^t Y_s dX_s$ es definida por

$$\int_0^t Y_s dX_s = \int_0^t Y_s \mu_s ds + \int_0^t Y_s \sigma_s dW_s. \quad (4.17)$$

A continuación se expone la **fórmula de Itô**, también conocida como la **fórmula de cambio de variables** o **regla de la cadena**; ésta es una de las principales herramientas del cálculo estocástico y da origen a otras fórmulas, por ejemplo, a la **fórmula de integración por partes**. Las demostraciones de la fórmula de Itô y la fórmula de integración por partes pueden ser consultadas, por ejemplo, en la sección 7 del capítulo II de [24].

Teorema 4.1. (Fórmula de Itô) Sean X un proceso de Itô satisfaciendo (4.15) y $f : \mathbb{R} \rightarrow \mathbb{R}$ una función dos veces continuamente diferenciable. Entonces $f(X)$ es también un proceso de Itô con la representación

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) dX_s + \frac{1}{2} \int_0^t f''(X_s) \sigma_s^2 ds. \quad (4.18)$$

La representación diferencial de (4.18) es

$$df(X_t) = f'(X_t) dX_t + \frac{1}{2} f''(X_t) (dX_t)^2, \quad (4.19)$$

de donde, para obtener la ecuación (4.18), se hace uso de (4.16) y de las igualdades

$$(dW_t)^2 = dt, \quad dt dW_t = 0 \quad \text{y} \quad (dt)^2 = 0.$$

Proposición 4.4. Sean X y Y procesos de Itô. Entonces, la **fórmula de integración por partes** es dada por

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + [X, Y]_t. \quad (4.20)$$

Debido a que los procesos de Itô son continuos, por la proposición 1.5 se sigue que si al menos uno de los procesos de Itô X o Y es de variación finita, entonces (4.20) se reduce a la fórmula usual de integración por partes

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s.$$

4.4. Ecuaciones diferenciales estocásticas

Una **ecuación diferencial estocástica** conducida por el proceso de Wiener, W , tiene la forma

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \quad (4.21)$$

donde las funciones $\mu(t, x)$ y $\sigma(t, x)$, llamadas **deriva** y **coeficiente de difusión** respectivamente, son conocidas y el proceso X es desconocido. Se asume que la condición inicial X_0 es $\mathcal{F}_0$ -medible y que pertenece a $L^2(\Omega, \mathcal{F}, P)$.

Definición 4.7. Un proceso estocástico X es llamado **una solución fuerte de la ecuación diferencial estocástica (4.21)**, si para toda $t > 0$ las integrales $\int_0^t \mu(s, X_s)ds$ y $\int_0^t \sigma(s, X_s)dW_s$ existen, y

$$X_t = X_0 + \int_0^t \mu(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s. \quad (4.22)$$

La solución fuerte X en (4.22) puede interpretarse como el estado de un sistema que evoluciona de manera determinista gobernado por la parte no aleatoria de la ecuación, pero alterado por un ruido aditivo dado por la integral estocástica.

Así como en el caso de ecuaciones diferenciales deterministas, en el caso estocástico también existen teoremas para la existencia y unicidad de soluciones fuertes de ecuaciones diferenciales estocásticas. La prueba del siguiente teorema de existencia y unicidad puede ser consultada, por ejemplo, en [14] págs. 104-107.

Teorema 4.2. (Existencia y Unicidad) Si $\mu(t, x)$ y $\sigma(t, x)$ son localmente Lipschitz en x uniformemente en t , y satisfacen la condición de crecimiento lineal en x , entonces existe una única solución fuerte X de la ecuación diferencial estocástica (4.21). Además, X es continuo.

El siguiente resultado requiere condiciones menos estrictas para garantizar la existencia y unicidad de soluciones de ecuaciones diferenciales estocásticas. Para efectos de este trabajo, se expone en su versión más simple, esto es, para ecuaciones diferenciales estocásticas con coeficientes homogéneos en el tiempo. Su demostración puede ser vista, por ejemplo, en [26] pág. 265.

Teorema 4.3. (Yamada-Watanabe) Si $\mu(x)$ satisface la condición de Lipschitz y $\sigma(x)$ satisface una condición de Hölder de orden $\alpha \geq 1/2$, entonces existe una única solución fuerte para la ecuación

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t. \quad (4.23)$$

Otro concepto de solución de una ecuación diferencial estocástica es dada por la definición de solución débil, la cual permite darle significado a la ecuación cuando ésta no tiene

1 soluciones fuertes. Las soluciones débiles son soluciones en distribución, pueden ser definidas sobre algún espacio de probabilidad filtrado y existir bajo condiciones más débiles sobre los coeficientes de la ecuación.

Definición 4.8. Si existe un espacio de probabilidad filtrado, un proceso de Wiener $\widetilde{W}$ y un proceso $\widetilde{X}$ adaptado a la filtración dada, tal que $\widetilde{X}_0$ tiene la distribución dada, las integrales $\int_0^t \mu(s, \widetilde{X}_s) ds$ y $\int_0^t \sigma(s, \widetilde{X}_s) d\widetilde{W}_s$ están definidas para toda t y

$$\widetilde{X}_t = \widetilde{X}_0 + \int_0^t \mu(s, \widetilde{X}_s) ds + \int_0^t \sigma(s, \widetilde{X}_s) d\widetilde{W}_s, \quad (4.24)$$

entonces $\widetilde{X}$ es llamado una **solución débil** a la ecuación diferencial estocástica (4.21).

Es claro de la definiciones 4.7 y 4.8 que cualquier solución fuerte es también una solución débil. Sin embargo, el recíproco no se cumple en general. Por ejemplo, la **ecuación diferencial estocástica de Tanaka**,

$$dX_t = \text{Sign}(X_t) dW_t, \quad (4.25)$$

donde $\text{Sign}(x)$ toma el valor 1 para $x \geq 0$ y -1 para $x < 0$, posee como única solución débil al proceso de Wiener, y no tiene solución fuerte. En efecto, sea X algún proceso de Wiener y considere el proceso

$$Y_t = \int_0^t \frac{1}{\text{Sign}(X_s)} dX_s = \int_0^t \text{Sign}(X_s) dX_s.$$

Note que el proceso $\{\text{Sign}(X_t)\}_{t \geq 0}$ es adaptado y $\int_0^T \text{Sign}^2(X_t) dt = T < \infty$, de modo que Y está bien definido. Así, por la proposición 4.3, Y es una martingala local continua con $[Y, Y]_t = \int_0^t \text{Sign}^2(X_s) ds = t$. Por tanto, por el teorema 3.2, Y es un proceso de Wiener. El teorema 3.2 implica también que cualquier solución débil de la ecuación (4.25) es un proceso de Wiener. El hecho de que la ecuación de Tanaka no tiene soluciones fuertes, puede ser visto, por ejemplo, en [26] pág. 150.

La demostración del siguiente resultado sobre existencia y unicidad de soluciones débiles puede ser vista, por ejemplo, en [28] pág. 160 (corolario 6.5.5).

Proposición 4.5. Si $\sigma(t, x)$ es positiva y continua, y para cualquier $T > 0$ existe K_T tal que para toda $x \in \mathbb{R}$

$$|\mu(t, x)| + |\sigma(t, x)| \leq K_T(1 + |x|),$$

entonces existe una única solución débil a la ecuación diferencial estocástica (4.21) iniciando en cualquier punto x en cualquier tiempo $t \geq 0$.

Puede probarse también que las soluciones de ecuaciones diferenciales estocásticas, usualmente llamadas **difusiones**, son procesos de Markov fuertes (véase por ejemplo, [23] págs. 20-21). De este modo, si la ecuación diferencial estocástica (4.21) posee una solución débil X , entonces X tiene una probabilidad de transición $P(s, x, t, y) = P(X_t \leq y | X_s = x)$. Es conocido (véase por ejemplo, [15] pág. 102) que si las derivadas parciales $\frac{\partial}{\partial x}\mu(t, x)$, $\frac{\partial}{\partial x}\sigma(t, x)$ y $\frac{\partial^2}{\partial x^2}\sigma(t, x)$ existen y $P(s, x, t, y)$ tiene una **densidad de transición** $p(s, x, t, y)$, es decir, $P(X_t \leq y | X_s = x) = \int_{-\infty}^y p(s, x, t, z) dz$, tal que $\frac{\partial}{\partial t}p(s, x, t, y)$, $\frac{\partial}{\partial y}p(s, x, t, y)$ y $\frac{\partial^2}{\partial y^2}p(s, x, t, y)$ existen, entonces para $t > s$, $p(s, x, t, y)$ satisface la ecuación diferencial parcial

$$\frac{\partial}{\partial t}p(s, x, t, y) = -\frac{\partial}{\partial y}[\mu(t, y)p(s, x, t, y)] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[\sigma^2(t, y)p(s, x, t, y)], \quad (4.26)$$

la cual es conocida ¹ como la **ecuación hacia adelante de Kolmogorov o ecuación de Fokker-Planck**.

Recíprocamente, si las condiciones de existencia y unicidad de soluciones se cumplen para la ecuación diferencial estocástica (4.21) y $p(s, x, t, y)$ como una función en t y y , satisface la ecuación (4.26), entonces $p(s, x, t, y)$ es la densidad de transición de la solución X de la ecuación (4.21) (véase por ejemplo, [17] pág. 369).

Definición 4.9. Sean X una solución de (4.21) y $f(t, x)$ una función dos veces continuamente diferenciable en x . Se define el operador diferencial L , llamado el **generador** de X , por

$$Lf(t, x) = (Lf)(t, x) = \frac{1}{2}\sigma^2(t, x)\frac{\partial^2}{\partial x^2}f(t, x) + \mu(t, x)\frac{\partial}{\partial x}f(t, x). \quad (4.27)$$

4.5. Tiempos de salida de un intervalo

Para lograr el objetivo que se pretende en este trabajo, se restringirá en esta sección al estudio de algunos resultados concernientes a ecuaciones diferenciales estocásticas con coeficientes homogéneos en el tiempo. ¹ Para ello, suponga que X es una difusión que resuelve la ecuación (4.23). En cuyo caso su probabilidad de transición $P(s, x, t, y) = P(0, x, t-s, y)$ ¹ depende sólo de $t-s$ (véase por ejemplo, [18] pág. 156). Por lo que en este caso, la probabilidad de transición es denotada por

$$P(t, x, y) = P(s, x, t+s, y) = P(0, x, t, y) = P(X_t \leq y | X_0 = x),$$

¹ su densidad, cuando existe, es escrita como $p(t, x, y)$. La ecuación de Fokker-Planck (4.26) queda expresada como

$$\frac{\partial}{\partial t}p(t, x, y) = -\frac{\partial}{\partial y}[\mu(y)p(t, x, y)] + \frac{1}{2}\frac{\partial^2}{\partial y^2}[\sigma^2(y)p(t, x, y)],$$

y el generador de X , L , dado en la ecuación (4.27) se reduce a

$$Lf(x) = \frac{1}{2}\sigma^2(x)f''(x) + \mu(x)f'(x). \quad (4.28)$$

Sea τ el primer tiempo de salida de X de un intervalo (a, b) , esto es,

$$\tau = \inf\{t > 0 : X_t \notin (a, b)\},$$

y sea τ_x el primer tiempo en el cual la difusión X alcanza el valor $x \in \{a, b\}$, es decir,

$$\tau_x := \inf\{t > 0 : X_t = x\}. \quad (4.29)$$

Observe que debido a la proposición 3.4, tanto τ como τ_x son tiempos de paro. Más aún, por la continuidad de X , es claro que $\tau = \tau_a \wedge \tau_b$.

Notación. En lo que sigue, se escribirá a la esperanza condicional $E(\cdot | X_0 = x)$ simplemente como $E_x(\cdot)$. Similarmente, a la probabilidad condicional $P(\cdot | X_0 = x)$ se le escribirá como $P_x(\cdot)$.

Proposición 4.6. Sean $\sigma(x) > 0$ continua sobre $[a, b]$ y $X_0 = x$, $a < x < b$. Entonces para cualquier función $f(x)$ dos veces continuamente diferenciable en $\mathbb{R}$, el proceso dado por

$$f(X_{t \wedge \tau}) - \int_0^{t \wedge \tau} Lf(X_s) ds \quad (4.30)$$

es una martingala. En particular,

$$E_x \left[f(X_{t \wedge \tau}) - \int_0^{t \wedge \tau} Lf(X_s) ds \right] = f(x). \quad (4.31)$$

Demostración. Empleando las ecuaciones (4.17) y (4.18), y sustituyendo t por $t \wedge \tau$ se obtiene que

$$f(X_{t \wedge \tau}) - \int_0^{t \wedge \tau} Lf(X_s) ds = f(x) + \int_0^{t \wedge \tau} f'(X_s) \sigma(X_s) dW_s. \quad (4.32)$$

Observe que la integral de Itô se puede escribir como $\int_0^t 1_{\{s \leq \tau\}} f'(X_s) \sigma(X_s) dW_s$. Dado que la filtración en consideración es continua por la derecha, el proceso $\{1_{\{s \leq \tau\}}\}_{s \geq 0}$ es adaptado. Ahora, note que para $s \leq \tau$, $X_s \in [a, b]$, y debido a la continuidad de $f'(x)\sigma(x)$ sobre $[a, b]$, es posible elegir una constante $K > 0$ tal que $|1_{\{s \leq \tau\}} f'(X_s) \sigma(X_s)| \leq K$ para $s \leq t$. Luego, $E \int_0^t [1_{\{s \leq \tau\}} f'(X_s) \sigma(X_s)]^2 ds \leq K^2 t < \infty$. Esto significa que el proceso $Y := \{1_{\{t \leq \tau\}} f'(X_t) \sigma(X_t)\}_{0 \leq t \leq T}$ pertenece a L_T^2 para cualquier T , y así, por lo visto en la prueba de 1) de la proposición 4.3 se tiene que Y es una martingala. Por ende el proceso dado en (4.32) es también una martingala. ▲

El siguiente resultado establece en particular que el tiempo de paro τ tiene esperanza finita, y por consiguiente τ es finito casi seguramente.

Proposición 4.7. Sean $\sigma(x) > 0$ continua sobre $[a, b]$ y $X_0 = x$, $a < x < b$. Entonces $E_x(\tau) = f(x)$ satisface la ecuación diferencial ordinaria

$$\frac{1}{2}\sigma^2(x)f''(x) + \mu(x)f'(x) = -1 \quad \text{o brevemente} \quad Lf = -1, \quad (4.33)$$

con $f(a) = f(b) = 0$.

Demostración. Sea $f(x)$ una solución de la ecuación (4.33). Usando la ecuación (4.31) junto con $Lf = -1$ y reordenando términos se obtiene que

$$E_x(t \wedge \tau) = f(x) - E_x[f(X_{t \wedge \tau})]. \quad (4.34)$$

Dado que X y f son continuos, $f(X_{t \wedge \tau}) \rightarrow f(X_\tau)$ cuando $t \rightarrow \infty$, pues $0 \leq t \wedge \tau \uparrow \tau$ cuando $t \rightarrow \infty$. Debido a la continuidad de f y al hecho de que $X_{t \wedge \tau} \in [a, b]$ para toda t , se puede elegir una constante $K > 0$ tal que $E_x[f(X_{t \wedge \tau})] \leq K$. Luego, por el teorema de convergencia monótona y por la ecuación (4.34) se tiene que $E_x(\tau) < \infty$. Consecuentemente, τ es finito casi seguramente. Además, por el teorema de convergencia dominada $E_x[f(X_{t \wedge \tau})] \rightarrow E_x[f(X_\tau)]$. Pero $X_\tau = a$ o $X_\tau = b$, de modo que $f(X_\tau) = 0$. Por lo tanto, nuevamente de la ecuación (4.34), se concluye que $E_x(\tau) = f(x)$. ▲

La proposición que se presenta a continuación proporciona una fórmula para calcular la probabilidad de que el proceso alcance b antes de alcanzar a , $P_x(\tau_b < \tau_a)$.

Proposición 4.8. Sean $\sigma(x) > 0$ continua sobre $[a, b]$ y $X_0 = x$, $a < x < b$. Entonces

$$P_x(\tau_b < \tau_a) = \frac{S(x) - S(a)}{S(b) - S(a)}, \quad (4.35)$$

donde

$$S(x) = \int_{x_1}^x \exp\left(-\int_{x_0}^u \frac{2\mu(y)}{\sigma^2(y)} dy\right) du, \quad (4.36)$$

en x_0 y x_1 siendo constantes indeterminadas. La función S es conocida como la **función escala**, y es una solución de la ecuación diferencial ordinaria

$$\frac{1}{2}\sigma^2(x)S''(x) + \mu(x)S'(x) = 0 \quad \text{o brevemente} \quad LS = 0. \quad (4.37)$$

Demostración. Observe que por la ecuación (4.31)

$$E_x\left[S(X_{t \wedge \tau}) - \int_0^{t \wedge \tau} LS(X_s) ds\right] = S(x).$$

Luego, dado que $LS = 0$, resulta que $E_x[S(X_{t \wedge \tau})] = S(x)$. Ahora, note que argumentos completamente análogos a los dados en la demostración de la proposición 4.7 conducen a

que $E[S(X_{t \wedge \tau})] \rightarrow E[S(X_\tau)]$. Así, $E[S(X_\tau)] = S(x)$. Por consiguiente, haciendo uso del hecho de que τ es finito y toma los valores τ_b con probabilidad $P_x(\tau_b < \tau_a)$ y τ_a con probabilidad complementaria, se obtiene que

$$\begin{aligned} S(x) &= E[S(X_\tau)] = S(b)P_x(\tau_b < \tau_a) + S(a)[1 - P_x(\tau_b < \tau_a)] \\ &\equiv [S(b) - S(a)]P_x(\tau_b < \tau_a) + S(a). \end{aligned}$$

Por lo tanto la ecuación (4.35) se cumple. ▲

Capítulo 5

Aplicaciones en la biología

5.1. Difusión de ramificación de Feller

En este apartado se describirá un crecimiento poblacional continuo en el que los individuos se multiplican y son estadísticamente independientes entre sí con una función de fertilidad independiente del tamaño de la población total; en cuyo caso, el desplazamiento medio infinitesimal y la varianza son necesariamente proporcionales al tamaño instantáneo de la población. De esta manera, si X_t es una variable aleatoria que representa al tamaño de la población en el instante $t \geq 0$, entonces el proceso $X = \{X_t\}_{t \geq 0}$ será un proceso de Markov para el cual existirán constantes α y β tales que

$$\lim_{t \rightarrow 0} t^{-1} E(X_t - x | X_0 = x) = \alpha x$$

y

$$\lim_{t \rightarrow 0} t^{-1} E[(X_t - x)^2 | X_0 = x] = \beta^2 x.$$

Note que $\mu(x) = \alpha x$ es de Lipschitz, y por lo visto en la sección 1.2, $\sigma(x) = \beta\sqrt{x}$ es de Hölder de orden $1/2$. Así, por el teorema de Yamada-Watanabe (teorema 4.3), el proceso X es la solución única de la ecuación diferencial estocástica

$$dX_t = \alpha X_t dt + \beta \sqrt{X_t} dW_t. \quad (5.1)$$

La ecuación de Fokker-Planck correspondiente es

$$\frac{\partial}{\partial t} p(t, x, y) = -\alpha \frac{\partial}{\partial y} [yp(t, x, y)] + \frac{\beta^2}{2} \frac{\partial^2}{\partial y^2} [yp(t, x, y)]. \quad (5.2)$$

En 1951 el matemático William Feller [13] realizó el análisis de este proceso de difusión X resolviendo directamente la ecuación diferencial parcial (5.2). En esta sección se desarrollará el tratamiento expuesto en la sección 13.1 de [18], en el cual la teoría del cálculo estocástico

1 fue usada para obtener información del proceso X directamente de la ecuación diferencial estocástica (5.1).

Antes de proseguir, es importante mencionar que aunque el crecimiento poblacional es, estrictamente hablando, un proceso discontinuo, para **poblaciones grandes** el modelo continuo es una aproximación muy razonable, siempre y cuando se use una escala de tiempo apropiada (véase por ejemplo, [3] pág. 262).

El primer resultado muestra, como es de esperarse, que la solución de (5.1) es no negativa y que una vez que toma el valor 0, permanece por siempre en 0.

Proposición 5.1. 1 $P(X_t \geq 0 \text{ para toda } t \geq 0) = 1$. Además, si $\tau_0 = \inf\{t > 0 : X_t = 0\}$, entonces $X_t = 0$ para toda $t \geq \tau_0$.

Demostración. Considere primero a (5.1) con $X_0 = 0$. Así, es claro que $X \equiv 0$ es una solución, y como la solución es única, entonces en este caso $X \equiv 0$ es la solución. Considere ahora a (5.1) con $X_0 > 0$ y al tiempo aleatorio $\tau_0 = \inf\{t > 0 : X_t = 0\}$. Note que por la proposición 3.4, τ_0 es un tiempo de paro con $X_{\tau_0} = 0$. De este modo, como X es un proceso de Markov fuerte, entonces X_t para $t > \tau_0$ es la solución a (5.1) con $X_{\tau_0} = 0$, y así, $X_t = 0$ para toda $t > \tau_0$. Por lo tanto, si $X_0 > 0$, entonces $X_t > 0$ para toda $t < \tau_0$ y $X_t = 0$ para toda $t \geq \tau_0$. ▲

El siguiente resultado dice que el crecimiento medio de la población es exponencial.

Proposición 5.2. Sean $X_0 > 0$ y $\alpha > 0$. Entonces $EX_t = X_0 e^{\alpha t}$ y $\{X_t e^{-\alpha t}\}_{t \geq 0}$ es una martingala no negativa que converge casi seguramente cuando $t \rightarrow \infty$ a una variable aleatoria Z con $E|Z| < \infty$.

Demostración. 1 Se probará primero que $EX_t = X_0 e^{\alpha t}$. Para ello se verá que X_t es integrable. Por 1) de la proposición 4.3 se tiene que $\{\int_0^t \sqrt{X_s} dW_s\}_{t \geq 0}$ es una martingala local. Así, se puede elegir una sucesión localizante τ_n tal que el proceso parado en τ_n , $\{\int_0^{t \wedge \tau_n} \sqrt{X_s} dW_s\}_{t \geq 0}$, es una martingala. Luego, usando la expresión integral de (5.1) se escribe

$$X_{t \wedge \tau_n} = X_0 + \alpha \int_0^{t \wedge \tau_n} X_s ds + \beta \int_0^{t \wedge \tau_n} \sqrt{X_s} dW_s,$$

y tomando esperanzas se obtiene que

$$EX_{t \wedge \tau_n} = X_0 + \alpha E \int_0^{t \wedge \tau_n} X_s ds,$$

pues debido a que el proceso $\{\int_0^{t \wedge \tau_n} \sqrt{X_s} dW_s\}_{t \geq 0}$ es una martingala, $0 = \beta E \int_0^{0 \wedge \tau_n} \sqrt{X_s} dW_s = \beta E \int_0^{t \wedge \tau_n} \sqrt{X_s} dW_s$ para toda $t \geq 0$. Como $t \wedge \tau_n \uparrow t$ cuando $n \uparrow \infty$ y dado que X es no

negativo, $0 \leq \int_0^{t \wedge \tau_n} X_s ds \uparrow \int_0^t X_s ds$ cuando $n \uparrow \infty$. De manera que, por el teorema de convergencia monótona,

$$E \int_0^{t \wedge \tau_n} X_s ds \rightarrow E \int_0^t X_s ds$$

cuando $n \uparrow \infty$. Por consiguiente,

$$EX_{t \wedge \tau_n} \rightarrow X_0 + \alpha \int_0^t EX_s ds$$

cuando $n \uparrow \infty$, donde se ha usado que, por el teorema de Fubini, $E \int_0^t X_s ds = \int_0^t EX_s ds$. Debido a que X es continuo, $\lim_{n \uparrow \infty} X_{t \wedge \tau_n} = X_t$, y por el lema de Fatou,

$$\begin{aligned} EX_t &= E \lim_{n \uparrow \infty} X_{t \wedge \tau_n} = E \liminf_{n \uparrow \infty} X_{t \wedge \tau_n} \\ &\leq \liminf_{n \uparrow \infty} EX_{t \wedge \tau_n} = \lim_{n \uparrow \infty} EX_{t \wedge \tau_n} = X_0 + \alpha \int_0^t EX_s ds. \end{aligned}$$

De este modo, la desigualdad de Gronwall (teorema 1.3) dice que la última desigualdad implica

$$EX_t \leq X_0 + \alpha \int_0^t X_0 e^{\alpha(t-s)} ds = X_0 e^{\alpha t}.$$

Se tiene así que $EX_t \leq X_0 e^{\alpha t} < \infty$ para toda $t \geq 0$, lo cual prueba que X_t es integrable. Se verá ahora que la martingala local $\{\int_0^t \sqrt{X_s} dW_s\}_{t \geq 0}$ es en realidad una martingala. En efecto, haciendo uso de 3) de la proposición 4.3 y el teorema de Fubini, se obtiene que

$$\begin{aligned} E \left[\int_0^\cdot \sqrt{X_s} dW_s, \int_0^\cdot \sqrt{X_s} dW_s \right]_t &= E \int_0^t X_s ds = \int_0^t EX_s ds \\ &\leq X_0 \int_0^t e^{\alpha s} ds = (X_0/\alpha) (e^{\alpha t} - 1) < \infty \end{aligned}$$

para toda $t \geq 0$. Así, por 1) de la proposición 3.15, $\{\int_0^t \sqrt{X_s} dW_s\}_{t \geq 0}$ es una martingala. En particular, $0 = E \int_0^0 \sqrt{X_s} dW_s = E \int_0^t \sqrt{X_s} dW_s$. Luego, tomando esperanza en cada lado de la expresión integral de (5.1) resulta que $EX_t = X_0 + E \int_0^t \alpha X_s ds$, de lo cual se tiene la ecuación diferencial ordinaria $\frac{d}{dt} EX_t = \alpha EX_t$ con condición inicial $EX_0 = X_0$. Por lo tanto $EX_t = X_0 e^{\alpha t}$.

Para probar que $Y := \{X_t e^{-\alpha t}\}_{t \geq 0}$ es una martingala no negativa que converge c.s. a una variable aleatoria integrable, observe que tanto $\{X_t\}_{t \geq 0}$ como $\{e^{-\alpha t}\}_{t \geq 0}$ son procesos de Itô

1 para los cuales la fórmula de integración por partes (proposición 4.4) da,

$$X_t e^{-\alpha t} = X_0 + \int_0^t X_s d e^{-\alpha s} + \int_0^t e^{-\alpha s} d X_s,$$

debido a que $[X, e^{-\alpha(\cdot)}]_t \equiv 0$, pues X es continuo y $\{e^{-\alpha t}\}_{t \geq 0}$ es de variación finita. Pero, de la ecuación (4.17) se obtiene que $\int_0^t e^{-\alpha s} d X_s := \int_0^t e^{-\alpha s} \alpha X_s ds + \int_0^t e^{-\alpha s} \beta \sqrt{X_s} d W_s$, y aplicando la proposición 1.6 se tiene que $\int_0^t X_s d e^{-\alpha s} = \int_0^t -\alpha X_s e^{-\alpha s} ds$. De manera que

$$Y_t = X_t e^{-\alpha t} = X_0 + \beta \int_0^t e^{-\alpha s} \sqrt{X_s} d W_s,$$

1 y así, nuevamente por la proposición 4.3, Y es una martingala local con

$$\begin{aligned} E[Y, Y]_t &= E \left[\beta \int_0^t e^{-\alpha s} \sqrt{X_s} d W_s, \beta \int_0^t e^{-\alpha s} \sqrt{X_s} d W_s \right]_t = E \int_0^t \beta^2 e^{-2\alpha s} X_s ds \\ &= \int_0^t \beta^2 e^{-2\alpha s} E X_s ds = \int_0^t \beta^2 e^{-2\alpha s} X_0 e^{\alpha s} ds = X_0 \beta^2 \int_0^t e^{-\alpha s} ds \\ &= (X_0 \beta^2 / \alpha) (1 - e^{-\alpha t}). \end{aligned}$$

Consecuentemente, $\sup_{t \geq 0} E[Y, Y]_t = X_0 \beta^2 \int_0^\infty e^{-\alpha s} ds = X_0 \beta^2 / \alpha < \infty$ pues $\alpha > 0$. Por lo tanto, por 2) de la proposición 3.15, Y es una martingala cuadrado integrable, y así, debido a 1) del corolario 3.4, converge casi seguramente a una variable aleatoria integrable Z . ▲

El resultado que sigue da la probabilidad de extinción de la población.

1 **Proposición 5.3.** Sea $X_0 = x > 0$. Entonces la probabilidad de extinción es $e^{(-2\alpha/\beta^2)x}$ si $\alpha > 0$ y 1 si $\alpha \leq 0$.

Demostración. Considere los tiempos de paro τ_0 y τ_b con $b > 0$ definidos por (4.29). De la ecuación (4.35) se obtiene que

$$1 P_x(\tau_0 < \tau_b) = 1 - \frac{S(x) - S(0)}{S(b) - S(0)} = \frac{S(b) - S(x)}{S(b) - S(0)},$$

donde $S(x)$ está dada por (4.36). En este caso particular resulta

$$S(x) = \int_{x_1}^x \exp \left(- \int_{x_0}^u \frac{2\alpha}{\beta^2} dy \right) du = -\frac{1}{c} e^{cx_0} (e^{-cx} - e^{-cx_1}),$$

1 donde $c = 2\alpha/\beta^2$ y $x_0, x_1 \geq 0$. Así,

$$P_x(\tau_0 < \tau_b) = \frac{e^{-cx} - e^{-cb}}{1 - e^{-cb}}.$$

Observe ahora que $\lim_{b \rightarrow \infty} P_x(\tau_0 < \tau_b) = e^{-cx}$ si $\alpha > 0$ ya que $e^{-cb} \rightarrow 0$ cuando $b \rightarrow \infty$, y $\lim_{b \rightarrow \infty} P_x(\tau_0 < \tau_b) = 1$ si $\alpha \leq 0$ pues por la regla de L' Hopital

$$\lim_{b \rightarrow \infty} \frac{e^{-cx} - e^{-cb}}{1 - e^{-cb}} = \lim_{b \rightarrow \infty} \frac{\frac{d}{db}(e^{-cx} - e^{-cb})}{\frac{d}{db}(1 - e^{-cb})} = \lim_{b \rightarrow \infty} \frac{ce^{-cb}}{ce^{-cb}} = 1.$$

Por lo tanto, la probabilidad de extinción es

$$P_x(\tau_0 < \tau_\infty) = \lim_{b \rightarrow \infty} P_x(\tau_0 < \tau_b) = \begin{cases} e^{(-2\alpha/\beta^2)x} & \text{si } \alpha > 0 \\ 1 & \text{si } \alpha \leq 0, \end{cases}$$

donde $\tau_\infty = \lim_{b \rightarrow \infty} \tau_b$ es el tiempo de explosión, el cual es infinito si la explosión no ocurre. La explosión ocurre si el proceso X toma el valor ∞ en un tiempo τ_∞ finito. ▲

5.2. Difusión de Wright-Fisher

En genética de poblaciones la atención se centra en la cuantificación de las frecuencias alélicas y genotípicas en generaciones sucesivas. En este desarrollo se abordará únicamente la cuantificación de frecuencias alélicas, que es la proporción de alelos de cada locus particular. En este punto es conveniente mencionar algunas terminologías básicas de la genética de poblaciones: todo organismo es inicialmente, en el momento de la concepción, solo una célula. Es esta célula, llamada cigoto (y otras formadas posteriormente que tienen la misma composición genética), la que contiene toda la información genética relevante sobre un individuo e influye en la de su descendencia. Dentro de cada célula hay un cierto número fijo de cromosomas, objetos filiformes que rigen las características heredables de un organismo. Dispuestos en orden lineal en ciertas posiciones (loci) en los cromosomas están los **genes**, las unidades fundamentales de la herencia. En cada posición (locus) hay varios tipos alternativos de genes que pueden ocurrir; las diversas alternativas se llaman **alelos**.

Considere, por simplicidad, una población con un número finito, $2N$, de alelos. Suponga que cada uno de estos alelos es o bien de tipo A_1 o bien de tipo A_2 y que no existe diferencia selectiva entre los alelos A_1 y A_2 . Debido a que el tamaño de la población es finito, las frecuencias alélicas pueden cambiar de generación en generación por un proceso aleatorio conocido como **deriva genética**, la cual es un caso particular de los errores de muestreo. Usando la teoría de la convergencia débil puede demostrarse (véase por ejemplo, el capítulo 10 de [11]) que si los alelos A_1 mutan a A_2 con una tasa $\alpha/2N$ y los alelos A_2 mutan a A_1 con una tasa $\beta/2N$, entonces es posible aproximar la frecuencia X_t del alelo A_1 en la generación t mediante la **difusión de Wright-Fisher**, la cual está dada por la ecuación diferencial estocástica

$$dX_t = [-\alpha X_t + \beta(1 - X_t)] dt + \sqrt{X_t(1 - X_t)} dW_t, \quad 0 < X_0 < 1. \quad (5.3)$$

Observe primero que en este caso particular $\mu(x) = -\alpha x + \beta(1-x)$ satisface la condición de Lipschitz para cualesquiera constantes α, β , y $\sigma(x) = \sqrt{x(1-x)}$ satisface la condición de Hölder de orden $1/2$ pues $\sigma(x) = h(g(x))$ con $h(x) = \sqrt{x}$ siendo Hölder de orden $1/2$ y $g(x) = x(1-x)$ siendo Lipschitz. Por lo tanto, por el teorema de Yamada-Watanabe (teorema 4.3), la difusión de Wright-Fisher existe.

Se realizará el desarrollo del tratamiento expuesto en la sección 13.2 de [18], considerando únicamente el caso de no mutación. En este caso, $\alpha = \beta = 0$, y así, la ecuación para la frecuencia X_t del alelo A_1 en la generación t es

$$dX_t = \sqrt{X_t(1-X_t)}dW_t, \quad 0 < X_0 = x < 1. \quad (5.4)$$

Note que cuando $X_t = 0$ o $X_t = 1$, todos los individuos en la generación t y en las generaciones posteriores son del mismo tipo. Genéticamente hablando, esto es debido a que en ausencia de mutación, una vez que todos los individuos son del genotipo A_1A_1 o bien del genotipo A_2A_2 , no existe variación y la evolución no es posible.

En este caso particular, la función escala está dada por

$$S(x) = \int_0^x \exp\left(-\int_0^u \frac{2\mu(y)}{\sigma^2(y)} dy\right) du = \int_0^x du = x.$$

Por consiguiente, la proposición 4.8 dice que para cualesquiera $0 \leq a < x < b \leq 1$

$$P_x(\tau_b < \tau_a) = \frac{S(x) - S(a)}{S(b) - S(a)} = \frac{x - a}{b - a}. \quad (5.5)$$

Como se observó antes, el proceso es parado cuando alcanza los valores 0 o 1 porque a partir de ese tiempo, τ , todos los individuos son del mismo tipo. Este fenómeno es llamado **fijación**. La probabilidad de que la fijación ocurra en A_1 al iniciar el proceso con la proporción $x = X_0$ del alelo A_1 es x , y por ende, con la probabilidad complementaria la fijación ocurre en A_2 , pues de la ecuación (5.5) se tiene que

$$P_x(\tau_1 < \tau_0) = x.$$

Debido a la proposición 4.7, el tiempo esperado para que la fijación ocurra es dada por la solución de la ecuación diferencial ordinaria con condiciones de frontera

$$\frac{x(1-x)}{2} f''(x) = -1, \quad f(0) = f(1) = 0. \quad (5.6)$$

Se encontrará la solución de la ecuación (5.6). Note que si $g(x) = f'(x)$, entonces

$$g(x) = \int \frac{-2}{x(1-x)} dx = -2 \int \left(\frac{1}{x} + \frac{1}{1-x} \right) dx = -2 [\log(x) - \log(1-x)] + c_1.$$

Luego, usando el hecho de que $\int \log(x)dx = x \log(x) - x$, se obtiene que

$$\begin{aligned} f(x) &= \int g(x)dx = -2 \left[\int \log(x)dx - \int \log(1-x)dx \right] + \int c_1 dx \\ &= -2 [x \log(x) + (1-x) \log(1-x) - 1] + c_1 x + c_2. \end{aligned}$$

Así, usando las condiciones iniciales $f(0) = f(1) = 0$, se tiene que $c_1 = 0$ y $c_2 = -2$. Por lo tanto, el tiempo medio esperado para que la fijación ocurra en A_1 cuando el proceso inicia con una proporción x de alelos de tipo A_1 es

$$f(x) = E_x(\tau) = -2 [x \log(x) + (1-x) \log(1-x)].$$

Conclusión

Esta tesis presenta una introducción a las aplicaciones del cálculo estocástico en la biología, en particular en la reproducción y evolución genética de poblaciones. Su propósito es contribuir en la difusión del uso de la teoría del cálculo estocástico en las ciencias biológicas; aplicaciones que han sido mucho menos difundidas que las financieras. Con el propósito de hacer de esta tesis un trabajo accesible, se incluyeron conceptos y resultados preliminares que van desde el cálculo avanzado, teoría de probabilidad, de procesos estocásticos y del cálculo estocástico. Se dan también referencias en donde se puede profundizar en dichas teorías.

En lo que respecta a los modelos aquí analizados, cabe destacar que el caso de la difusión de ramificación de Feller, aparte de ser de interés por sí mismo, cuando el tamaño de la población es grande, representa una muy buena aproximación al proceso de ramificación simple (véase por ejemplo, [13] pág. 237). También, en lo que respecta a la difusión de Wright-Fisher, puede verse que, en presencia de la deriva genética en una población finita apropiada, esta aproximación es una muy buena aproximación al correspondiente modelo discreto para el cálculo de frecuencias alélicas, debido a que, por ejemplo, la probabilidad de la fijación y el tiempo esperado para que ésta ocurra son iguales en ambos modelos (véase [12] pág. 160).

Bibliografía

- [1] Apostol, T. M. (1974). *Mathematical Analysis*, Second Edition. USA: Addison-Wesley.
- [2] Ash, R. B. (2000). *Probability and Measure Theory*, Second Edition. USA: Academic Press.
- [3] Athreya, K. B., Ney, P. E. (1972). *Branching Processes*. Germany: Springer-Verlag.
- [4] Billingsley P. (1995). *Probability and Measure*, Third Edition. USA: Wiley.
- [5] Brzeźniak, Z., Zastawniak, T. (1999). *Basic Stochastic Processes: A Course Through Exercises*. Great Britain: Springer-Verlag.
- [6] Capasso, V., Bakstein D. (2005). *An introduction to continuous-time stochastic processes: theory, models, and applications to finance, biology, and medicine*. USA: Birkhäuser.
- [7] Carothers, N. L. (2000). *Real Analysis*. USA: Cambridge University Press.
- [8] Casella, G., Berger, R. L. (2002). *Statistical Inference*, Second Edition. USA: Duxbury-Wadsworth.
- [9] Dudley, R. M. (2004). *Real Analysis and Probability*. USA: Cambridge University Press.
- [10] Durrett, R. (2010). *Probability: Theory and Examples*, Fourth Edition. USA: Cambridge University Press.
- [11] Ethier, S. N., Kurtz, T. G. (1986). *Markov Processes: Characterization and Convergence*. USA: Wiley.
- [12] Ewens, W. J. (2004). *Mathematical Population Genetics: I. Theoretical Introduction*, Second Edition. USA: Springer-Verlag.
- [13] Feller, W. (1951). *Diffusion Processes in Genetics*, Second Berkely Symposium on Math. Stat. and Probab., 227-246.
- [14] Friedman, A. (1975). *Stochastic Differential Equations and Applications*, Volume 1. USA: Academic Press.

- [15] Gihman, I. I., Skorohod, A. V. (1972). *Stochastic Differential Equations*. USA: Springer-Verlag.
- [16] Kallianpur, G. (1980). *Stochastic Filtering Theory*. USA: Springer-Verlag.
- [17] Karatzas, I., Shreve, S. E. (1998). *Brownian Motion and Stochastic Calculus*, Second Edition. USA: Springer-Verlag.
- [18] Klebaner, F. C. (2005). *Introduction to Stochastic Calculus with Applications*, Second Edition. Singapore: Imperial College Press.
- [19] Knapp, W. Anthony. (2005). *Basic Real Analysis*. USA: Birkhäuser.
- [20] Kuo, H. H. (2006). *Introduction to Stochastic Integration*. USA: Springer-Verlag.
- [21] Liptser, R. S., Shiriyayev, A. N. (1989). *Theory of Martingales*. Russia: Kluwer.
- [22] Meyer, M. (2001). *Continuous Stochastic Calculus with Applications to Finance*. USA: Chapman & Hall/CRC.
- [23] Pinsky, R. G. (1995). *Positive Harmonic Functions and Diffusion*. USA: Cambridge University Press.
- [24] Protter, P. E. (2005). *Stochastic Integration and Differential Equations*. Second Edition. Germany: Springer-Verlag.
- [25] Revuz, D., Yor, M. (2005). *Continuous Martingales and Brownian Motion*, Third Edition. USA: Springer-Verlag.
- [26] Rogers, L. C. G., Williams, D. (2000). *Diffusions, Markov Processes and Martingales*, Volume 2: Itô Calculus. Second Edition. USA: Cambridge University Press.
- [27] Rudin, W. (1976). *Principles of Mathematical Analysis*, Third Edition. USA: McGraw-Hill.
- [28] Stroock, W. D., Varadhan, S. R. S. (1979). *Multidimensional Diffusion Processes*. Germany: Springer-Verlag.
- [29] Uribe Bravo, G. (2014). *Notas del Curso Avanzado de Probabilidad e Integración Estocástica*. Instituto de Matemáticas, UNAM.

Aplicaciones del Cálculo Estocástico en la Biología

ORIGINALITY REPORT

11%

SIMILARITY INDEX

PRIMARY SOURCES

1	sahuarus.unison.mx Internet	1641 words — 8%
2	hdl.handle.net Internet	80 words — < 1%
3	lya.fciencias.unam.mx Internet	57 words — < 1%
4	ia800903.us.archive.org Internet	45 words — < 1%
5	1library.co Internet	31 words — < 1%
6	dokumen.pub Internet	28 words — < 1%
7	pt.scribd.com Internet	28 words — < 1%
8	documentop.com Internet	24 words — < 1%
9	vdocumento.com Internet	23 words — < 1%
10	www.coursehero.com Internet	22 words — < 1%
11	www.scribd.com	

21 words — < 1%

12 "Financial Mathematics, Derivatives and Structured Products", 'Springer Science and Business Media LLC', 2019

Internet

18 words — < 1%

13 ANULEKHA DHARA, APARNA MEHRA. "SEQUENTIAL LAGRANGE MULTIPLIER CONDITIONS FOR MINIMAX PROGRAMMING PROBLEMS", Asia-Pacific Journal of Operational Research, 2011

Crossref

17 words — < 1%

14 dedzago.github.io

Internet

15 words — < 1%

15 www.12000.org

Internet

15 words — < 1%

16 uu.diva-portal.org

Internet

14 words — < 1%

17 Medina, Ranu Castañeda. "Control Óptimo Estocástico de Una Cuenta Individual de Capitalización en el Sistema Privado de Pensiones Del Perú", Pontificia Universidad Catolica del Peru - CENTRUM Catolica (Peru), 2022

ProQuest

13 words — < 1%

18 idoc.tips

Internet

13 words — < 1%

19 mafalda.univalle.edu.co

Internet

13 words — < 1%

20 repositorio.utp.edu.co

Internet

13 words — < 1%

21 mitschriebwiki.nomeata.de

12 words — < 1 %

22 www.slideshare.net
Internet

12 words — < 1 %

23 www.yumpu.com
Internet

12 words — < 1 %

24 archive.org
Internet

11 words — < 1 %

25 plorans.ugr.es
Internet

11 words — < 1 %

26 uvadoc.uva.es
Internet

11 words — < 1 %

27 Guanxing Fu, Paulwin Graewe, Ulrich Horst,
Alexandre Popier. "A Mean Field Game of Optimal
Portfolio Liquidation", Mathematics of Operations Research,
2021
Crossref

10 words — < 1 %

28 José Oscar Romero Martínez. "Estudio de la clase
de matrices $\{K, s+1\}$ -potentes", 'Universitat
Politecnica de Valencia', 2015
Internet

10 words — < 1 %

29 docshare.tips
Internet

10 words — < 1 %

30 doku.pub
Internet

10 words — < 1 %

31 www.member.uni-oldenburg.de
Internet

10 words — < 1 %

32 www.menaldi.wayne.edu
Internet

10 words — < 1 %

EXCLUDE QUOTES

ON

EXCLUDE BIBLIOGRAPHY

ON

EXCLUDE SOURCES

OFF

EXCLUDE MATCHES

< 10 WORDS