Universidad Nacional Mayor de San Marcos
Universidad del Perú. Decana de América
Facultad de Ciencias Matemáticas
Escuela Profesional de Matemática

Una generalización de la integral clásica para resolver ecuaciones diferenciales estocásticas

TESIS

Para optar el Título Profesional de Licenciado en Matemática

AUTOR

Elmer LÉVANO HUAMACCTO

ASESOR

Víctor Emilio CARRERA BARRANTES

Lima, Perú

2020

Referencia bibliográfica

Lévano, E. (2020). *Una generalización de la integral clásica para resolver ecuaciones diferenciales estocásticas*. Tesis para optar el título profesional de Licenciado en Matemática. Escuela Profesional de Matemática, Facultad de Ciencias Matemáticas, Universidad Nacional Mayor de San Marcos, Lima, Perú.

Hoja de metadatos complementarios

Código ORCID del autor (dato opcional): 0000-0001-7660-9179

Código ORCID del asesor o asesores (dato obligatorio): 0000-0002-2048-9494

DNI del autor: 41986595

Grupo de investigación: Ninguno

Institución que financia parcial o totalmente la investigación: Autofinanciar

Ubicación geográfica donde se desarrolló la investigación. Debe incluir localidades y/o coordenadas geográficas: Psj. Senda Blanca 104, Lima.

Año o rango de años que la investigación abarcó: Un año.
Inicio: enero 2019
Fin: diciembre 2019

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
(Universidad del Perú, DECANA DE AMÉRICA)
FACULTAD DE CIENCIAS MATEMATICAS
Ciudad Universitaria – Av. Venezuela S/N cuadra 34
Teléfono: 619-7000, Anexo 1610
Correo Postal: 05-0021, E-mail: eapmat@unmsm.edu.pe
Lima – Perú

Escuela Profesional de Matemática

ACTA DE SUSTENTACIÓN DE TESIS PARA OPTAR EL TÍTULO PROFESIONAL DE LICENCIADO EN MATEMÁTICA

En la UNMSM - Ciudad Universitaria - Facultad de Ciencias Matemáticas, siendo las ...12:00.... horas del Lunes 10 de febrero de 2020, se reunieron los docentes designados como Miembros del Jurado Evaluador de Tesis: Dr. Jorge Alberto Coripaco Huarcaya (PRESIDENTE), Mg. Willy David Barahona Martínez (MIEMBRO), Lic. Víctor Emilio Carrera Barrantes (MIEMBRO ASESOR), para la sustentación de la tesis titulada: «UNA GENERALIZACIÓN DE LA INTEGRAL CLÁSICA PARA RESOLVER ECUACIONES DIFERENCIALES ESTOCÁSTICAS», presentado por el señor Bachiller ELMER LÉVANO HUAMACCTO, para optar el Título Profesional de Licenciado en Matemática.

Luego de la exposición del tesista, el Presidente del Jurado invitó a dar respuestas a las preguntas que le formulen.

Hecha la evaluación correspondiente por los Miembros del Jurado, el tesista mereció la aprobación unánime obteniendo como calificativo promedio la nota de:

..Dieciocho (sobresaliente)........(18).

A continuación, el Presidente del Jurado, Dr. Jorge Alberto Coripaco Huarcaya, manifestó que el señor Bachiller ELMER LÉVANO HUAMACCTO, en vista de haber aprobado la sustentación de su tesis, será propuesto para que se le otorgue el Título Profesional de Licenciado en Matemática.

Siendo las ..12:45..... horas se levantó la sesión firmando para constancia la presente Acta en tres (3) copias originales.

DR. JORGE ALBERTO CORIPACO HUARCAYA
PRESIDENTE

MG. WILLY DAVID BARAHONA MARTÍNEZ
MIEMBRO

LIC. VÍCTOR EMILIO CARRERA BARRANTES
MIEMBRO ASESOR

Dedico esta tese de forma póstuma, a meu orientador e amigo Prof. Dr. Raúl Moisés Izaguirre Maguiña, pelo seus conselhos em todas essas reuniões de café, abençoado por compartilhar seus valiosos conhecimentos por onde quer que o senhor esteja.

AGRADECIMIENTOS

A mi familia, en especial a mi hermano Carlos, por los grandes sacrifícios personales por la familia;

a mi orientador, por dirigir con grandes aciertos esta tesis Prof. Lic. Víctor Emilio Carrera Barrantes;

RESUMEN

Una Generalización de la Integral Clásica para Resolver Ecuaciones Diferenciales Estocásticas

Elmer Lévano Huamaccto

Asesor: Lic. Víctor Emilio Carrera Barrantes
Título a obtener: Licenciado en Matemática

Este trabajo presenta resultados de importancia sobre la teoría de integración y ecuaciones diferenciales. Sistemas afectados por ruidos estocásticos son tratados por ecuaciones diferenciales estocásticas tanto en dimensión finita como infinita. En particular, mostraremos la construcción de la integral estocástica para abordar las ecuaciones diferenciales estocásticas. También presentamos algunos ejemplos, principalmente la formulación de problemas estocásticos y las herramientas para resolverlos. Cada ejemplo pueden diferir considerablemente de su contraparte determinista.

Palabras Claves:Stieltjes Análisis matemático;
teoría de la medida;
integral estocástica;
procesos de difusión;
ecuaciones diferenciales estocásticas.

ABSTRACT

A Generalization of Classical Integral to Solve Stochastic Differential Equations

Elmer Lévano Huamaccto

February - 2020

Adviser: Lic. Víctor Emilio Carrera Barrantes

Degree to obtain: Math degree

In this work presents important results about integration theory and differential equations. Systems affected by stochastic noise are treated by stochastic differential equations in both finite and infinite dimensions. In particular, we show the construction of the stochastic integral to deal with stochastic differential equations. We also present some examples, mainly the formulation of stochastic problems and the tools to solve them. Each example can differ considerably from its deterministic counterpart.

Key Words: Mathematical analysis;
measure theory;
stochastic integral;
stochastic processes;
differential stochastic equation.

Índice general

Introducción

Uno de los problemas que dio origen a las matemáticas, a lo largo de la historia, fue el problema del conteo [14]. Cuando se requiere contar ciertas cantidades de formas diferentes. Una manera eficiente de contar, es agrupar por cantidades exactas o por tamaño lo que se está contando. Este problema nos coloca en la necesidad de asignar un número para describir, por ejemplo, una área, el volumen y otros conceptos que surgen por la misma necesidad matemática al combinar mas datos. De esta forma, el acto de contar llevo a la humanidad a audaces descubrimientos en todas las áreas de las matemática, entre los que se destaca el cálculo.

En el cálculo de Newton-Leibniz, aprendemos el significado de la diferenciación y la integración de funciones deterministas. En ese ámbito, un concepto fundamental es la regla de la cadena. Esta regla resuelve la derivada de la composición de funciones diferenciables. En ese sentido, surge una pregunta natural, existe alguna representación de la regla de la cadena cuando estas funciones pierden la propiedad de la diferenciabilidad? La respuesta es afirmativa, debido a que la regla de la cadena puede ser escrita en forma de integrales indefinidas, y así se produce una expresión alternativa a la regla original. Primero recordamos la relación entre la derivada y la integral clásica en el cálculo de Leibniz-Newton. La antiderivada o primitiva de un función es un concepto que envuelve la derivada, en ese sentido podemos aprovechar muy poco en nuestra regla de la cadena alternativa.

Una función $f : [a, b] \to \mathbb{R}$ es llamada Riemann integrable si existe el siguiente limite

$$\int_a^b f(t)dt = \lim_{\|\Delta_n\|\to 0} \sum_{i=1}^{n} f(\tau_i)(t_i - t_{i-1}),$$

en donde, $\Delta_n = \{t_0, t_1, \cdots, t_{n-1}, t_n\}$ es una partición del intervalo $[a, b]$ con la convención que $a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$, $\|\Delta_n\| = \max_{1\leq i\leq n}(t_i - t_{i-1})$, y τ_i es un punto en el interior del intervalo $[t_{i-1}, t_i]$. Si $f : [a, b] \to \mathbb{R}$ es una función continua sobre $[a, b]$ entonces esta es Riemann integrable. Un resultado conocido en la literatura especializada es cuando una función limitada en $[a, b]$ es Riemann integrable si, solamente si es continua por trechos y los puntos de discontinuidad son de medida nula (continua casi seguramente).

Por otro lado, en el cálculo avanzado, la integral de Riemann-Stieltjes es definida a través de los mismos procedimientos usados en la integral de Riemann, osea, "Partición-Evaluación-Suma-Limite". Cuando se trata con funciones aleatorias, como por ejemplo el movimiento browniano, la regla de la cadena definida en el cálculo de Leibniz-Newton no necesariamente se cumple.

Un objeto con movimiento browniano se mueve de forma rápida e irregular, de esta forma se puede decir que este objeto define una trayectoria que es vista como una función no diferenciable, osea, perdemos la diferenciabilidad de la trayectoria en casi todo punto. Como anteriormente fue expuesto, la diferenciabilidad es un requisito necesario para producir la regla de la cadena, en ese sentido, colocamos nuestros ojos a la forma integral de la regla.

En 1944, Kiyosi Itô publico su mas celebre articulo, “Stochastic Integral” en la Imperial Academy (Tokyo) [12]. Este fue el inicio del cálculo de Itô o también conocido como “cálculo estocástico”, que a decir verdad es una extención del cálculo de Leibniz-Newton para funciones aleatorias. En ese artículo, Itô también introdujo un resultado sorprendente, conocido en la actualidad como la formula de Itô. Esta formula en realidad es la regla de la cadena para el cálculo de estocástico. Como en el cálculo estocástico se trabajan con funciones aleatorias (trayectorias no diferenciables) esta formula solo puede ser escrita en forma de integrales.

El método usado por Itô para definir la integral estocástica es en realidad una combinación de técnicas, la integral de Riemann-Stieltjes (para el integrador) y la integral de Lebesgue (para el integrando).

El cálculo de Itô tiene muchas aplicaciones actualmente, entre las cuales se destacan las Ecuaciones Diferenciales Estocásticas (EDE).

Considere una Ecuación Diferencial Ordinaria (EDO) de la siguiente forma

$$x'(t) = f(x(t), t), \quad t \in (0, T), \quad x(0) = x_0,$$

en donde $f : \mathbb{R}^n \times [0, T] \to \mathbb{R}^n$ es una función con suficiente regularidad y x_0 es una constante fija. En ese sentido, la solución $x(t)$ buscada es una función $x : [0, T] \to \mathbb{R}^n$ satisfaciendo las igualdades anteriores. Es común escribir la EDO en su forma diferencial

$$dx - f(x, t)dt = 0$$

Si estamos interesados en que un modelo matemático describa también la acción de algún efecto aleatorio que influencie al sistema, tal como algún tipo de perturbación o ruido. En ese sentido, somos llevados a las ecuaciones diferenciales estocásticas

$$dx_t = b(x_t, t)dt + \sigma(x_t, t)dW_t, \quad x_0 = \xi_0,$$

en donde $b : \mathbb{R}^n \times [0, T] \to \mathbb{R}^n$ y $\sigma : \mathbb{R}^n \times [0, T] \to \mathbb{R}^{n\times r}$ son funciones afectadas por el proceso $x_t \in \mathbb{R}^n$, $\{W_r\}_{r\geq 0}$ es un proceso de Wiener (movimiento browniano). En este caso la solución buscada es un proceso estocástico $\{x_t\}_{t\in[0,T]}$. Una interpretación de la solución de la EDE puede ser escrita de la siguiente forma

$$x_t = x_0 + \int_0^t b(x_s, s)ds + \int_0^t \sigma(x_s, s)dW_s.$$

Mirando cuidadosamente el termino $\int_0^t \sigma(x_s, s)dW_s$ en la ecuación anterior, es notoria la necesidad de definir matemáticamente esa integral. Esta forma nos conduce a dar un sentido a la integral $\int_0^t \sigma(x_s, s)dW_s$ que forma parte de la solución de la EDE. Este problema será abordado en el Capítulo 2. También en este capítulo tratamos algunas propiedades de la integral estocástica. Las EDE son usadas para representar diversos modelos matemáticos, y por eso su importancia, como por ejemplo en la economía el modelo “Stock Prices”, en la física el modelo “Thermal Fluctuations”, en la biología los diversos modelos de crecimiento poblacional, etc.

Organización del texto

Para facilitar a lectura y la comprención de este trabajo, él fue estructurado de la siguiente forma que pasamos a describir.

- En el Capítulo 1, colocamos todos los conceptos necesarios para una buena comprensión de este documento: la integral de Riemann-Stieltjes sobre un espacio euclidiano. Ademas en este capítulo, se presentan todas las propiedades de la integral de Riemann-Stieltjes que serán necesarias al momento de definir la integral estocástica.
- El Capítulo 2 presenta el resultado principal de esta tesis. La integral estocástica o también conocida como integral de Itô y sus propiedades son las herramientas necesarias del cálculo estocástico o también conocido como cálculo de Itô. La formula de Itô es una propiedad fundamental para resolver ecuaciones diferenciales estocásticas.
- El Capítulo 3, son usadas las herramientas definidas en el capítulo anterior para resolver, a manera de ejemplo, diversas ecuaciones diferenciales estocásticas. En este capítulo también se desarrollan los teoremas clásicos de existencia y unicidad de las soluciones de una ecuación diferencial estocástica.
- Finalmente, en el Capítulo 4, presentamos las conclusiones y direccionamiento de trabajos futuros a partir de este manuscrito.

Capítulo 1

Fundamentos Teóricos

En este capítulo, se definen algunos elementos básicos para el desenvolvimiento de este trabajo entre los cuales podemos destacar: teoría de la medida, análisis funcional, teoría de probabilidades, procesos estocásticos, integral estocástica, ecuaciones diferenciales estocásticas, entre otros tópicos que serán fundamentales en el desarrollo de los modelos matemáticos tratados en el interior. En caso sea necesario de algún esclarecimiento, colocamos en la parte final la bibliografía correspondiente para ser consultada.

Primeramente, presentamos algunos conceptos básicos de la teoría de probabilidades.

1.1. Elementos de la teoría de probabilidad

Definición 1.1.1. *Sea Ω un conjunto no vació. Una colección $\mathcal{F}$ de subconjuntos de Ω es denominado una σ-álgebra si satisface las siguientes propiedades:*

i) $\Omega \in \mathcal{F}$;

ii) Si $A \in \mathcal{F}$, *entonces* $A^c \in \mathcal{F}$;

iii) Si $A_1, A_2, A_3 \cdots \in \mathcal{F}$, *entonces* $\bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}$ *(unión numerable).*

En seguida presentamos algunos ejemplos de σ-álgebras.

Ejemplo 1.1.1. *(a) Sea $\mathcal{F} = \mathcal{P}(\Omega)$, el conjunto de partes. Entonces veamos si $\mathcal{F}$ es una σ-álgebra. Para esto debemos corroborar cada condición dada en la Definición 1.1.1. Claramente $\mathcal{P}(\Omega) \in \mathcal{F}$, con esto la condición (i) es satisfecha. Enseguida, tomamos $A \in \mathcal{F}$, como sabemos que $A^c \in \mathcal{P}(\Omega)$, entonces $A^c \in \mathcal{F}$ y así la condición (ii) es satisfecha. Si $A_1, A_2, A_3 \cdots \in \mathcal{P}(\Omega)$ entonces $\cup_{i=1}^{\infty} A_i \in \mathcal{P}(\Omega) = \mathcal{F}$ y así la condición (iii) es satisfecha. Por lo tanto $\mathcal{F}$ es una σ-álgebra.*

(b) La σ-álgebra trivial se define como $\mathcal{F} = \{\emptyset, \Omega\}$. De esta forma es sencillo verificar todas las condiciones de la Definición 1.1.1.

Sean $\mathcal{F}_1, \mathcal{F}_2$ dos σ-álgebras en Ω. Se dice que $\mathcal{F}_1$ es una sub-σ-álgebra de $\mathcal{F}_2$ si, y solo si $\mathcal{F}_1 \subset \mathcal{F}_2$. Por ejemplo $\mathcal{F}_1 = \{\emptyset, \Omega\}$ y $\mathcal{F}_2 = \mathcal{P}(\Omega)$.

Definición 1.1.2. *El par $(\Omega, \mathcal{F})$ es denominado espacio medible si, y solo si $\mathcal{F}$ es una σ-álgebra. Cada elemento A de la colección $\mathcal{F}$ será denominado conjunto medible.*

Ejemplo 1.1.2. *(a) Todo subconjunto abierto de $\mathbb{R}^n$ es medible.*

(b) Sea $\Omega = \{x_1, x_2, \ldots\}$ un conjunto finito o numerable y sea $p_1, p_2, \ldots$ una sucesión de números no negativos. Sea $\mathcal{F} = \mathcal{P}(\Omega)$ y para cada $A \subset \Omega$ definimos

$$\mu(A) = \sum_{x_i \in A} p_i.$$

Así μ es una medida en $(\Omega, \mathcal{P}(\Omega))$ un espacio medible para la que $\mu(\{x_i\}) = p_i$, $i = 1, 2, \ldots$

Lema 1.1.1. *[17, pag. 7, cap. 2] Sea Ω un conjunto no vació. Si $A \in \mathcal{F}$ entonces existe una única σ-álgebra menor, denotada por $\sigma(A)$, conteniendo A.*

Definición 1.1.3. *Sea Ω un conjunto no vació y $\mathcal{T} \subset \mathcal{F}$ es una topológia en Ω si, y solo si*

i) $\phi, \Omega \in \mathcal{T}$;

ii) $U_1, U_2, \cdots, U_n \in \mathcal{T}$ *entonces* $\displaystyle\bigcap_{i=1}^{n} U_i \in \mathcal{F}$;

iii) $U_1, U_2, \cdots, U_n \in \mathcal{T}$ *entonces* $\displaystyle\bigcup_{i=1}^{n} U_i \in \mathcal{F}$.

$(\Omega, \mathcal{T})$ *es un espacio topológico y los elementos de $\mathcal{T}$ son denominados abiertos.*

Definición 1.1.4. $(\mathcal{M}, d)$ *es un espacio métrico si, y solo si $\mathcal{M}$ es un conjunto no vació y $d : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$, tal que satisfaz las siguientes condiciones*

i) $d(x, y) \geq 0$ *para todo $x, y \in \mathcal{M}$ y será "$=$" si $x = y$;*

ii) $d(x, y) = d(y, x)$ *para cualquier $x, y \in \mathcal{M}$;*

iii) $d(x, y) \leq d(x, z) + d(z, y)$ *para cualquier $x, y, z \in \mathcal{M}$.*

$(\Omega, \mathcal{M})$ *es un espacio topológico y los elementos de $\mathcal{M}$ son denominados de abiertos.*

A continuación damos unos pocos ejemplos de espacios métricos.

Ejemplo 1.1.3. *(a) $\mathbb{R}^n$ forma un espacio métrico vía la función distancia $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ tal que $d(x, y) = \|y - x\|$.*

(b) El conjunto de las funciones continuas $C[a, b]$ en el intervalo $[a, b]$, es un espacio métrico vía la distancia $d : C[a, b] \times C[a, b] \to \mathbb{R}_+$ definida por $d(f, g) = \sup \|f - g\|$.

Definición 1.1.5. *Dado un espacio topológico $(\Omega, \mathcal{M})$ y $U \subset \mathcal{M}$. Por Lema 1.1.1, existe una σ-álgebra mínima, denotada por $\mathcal{B}(U)$, conteniendo U. $\mathcal{B}(U)$ es también una σ-álgebra del espacio topológico. De esta forma, cada elemento $B \in \mathcal{B}(U)$ es denominado conjunto de Borel.*

Ejemplo 1.1.4. *Los subconjuntos abiertos en $\mathbb{R}^n$ forman conjuntos de Borel.*

Un espacio métrico es dicho "completo" cuando todas las sucesiones de Cauchy convergen para um limite que pertenece al espacio. Por otro lado, para estudiar a convergencia de series de funciones, es necesario exigir que el espacio métrico tenga propiedades algébricas que nos permitan la suma de funciones, entonces este es un motivo para estudiar espacios vectoriales normados.

Definición 1.1.6. *Sea $\mathcal{V}$ un espacio vectorial en $\mathbb{R}$ y $\|\cdot\| : \mathcal{V} \to \mathbb{R}$ satisfaciendo las siguientes condiciones:*

i) $\|x\| \geq 0$ *para cualquier* $x \in \mathcal{V}$ *y será "$=$" si* $x = 0$*;*

ii) $\|\lambda x\| = |\lambda| \, \|x\|$ *para cualquier* $x \in \mathcal{V}$ *e* $\lambda \in \mathbb{R}$*;*

iii) $\|x + y\| \leq \|x\| + \|y\|$ *para cualquier* $x, y \in \mathcal{V}$*.*

De esta forma, $(\mathcal{V}, \|\cdot\|)$ es un espacio normado y la métrica es dada por $d(x, y) = \|x-y\|$. Se dice que un espacio es de Banach si es un espacio vectorial normado completo.

Ejemplo 1.1.5. $\mathbb{R}^n$ *es un espacio normado vía* $\|x\| = d(x, x)$ *para todo* $x \in \mathbb{R}^n$*.*

Definición 1.1.7. *Sea $\mathcal{V}$ un espacio vectorial en $\mathbb{R}$ y $\langle, \rangle : \mathcal{V}\times\mathcal{V} \to \mathbb{R}$ satisfaz las siguientes condiciones:*

i) $\langle x, x\rangle \geq 0$ *e igual a cero si, y solo si* $x = 0$ *para cualquier* $x \in \mathcal{V}$*;*

ii) $\langle x, y\rangle = \langle y, x\rangle$ *para cualquier* $x, y \in \mathcal{V}$*;*

iii) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$ *para cualquier* $x, y, z \in \mathcal{V}$*;*

iv) $\langle \lambda x, y\rangle = \lambda \langle x, y\rangle$ *para cualquier* $x, y \in \mathcal{V}$ *y* $\lambda \in \mathbb{R}$*.*

$(\mathcal{V}, \langle\cdot,\cdot\rangle)$ es un espacio dotado de un producto interno, osea, con nociones de distancia y ángulo. Se dice espacio de Hilbert a todo espacio vectorial dotado de un producto interno.

Ejemplo 1.1.6. $\mathbb{R}^n$ *es un espacio de Hilbert si es completo con respecto a la norma* $\|x\| = \sqrt{\langle x, x\rangle}$*.*

Definición 1.1.8. *Sean $(\Omega_1, \mathcal{F}_{\Omega_1})$ y $(\Omega_2, \mathcal{F}_{\Omega_2})$ dos espacios medibles. Se dice que $f : \Omega_1 \to \Omega_2$ es una función medible si, y solo si las pre-imágenes son conjuntos medibles, esto es, $f^{-1}(A_2) \in \mathcal{F}_{\Omega_1}$ para cualquier $A_2 \in \mathcal{F}_{\Omega_2}$.*

Definición 1.1.9. *Sea $X : (\Omega, \mathcal{F}) \to (R, \mathcal{M})$ una función, en que, $(\Omega, \mathcal{F})$ es un espacio medible y $(R, \mathcal{M})$ es un espacio topológico. Se dice que X es $\mathcal{F}$-medible si $X^{-1}(U) \in \mathcal{F}$ para todo $U \in \mathcal{M}$.*

Ejemplo 1.1.7. *La función característica definida sobre un conjunto medible Ω dada por*

$$\mathbb{1}_\Omega(x) = \begin{cases} 0 \, si & x \notin \Omega \\ 1 \, si & x \in \Omega \end{cases}$$

es una función medible, en efecto, para cualquier abierto G en $\mathbb{R}$, si estudiamos

$$\mathbb{1}_\Omega^{-1}(G) = \begin{cases} \emptyset \, si & 0 \notin G,\, 1 \notin G \\ \mathbb{R} \setminus \Omega \, si & 0 \in G,\, 1 \notin G \\ \Omega \, si & 0 \notin G,\, 1 \in G \\ \mathbb{R} \, si & 0 \in G,\, 1 \in G \end{cases}$$

y claramente en cualquier caso es una función medible.

Definición 1.1.10. *Considere un espacio medible* $(\Omega, \mathcal{F})$. *Se dice que una aplicación* $\mu : \mathcal{F} \to [0, +\infty)$ *es una medida si las siguientes condiciones son verificadas:*

i) $\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$, *si* $A_i \in \mathcal{F}$; $i \in \{1, 2, \cdots\}$ *y* $A_i \cap A_j = \emptyset$ *para* $i \neq j$;

ii) $\mu(A) < +\infty$ *para cualquier* $A \in \mathcal{F}$.

En este caso, la terna $(\Omega, \mathcal{F}, \mu)$ *es denominada espacio de medida.*

Denotamos una medida de probabilidad por $\mathbf{P}$, o también conocida como "ley de probabilidad", o simplemente "probabilidad", a la función real $\mathbf{P} : \mathcal{F} \to [0, 1]$ que satisface los siguientes axiomas (axiomas de Kolmogorov, véase [2]):

- $\mathbf{P}(\Omega) = 1$;
- $\mathbf{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbf{P}(A_i)$, si $A_i \in \mathcal{F}$; $i \in \{1, 2, \cdots\}$ y $A_i \cap A_j = \emptyset$ para $i \neq j$;
- $\mathbf{P}(A) \geq 0$, para todo $A \in \mathcal{F}$.

La terna $(\Omega, \mathcal{F}, \mathbf{P})$ es denominada "espacio de probabilidad". Se dice que un espacio de probabilidad es completo, si para cualesquiera $B \in \mathcal{F}$ con $\mathbf{P}(B) = 0$, entonces para todo $A \subset B$ se tiene que $A \in \mathcal{F}$.

Una probabilidad condicional es una nueva medida y es definida de la siguiente manera.

Definición 1.1.11. *Dado* $A, B \in \mathcal{F}$ *en un espacio de probabilidad* $(\Omega, \mathcal{F}, \mathbf{P})$, *se define la probabilidad condicional de* A *dado que ocurrió* B, *o simplemente probabilidad condicional de* A *dado* B, *por*

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

En la Definición 1.1.11 surgen algunas dificultades, por ejemplo cuando B tiene un solo elemento, entonces $\mathbf{P}(B) = 0$, pero nos gustaría asignar un significado a una probabilidad condicional como $\mathbf{P}(A|X = 1/2)$. A rigor se necesita de una nueva definición de probabilidad condicional que supere la dificultad anterior.

Definición 1.1.12. *[8, def. 3.16, pag. 306] Sea* $(\Omega, \mathcal{F}, \mathbf{P})$ *un espacio de probabilidad,* $(E, \mathcal{B})$ *un espacio medible y* $\tau : \Omega \to E$ *una función medible. Una probabilidad condicional regular con respecto a* τ *es definida como una aplicación* $\nu : E \times \mathcal{F} \to [0, 1]$ *tal que*

$$\mathbf{P}\left(A \cap \tau^{-1}(B)\right) = \int_B \nu(x, A) \mathbf{P}(\tau^{-1}(dx))$$

para todo $A \in \mathcal{F}$ *y* $B \in \mathcal{B}$.

Ejemplo 1.1.8. *Para continuar con nuestro la motivación anterior, consideramos el evento* $\{X = x_0\}$ *y escribimos*

$$\mathbf{P}(A|X = x_0) = \nu(x_0, A) = \lim_{\epsilon \to 0^+} \frac{\mathbf{P}(A \cap \{x_0 - \epsilon < X < x_0 + \epsilon\})}{\mathbf{P}(\{x_0 - \epsilon < X < x_0 + \epsilon\})},$$

(en donde $x_0 = 1/2$*) si este limite existe, es una condición de probabilidad regular para la v.a.r* X, *restringido a* $\sup X$.

De este punto en adelante, considere la terna $(\Omega, \mathcal{F}, \mathbf{P})$ como un espacio de probabilidad completo. Se dice que un evento $A \in \mathcal{F}$ ocurre casi seguramente (abreviadamente, c.s), si $\mathbf{P}(A) = 1$.

1.2. Variables aleatorias

Sea $(\Omega, \mathcal{F}_t, \mathbf{P})$ un espacio de probabilidad y de acuerdo con la Definición 1.1.8, se dice que la aplicación $X : \Omega \to E$ es una variable aleatoria (abreviadamente, v.a.) si es una función medible de $(\Omega, \mathcal{F})$ en $(E, \mathcal{B}(E))$, donde E es un conjunto no vació. Por otro lado, se dice que X es una variable aleatoria real (abreviadamente, v.a.r.) a toda aplicación de $(\Omega, \mathcal{F})$ en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ de forma que, $X^{-1}(B) \in \mathcal{F}$ para todo $B \in \mathcal{B}(\mathbb{R})$, en donde, $\mathcal{B}(\mathbb{R})$ representa la menor σ-álgebra generada por los conjuntos abiertos en $\mathbb{R}$, comúnmente denominados de "Borelianos".

El lector debe notar la diferencia entre v.a. y vector aleatorio. El vector de una v.a. toma valores em $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.

Un hecho importante a realzar, es que las funciones medibles (continuas) transforman v.a. en v.a. Esta afirmación es reforzada en el siguiente lema.

Lema 1.2.1. *[19, sec. 2.6, cap. 2] Sean $(X_1, \ldots, X_n)$ v.a.r., y una aplicación $\psi : \mathbb{R}^n \to \mathbb{R}$ tal que*

i) si ψ es una función continua, entonces $\psi(X_1, \ldots, X_n)$ es una v.a.r.

ii) si ψ es una función medible, entonces $\psi(X_1, \ldots, X_n)$ es una v.a.r.

Enseguida veamos la ley de probabilidad de las v.a.r. Sea X una v.a.r. entonces la ley de probabilidad de X es definida de la siguiente manera

$$\mathbf{P}_x(B) := \mathbf{P}(X^{-1}(B)) \text{ para todo } B \in \mathcal{B}(\mathbb{R}).$$

De esta forma, se define $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbf{P}_x)$ como um nuevo espacio de probabilidad, en donde $\mathbf{P}_x$ es una nueva medida de probabilidad sobre $\mathcal{B}(\mathbb{R})$ y es denominado comúnmente como distribución de X. Así cualquier espacio medible $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ puede convertirse en un espacio de probabilidad ayudado por la v.a., osea, para cada Boreliano B se asocia una distribución $F(x)$ la cual puede ser definida de la siguiente forma $F(x) = \mathbf{P}(X \leq x)$ para cualquier número real x.

Enseguida veamos algunas definiciones y propiedades de la esperanza matemática de una v.a. La esperanza matemática (o media matemática) de una v.a.r. X es definida como

$$\mathbf{E}[X] = \int_\Omega X(\omega) d\mathbf{P}(\omega) = \int_\mathbb{R} x\mathbf{P}_x(dx). \tag{1.2.1}$$

La penúltima integral es una integral de Lebesgue[1] en el espacio $(\Omega, \mathcal{F}, \mathbf{P})$; la última integral es una integral de Lebesgue en el espacio $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbf{P}_x)$. Mediante el Lema 1.2.1 este concepto puede ser generalizado para la esperanza matemática de la composición de

[1] La integral de Lebesgue para una función simple f, es definida de la siguiente forma

$$\int_\Omega f d\mu := \sum_{i=1}^n \lambda_i \mu(B_i)$$

en donde B_i son subconjunto medibles en Ω para cada $i = 1, \cdots, n$.

una función con una v.a. Si g es una función real tal que la composición $g(X)$ es una v.a., entonces se define la esperanza matemática de $g(X)$ como sigue

$$\mathbf{E}[g(X)] = \int_{-\infty}^{\infty} g(x)dF(x),$$

en donde $F(x)$ es una función distribución de X, y la integral es de tipo Riemann-Stieltjes, sobre la que asumimos su existencia.

El valor de $\mathbf{E}[(X-b)^k]$, en caso exista, es denominado "k-ésimo momento" de X en torno de $b \in \mathbb{R}$, $k = 1, 2, 3, \ldots$. El k-ésimo momento en torno de cero, $\mathbf{E}[X^k]$, es denominado simplemente de k-ésimo momento de X o simplemente "momento" de orden k de X. El segundo momento en torno de la media es denominado "varianza" de X, y se denota por $\mathrm{Var}[X]$, tal que $\mathrm{Var}[X] = \mathbf{E}[(X - \mathbf{E}[X])^2]$. Se dicen que dos v.a. en un mismo espacio de probabilidad son "independientes" si, y solo si cualquier evento determinado por cualquier grupo de esas v.a. no tienen elementos en común, véase [19, def. 2.8, cap. 2, pag. 61].

Definición 1.2.1. *Sea X una v.a.r. Si $\mathbf{E}[|X|]$ es finita, se dice que X es integrable y el conjunto de funciones integrables se denota por $L^1(\Omega;\mathbb{R})$. Se denota el conjunto de v.a. integrables $X : [a,b] \to \mathbb{R}^n$ por $L^1(a,b;\mathbb{R}^n)$.*

Ejemplo 1.2.1. *Considere $f : [0,1] \to \mathbb{R}$ una función tal que*

$$f(x) = \frac{1}{x},$$

en donde el valor de $f(0)$ es insignificante. Usando el la convergencia monótona se tiene que

$$\int_{[0,1]} |f| d\mu = \lim_{a\to 0^+} \int_{[a,1]} \frac{1}{x} d\mu(x) = \lim_{a \to 0^+} [\log x]_a^1 = \infty.$$

Entonces podemos concluir que $f \notin L^1$.

La esperanza condicional. Primero presentamos la motivación matemática para definir la esperanza condicional. Sea X una v.a.r. en $L^1(\Omega;\mathbb{R})$ y sea $\mathcal{G}$ una σ-álgebra contenida en $\mathcal{F}$. De esta forma se calcula

$$\int_G X d\mathbf{P} \tag{1.2.2}$$

para cualquier G en $\mathcal{G}$. En ese sentido, con (1.2.2) bien definido, nos preguntamos si X es $\mathcal{G}$-medible? La respuesta en general es no, y para que esto pueda ocurrir es necesario una nueva definición dada en seguida.

Definición 1.2.2. *Sea X una v.a.r. en $L^1(\Omega;\mathbb{R})$ y $\mathcal{G}$ una σ-álgebra contenida en $\mathcal{F}$. Se dice que $Y : (\Omega, \mathcal{G}, \mathbf{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ ($\mathcal{G}$-medible) es la esperanza de X condicionado a la σ-álgebra $\mathcal{G}$ para cualquier $G \in \mathcal{G}$, se tiene*

$$\int_G X(\omega) d\mathbf{P}(\omega) = \int_G Y(\omega) d\mathbf{P}(\omega). \tag{1.2.3}$$

De aquí en adelante, denotamos $Y = \mathbf{E}[X|\mathcal{G}]$ una v.a.

Observación 1. *Existe un resultado de gran importancia, pero que no involucra el objetivo de este trabajo, "la existencia y unicidad de la esperanza condicional" puede ser revisada con mas detalle en [17, Pag. 85].*

Propiedades de la esperanza condicional. [19, sec. 4.5, pag. 175] En la siguientes propiedades, X, Y son v.a. integrables en $(\Omega, \mathcal{F}, \mathbf{P})$ y $\mathcal{G}$ una sub σ-álgebra de $\mathcal{F}$ y considere a, b constantes.

i) $\mathbf{E}[\mathbf{E}[X|\mathcal{G}]] = \mathbf{E}[X]$;

ii) Si X es $\mathcal{G}$-medible entonces $\mathbf{E}[X|\mathcal{G}] = X$;

iii) $\mathbf{E}[aX + bY|\mathcal{G}] = a\mathbf{E}[X|\mathcal{G}] + b\mathbf{E}[Y|\mathcal{G}]$.

Para la esperanza condicional también vale la desigualdad de Jensen.

Lema 1.2.2. *[19, cap. 3, pag. 116] Si $\varphi : \mathbb{R} \to \mathbb{R}$ función convexa y además se tiene que $\mathbf{E}[|X|] < \infty$, $\mathbf{E}[\varphi(X)] < \infty$, entonces con probabilidad 1 tenemos que*

$$\varphi(\mathbf{E}[X|\mathcal{F}]) \leq \mathbf{E}[\varphi(X)|\mathcal{F}].$$

1.3. Procesos estocásticos

Un proceso estocástico es una colección de v.a. a las que se les asocia el tiempo t, de esta forma se puede denotar la colección $\{X_t : t \geq 0\}$ o simplemente $\{x_t\}_{t\geq 0}$. Un proceso estocástico es entonces una aplicación de dos variables

$$X : [0, \infty) \times \Omega \to \mathbb{R}^n,$$

tal que, para cada $t \geq 0$, la función $\omega \mapsto X_t(\omega)$ es una v.a., y para cada $\omega \in \Omega$, la función $t \mapsto X_t(\omega)$ es una trayectoria continua casi seguramente[2]. Denotamos un proceso estocástico, en conformidad con la literatura clásica, por x_t.

Una familia de σ-álgebras $\{\mathcal{F}_t\}_{t\geq 0}$ es una filtración, si para cada $0 \leq s \leq t$ se tiene que $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$. De esta forma podemos definir $(\Omega, \mathcal{F}, \mathbf{P}, \{\mathcal{F}_t\}_{t\geq 0})$ un espacio de probabilidad filtrado. Se dice que un proceso x_t es adaptado a una filtración cuando el proceso es $\mathcal{F}_t$-medible para cada $t \geq 0$. Cualquier proceso puede conformar una filtración de la siguiente forma $\mathcal{F}_t = \sigma(x_s : 0 \leq s \leq t)$, esta filtración es conocida como "filtración natural".

Es importante conocer algunos conceptos de comparación entre los procesos estocásticos. Dos procesos x_t e y_t se dicen equivalente o uno versión del otro o modificación, si para cada $t \geq 0$ se tiene $\mathbf{P}(x_t = y_t) = 1$. En ese caso se dice que la v.a. x_t es igual a y_t casi seguramente y se denota de la siguiente forma

$$x_t = y_t, \text{ c.s.}$$

Un tipo de igualdad mas fuerte es cuando dos procesos son indistinguibles, esto es,

$$\mathbf{P}(x_t = y_t \text{ para cada } t \geq 0) = 1.$$

Esto significa que con probabilidad uno las trayectorias de los procesos son casi iguales. Cuando los procesos son continuos, i.e., sus trayectorias son funciones continuas, ambas nociones de igualdad coinciden.

En seguida, veamos una propiedad importante de algunos procesos, los incrementos independientes.

[2]es común suponer que esta trayectoria es tipo *càdlàg*, del francés, continua a la derecha y con limite a la izquierda [?].

Definición 1.3.1. *Un proceso* $\{N_t\}_{t\geq 0}$ *tiene incrementos independientes si para cualquier* $t_0, t_1, \cdots, t_n$ *con* $0 < t_0 < t_1 < \cdots < t_n$ *se tiene que*

$$N_{t_0}, N_{t_1} - N_{t_0}, \cdots, N_{t_n} - N_{t_{n-1}}$$

son v.a. independientes en algún espacio de probabilidad.

Otra propiedad importante de algunos procesos estocásticos es la de tener incrementos estacionarios.

Definición 1.3.2. *Un proceso* $\{N_t\}_{t\geq 0}$ *tiene incrementos estacionarios si para cualquier* t_1, t_3 *se tiene que*

$$\mathbf{P}(N_{t_1+t} - N_{t_1} = n) = \mathbf{P}(N_{t_2+t} - N_{t_2} = n), \;\; n \in \mathbb{R}.$$

Definición 1.3.3. *Un proceso estocástico* $\{x_t\}_{t\geq 0}$ *es denominado adaptado a una filtración* $\{\mathcal{F}_t\}_{t\geq 0}$ *si para cada* $t \geq 0$ *la v.a.* x_t *es medible en relación a la* σ*-álgebra* $\mathcal{F}_t$*.*

Ejemplo 1.3.1. *Un proceso estocástico en la recta es una familia de v.a.* $\{x_t\}_{t\geq 0}$ *definidas sobre* $(\Omega, \mathcal{F}, \mathbf{P})$ *con espacio de estado* (E, μ) *es una transformación* $X : [0,\infty) \times \Omega \to E$ *tal que para todo* $t \in [0,\infty)$ *la transformación* $x_t : \Omega \to E$ *es* $\mathcal{F}_t$*-medible i.e.,* x_t *es un proceso adaptado.*

Observación 2. *Todo proceso* x_t *es adaptado a su filtración natural* $\mathcal{F}_t = \sigma(x_s, 0 \leq s \leq t)$*. De esta forma, es intuitivo decir que un proceso estocástico* x_t *es adaptado a su filtración* $\mathcal{F}_t$*, significa que, para cada* $t \geq 0$ *la v.a.* x_t *va depender solo de la información disponible en la* σ*-álgebra* $\mathcal{F}_t$*.*

Para finalizar, esta breve sección veamos algunos de los procesos estocásticos mas importantes.

1.3.1. Procesos de Markov

Un proceso estocástico x_t es de Markov si para cada $0 \leq s \leq t$ y $A \in \mathcal{B}(\mathbb{R})$, con probabilidad uno se tiene que

$$\mathbf{P}(x_t \in A|\mathcal{F}_s) = \mathbf{P}(x_t \in A|x_s),$$

este resultado nos dice que el estado de un proceso en un tiempo futuro t es independiente del estado del proceso en un tiempo pasado s conocido. Esta propiedad se puede encontrar en los sistemas dinámicos deterministas, donde su evolución queda totalmente determinada por el estado inicial y la ley de su movimiento.

Se denota por $L^p(\Omega;\mathbb{R})$ o simplemente L^p el conjunto de las v.a.r. integrables definidas sobre $(\Omega;\mathbb{R})$ para todo $1 \leq p < \infty$, esto es, el espacio vectorial de v.a.r. tales que

$$L^p = \{X : X \text{ es medible y } |X|^p \in L^1\},$$

con norma

$$\|X\|_{L^p} = \left(\int_\Omega |X|^p d\mathbf{P}\right)^{1/p}. \tag{1.3.4}$$

El próximo resultado verifica el comentado anterior.

Lema 1.3.1. *[6, pag. 93, Teo. 4.7] L^p es un espacio vectorial y* (1.3.4) *representa una norma para todo $1 \leq p < \infty$.*

Lema 1.3.2. *[6, pag. 93, Teo. 4.8] L^p es un espacio de Banach para todo $1 \leq p < \infty$.*

Ejemplo 1.3.2. *Sea $f : [1, \infty) \to \mathbb{R}$ una función tal que*

$$f(x) = \frac{1}{x}.$$

Veamos que $f \in L^2$,

$$\int_{[1,\infty)} |f|^2 dm = \lim_{b\to\infty} \int_{[1,b]} \frac{1}{x^2} dm(x) = \lim_{b\to\infty} \left[-\frac{1}{x} \right]_1^b = 1.$$

Por otro lado, mirando el Ejemplo 1.2.1 podemos decir que $f \notin L^1$ y de forma general

$$\int_{[1,\infty)} \frac{1}{x^r} dm(x) = \begin{cases} 1/(r-1), & r > 1, \\ \infty, & r \leq 1. \end{cases}$$

Por lo tanto $f(x) = 1/x^r \in L^p$ si solo si $p > 1/r$.

El siguiente resultado muestra que $\mathbf{E}[X|\mathcal{G}]$ es una proyección de X sobre un subespacio $L^2(\Omega, \mathcal{G}, \mathbf{P})$, i.e., la distancia mínima entre X y $L^2(\Omega, \mathcal{G}, \mathbf{P})$ en la norma dada por el producto interno.

La noción de la esperanza condicional en relación a una v.a. se torna un caso particular de la esperanza condicional en relación a una σ-álgebra tal que, si $Y = \mathbf{E}[X|\mathcal{G}]$, $\mathbf{E}[X|Y] = \mathbf{E}[X|\sigma(Y)]$. De esta forma, podemos decir que, la característica mas importante de la v.a. Y no es el valor que asume en $\omega \in \Omega$, mas si la información que la ocurrencia de cada valor de Y puede proporcionar al respecto de la ocurrencia de los eventos $\omega \in \Omega$. Esta información es capturada por los σ-álgebras generados por $\sigma(Y)$.

Teorema 1.3.1 (Propiedades de la esperanza condicional)**.** *Algunas propiedades de la esperanza condicional*

1. $\mathbf{E}[\mathbf{E}[X|\mathcal{G}]] = \mathbf{E}[X]$ *(invariancia);*
2. *Si X es $\mathcal{G}$-medible entonces $\mathbf{E}[X|\mathcal{G}] = X$ (proyección);*
3. $\mathbf{E}[aX + bY|\mathcal{G}] = a\mathbf{E}[X|\mathcal{G}] + b\mathbf{E}[Y|\mathcal{G}]$ *(linealidad);*
4. *Si $X \leq Y$ c.s., entonces $\mathbf{E}[X|\mathcal{G}] \leq \mathbf{E}[Y|\mathcal{G}]$ c.s. (monoticidad);*
5. *Si $\lim_{n\to\infty} X_n = X$ con probabilidad 1, $|X_n| \leq Y$, y Y es integrable, entonces $\lim_{n\to\infty} \mathbf{E}[X_n|\mathcal{G}] = \mathbf{E}[X|\mathcal{G}]$ (convergencia dominada);*
6. *Si $X_n \geq 0$ y $X_n \uparrow X$ con $\mathbf{E}[X] < \infty$, entonces $\mathbf{E}[X_n|\mathcal{G}] \uparrow \mathbf{E}[X|\mathcal{G}]$ (convergencia monótona);*
7. *Si X es $\mathcal{G}$-medible y XY es integrable, entonces $\mathbf{E}[XY|\mathcal{G}] = X\mathbf{E}[Y|\mathcal{G}]$;*
8. *Si X es independiente de $\mathcal{G}$, entonces $\mathbf{E}[X|\mathcal{G}] = \mathbf{E}[X]$;*
9. *Si $\mathcal{G}_1 \subset \mathcal{G}_2$ entonces $\mathbf{E}[\mathbf{E}[X|\mathcal{G}]]$ (filtración).*

Demostración. 1. Basta considerar $G = \Omega$ en la Definición 1.2.2.

2. Si definimos $\mathbf{E}[X|\mathcal{G}] \equiv X$ tenemos que $\mathbf{E}[X|\mathcal{G}]$ es $\mathcal{G}$-medible y satisface la Definición 1.2.2. Luego el resultado se obtiene por unicidad.

3. $a\mathbf{E}[X|\mathcal{G}] + b\mathbf{E}[Y|\mathcal{G}]$ es $\mathcal{G}$-medible y satisface la Definición 1.2.2. Luego el resultado se obtiene por unicidad.

4. $$\int_A \mathbf{E}[X|\mathcal{G}]d\mathbf{P} = \int_A X d\mathbf{P} \leq \int_A Y d\mathbf{P} = \int_A \mathbf{E}[Y|\mathcal{G}]d\mathbf{P}$$
para todo $A \in \mathcal{G}$. Luego, si consideramos $A_n = \{\mathbf{E}[X|\mathcal{G}] - \mathbf{E}[Y|\mathcal{G}] > \frac{1}{n}\}$, tenemos que $\mathbf{P}(A_n) = 0$ para todo n, y por tanto
$$\mathbf{P}(\mathbf{E}[X|\mathcal{G}] > \mathbf{E}[Y|\mathcal{G}]) = 0.$$

5. Sea $Z_n = \sup_{k \geq n} |X_k - X|$. Entonces $Z_n \downarrow 0$ con probabilidad 1 y, por la propiedad 3 y 4, $|\mathbf{E}[X|\mathcal{G}] - \mathbf{E}[Y|\mathcal{G}]| \leq \mathbf{E}[Z_n|\mathcal{G}]$. Por lo tanto, basta demostrar que $\mathbf{E}[Z_n|\mathcal{G}] \downarrow 0$ con probabilidad 1. Como $Z_n \geq 0$ es decreciente, y de acuerdo con 4, $\mathbf{E}[Z_n|\mathcal{G}]$ también es decreciente y, por lo tanto, tiene un limite para Z. Luego, notamos que $0 \leq Z \leq \mathbf{E}[Z_n|\mathcal{G}]$ para todo n. Se desea demostrar que $Z = 0$ con probabilidad 1. Como Z es no negativo, es suficiente probar que $\mathbf{E}[Z] = 0$. Mas $0 \leq Z_n \leq 2Y$, entonces por la convergencia dominada común ([18, pag. 44]), tenemos que
$$\mathbf{E}[Z] = \int Z d\mathbf{P} \leq \int \mathbf{E}[Z_n|\mathcal{G}]d\mathbf{P} = \mathbf{E}[Z_n] \to 0.$$

6. Sea $Y_n = X - X_n$. Usando linealidad, podemos demostrar que $Z_n \equiv \mathbf{E}[Y_n|\mathcal{G}] \downarrow 0$. Por otro lado, $Y_n \downarrow 0$, se tiene por la propiedad 4, que también Z_n es decreciente, luego tenemos un limite Z, $0 \leq Z \leq Z_n$ para todo n. Notemos que Y_n es dominada por X que es integrable y por lo tanto, por la propiedad 1 y el teorema de convergencia dominada usual, tenemos
$$\mathbf{E}[Z] \leq \mathbf{E}[Z_n] = \mathbf{E}[Y_n] \to 0.$$
Luego, $Z = 0$ c.s.

7. Como $X\mathbf{E}[Y|\mathcal{G}]$ es $\mathcal{G}$-medible, es suficiente demostrar que
$$\int_A X\mathbf{E}[Y|\mathcal{G}] = \int_A XY d\mathbf{P} \ \text{ para todo } A \in \mathcal{G}. \tag{1.3.5}$$
Inicialmente, supongamos que $X = \mathbb{1}_B$ con $B \in \mathcal{G}$. En este caso, si $A \in \mathcal{G}$, entonces, como $A \cup B \in \mathcal{G}$,
$$\int_A \mathbb{1}_B \mathbf{E}[Y|\mathcal{G}]d\mathbf{P} = \int_{A\cap B} \mathbf{E}[Y|\mathcal{G}]d\mathbf{P} = \int_{A\cap B} Y d\mathbf{P} = \int_A \mathbb{1}_A Y d\mathbf{P}.$$
Luego, (1.3.5) es valido. De esta forma, podemos extender esto para X, función simple $\mathcal{G}$-medible, $X = \sum_{i=1}^n a_i \mathbb{1}_{B_i}$, por linealidad. Para X $\mathcal{G}$-medible general, tenemos una sucesión de funciones simples $\mathcal{G}$-medible tal que $|X_n| \leq X$ y $\lim_n X_n = X$. Como $|X_nY| \leq |XY|$ y $|XY|$ es integrable, por la propiedad 5 tenemos que $\lim_n \mathbf{E}[X_nY|\mathcal{G}]$. Por otro lado, $\mathbf{E}[X_nY|\mathcal{G}] = X_n\mathbf{E}[Y|\mathcal{G}]$, por ser X_n son funciones simples, y $\lim_n X_n\mathbf{E}[Y|\mathcal{G}] = X\mathbf{E}[Y|\mathcal{G}]$. Por lo tanto, la propiedad 7 es valida.

8. X es independiente de $\mathcal{G}$ significa que la v.a. X es independiente de la información dada por $\mathcal{G}$. Como $\mathbf{E}[X]$ es una constante, y por lo tanto $\mathcal{G}$-medible, es suficiente demostrar que, para todo $A \in \mathcal{G}$, se tiene que $\int_A X d\mathcal{P} = \int_A \mathbb{1}_A \mathbf{E}[X] d\mathbf{P}$. Veamos

$$\int_A X d\mathbf{P} = \int_A \mathbb{1}_A d\mathbf{P} = \mathbf{E}[\mathbb{1}_A X] = \mathbf{E}[\mathbb{1}_A]\mathbf{E}[X] = \mathbf{P}(A)\mathbf{E}[X] = \int_A \mathbf{E}[X] d\mathbf{P},$$

en donde usamos la independencia entre X y la v.a. $\mathcal{G}$-medible.

9. Notar que $\mathbf{E}[X|\mathcal{G}_1]$ es $\mathcal{G}_2$-medible, pues $\mathcal{G}_1 \subset \mathcal{G}_2$. Luego por Propiedad 2, $\mathbf{E}[\mathbf{E}[X|\mathcal{G}_1]|\mathcal{G}_2] = \mathbf{E}[X|\mathcal{G}_1]$. Para probar que $\mathbf{E}[\mathbf{E}[X|\mathcal{G}_2]|\mathcal{G}_1] = \mathbf{E}[X|\mathcal{G}_1]$, debemos notar que $\mathbf{E}[X|\mathcal{G}_1]$ es $\mathcal{G}_1$-medible y si $A \in \mathcal{G}_1 \subset \mathcal{G}_2$ entonces

$$\int_A \mathbf{E}[X|\mathcal{G}_1] d\mathbf{P} = \int_A X d\mathbf{P} = \int_A \mathbf{E}[X|\mathcal{G}_2] d\mathbf{P}.$$

□

1.3.2. Procesos Martingala

Definición 1.3.4. *Sea $(\Omega, \mathcal{F}, \mathbf{P})$ un espacio de probabilidad y $\{\mathcal{F}_t\}_{t\geq 0}$ una filtración. Un proceso estocástico $\{x_t\}_{t\geq 0}$ a valores reales definido en $(\Omega, \mathcal{F}, \mathbf{P})$, es denominado martingala en relación a la filtración $\mathcal{F}_t$ si*

i) $\mathbf{E}[|x_t|] < \infty$ *para todo t;*

ii) $\{x_t\}_{t\geq 0}$ *es adaptado a la filtración $\mathcal{F}_t$;*

iii) $\mathbf{E}[x_t|\mathcal{F}_s] = x_s$ *c.s. para todo t.*

Ejemplo 1.3.3. *Sea $\{N_t\}_{t\geq 0}$ es un proceso de Poisson con intensidad λ, entonces el proceso de Poisson compensado $\{N_t - \lambda_t\}_{t\geq 0}$ es un martingala.*

Cuando la filtración no es mencionada, queda sobrentendido que estamos tratando con una filtración natural generada por el propio proceso, osea, $\mathcal{F}_t = \sigma(x_s, 0 \leq s \leq t)$. La condición (iii) expresa nos dice que teniendo toda la información sobre el proceso hasta el tiempo s, la mejor estimativa para el estado futuro del proceso en el tiempo $t \geq s$ es el último valor observado x_s.

1.3.3. Movimiento Browniano

Uno de los grandes aciertos de la matemática en el último siglo, fue la creación de la estructura matemática para los procesos estocásticos, es decir, la colección de v.a., resulto ser exitosa para modelar matemáticamente la posición de una partícula aleatoria (Browniana) en un determinado tiempo t.

La definición del movimiento Browniano en una dimensión es como sigue.

Definición 1.3.5. *Un movimiento Browniano unidimensional es un proceso estocástico $\{B_t : t \geq 0\}$ o también denotado por $\{B_t\}_{t\geq 0}$, tal que*

1. $B_0 = 0$ *c.s.;*

2. $t \mapsto B_t$ *son trayectorias continuas;*

3. *el proceso tiene incrementos independientes;*

4. *si la variable $B_t - B_s$ tiene distribución $N(\epsilon, t-s)$ para $0 \leq s < t$, y si $\epsilon = 0$ el proceso es denominado de movimiento Browniano estándar.*

Definición 1.3.6. *Un proceso $\{B_t\}_{t\geq 0} = \left(B_t^{(1)}, \ldots, B_t^{(n)}\right)_{t\geq 0}$ en $\mathbb{R}^n$ es denominado un movimiento Browniano n-dimensional si*

i) Para cada $k = 1, \ldots, n$ cada componente $\{B_t^{(k)}\}_{t\geq 0}$ es un movimiento Browniano unidimensional.

ii) Los procesos $\{B_t^{(k)}\}_{t\geq 0}$, $k = 1, \ldots, n$ son independientes, i.e., para cualquier sucesión finita de tiempos

$$\begin{array}{cccc} t_{11} & t_{12} & \cdots & t_{1,k_1} \\ t_{21} & t_{22} & \cdots & t_{2,k_2} \\ \vdots & \vdots & \vdots & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{n,k_n} \end{array}$$

los n vectores

$$\begin{aligned} X_1 &= \left(B_{t_{11}}^{(1)}, B_{t_{12}}^{(1)}, \ldots, B_{t_{1,k_1}}^{(1)}\right) \\ X_2 &= \left(B_{t_{21}}^{(2)}, B_{t_{22}}^{(2)}, \ldots, B_{t_{2,k_2}}^{(2)}\right) \\ &\vdots \\ X_n &= \left(B_{t_{n1}}^{(n)}, B_{t_{n2}}^{(n)}, \ldots, B_{t_{n,k_n}}^{(n)}\right) \end{aligned}$$

son independientes.

Observación 3. *Características del movimiento browniano:*

- *El movimiento browniano también es conocido como proceso de wiener y denotado por $\{W_t\}_{t\geq 0}$;*
- *por definición las trayectorias del movimiento browniano son continuas c.s.;*
- *las trayectorias no son diferenciables en ningún punto;*
- *el movimiento browniano es un proceso de Markov;*
- *el movimiento browniano es un proceso martingala.*

El movimiento browniano visto como una v.a., B_t tiene distribución normal $N(0, t)$ y por lo tanto $\mathbf{E}[B_t] = 0$ y $\mathrm{Var}[B_t] = \mathbf{E}[B_t^2] = t$. En particular, se tiene que $\mathbf{E}[B_t - B_s] = t - s$ para cualquier $0 \leq s < t$.

Ahora presentamos la Desigualdad L^p de Dood de gran importancia en sistemas en tiempo continuo y también existe para un tiempo discreto.

Teorema 1.3.2. *[16, pag. 11, teo. 20] Sea $\{x_t\}_{t\geq 0}$ un proceso estocástico con trayectorias continuas c.s.*

(a) Si $\{x_t\}_{t\geq 0}$ *es un submartingala, entonces para todo* $t \geq 0$ *y* $\lambda > 0$*,*

$$\mathbf{P}\left(\sup_{0\leq s\leq t} x_s \geq \lambda\right) \leq \frac{1}{\lambda}\mathbf{E}[\text{máx}\{0, x_t\}].$$

(b) Si $\{x_t\}_{t\geq 0}$ *es un martingala y* $p > 1$*, entonces,*

$$\mathbf{E}\left[\sup_{0\leq s\leq t} |x_s|^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathbf{E}[|x_t|^p].$$

Capítulo 2

Integración Estocástica

Este capítulo tiene por objetivo definir la integral estocástica o también conocida como integral de Itô de un proceso estocástico x_t con respecto al proceso de Wiener[1] W_t, i.e., un ente matemático de la siguiente forma

$$\int x_t dW_t = \int x_t(\omega) dW_t(\omega).$$

Este tipo de expresiones aparecen con frecuencia en la literatura de ecuaciones diferenciales estocástica, tema del próximo capítulo.

2.1. Integral estocástica para procesos elementales

Definimos la integral estocástica en relación al proceso de Wiener en varios pasos. Primero para procesos simples y después, por aproximación, para procesos más generales. Cabe resaltar, que asumiremos el proceso $\{x_t\}_{t\geq 0}$ visto como una función medible y adaptada.

Definición 2.1.1. *Un proceso estocástico* $\varphi : [\alpha, \beta] \times \Omega \to \mathbb{R}$ *es denominado simple si es constante por partes con respecto a la variable* t*, i.e., si existe una partición* P_n*,*

$$\alpha = t_0 < t_1 < \cdots < t_{n-1} < t_n = \beta,$$

del intervalo $[\alpha, \beta]$ *tal que*

$$\varphi(t) = \varphi(t_i), \;\; t_i \leq t < t_{i+1}, \;\; 0 \leq i \leq n-1.$$

En este capítulo, $\varphi(t)$ *representa una v.a., definida en* Ω*, que traslada* ω *en* $\varphi(t, \omega)$*. Evitamos la notación* φ_t*, la cual la reservamos solo para los procesos estocásticos* x_t*,* y_t *y* W_t *en* $\mathbb{R}^n$*. La única razón de esto, es para evitar una confusión cuando tratemos mas adelante sucesión de procesos simples.*

Observación 4. *La partición dada en la Definición 2.1.1 es fijada de antemano y para cada subintervalo, el proceso es igual a una v.a. No todos los procesos cumplen esta definición, por ejemplo los proceso de Poisson*[2]*.*

[1] En 1923 el matemático Norbert Wiener demostró la existencia de un proceso con las condiciones del movimiento browniano. Es por esto, que el movimiento browniano también se conoce como proceso de Wiener y se denota por $\{W_t\}_{t\geq 0}$

[2] Considere una sucesión de v.a. estrictamente crecientes $0 = t_0 < t_1 < \cdots$ definidas en un espacio de probabilidad. Un proceso $\{N_t\}_{t\geq 0}$ adaptado a una filtración $\{\mathcal{F}_t\}_{t\geq 0}$ y no explosivo, es chamado proceso de Poisson si:

Para un proceso simple φ definido en $[\alpha, \beta]$ se tiene

$$\varphi(t) = \sum_{i=0}^{n-1} \varphi(t_i) \mathbb{1}_{[t_i, t_{i+1})}(t), \tag{2.1.1}$$

en donde $\mathbb{1}_{[t_i, t_{i+1})}(t)$ es una función indicadora del intervalo $[t_i, t_{i+1})$, así podemos definir

$$\int_\alpha^\beta \varphi(t) dW_t = \sum_{i=0}^{n-1} \varphi(t_i)(W_{t_{i+1}} - W_{t_i}). \tag{2.1.2}$$

El próximo ejemplo presenta la dificultad que surge cuando intentamos extender la Definición 2.1.1 para procesos mas generales de acuerdo con la integral de *Riemann-Stieltjes*. Esta dificultad está relacionada con el hecho que las trayectorias del movimiento browniano son de variación ilimitada.

Ejemplo 2.1.1. *Sea $\{W_t\}_{t\geq 0}$ un proceso de Wiener definida en un intervalo $[\alpha, \beta]$. Considere los siguientes procesos para cada $n = 1, 2, \dots,$ tal que*

$$\varphi_n(t) = \sum_{i=0}^{m-1} W_{\frac{i}{2^n}} \mathbb{1}_{[\frac{i}{2^n}, \frac{i+1}{2^n})}(t) + W_{\frac{m}{2^n}} \mathbb{1}_{[\frac{m}{2^n}, \beta]}(t), \quad si \ \frac{m}{2^n} \leq \beta < \frac{m+1}{2^n};$$

$$\psi_n(t) = \sum_{i=0}^{m-1} W_{\frac{i+1}{2^n}} \mathbb{1}_{[\frac{i}{2^n}, \frac{i+1}{2^n})}(t) + W_\beta \mathbb{1}_{[\frac{m}{2^n}, \beta]}(t), \quad si \ \frac{m}{2^n} \leq \beta < \frac{m+1}{2^n}.$$

De acuerdo con (2.1.2)*, tenemos que*

$$\begin{aligned} \mathbf{E}\left[\int_0^\beta \varphi_n(t) dW_t\right] &= \sum_{i=0}^{m-1} \mathbf{E}\left[W_{\frac{i}{2^n}}\left(W_{\frac{i+1}{2^n}} - W_{\frac{i}{2^n}}\right)\right] + \mathbf{E}\left[W_{\frac{m}{2^n}}\left(W_\beta - W_{\frac{m}{2^n}}\right)\right] \\ &= \sum_{i=0}^{m-1} \mathbf{E}\left[W_{\frac{i}{2^n}}\right] \mathbf{E}\left[\left(W_{\frac{i+1}{2^n}} - W_{\frac{i}{2^n}}\right)\right] + \mathbf{E}\left[W_{\frac{m}{2^n}}\right] \mathbf{E}\left[\left(W_\beta - W_{\frac{m}{2^n}}\right)\right] \\ &= 0 \end{aligned}$$

esto último resultado es debido a la independencia de los incrementos del movimiento browniano, los cuales tienen media cero. Por otro lado, tenemos lo siguiente

$$\begin{aligned} \mathbf{E}\left[\int_0^\beta \psi_n(t) dW_t\right] &= \sum_{i=0}^{m-1} \mathbf{E}\left[W_{\frac{i+1}{2^n}}\left(W_{\frac{i+1}{2^n}} - W_{\frac{i}{2^n}}\right)\right] + \mathbf{E}\left[W_\beta\left(W_\beta - W_{\frac{m}{2^n}}\right)\right] \\ &= \sum_{i=0}^{m-1} \mathbf{E}\left[W^2_{\frac{i}{2^n}}\right] \mathbf{E}\left[W_{\frac{i+1}{2^n}} W_{\frac{i}{2^n}}\right] + \mathbf{E}\left[W^2_\beta\right] \mathbf{E}\left[W_\beta W_{\frac{m}{2^n}}\right] \\ &= \sum_{i=0}^{m-1}\left(\frac{i+1}{2^n} - \frac{i}{2^n}\right) + \beta - \frac{m}{2^n} \\ &= \beta \end{aligned}$$

- $N_0 = 0$;
- Tiene incrementos independientes, ver Definición 1.3.1;
- Tiene incrementos estacionarios, ver Definición 1.3.2.

tenemos una suma telescópica.

En este ejemplo queda expuesto que no todos los procesos simples son buenas aproximaciones para un proceso de Wiener, por que sus integrales no son próximas entre si, no importando el tamaño elegido de n.

Ahora veamos como deben ser esos procesos simples en donde podamos definir la integral estocástica. El Ejemplo 2.1.1 revelo que estos procesos simples están de cierta forma sujetos a alguna restricción. De forma intuitiva, exijamos para cada t, que x_t dependa exclusivamente del comportamiento del movimiento browniano hasta t (de esta forma dejamos en claro que no dependerá del comportamiento futuro del movimiento browniano). Esta razonamiento intuitivo es formalizado en términos de σ-álgebras. Consideremos una filtración $\{\mathcal{F}_t\}_{t\geq 0}$ de un espacio de probabilidades, tal que

(i) $\sigma(W_s, 0 \leq s \leq t) \subset \mathcal{F}_t$ para todo $t \geq 0$;

(ii) $\sigma(W_{t+\lambda} - W_t, \lambda \geq 0)$ es independiente de $\mathcal{F}_t$ para todo $t \geq 0$.

Observación 5. *También se pueden tomar, en particular la filtración natural del movimiento browniano. En ese sentido las condiciones (i) y (ii) se vuelven mas generales.*

Definición 2.1.2. *Dado un intervalo $[\alpha, \beta]$, un espacio de probabilidad donde están definidos los procesos $\{W_t\}_{t\geq 0}$ y una filtración $\{\mathcal{F}_t\}_{t\geq 0}$ satisfaciendo (i) y (ii). Considere el conjunto de procesos estocásticos $\{x_t\}_{t\in[\alpha,\beta]}$ que satisfacen:*

(A1) $\{x_t\}_{t\in[\alpha,\beta]}$ *es un proceso $\mathcal{F}_t$-medible;*

(A2) $\{x_t\}_{t\in[\alpha,\beta]}$ *es un proceso mensurable, i.e., $(t,\omega) \mapsto x_t(\omega)$ es $\mathcal{B} \times \mathcal{F}$-medible;*

(A3) $\mathbf{E}\left[\int_\alpha^\beta |x_t|^2 dt\right] < \infty.$

Este conjunto de procesos es denotado por $\mathbb{M}^2 = \mathbb{M}^2(\alpha, \beta)$. Denominamos de elementales[3] *a los procesos simples en $\mathbb{M}^2$ y denotados por $\mathbb{M}_0^2$.*

Del Ejemplo 2.1.1 podemos notar que $\varphi_n(t) \in \mathbb{M}^2$ y $\psi_n(t) \notin \mathbb{M}^2$.

Definición 2.1.3. *Consideremos un procesos elemental $\varphi(t) = \sum_{i=0}^{n-1} \varphi(t_i)\mathbb{1}_{[t_i,t_{i+1})}(t)$, donde $\alpha = t_0 < t_1 < \cdots < t_n = \beta$ es una partición del intervalo $[\alpha, \beta]$. Definimos la integral de Itô de φ en el intervalo $[\alpha, \beta]$ de la siguiente forma*

$$\int_\alpha^\beta \varphi(t) dW_t = \sum_{i=0}^{n-1} \varphi(t_i)(W_{t_{i+1}} - W_{t_i}). \tag{2.1.3}$$

Observación 6. *1. Debemos notar que la integral dada en la Definición 2.1.3 es una v.a. definida en un algún espacio de probabilidad $(\Omega, \mathcal{F}, \mathbf{P})$.*

2. Mediante un procedimiento llamado 'localización', es posible extender la Definición 2.1.3 para procesos medibles y adaptados que cumplen la condición mas relajada de la condición (A3) en la Definición 2.1.2

$$\mathbf{P}\left(\int_\alpha^\beta |x_t|^2 dt < \infty\right) = 1.$$

[3] En otros texto, denotan por $L^2(\mathbf{P})$ al espacio de v.a., que son cuadrado integrables i.e., $\|X\|_{L^2(\mathbf{P})} = \mathbf{E}[|X|^2]^{1/2} < \infty$ y por $L^2(\mathbf{P} \times dt)$ el espacio de procesos estocásticos cuadrado integrables i.e., $\|X\|_{L^2(\mathbf{P}\times dt)} = \mathbf{E}[\int |X|^2 dt]^{1/2} < \infty$.

Esta forma no será tratada en este texto, para mayores detalles consultar [1, Cap 4][4].

Lema 2.1.1 (Propiedades de la Integral Estocástica para Procesos Elementales)**.** *Considere $\varphi(t)$ y $\psi(t)$ en $\mathbb{M}^2(\alpha,\beta)$, a y b números reales. Entonces:*

(a) $\displaystyle\int_\alpha^\beta (a\varphi(t)+b\psi(t))\,dW_t = a\int_\alpha^\beta \varphi(t)dW_t + b\int_\alpha^\beta \psi(t)dW_t.$

(b) $\displaystyle\mathbf{E}\left[\int_\alpha^\beta \varphi(t)dW_t\right] = 0.$

(c) $\displaystyle\mathbf{E}\left[\left(\int_\alpha^\beta \varphi(t)dW_t\right)^2\right] = \mathbf{E}\left[\int_\alpha^\beta \varphi^2(t)dt\right].$

Demostración.

(a) Usando la Definición 2.1.3, se tiene que

$$\begin{aligned}\int_\alpha^\beta (a\varphi(t)+b\psi(t))\,dW_t =& \sum_{i=0}^{n-1}(a\varphi(t_i)+b\psi(t_i))(W_{t_{i+1}}-W_{t_i})\\ =& a\sum_{i=0}^{n-1}\varphi(t_i)(W_{t_{i+1}}-W_{t_i}) + b\sum_{i=0}^{n-1}\psi(t_i)(W_{t_{i+1}}-W_{t_i})\\ =& a\int_\alpha^\beta \varphi(t)dW_t + b\int_\alpha^\beta \psi(t)dW_t.\end{aligned}$$

(b) Por la linealidad de la esperanza matemática, tenemos

$$\mathbf{E}\left[\int_\alpha^\beta \varphi^2(t)dt\right] = \sum_{i=0}^{n-1}\mathbf{E}[\varphi^2(t_i)](t_{i+1}-t_i) \tag{2.1.4}$$

De la hipótesis (e) en la Definición 2.1.2, (2.1.4) es acotada. Entonces para todo i, $\mathbf{E}[\varphi^2(t_i)]$ es acotado, y por la desigualdad de Cauchy Schwarz, $\mathbf{E}[\varphi(t_i)]$ es acotado. Nuevamente por la esperanza matemática, tenemos que

$$\mathbf{E}\left[\int_\alpha^\beta \varphi(t)dW_t\right] = \sum_{i=0}^{n-1}\mathbf{E}[\varphi(t_i)(W_{t_{i+1}}-W_{t_i})].$$

Desde que $\varphi\in\mathbb{M}^2$, tenemos que $\varphi(t_i)$ es $\mathcal{F}_{t_i}$-medible. Por otro lado, $(W_{t_{i+1}}-W_{t_i})$ es independiente de $\mathcal{F}_{t_i}$ y, así, $\varphi(t_i)$ y $(W_{t_{i+1}}-W_{t_i})$ son independientes. Po lo tanto, para cada $i=0,\ldots,n-1$ tenemos

$$\mathbf{E}[\varphi(t_i)(W_{t_i+1}-W_{t_i})] = \underbrace{\mathbf{E}[\varphi(t_i)]}_{<\infty}\,\underbrace{\mathbf{E}[W_{t_{i+1}}-W_{t_i}]}_{=0} = 0.$$

Sumando en i, tenemos el resultado deseado.

[4]Se denota por $\mathcal{L}^2_{Loc}$ el espacio de procesos con la condición mas débil.

(c) Se sabe que para cada i, $\varphi^2(t_i)$ y $\|W_{t_{i+1}} - W_{t_i}\|^2$ son independientes con esperanza matemática finita. Por la desigualdad de Cauchy-Schwarz se sigue que, para cualquier i y j,

$$\mathbf{E}[\varphi(t_i)\varphi(t_j)(W_{t_i+1} - W_{t_i})] < \infty.$$

Si $i < j$, entonces $(W_{t_i+1} - W_{t_i})$ y $\varphi(t_i)\varphi(t_j)(W_{t_i+1} - W_{t_i})$ son independientes, en donde

$$\begin{aligned}
&\mathbf{E}[\varphi(t_i)\varphi(t_j)(W_{t_i+1} - W_{t_i})(W_{t_j+1} - W_{t_j})] \\
&\qquad = \underbrace{\mathbf{E}[\varphi(t_i)\varphi(t_j)(W_{t_i+1} - W_{t_i})]}_{<\infty}\underbrace{\mathbf{E}[(W_{t_j+1} - W_{t_j})]}_{=0} = 0.
\end{aligned}$$

Luego, los términos cruzados se anula, entonces se tiene

$$\begin{aligned}
\mathbf{E}\left[\left(\int_\alpha^\beta \varphi(t)dW_t\right)^2\right] &= \sum_{i=0}^{n-1} \mathbf{E}[\varphi^2(t_i)(W_{t_{i+1}} - W_{t_i})^2] \\
&= \sum_{i=0}^{n-1} \mathbf{E}[\varphi^2(t_i)]\mathbf{E}[(W_{t_{i+1}} - W_{t_i})^2] \\
&= \sum_{i=0}^{n-1} \mathbf{E}[\varphi^2(t_i)](t_{i+1} - t_i) \\
&= \mathbf{E}\left[\int_\alpha^\beta \varphi^2(t)dt\right]
\end{aligned}$$

Luego por (2.1.4), tenemos el resultado deseado.

■

La propiedad (c), es conocida en la literatura especializada como la isometría de Itô[5]. La isometría juega un papel importante en la definición de la integral estocástica como veremos enseguida. Un proceso estocástico puede ser visto como una v.a. definida en el espacio $([\alpha,\beta]\times\Omega, \mathcal{B}\times\mathcal{F}, Leb\times\mathbf{P})$. Luego, por el item (A3) en la Definición 2.1.2 tenemos que $\varphi \in L^2([\alpha,\beta]\times\Omega, \mathcal{B}\times\mathcal{F}, Leb\times\mathbf{P})$ i.e., la integral estocástica asigna a cada elemento del espacio $\mathbb{M}^2(\alpha,\beta)$ una v.a. en el espacio $L^2([\alpha,\beta]\times\Omega, \mathcal{B}\times\mathcal{F}, Leb\times\mathbf{P})$. De esta forma tenemos un mapeo lineal $I : \mathbb{M}_0^2(\alpha,\beta) \to L^2([\alpha,\beta]\times\Omega, \mathcal{B}\times\mathcal{F}, Leb\times\mathbf{P})$ que resulta ser continuo por la isometría de Itô. Por lo tanto, la isometría de Itô garantiza que las v.a.,

$$Y \equiv \int_\alpha^\beta \varphi(t)dW_t \in L^2(\Omega,\mathcal{F},\mathbf{P}),$$

y

$$\varphi \in L^2([\alpha,\beta]\times\Omega, \mathcal{B}\times\mathcal{F}, Leb\times\mathbf{P})$$

sean iguales en norma en sus respectivos espacios, osea, por un lado tenemos

$$\|Y\|^2_{L^2(\Omega)} = \int_\Omega Y^2 d\mathbf{P} = \int_\Omega \left(\int_\alpha^\beta \varphi(t)dt\right)^2 d\mathbf{P} = \mathbf{E}\left[\left(\int_\alpha^\beta \varphi(t)dW_t\right)^2\right] \tag{2.1.5}$$

[5] La isometría de Itô puede ser escrita también como $\|I(X)\|_{L^2(\mathbf{P})} = \|X\|_{L^2(\mathbf{P}\times dt)}$.

y para la otra v.a. se tiene

$$\|\varphi\|^2_{L^2([\alpha,\beta]\times\Omega)} = \int_{[\alpha,\beta]\times\Omega} \varphi^2(t) d(Leb \times \mathbf{P}) =$$

$$\int_{\Omega} \left(\int_{\alpha}^{\beta} \varphi^2(t)dt \right) d\mathbf{P} = \mathbf{E}\left[\int_{\alpha}^{\beta} \varphi^2(t)dt \right], \quad (2.1.6)$$

después, utilizamos el Teorema de Fubini para justificar la igualdad entre (2.1.5) y (2.1.6).

2.2. Aproximación mediante procesos elementales

El siguiente resultado muestra que los procesos elementales son densos en $\mathbb{M}^2$, considerando la norma antes ya definida. Este hecho y la isometría de Itô, nos permite extender la Definición 2.1.3 para cualquier elemento de la clase $\mathbb{M}^2$.

Lema 2.2.1. *Sea $\{x_t\}_{t\in[\alpha,\beta]}$ un proceso en $\mathbb{M}^2$. Entonces existe una sucesión $\{\varphi_n(t)\}_{n=1}^{\infty}$ de procesos elementales, tales que,*

$$\lim_{n\to\infty} \mathbf{E}\left[\int_{\alpha}^{\beta} |x_t - \varphi_n(t)|^2 dt \right] = 0. \quad (2.2.7)$$

La demostración del Lema 2.2.1 es dividida en tres pasos.

Paso 1 Sea $g \in \mathbb{M}^2$ acotado tal que $g(\cdot,\omega)$ es continua para todo ω. Entonces existe una sucesión de funciones elementales $\{\varphi_n\}_{n\in\mathbb{N}}$ tal que

$$\lim_{n\to\infty} \mathbf{E}\left[\int_{\alpha}^{\beta} |g - \varphi_n|^2 dt \right] = 0.$$

Demostración. Considere $M > 0$ tal que $|g(t,\omega)| < M$ para todo t y ω. Para cada n tomemos una partición P_n, tal que, $|P_n| \to 0$ cuando $n \to \infty$. Luego, para cada n definimos

$$\varphi_n(t,\omega) = \sum_{j=0}^{m_n-1} g(t_j^n,\omega) \mathbb{1}_{[t_j^n, t_{j+1}^n)}(t).$$

Entonces φ_n es un proceso elemental, desde que $g \in \mathbb{M}^2$. Ademas, tenemos que

$$\lim_{n\to\infty} \int_{\alpha}^{\beta} |g(t,\omega) - \varphi_n(t,\omega)|^2 dt = 0, \text{ para cada } \omega, \quad (2.2.8)$$

desde que $g(\cdot,\omega)$ es continua para cada $\omega \in \Omega$. Debemos notar que, para cualquier t, ω y n se tiene $|g(t,\omega) - \varphi_n(t,\omega)| < 2M$ y consecuentemente

$$\int_{\alpha}^{\beta} |g(t,\omega) - \varphi_n(t,\omega)|^2 dt < 4M^2(\beta - \alpha). \quad (2.2.9)$$

Por lo tanto, de (2.2.8) y (2.2.9) y por el teorema de la Convergencia Dominada, tenemos

$$\lim_{n\to\infty} \mathbf{E}\left[\int_{\alpha}^{\beta} |g - \varphi_n|^2 dt \right] = 0.$$

■

Paso 2 Sea $h \in \mathbb{M}^2$ una función acotada. Entonces existe una sucesión de funciones acotadas $\{g_n\}_{n\in\mathbb{N}}$ en $\mathbb{M}^2$ tal que $g_n(\cdot,\omega)$ es continua para todo ω y n, ademas

$$\lim_{n\to\infty} \mathbf{E}\left[\int_\alpha^\beta |h-g_n|^2 dt\right] = 0.$$

Demostración. sea $M > 0$, tal que $|h(t,\omega)| < M$ para cualquier t y ω. Para cada n, consideremos una función continua $\varphi_n : \mathbb{R} \to \mathbb{R}$ tal que $\varphi_n(r) = 0$ si $r < -\frac{1}{n}$ y $r \geq 0$, $\displaystyle\int_{-\infty}^{+\infty} \varphi_n(r)dr = 1$. La sucesión de funciones φ_n es denominada de núcleos de Dirac. Luego, para cada n definimos

$$g_n(t,\omega) = \int_0^t \varphi_n(s-t)h(s,\omega)ds.$$

Entonces $g_n(\cdot,\omega)$ es continua para cualquier ω y n, pues el integrando anterior es limitado. También tenemos que $|g_n(t,\omega)| \leq M$, y podemos calcular un limitante superior

$$\begin{aligned}\left|\int_0^t \varphi_n(s-t)h(s,\omega)ds\right| &\leq \int_0^t \varphi_n(s-t)h(s,\omega)ds\\ &\leq \int_0^t \varphi_n(s-t)Mds\\ &\leq M\int_{-\infty}^{\infty} \varphi_n(s-t)ds\\ &= M.\end{aligned}$$

Usando las propiedades de h en $\mathbb{M}^2$ y desde que la integral que define $g_n(t,\cdot)$ no envuelve valores de $h(s,\cdot)$ para $s > t$, tenemos que $g_n \in \mathbb{M}^2$. Ademas, para cada ω, se tiene que

$$\lim_{n\to\infty} \int_\alpha^\beta |h(t,\omega) - g_n(t,\omega)|^2 dt = 0, \tag{2.2.10}$$

desde que φ_n representa aproximaciones de la distribución delta de Dirac. Podemos notar también que para cualquier t, ω y n, se tiene que $|h(t,\omega) - g_n(t,\omega)| \leq 2M$ de tal manera que

$$\int_\alpha^\beta |h(t,\omega) - g_n(t,\omega)|^2 \leq 4M^2(\beta-\alpha). \tag{2.2.11}$$

Luego, por (2.2.10), (2.2.11) y el teorema de Convergencia Dominada, podemos concluir que

$$\lim_{n\to\infty} \mathbf{E}\left[\int_\alpha^\beta |h-g_n|^2 dt\right] = 0.$$

■

Paso 3 Sea $\{x_t\}_{t\geq 0}$ un proceso en $\mathbb{M}^2$. Entonces existe una sucesión $\{h_n\}_{n\in\mathbb{N}}$ de funciones acotadas en $\mathbb{M}^2$, tal que

$$\lim_{n\to\infty} \mathbf{E}\left[\int_\alpha^\beta |x_t - h_n(t)|^2 dt\right] = 0.$$

Demostración. Definamos un proceso de la siguiente manera

$$h_n(t,\omega) = \begin{cases} -n, & \text{si} \quad x_t(\omega) < -n, \\ x_t(\omega), & \text{si} \quad -n \leq x_t(\omega) \leq n, \\ n, & \text{si} \quad x_t(\omega) > n. \end{cases}$$

Entonces, para cada n, podemos notar que h_n es acotado y ademas pertenece a $\mathbb{M}^2$. Para cada par $(t\,\omega) \in [\alpha,\beta] \times \Omega$ se tiene que

$$\lim_{n\to\infty} h_n(t,\omega) = x_t(\omega), \quad \text{y} \quad |h_n(t,\omega)| \leq |x_t(\omega)|, \text{ para todo } n.$$

Por lo tanto, la demostración sigue del Teorema de Convergencia Dominada en espacios $L^2([\alpha,\beta]\times\Omega)$. ■

De esta forma, si concatenamos secuencialmente los pasos 1,2 y 3, la prueba del Lema 2.2.1 esta completa.

2.3. Integral estocástica para procesos mas generales

Ahora, podemos definir la integral estocástica para toda la clase $\mathbb{M}^2$.

Definición 2.3.1. *Sea $\{x_t\}_{t\in[\alpha,\beta]}$ un proceso en $\mathbb{M}^2(\alpha,\beta)$. Si consideramos una sucesión $\{\varphi_n\}_{n\in\mathbb{N}}$ de procesos elementales, tal que*

$$\lim_{n\to\infty} \mathbf{E}\left[\int_\alpha^\beta |x_t - \varphi_n(t)|^2 dt\right] = 0$$

y definimos

$$\int_\alpha^\beta x_t dW_t \equiv \lim_{n\to\infty} \int_\alpha^\beta \varphi_n(t) dW_t,$$

como un limite en $L^2(\Omega)$.

El limite anterior existe, desde que la isometría de Itô fue dada para procesos elementales, de tal forma, se tiene que

$$\mathbf{E}\left[\left(\int_\alpha^\beta \varphi_n(t)dW_t - \int_\alpha^\beta \varphi_m(t)dW_t\right)^2\right] = \mathbf{E}\left[\int_\alpha^\beta |\varphi_n(t) - \varphi_m|^2 dt\right] \to 0$$

cuando $n,m\to\infty$. Por lo tanto, $\int_\alpha^\beta \varphi_n(t)dtdW_t$ es una sucesión de cauchy en el espacio completo $L^2(\Omega)$. Ademas, notamos que el limite no depende de una selección particular de la sucesión $\{\varphi_n\}_{n\in\mathbb{N}}$. De esta forma cualquier subsucesión debe ser convergente para el mismo limite. De forma sucinta, podemos decir que $\{\varphi_n\}_{n\in\mathbb{N}}$ es tomada discretizando el el proceso $\{x_t\}_{t\in[\alpha,\beta]}$ en particiones del intervalo $[\alpha,\beta]$ que tienen norma tendiendo a cero para puntos evaluados en el extremo izquierdo de cada intervalo i.e., la integral estocástica corresponderá a la selección $\xi_j^n = t_j^n$, el extremo izquierdo del intervalo $[t_j^n, t_{j+1}^n] \subset [\alpha,\beta]$. Otra noción de integración estocástica es dada por la integral de *Stratonovic*, que no será estudiada con detalle en este documento, en este caso, la selección es un punto medio $\xi_j^n = \frac{t_j^n + t_{j+1}^n}{2}$. En muchas situaciones el uso de la integral de *Stratonovic* es mas conveniente, para mas detalles el lector puede consultar [4, Pag. 32-24].

Lema 2.3.1 (Propiedades de la Integral Estocástica para Procesos Generales en $\mathbb{M}^2$)**.** *Sean $\{x_t\}_{t\in[\alpha,\beta]}$ y $\{y_t\}_{t\in[\alpha,\beta]}$ dos procesos en $\mathbb{M}^2(\alpha,\beta)$, con a y b dos números reales, entonces:*

(a) $\displaystyle\int_\alpha^\beta (ax_t + by_t)\, dW_t = a\int_\alpha^\beta x_t dW_t + b\int_\alpha^\beta y_t dW_t.$

(b) $\displaystyle\mathbf{E}\left[\int_\alpha^\beta x_t dW_t\right] = 0.$

(c) $\displaystyle\mathbf{E}\left[\left(\int_\alpha^\beta x_t dW_t\right)^2\right] = \mathbf{E}\left[\int_\alpha^\beta x_t^2 dt\right].$

(d) $\displaystyle\mathbf{E}\left[\int_\alpha^\beta x_t dW_t \int_\alpha^\beta y_t dW_t\right] = \mathbf{E}\left[\int_\alpha^\beta x_t y_t dt\right].$

Demostración. (a), (b) y (c) siguen el mismo procedimiento que el Lema 2.1.1 junto a la desigualdad de Cauchy-Schwarz de la siguiente forma

$$\mathbf{E}[Y_n] - \mathbf{E}[Y] = \mathbf{E}[Y_n - Y] \leq (\mathbf{E}[(Y_n - Y)^2])^{1/2}\mathbb{1}^{1/2} \to 0.$$

De forma análoga $\mathbf{E}[Y_n] - \mathbf{E}[Y] \to 0$, entonces concluimos que $\mathbf{E}[Y_n] \to \mathbf{E}[Y]$ y se prueba (b). Para (d), podemos usar la siguiente identidad $2ab = (a+b)^2 - a^2 - b^2$ junto con la isometría de Itô. ■

Ahora queremos calcular la siguiente integral estocástica

$$\int_\alpha^\beta W_t dW_t.$$

Para esto vamos a necesitar el siguiente resultado.

Lema 2.3.2 (Variación Cuadrática de los Procesos de Wiener)**.** *Sea $[\alpha,\beta] \subset [0,\infty)$, y suponga que $P_n = \{\alpha = t_0^n < t_1^n < \ldots < t_{m_n}^n = \beta\}$ son particiones de $[\alpha,\beta]$, con $|P_n| \to 0$ cuando $n \to \infty$. Entonces*

$$\sum_{k=0}^{m_n-1} (W_{t_{k+1}^n} - W_{t_k^n})^2 \to \beta - \alpha$$

en $L^2(\Omega)$ cuando $n \to \infty$.

Demostración. Consideremos $\Delta_k^n(W) = W_{t_{k+1}^n} - W_{t_k^n}$ y $\Delta_k^n = t_{k+1}^n - t_k^n$. Sea $Q_n = \sum_{k=0}^{m_n-1} \Delta_k^n(W)^2$. Entonces

$$Q_n - (\beta - \alpha) = \sum_{k=0}^{m_n-1} (\Delta_k^n(W)^2 - \Delta_k^n),$$

y aplicando la esperanza matemática

$$\mathbf{E}[(Q_n - (\beta-\alpha))^2] = \sum_{k=0}^{m_n-1}\sum_{j=0}^{m_n-1} \mathbf{E}[(\Delta_k^n(W)^2 - \Delta_k^n)(\Delta_j^n(W)^2 - \Delta_j^n)].$$

Para todo $k \neq j$, de acuerdo con la independencia de los incrementos, el termino duplo en la sumatoria es

$$\mathbf{E}[(\Delta_k^n(W)^2 - \Delta_k^n)(\Delta_j^n(W)^2 - \Delta_j^n)] = 0,$$

pues $\mathbf{E}[\Delta_k^n(W)^2] = \Delta_k^n$. Por lo tanto

$$\begin{aligned}\mathbf{E}[(Q_n - (\beta - \alpha))^2] &= \sum_{k=0}^{m_n-1} \mathbf{E}[((\Delta_k^n(W))^2 - \Delta_k^n)^2] \\ &= \sum_{k=0}^{m_n-1} \mathbf{E}[(Y^2 - 1)^2](\Delta_k^n)^2,\end{aligned}$$

en donde

$$Y = Y_k^n \equiv \frac{\Delta_k^n(W)}{\sqrt{\Delta_k^n}} \sim N(0,1).$$

Consecuentemente, si $C \equiv \mathbf{E}[(Y^2 - 1)^2]$, tenemos que

$$\begin{aligned}\mathbf{E}[(Q_n - (\alpha - \beta))^2] &= C \sum_{k=0}^{m_n-1} (\Delta_k^n)^2 \\ &\leq C|P_n|(\beta - \alpha) \to 0 \text{ cuando } n \to \infty.\end{aligned}$$

■

Finalmente para calcular la integral estocástica de un proceso de Wiener es necesario probar las hipótesis de la Definición 2.1.2 sobre $\{W_t\}_{t\in[\alpha,\beta]}$, osea,

- W_t es medible.
- W_t es un proceso adaptado.
- $\mathbf{E}\left[\int_\alpha^\beta |W_t|^2 dt\right] < \infty$.

Omitimos la demostración, pero el lector interesado puede consultar [8]. Luego, se tiene

$$\begin{aligned}\int_\alpha^\beta W_t dW_t &= \lim_{n\to\infty} \sum_{k=0}^{m_n-1} W_{t_k^n}(W_{t_{k+1}^n} - W_{t_k^n}), \\ &= \frac{1}{2}\lim_{n\to\infty} \sum_{k=0}^{m_n-1} (W_{t_{k+1^n}}^2 - W_{t_k^n}^2) - \frac{1}{2}\lim_{n\to\infty} \sum_{k=0}^{m_n-1} (W_{t_{k+1^n}} - W_{t_k^n})^2, \\ &= \frac{1}{2}(W_\beta^2 - W_\alpha^2) - \frac{1}{2}(\beta - \alpha),\end{aligned}$$

en donde $W_{t_k^n}(W_{t_{k+1}^n} - W_{t_k^n}) = \dfrac{1}{2}(W_{t_{k+1^n}}^2 - W_{t_k^n}^2) + \dfrac{1}{2}(W_{t_{k+1^n}} - W_{t_k^n})^2$.

De esta forma, se puede observar que el termino extra $-\frac{1}{2}(\beta - \alpha)$ nos dice que la integral estocástica no tiene el mismo comportamiento que la integral de Riemann.

2.4. Integral estocástica indefinida

Definición 2.4.1. *Sea* $\{x_t\}_{t\in[0,T]} \in \mathbb{M}^2(0,T)$ *un proceso, definimos o proceso estocástico* $\{I_t\}_{t\in[0,T]}$ *tal que*

$$I_t = \int_0^t x_s dW_s,$$

denominado de integral indefinida de $\{x_t\}_{t\in[0,T]}$. *Este proceso es conocido como integral indefinida. Se asume* $I_0 = 0$ *a partir de esta definición.*

A continuación presentamos dos propiedades importante de este proceso $\{x_t\}_{t\in[0,T]}$: es un martingala y tiene trayectorias continuas.

Teorema 2.4.1. *Sea* $\{x_t\}_{t\in[0,T]} \in \mathbb{M}^2(0,T)$ *un proceso, entonces la integral indefinida* $\{I_t\}_{t\in[0,T]}$ *es un martingala en relación a la filtración* $\{\mathcal{F}_t\}_{t\geq 0}$ *implícita en la definición de la clase* $\mathbb{M}^2$.

Demostración. De acuerdo con la definición de Martingala, necesitamos probar que:

(a) $\mathbf{E}[|I_t|] < \infty$ para todo $t \in [0,T]$;

(b) $\{I_t\}_{t\in[0,T]}$ es adaptado a la filtración $\{\mathcal{F}_t\}_{t\in[0,T]}$;

(c) $\mathbf{E}[I_t|\mathcal{F}_s] = I_s$ c.s. para todo $t \geq s$.

La parte (a), se utiliza primero la desigualdad de Cauchy-Schwarz, la isometria de Itô, y del hecho que $\{x_t\}_{t\in[0,T]} \in \mathbb{M}^2$, entonces

$$(\mathbf{E}[|I_t|])^2 \leq \mathbf{E}[|I_t|^2] = \mathbf{E}\left[\left(\int_0^t x_s dW_s\right)^2\right] = \mathbf{E}\left[\int_0^t x_s^2 ds\right] < \infty.$$

Para la parte (b), tomemos para cada t, $\displaystyle\int_0^t x_s dW_s$, un proceso $\mathcal{F}_t$-medible (ver como el limite de v.a. $\mathcal{F}_t$-medible). Finalmente, la parte (c), en primer lugar considerar integrandos elementales, y después extender vía limites para el resultado general. Considere φ un proceso elemental en $\mathbb{M}^2[0,T]$. Para $t > s$, tenemos que

$$\begin{aligned}\mathbf{E}\left[\int_0^t \varphi(r) dW_r|\mathcal{F}_s\right] &= \mathbf{E}\left[\int_0^s \varphi(r) dW_r|\mathcal{F}_s\right] + \mathbf{E}\left[\int_s^t \varphi(r) dW_r|\mathcal{F}_s\right] \\ &= \int_0^s \varphi(r) dW_r + \mathbf{E}\left[\int_s^t \varphi(r) dW_r|\mathcal{F}_s\right] \qquad (2.4.12)\end{aligned}$$

desde que $\int_0^s \varphi(r) dW_r$ es $\mathcal{F}_s$-medible. Por lo tanto, solo nos queda por probar que la segunda del lado derecho de (2.4.12) se anula. Restringiéndonos al intervalo $[s,t]$, φ adopta la forma $\varphi(r) = \displaystyle\sum_{k=0}^{n-1} \varphi(t_k) I_{[t_k,t_{k+1})}$ en donde $\{s = t_0 < t_1 < \cdots < t_n = t\}$ es una partición del intervalo $[s,t]$. Entonces usando las propiedades de la esperanza matemática condicional, obtenemos

$$\begin{aligned}
\mathbf{E}\left[\int_s^t \varphi(r)dW_r|\mathcal{F}_s\right] &= \sum_{k=0}^{n-1}\mathbf{E}[\varphi(t_k)(W_{t_{k+1}}-W_{t_k})|\mathcal{F}_s] \\
\text{(item 9 Teorema 1.3.1)} &= \sum_{k=0}^{n-1}\mathbf{E}\left[\mathbf{E}[\varphi(t_k)(W_{t_{k+1}}-W_{t_k})|\mathcal{F}_{t_k}]|\mathcal{F}_s\right] \\
\text{(item 7 Teorema 1.3.1)} &= \sum_{k=0}^{n-1}\mathbf{E}\left[\varphi(t_k)\mathbf{E}[(W_{t_{k+1}}-W_{t_k})|\mathcal{F}_{t_k}]|\mathcal{F}_s\right] \\
\text{(item 8 Teorema 1.3.1)} &= \sum_{k=0}^{n-1}\mathbf{E}\left[\varphi(t_k)\mathbf{E}[(W_{t_{k+1}}-W_{t_k})]|\mathcal{F}_s\right] \\
&= 0.
\end{aligned}$$

■

Teorema 2.4.2. *Sea* $\{x_t\}_{t\in[0,T]}\in\mathbb{M}^2(0,T)$ *un proceso, entonces la integral indefinida* $\{I_t\}_{t\in[0,T]}$ *admite una modificación con trayectorias continuas i.e., existe un proceso* $\{J_t\}_{t\in[0,T]}$ *con trayectorias continuas c.s., tal que* $\mathbf{P}(I_t=J_t)=1$*, para todo* $t\in[0,T]$*.*

Demostración. Primeramente, es fácil notar que, si $\varphi\in\mathbb{M}^2(0,T)$ entonces el proceso $\left\{\int_0^t\varphi(s)dW_s\right\}_{t\in[0,T]}$ tiene trayectorias continuas c.s. Supongamos que el proceso $\varphi(s)=\sum_{k=0}^{m-1}\varphi(t_k)\mathbb{1}_{[t_k,t_{k+1})}(s)$. Entonces, si $t_m\le t<t_{m+1}$ tenemos que

$$\int_0^t\varphi(s)dW_s=\sum_{k=0}^{m-1}\varphi(t_k)(W_{t_{k+1}}-W_{t_k})+W_t-W_{t_m}.$$

Así, si $t\neq t_m$ la continuidad en t sigue directamente de la continuidad de las trayectorias del movimiento browniano. Si$t=t_m$, entonces es fácil verificar que de la expresión anterior, que

$$\lim_{r\to t_m^-}\int_0^r\varphi(s)dW_s=\lim_{r\to t_m^+}\int_0^r\varphi(s)dW_s=\int_0^{t_m}\varphi(s)dW_s.$$

Por lo tanto, este resultado es valido para integrando elementales. Si $\{\psi_n\}_{n\in\mathbb{N}}$ es una sucesión de procesos elementales en $\mathbb{M}^2(0,T)$ tal que

$$\lim_{n\to\infty}\mathbf{E}\left[\int_0^T|x_t-\psi_n(t)|^2dt\right]=0.$$

Denotemos $I_t^n=\int_0^t\psi_n(t)dW_t$. De la observación anterior, podemos decir que los procesos $\{I_t^n\}_{t\in[0,T]}$ tienen trayectorias continuas c.s. Ademas, por el Teorema 2.4.1, $\{I_t^n\}_{t\in[0,T]}$ es un martingala para cada n y consecuentemente $(I_t^n-I_t^m)_{t\in[0,T]}$ es un martingala para cualquier n, m. Por lo tanto, por el Corolario 2.5.20 y por la isometria de Itô se tiene que, para todo $\epsilon > 0$,

$$\begin{aligned}\mathbf{P}\left(\sup_{0\le t\le T}|I_t^n - I_t^m| > \epsilon\right) &\le \frac{1}{\epsilon^2}\mathbf{E}[|I_t^n - I_t^m|^2] \\ &= \frac{1}{\epsilon^2}\mathbf{E}\left[\int_0^T |\psi_n(t) - \psi_m(t)|^2 dt\right] \to 0 \text{ n,m} \to \infty.\end{aligned}$$

En particular, para $\epsilon = \dfrac{1}{2^k}$ existe un n_k tal que

$$\mathbf{P}\left(\sup_{0\le t\le T}|I_t^{n_k} - I_t^m| > \frac{1}{2^k}\right) < \frac{1}{k^2}$$

Para todo $m \ge n_k$, se selecciona los n_k que tengan la siguiente característica $n_k < n_{k+1}$. Así podemos obtener una subsucesión $\{n_k\}$ tal que

$$\mathbf{P}\left(\sup_{0\le t\le T}|I_t^{n_k} - I_t^{n_{k+1}}|. > \frac{1}{2^k}\right) < \frac{1}{k^2}$$

Se sabe que $\sum \frac{1}{k^2}$ converge, se sigue por el Teorema de Borel-Cantelli que

$$\mathbf{P}\left(\sup_{0\le t\le T}|I_t^{n_k} - I_t^{n_{k+1}}| > \frac{1}{2^k} \text{ para infinitos valores de } k\right) = 0.$$

Para cualquier $j > k$, usando la desigualdad triangular, tenemos

$$\sup_{0\le t\le T}|I_t^{n_k}(\omega) - I_t^{n_j}(\omega)| \le \sum_{i=0}^{j-k-1}\sup_{0\le t\le T}|I_t^{n_{k+i}}(\omega) - I_t^{n_{k+i+1}}(\omega)| \le \sum_{i=0}^{j-k-1}\frac{1}{2^{k+i}} \le \frac{1}{2^{k-1}}$$

Para todo $k \ge k_0(\omega)$. Luego, $\{I_t^{n_k}(\omega)\}_{n\in\mathbb{N}}$ es una sucesión de Cauchy (uniformemente en t) y entonces converge uniformemente en t, para casi todo $\omega \in \Omega$. De forma análoga el limite denotado por $\{J_t(\omega)\}$, es continuo para $t \in [0, T]$, para casi todo $\omega \in \Omega$. Ademas se tiene que $I_t^{n_k} \to I_t$ en $L^2(\Omega)$, cuando $k \to \infty$ y para todo $t \in [0, T]$, y consecuentemente para cada t, existe una subsucesión de $\{I_t^{n_k}\}_{k\in\mathbb{N}}$ que converge c.s. hacia I_t. Por lo tanto $J_t = I_t$ c.s. para todo $t \in [0, T]$. Esto completa la demostración. ■

2.5. Formula de Itô

De forma similar a la definición de la clase $\mathbb{M}^2$, definimos ahora la clase $\mathbb{M}^1$.

Definición 2.5.1. *Sea un proceso $\{x_t\}_{t\in[\alpha,\beta]}$ que satisface las condiciones (1) y (2) de la Definición 2.1.2 y ademas*

$$\mathbf{E}\left[\int_\alpha^\beta |x_t|dt\right] < \infty.$$

Definición 2.5.2. *Sea $\{x_t\}_{t\in[0,T]}$ un proceso que satisface*

$$x_\tau = x_s + \int_s^\tau b(t)dt + \int_s^\tau \sigma(t)dW_t$$

para cualquier tiempo $0 \leq s \leq \tau \leq T$*, en donde* $b \in \mathbb{M}^1(0,T)$ *y* $\sigma \in \mathbb{M}^2(0,T)$*. En esos casos, decimos que el procesos* $\{x_t\}_{t\in[0,T]}$ *tiene diferencial estocástico de la forma*

$$dx_t = b(t)dt + \sigma(t)dW_t.$$

Ejemplo 2.5.1. *Sea* $\{W_t\}_{t\in[\alpha,\beta]}$ *un proceso de Wiener, y del Ejemplo 3.3.4 podemos escribir*

$$\int_\alpha^\beta W_t dW_t = \frac{1}{2}(W_\beta^2 - W_\alpha^2) - \frac{1}{2}(\beta - \alpha),$$

por lo tanto

$$W_\beta^2 = W_\alpha^2 + 2\int_\alpha^\beta W_t dW_t + (\beta - \alpha),$$

así tenemos

$$d(W_t^2) = 2W_t dW_t + dt.$$

Notamos que, usando la regla de la cadena usual (cálculo de Newton-Leibniz) el término dt no aparece.

Teorema 2.5.1 (Formula de Itô)**.** *Sea* $\{x_t\}_{t\in[0,T]}$ *un proceso tal que*

$$dx_t = b(t)dt + \sigma(t)dW_t,$$

en donde $b \in \mathbb{M}^1(0,T)$ *y* $\sigma \in \mathbb{M}^2(0,T)$*. Sea* $h : [0,T]\times\mathbb{R} \to \mathbb{R}$ *una función continua tal que* $\dfrac{\partial h}{\partial x}$*,* $\dfrac{\partial^2 h}{\partial x^2}$ *y* $\dfrac{\partial h}{\partial t}$ *existen y son continuas. Definimos un proceso estocástico* $y_t \equiv h(t, x_t)$*. Entonces*

$$dy_t = \frac{\partial h}{\partial t}(t,x_t)dt + \frac{\partial h}{\partial x}(t,x_t)dx_t + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}(t,x_t)\sigma^2(t)dt \tag{2.5.13}$$

o equivalentemente

$$dy_t = \left(\frac{\partial h}{\partial t}(t,x_t) + \frac{\partial h}{\partial x}(t,x_t)b(t) + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}(t,x_t)\sigma^2(t)\right)dt + \frac{\partial h}{\partial x}(t,x_t)\sigma(t)dW_t. \tag{2.5.14}$$

Demostración. Dividimos la demostración en varios pasos. Primer paso consiste en verificar (2.5.13) en dos pasos particulares.

Paso 1(Casos elementales)

(i) $d(W_t^2) = 2W_t dW_t + dt$. Consideremos $h(t,x) = x^2$. Sigue como Ejemplo 2.5.1.

(ii) $d(tW_t) = W_t dt + t dW_t$. Consideremos $h(t,x) = tx$. En seguida verifiquemos

$$\int_s^r t dW_t = \lim_{n\to\infty} \sum_{k=0}^{m_n - 1} t_k^n (W_{t_{k+1}^n} - W_{t_k^n}), \tag{2.5.15}$$

en donde $P_n\{s = t_0^n < t_1^n < \cdots t_{m_n}^n = r\}$ son particiones del intervalo $[r,s]$, $|P_n| \to 0$ y el limite considerado en probabilidad. Para cada ω para el que $t \to W_t(\omega)$ es continua y tenemos también que

$$\int_s^r W_t dt = \lim_{n\to\infty} \sum_{k=0}^{m_n - 1} W_{t_{k+1}^n}(\omega)(t_{k+1}^n - t_k^n). \tag{2.5.16}$$

Esta última integral es del tipo Riemann en ese sentido podemos evaluar la función en cada extremos derecho del intervalo. Podemos escribir (2.5.16) como una igualdad entre v.a.:

$$\int_s^r W_t dt = \lim_{n\to\infty} \sum_{k=0}^{m_n-1} W_{t_{k+1}^n}(t_{k+1}^n - t_k^n), \tag{2.5.17}$$

en donde se usa que la convergencia c.s. implica convergencia en probabilidad. Sumando las ecuaciones (2.5.15) y (2.5.17) y usando la linealidad del limite, tenemos que

$$\begin{aligned}\int_s^r t dW_t + \int_s^r W_t dt &= \lim_{n\to\infty} \sum_{k=0}^{m_n-1} t_k^n \left(W_{t_{k+1}^n} - W_{t_k^n} \right) + W_{t_{k+1}^n}(t_{k+1}^n - t_k^n) \\ &= \lim_{n\to\infty} \sum_{k=0}^{m_n-1} t_{k+1}^n W_{t_{k+1}^n} - t_k^n W_{t_k^n} \\ &= rW_r sW_s,\end{aligned}$$

luego usando la propiedad de la suma telescópica i.e., $rW_r = sW_s + \int_s^r W_t dt + \int_s^r t dW_t$, tal que en notación diferencial, $d(tdW_t) = W_t dt + tdW_t$.

Paso 2 (Regla del producto de Itô) Suponga que $\{x_t\}_{t\in[0,T]}$ y $\{z_t\}_{t\in[0,T]}$ son procesos estocásticos tales que

$$\begin{cases} dx_t = F_1 dt + G_1 dW_t \\ dz_t = F_2 dt + G_2 dW_t \end{cases}$$

en donde $F_i \in \mathbb{M}^1(0,T)$ y $G_i \in \mathbb{M}^2$, $i = 1, 2$. Entonces

$$d(x_t z_t) = z_t dx_t + x_t dz_t + G_1 G_2 dt. \tag{2.5.18}$$

La Ecuación 2.5.18 es equivalente a

$$d(x_t z_t) = (x_t F_2 + z_t F_1 + G_1 G_2)dt + (x_t G_2 + z_t G_1)dW_t.$$

Luego, probaremos que

$$x_r z_f r - x_0 z_0 = \int_0^r (x_t F_2(t) + z_t F_1(t) + G_1(t) G_2(t))dt + \int_0^r (x_t G_3(t) + z_t G_1(t))dW_t. \tag{2.5.19}$$

Inicialmente, supongamos primero que F_i y G_i, $i = 1, 2$ son procesos constantes en el tiempo. En esos casos, podemos escribir

$$(**) \begin{cases} x_t = x_0 + F_1 dt + G_1 dW_t \\ z_t = y_0 + F_2 dt + G_2 dW_t \end{cases}$$

De (2.5.19) podemos escribir

$$\int_0^r ((x_0 + F_1 t + G_1 W_t)F_2 + (z_0 + F_2 t + G_2 W_t)F_1 + G_1 G_2)dt+$$
$$\int_0^r ((x_0 + F_1 t + G_1 W_t)G_2 + (z_0 + F_2 t + G_2 W_t)G_1)dW_t =$$
$$x_0(F_2 r + G_2 W_r) + z_0(F_1 r + G_1 W_r) + F_1 F_2 r^2 +$$
$$(F_1 G_2 + G_1 F_2)\left(\int_0^r W_t dt + \int_0^r t dW_t\right) + 2G_1 G_2 \int_0^r W_t dW_t + G_1 G_2 r. \quad (2.5.20)$$

Del Paso 1, tenemos que $2\int_0^r W_t dt = W_r^2 + r$ y $\int_0^r W_t dt + \int_0^r t dW_t = rW_r$. Luego, la expresión anterior se transforma en

$$= x_0(F_2 r + G_2 W_r) + z_0(F_1 r + G_1 W_r) + F_1 F_2 r^2 + (F_1 G_2 + G_1 F_2) r W_r + G_1 G_2 W_r^2 = x_r z_r - x_0 z_0$$

en donde la última igualdad es verificada directamente a partir de (**). Para el caso en donde F_i, G_i, $i = 1, 2$ son procesos simples, entonces aplicamos el resultado que acabamos de probar en cada subintervalo en donde F_i, G_i, $i = 1, 2$ son constantes y sumamos las integrales resultantes, obteniendo (2.5.18) también en este caso. El caso general, sigue por aproximación de procesos simples vía limites.

Paso 3 (Caso particular $h(t, x) = x^m$) Suponga que $dx_t = F dt + G dW_t$, entonces

$$d(x_t^m) = m x_t^{m-1} dx_t + \frac{1}{2} m(m-1) x_t^{m-2} G^2 dt.$$

La prueba sigue por inducción. El resultado es claro para $m = 0, 1$ y para $m = 2$ se sigue de la regla del producto. Suponiendo que el resultado es valido para $m - 1$, se tiene

$$\begin{aligned} d(x_t^{m-1}) &= (m-1) x_t^{m-2} dx_t + \frac{1}{2}(m-1)(m-2) x_t^{m-3} G^2 dt \\ &= \left((m-1) x_t^{m-2} F + \frac{1}{2}(m-1)(m-2) x_t^{m-3} G^2\right) dt + (m-1) x_t^{m-2} G dW_t, \end{aligned}$$

Ahora verifiquemos la regla para m. Por la regla del producto tenemos

$$\begin{aligned} d(x_t^m) &= d(x_t x_t^{m-1}) \\ &= x_t d(x_t^{m-1}) + x_t^{m-1} dx_t + (m-1) x_t^{m-2} G G^2 dt \\ &= x_t \left((m-1) x_t^{m-2} dx_t + \frac{1}{2}(m-1)(m-2) x_t^{m-3} G^2 dt\right) + x_t^{m-1} dx_t + (m-1) x_t^{m-2} G^2 dt \\ &= m x_t^{m-1} dx_t + \frac{1}{2} m(m-1) x_t^{m-2} G^2 dt, \end{aligned}$$

en donde la última usamos la relación $m - 1 + \frac{1}{2}(m-1)(m-2) = \frac{1}{2} m(m-1)$.

Paso 4 (Caso particular $h(t, x) = p(x)q(t)$) Del paso anterior, y usando la linealidad de los operadores diferenciales, tenemos

$$d(p(x_t)) = p'(x_t) dx_t + \frac{1}{2} p''(x_t) G^2 dt.$$

Ademas, si $q(t) = q(0) + \int_0^t q'(s)ds$ significa que $dq = q'dt$. Por la regla del producto, tenemos que,

$$\begin{aligned} d(h(x_t,t)) &= d(p(x_t)q(t)) \\ &= p(x_t)dq(t) + q(t)d(p(t)) \\ &= p(x_t)q'(t)dt + q(t)\left(p'(x_t)dx_t + \frac{1}{2}p''(x_t)G^2dt\right) \\ &= \frac{\partial h}{\partial t}(x_t,t) + \frac{\partial h}{\partial x}(x_t,t) + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}G^2dt. \end{aligned}$$

Paso 5 (Caso particular $h(t,x) = \sum_{k=1}^n p_k(x)q_k(t)$, p_k y q_k polinomios) Esto es deducible del paso anterior haciendo uso de la linealidad. Así podemos decir que la formula de Itô es valida para funciones h polinomiales en las variables x y t.

Paso 6 (Caso general) Sea $h : \mathbb{R} \times [0,T] \to \mathbb{R}$ como en la hipótesis del teorema, de acuerdo con una versión mas fuerte del Teorema de Aproximación de Weierstrass, existe una sucesión $\{n_n\}_{n\in\mathbb{N}}$ de funciones polinomiales tales que

$$h_n \to h,\ \frac{\partial h_n}{\partial x} \to \frac{\partial h}{\partial x},\ \frac{\partial^2 h_n}{\partial x^2} \to \frac{\partial^2 h}{\partial x^2},\ \text{y}\ \frac{\partial h_n}{\partial t} \to \frac{\partial h}{\partial t}$$

uniformemente en subconjuntos compactos de $\mathbb{R} \times [0,T]$. Luego del Paso 5, para cada $0 \leq r \leq T$ tenemos que

$$h_n(x_r,r) = h_n(x_0,0) + \int_0^r \left(\frac{\partial h_n}{\partial t}(x_t,t) + \frac{\partial h_n}{\partial x}(x_t,t)F(t) + \frac{1}{2}\frac{\partial^2 h_n}{\partial x^2}(x_t,t)G^2(t)\right)dt + \int_0^r \frac{\partial h_n}{\partial x}(x_t,t)G(t)dW_t.$$

Así, el resultado deseado es obtenido vía limite de la expresión anterior cuando $n \to \infty$. ■

Observación 7. *En alusión a la formula de Taylor, también se puede escribir la formula de Itô de la siguiente manera*

$$dy_t = \frac{\partial h}{\partial t}(t,x_t)dt + \frac{\partial h}{\partial x}(t,x_t)dx_t + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}(t,x_t)(dx_t)^2,$$

en donde $(dx_t)^2$ es calculado formalmente a través de la siguiente tabla de multiplicación que es formalmente fundamentado mediante el Lema 3.3.3. De esa forma, también

$\times$	dt	dW_t
dt	0	0
dW_t	0	dt

podemos escribir (2.5.14) *en su forma integral para cualquier $0 \leq s \leq \tau \leq T$, tal que*

$$y_t = y_s + \int_s^\tau \left(\frac{\partial h}{\partial t}(t,x_t) + \frac{\partial h}{\partial x}(t,x_t)b(t) + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}(t,x_t)\sigma^2(t)\right)dt + \int_s^\tau \frac{\partial h}{\partial x}(t,x_t)\sigma(t)dW_t$$

2.6. Formula de Itô en $\mathbb{R}^n$

Sea $\{W_t\}_{t\geq 0} = \left(W_t^{(1)}, W_t^{(2)}, \cdots, W_t^{(m)}\right)$ um proceso de Wiener en $\mathbb{R}^m$ definido sobre un espacio de probabilidad. Considere la filtración $\{\mathcal{F}_t\}_{t\geq 0}$ tal que

(i) $\sigma(W_s, 0 \leq s \leq t) \subset \mathcal{F}_t$ para todo $t \geq 0$.

(ii) $\sigma(W_{t+\lambda} - W_t, \lambda \geq 0)$ es independiente de $\mathcal{F}_t$ para todo $t \geq 0$.

Definición 2.6.1. *Un proceso $\{x_t\}_{t\in[\alpha,\beta]}$ en $\mathbb{R}^n$ pertenece a $\mathbb{M}^2_{n\times m}(\alpha,\beta)$ si cada componente $(x_t^{ij} | i = 1, \ldots, n; j = 1, \ldots, m)_{t\in[\alpha,\beta]}$ pertenece a la clase $\mathbb{M}^2(\alpha,\beta)$.*

Definición 2.6.2. *Sea $\{x_t\}_{t\in[\alpha,\beta]}$ un proceso estocástico en $\mathbb{M}^2_{n\times m}(\alpha,\beta)$. Entonces, usando la notación matricial, definimos la siguiente integral*

$$\int_\alpha^\beta x_t dW_t = \int_\alpha^\beta \begin{pmatrix} x_t^{(11)} & \cdots & x_t^{(1m)} \\ \vdots & \ddots & \vdots \\ x_t^{(n1)} & \cdots & x_t^{(nm)} \end{pmatrix} \begin{pmatrix} dW_t^{(1)} \\ \vdots \\ dW_t^{(m)} \end{pmatrix}$$

como un vector aleatorio en $\mathbb{R}^n$ con i-esima componente

$$\sum_{k=1}^m \int_\alpha^\beta x_t^{(ik)} dW_t^{(k)}.$$

De forma análoga a la a las propiedades dadas para procesos unidimensionales, tenemos el siguiente resultado.

Teorema 2.6.1. *Sea $\{x_t\}_{t\in[\alpha,\beta]} \in \mathbb{M}^2_{n\times m}(\alpha,\beta)$, entonces*

(i) $\mathbf{E}\left[\int_\alpha^\beta x_t dW_t\right] = 0 \in \mathbb{R}^n$;

(ii) $\mathbf{E}\left[\left|\int_\alpha^\beta x_t dW_t\right|^2\right] = \mathbf{E}\left[\int_\alpha^\beta \|x_t\|^2 dt\right]$.

en donde $\|x_t\|^2 = \sum_{i=1}^n \sum_{j=1}^m |x_t^{(ij)}|^2$.

Demostración. Análogamente a la prueba de los items (b) y (c) del Lema 2.3.1. □

Definición 2.6.3. *Si $\{x_t\}_{t\in[0,T]}$ es un procesos en $\mathbb{R}^n$ tal que*

$$x_\tau = x_s + \int_s^\tau b(t)dt + \int_s^\tau \sigma(t)dW_t$$

para cualquier $0 \leq s \leq t \leq T$, en donde $b(t) = (b^1(t), \ldots, b^n(t)) \in \mathbb{M}^1_n(0,T)$ y $\sigma(t) = (\sigma^{ij}(t) | i = 1, \ldots, n; j = 1, \ldots, m) \in \mathbb{M}^2_{n\times m}(0,T)$, entonces se dice que le proceso $\{x_t\}_{t\in[0,T]}$ tiene diferencial estocástico, para $i = 1, \ldots, n$

$$dx_t^{(i)} = b^{(i)}(t)dt + \sigma^{(ik)}(t)dW_t^{(k)}.$$

Teorema 2.6.2. *[4, pag. 48, sec. 4.2, cap. 4][Formula de Itô] Suponga que*

$$dx_t = b(t)dt + \sigma(t)dW_t,$$

como antes definido. Sea $h : [0, T] \times \mathbb{R}^n \to \mathbb{R}$ *una función continua tal que* $\dfrac{\partial h}{\partial x_i}$, $\dfrac{\partial^2 h}{\partial x_i \partial x_j}$, $i, j = 1, \ldots, n$ *y* $\dfrac{\partial h}{\partial t}$ *existen y son continuas. Definimos un proceso estocástico unidimensional* $y_t \equiv h(t, x_t^{(1)}, \ldots, x_t^{(n)})$. *Entonces*

$$dy_t = \frac{\partial h}{\partial t}(t, x_t)dt + \sum_{i=1}^{n} \frac{\partial h}{\partial x_i}(t, x_t^{(i)})dx_t^{(i)} + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 h}{\partial x_i \partial x_j}(t, x_t) \sum_{k=1}^{m} \sigma^{(ik)}(t)\sigma^{(jk)}(t)dt.$$

Observación 8. *En alusión a la formula de Taylor, también se puede escribir la formula de Itô de la siguiente manera*

$$dy_t = \frac{\partial h}{\partial t}(t, x_t)dt + \sum_{i=1}^{n} \frac{\partial h}{\partial x_i}(t, x_t)dx_t^{(i)} + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 h}{\partial x_i \partial x_j}(t, x_t)(dx_t^{(i)})(dx_t^{(j)}),$$

en donde $(dx_t^{(i)})(dx_t^{(j)})$ *es calculado formalmente a través de la siguiente tabla de multiplicación*

$\times$	dt	$dW_t^{(\ell)}$
dt	0	0
$dW_t^{(k)}$	0	$\delta_{k\ell}dt$

Capítulo 3

Aplicaciones: Ecuaciones Diferenciales Estocásticas

Una ecuación diferencial estocástica es una ecuación diferencial de la forma

$$dx_t = b(t, x_t)dt + \sigma(t, x_t)dW_t, \quad x_0 = Y, \tag{3.0.1}$$

definida para todo $t \in [0, T]$ y con una variable aleatoria Y como condición inicial que se considera $\mathcal{F}_0$-medible e independiente del proceso de Wiener. La solución de esta ecuación diferencial es denotado por x_t y la función $b : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ es llamada de coeficiente de tendencia (*drift*) y la función matricial $\sigma(\cdot)^\intercal\sigma(\cdot)$ es llamada de coeficiente de difusión en donde $\sigma : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n\times m}$. La solución de (3.0.1) en caso exista, se puede escribir de la siguiente manera

$$x_t = x_0 + \int_0^t b(s, x_s)ds + \int_0^t \sigma(s, x_s)dW_s \tag{3.0.2}$$

en donde la primera integral es de Riemann y la segunda es una integral estocástica.

El proceso (3.0.2) esta conducido por una parte determinista que representa la integral de Riemann y una parte afectada (perturbada) por un ruido aditivo el cual esta dado por la integral estocástica. La asociación del ruido a la distribución normal es fundamentada por el Teorema Central del Limite, pues este ruido es considerado como suma resultante de varios pequeñas perturbaciones independientes. En la literatura estas soluciones son conocidas como procesos de Itô, siempre y cuando exista tal solución.

Siguiendo los fundamentos dados en el capítulo anterior, definimos el proceso solución de la ecuación de (3.0.1).

Definición 3.0.1. *Se dice que un proceso estocástico $\{x_t\}_{t\in[0,T]}$ en $\mathbb{R}^n$ es una solución de la ecuación diferencial estocástica* (3.0.1) *si:*

i) $\{x_t\}_{t\in[0,T]}$ *es adaptado a una filtración* $\{\mathcal{F}_t\}_{t\geq 0}$*;*

ii) $b(t, x_t) \in \mathbb{M}^1_n(0, T)$*;*

iii) $\sigma(t, x_t) \in \mathbb{M}^2_{n\times m}(0, T)$*;*

iv) $x_t = x_0 + \int_0^t b(s, x_s)ds + \int_0^t \sigma(s, x_s)dW_s$ *c.s. para todo* $t \in [0, T]$.

La condición (i) captura la idea de que el proceso $\{x_t\}_{t\in[0,T]}$ puede ser construido a partir del proceso $\{W_t\}_{t\in[0,T]}$ y de la condición inicial x_0, a través de los coeficientes $b(\cdot)$ y $\sigma(\cdot)$, de tal manera que, para un determinado t, el valor de x_t dependa solo de la condición inicial x_0 y de los valores de W_s para $0 \le s \le t$, de forma coherente desde el punto de vista de un proceso dinámico.

La Definición 3.0.1 anterior se refiere a la solución fuerte, véase [8, Sec. 5.2]. Notamos que dado un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbf{P})$ donde se define el proceso de Wiener $\{W_t\}_{t\in[0,T]}$ el cual genera una filtración, y la solución se construye a partir de este hecho. Existe también la noción de solución débil, en ese caso, son dadas apenas las funciones $b(\cdot)$ y $\sigma(\cdot)$, y la solución consiste en un espacio de probabilidad con una filtración satisfaciendo las condiciones usuales, donde esta definido un proceso de Wiener $\{W_t\}_{t\ge 0}$ y un proceso adaptado a esta filtración $\{x_t\}_{t\in[0,T]}$ satisfaciendo los itens de la Definición 3.0.1. Para mayor detalle sobre soluciones débiles el lector puede consultar [8, Sec. 5.3].

Ejemplo 3.0.1. *[4, eje. 5.2.3, sec. 5.3, cap. 5] Considere la EDE denominada ecuación de Tanaka*

$$dx_t = -sgn(x_t)dW_t,\ x_0 = 0.$$

La función $sgn(\cdot)$ no satisface la condición de continuidad de Lipschitz requerida para los teoremas que garantizan la existencia y la unicidad de soluciones. Esta ecuación no tiene una solución fuerte, pero si tiene solución débil, en donde, el proceso x_t y la versión del movimiento Browniano se especifican como parte de la solución, en lugar de que el movimiento Browniano se dé a priori.

En este trabajo, solo manipularemos soluciones fuertes de las ecuaciones diferenciales estocásticas.

3.1. Existencia y Unicidad de Soluciones

Primeramente, probaremos dos resultados simples que son útiles en el desarrollo del teorema de existencia y unicidad.

Lema 3.1.1. *Consideremos $\varphi : [0,T] \to \mathbb{R}$ una función continua y no negativa tal que $\varphi(t) \le C\int_0^t \varphi(s)ds$ para todo $0 \le t \le T$, en donde C representa una constante arbitraria. Entonces $\varphi \equiv 0$.*

Demostración. Sea $\psi : [0,T] \to \mathbb{R}$, tal que

$$\psi(t) = e^{-Ct}\int_0^t \varphi(s)ds.$$

Derivando, tenemos que

$$\psi'(t) = e^{-Ct}\left(\varphi(t) - C\int_0^t \varphi(s)ds\right) \le 0, \ \text{ para todo } 0 \le t \le T.$$

Luego $\psi(t) \le \psi(0)$, osea, $e^{-Ct}\int_0^t \varphi(s)ds \le e^{-0}\int_0^0 \varphi(s)ds = 0$, si $t \in [0,T]$. Por lo tanto, $\int_0^t \varphi(s)ds = 0$ para cualquier $t \in [0,T]$. Como φ es una función no negativa y $\varphi(t) \le C\int_0^t \varphi(t)ds = 0$ para todo $t \in [0,T]$, se obtiene entonces que $\varphi \equiv 0$. ■

Lema 3.1.2. *Consideremos u y v dos vectores en $\mathbb{R}^n$ entonces $\|u+v\|^2 \leq 2\|u\|^2 + 2\|v\|^2$.*

Demostración. Usando la desigualdad de Cauchy-Schwarz en $\mathbb{R}^2$, para cualesquier par de números reales a y b, se tiene que, $(a+b)^2 \leq 2a^2 + 2b^2$, en ese sentido, es suficiente considerar los vectores (a, b) y $(1, 1)$. Por lo tanto, procediendo por la desigualdad triangular

$$(\|u+v\|)^2 \leq (\|u\| + \|v\|)^2 \leq 2\|u\|^2 + 2\|v\|^2$$

se obtiene el resultado deseado. ■

Teorema 3.1.1 (Existencia y unicidad)**.** *Consideremos $b : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ y $\sigma : [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n \times m}$ funciones medibles satisfaciendo las siguientes condiciones:*

i) $\|b(t,x) - b(t,y)\| + \|\sigma(t,x) - \sigma(t,y)\| \leq C\|x-y\|$, *para todo* $x, y \in \mathbb{R}^n$, $t \in [0,T]$,

ii) $\|b(t,x)\|^2 + \|\sigma(t,x)\|^2 \leq C^2(1 - \|x\|^2)$, *para todo* $x \in \mathbb{R}^n$, $t \in [0,T]$,

donde C es una constante y $\|\cdot\|$ representa una norma Euclidiana en cada espacio correspondiente. Sea Y una v.a. n-dimensional independiente de los procesos de Wiener $\{W_t\}_{t\in[0,T]}$ dado, tal que,

$$\mathbf{E}[\|Y\|^2] < \infty.$$

Entonces la ecuación diferencial estocástica (3.0.1) *admite una única*[1] *solución $\{x_t\}_{t\in[0,T]} \in \mathbb{M}_n^2(0,T)$ con trayectorias continuas c.s.*

Unicidad. Supongamos que $\{x_t\}_{t\in[0,T]}$ y $\{\tilde{x}_t\}_{t\in[0,T]}$ son dos procesos solución de (3.0.1) en $\mathbb{M}_2^2(0,T)$ con trayectorias continuas. Entonces, para todo $0 \leq t \leq T$,

$$x_t - \tilde{x}_t = \int_0^t (b(s,x_s) - b(s,\tilde{x}_s))ds - \int_0^t (\sigma(s,x_s) - \sigma(s,\tilde{x}_s))dW_s.$$

Por el Lema 3.1.2, en donde,

$$u = \int_0^t (b(s,x_s) - b(s,\tilde{x}_s))ds, \text{ y } v = \int_0^t (\sigma(s,x_s) - \sigma(s,\tilde{x}_s))dW_s$$

tenemos que

$$\mathbf{E}[\|x_t - \tilde{x}_t\|^2] \leq 2\mathbf{E}\left[\left\| \int_0^t (b(s,x_s) - b(s,\tilde{x}_s))ds \right\|^2\right] + \\ 2\mathbf{E}\left[\left\| \int_0^t (\sigma(s,x_s) - \sigma(s,\tilde{x}_s))dW_s \right\|^2\right]. \quad (3.1.3)$$

[1] La unicidad es entendida aquí de la siguiente forma, si $\{x_t\}_{t\in[0,T]}$ y $\{\tilde{x}_t\}_{t\in[0,T]}$ son dos procesos en $\mathbb{M}^2(0,T)$, y ambos son soluciones de (3.0.1) con trayectorias continuas, entonces $\mathbf{P}(x_t = \tilde{x}_t,\ \text{para todo } 0 \leq t \leq T) = 1$. En este caso ambos procesos son denominados indistinguibles.

Para toda función $h : [0,T] \to \mathbb{R}^n$, se tiene que $\left\| \displaystyle\int_0^t h(s)ds \right\|^2 \leq t \displaystyle\int_0^t \|h(s)ds\|^2 ds$. Veamos,

$$\begin{aligned}
\left\| \int_0^t h(s)ds \right\|^2 &= \left\| \left(\int_0^t h_1(s)ds, \dots, \int_0^t h_n(s)ds \right) \right\|^2 \\
&= \left(\int_0^t h_1(s)ds \right)^2 + \cdots + \left(\int_0^t h_n(s)ds \right)^2 \\
\text{Cauchy-Schwarz } &\leq t \int_0^t h_1^2(s)ds + \cdots + t \int_0^t h_n^2(s)ds \\
&= t \int_0^t \|h(s)\|^2 ds,
\end{aligned}$$

en donde, se considera que $\int_0^t h_i(s)ds$ es el producto interno de h_i con la función contante 1, en $L^2(0,t)$. Luego por lo dicho anteriormente, $h(t) = b(t,x_t) - b(t,\tilde{x}_t)$, estimamos

$$\begin{aligned}
\mathbf{E}\left[\left\| \int_0^t b(t,x_s) - b(t,\tilde{x}_s)ds \right\|^2 \right] &\leq T\mathbf{E}\left[\int_0^t \left\| b(t,x_s) - b(t,\tilde{x}_s) \right\|^2 ds \right] \\
\text{por (i) y Fubini } &\leq TC^2 \int_0^t \mathbf{E}\left[\|x_s - \tilde{x}_s\|^2 \right] ds, \qquad (3.1.4)
\end{aligned}$$

ademas por la isometría de Itô, se tiene que

$$\begin{aligned}
\mathbf{E}\left[\left\| \int_0^t \sigma(t,x_s) - \sigma(t,\tilde{x}_s)dW_s \right\|^2 \right] &= \mathbf{E}\left[\int_0^t \left\| \sigma(t,x_s) - \sigma(t,\tilde{x}_s) \right\|^2 ds \right] \\
&\leq C^2 \int_0^t \mathbf{E}\left[\|x_s - \tilde{x}_s\|^2 \right] ds. \qquad (3.1.5)
\end{aligned}$$

De (3.1.3), (3.1.4) y (3.1.5) podemos concluir que, para una constante $K = 2C^2(T+1)$, se tiene que

$$\mathbf{E}[\|x_t - \tilde{x}_t\|^2] \leq K \int_0^t \mathbf{E}[\|x_s - \tilde{x}_s\|^2] ds \qquad (3.1.6)$$

para todo $0 \leq t \leq T$. Definido $\varphi(t) = \mathbf{E}[\|x_t - \tilde{x}_t\|^2]$, la ecuación (3.1.6) nos dice que $\varphi(t) \leq K \int_0^t \varphi(s)ds$ para todo $0 \leq t \leq T$. Luego, en concordancia con el Lema 3.1.1, tenemos que $\varphi \equiv 0$, i.e., $\mathbf{E}[\|x_t - \tilde{x}_t\|^2] = 0$ para todo $0 \leq t \leq T$. Ahora, tomamos la intersección de conjuntos numerables con probabilidad 1, obtenemos

$$\mathbf{P}\left(\|x_t - \tilde{x}_t\| = 0, \text{ para todo } t \in [0,T] \cap \mathbb{Q} \right) = 1,$$

y entonces, usando la continuidad c.s. de las trayectorias o caminos, tal que

$$\mathbf{P}\left(\|x_t - \tilde{x}_t\| = 0, \text{ para todo } t \in [0,T] \right) = 1.$$

De esta manera, conseguimos probar la unicidad a menos de la indistinguibilidad. ■

Existencia La demostración esta basada en aproximaciones sucesivas y es muy próxima a la demostración correspondiente de las ecuaciones diferenciales ordinarias.

Consideramos la sucesión $\{x_t^{(k)}\}_{t\in[0,T],k\in\mathbb{N}}$ de procesos estocásticos definido de la siguiente manera

$$\begin{cases} x_t^{(0)} = Y, \\ x_t^{(k+1)} = Y + \int_0^t b(s, x_s^{(k)})ds + \int_0^t \sigma(s, x_s^{(k)})dW_s, \end{cases} \tag{3.1.7}$$

para cada $k = 0, 1, \ldots$ y para todo $0 \leq t \leq T$.

El objetivo es demostrar que, cuando $k \to \infty$, los procesos $\{x_t^{(k)}\}_{t\in[0,T]}$ convergen (uniformemente en t, c.s., i.e., para casi todo $\omega \in \Omega$ fijo la convergencia es uniforme en t) a un proceso $\{x_t\}_{t\in[0,T]} \in \mathbb{M}^2$ con trayectorias continuas y ademas satisface la ecuación diferencial estocástica (3.0.1).

Afirmación 1: Para $k = 1, 2, \ldots$, el proceso $\{x_t^{(k)}\}_{t\in[0,T]}$ está bien definido en $\mathbb{M}_n^2(0, T)$ y tiene trayectorias continuas c.s. Ademas, se tiene que

$$\mathbf{E}\left[\|x_t^{(k)} - x_t^{(k-1)}\|^2\right] \leq \frac{(Mt)^k}{k!}, \tag{3.1.8}$$

para todo $t \in [0, T]$, en donde M representa una constante que depende solo de C, T y $\mathbf{E}[\|Y\|^2]$.

Demostración. Por inducción. Tener en cuenta que $\{x_t^{(0)}\}_{t\in[0,T]}$, como b y σ son funciones medibles y por la hipótesis (ii) del teorema, podemos garantizar que los procesos $\{b(t, x_t)\}_{t\in[0,T]}$ y $\{\sigma(t, x_t)\}_{t\in[0,T]}$ pertenezcan a $\mathbb{M}_n^2(0, T)$ y $\mathbb{M}_{n\times m}^2(0, T)$ respectivamente. En ese contexto, $\{x_t^{(1)}\}_{t\in[0,T]}$ está bien definido. Ademas, por el Lema 3.1.2, tenemos

$$\begin{aligned} \mathbf{E}[\|x_t^{(1)} - x_t^{(0)}\|^2] &= \mathbf{E}\left[\left\| \int_0^t b(s, x_s)ds + \int_0^t \sigma(s, x_s)dW_s \right\|^2\right] \\ &\leq 2\mathbf{E}\left[\left\| \int_0^t b(s, x_s)ds \right\|^2\right] + 2\mathbf{E}\left[\left\| \int_0^t \sigma(s, x_s)dW_s \right\|^2\right], \end{aligned}$$

y entonces, por la desigualdad de Cauchy-Schwarz, la isometría de Itô y por la hipótesis (ii) del teorema, se tiene que

$$\begin{aligned} \mathbf{E}[\|x_t^{(1)} - x_t^{(0)}\|^2] &\leq 2T\mathbf{E}\left[\int_0^t C^2(1 + \|Y\|^2)ds\right] + 2\mathbf{E}\left[\int_0^t C^2(1 + \|Y\|^2)ds\right], \\ &= 2TtC^2(1 + \mathbf{E}[\|Y\|^2]) + 2tC^2(1 + \mathbf{E}[\|Y\|^2]), \\ &\leq Mt \end{aligned}$$

y por el hecho que $M \geq 2C^2(1+T)(1+\mathbf{E}[\|Y\|^2])$. Esto implica que

$$\begin{aligned}
\mathbf{E}\left[\int_0^t \|x_t^{(1)}\|^2 dt\right] &= \int_0^T \mathbf{E}[\|x_t^{(1)}\|^2]dt \\
\text{Lema 3.1.2} \ &= 2\int_0^T \mathbf{E}[\|x_t^{(1)} - x_t^{(0)}\|^2]dt + 2\int_0^T \mathbf{E}[\|x_t^{(0)}\|^2]dt \\
&\leq MT^2 + 2T\mathbf{E}[\|Y\|^2] \\
&< \infty.
\end{aligned}$$

y entonces $\{x_t^{(1)}\}_{t\in[0,T]} \in \mathbb{M}^2$. La continuidad de las trayectorias esta garantizada por la hipótesis (ii) y por el Teorema 2.4.2. De esta manera, el resultado es valido para $k = 1$.

En seguida, consideramos el resultado valido hasta k, y probamos la validad hasta $k+1$. Como $\{x_t^{(k)}\}_{t\in[0,T]} \in \mathbb{M}_n^2$. Como b y σ son funciones medibles y por la hipótesis (ii) tenemos garantizado que $\{x_t^{(k+1)}\}_{t\in[0,T]}$ esta bien definido. Enseguida, usando estimativas semejantes a las usadas en la prueba de unicidad tenemos que

$$\begin{aligned}
\mathbf{E}[\|x_t^{(k+1)} - x_t^{(k)}\|^2] &= \mathbf{E}\left[\left\| \int_0^t b(s, x_s^{(k)}) - b(s, x_s^{(k-1)})ds + \int_0^t \sigma(s, x_s^{(k)}) - \sigma(s, x_s^{(k-1)})dW_s \right\|^2\right] \\
&\leq 2TC^2 \int_0^t \mathbf{E}[\|x_s^{(k)} - x_s^{(k-1)}\|^2]ds + 2C^2 \int_0^t \mathbf{E}[\|x_s^{(k)} - x_s^{(k-1)}\|^2]ds.
\end{aligned}$$

Por hipótesis de inducción se, tiene que

$$\mathbf{E}[\|x_t^{(k+1)} - x_t^{(k)}\|^2] \leq 2C^2(1+T)\int_0^t \frac{M^k s^k}{k!}ds \leq \frac{M^{k+1}t^{k+1}}{(k+1)!}$$

como $M \geq 2C^2(1+T)$. Como antes, esto va implicar que $\{x_t^{(k+1)}\}_{t\in[0,T]} \in \mathbb{M}^2$, por lo tanto la inducción esta completa. ■

Afirmación 2: Con probabilidad 1, $\{x_t^{(k)}\}_{t\in[0,T]} \in \mathbb{M}^2$ converge uniformemente en t, a un proceso $\{x_t\}_{t\in[0,T]} \in \mathbb{M}_n^2(0,T)$ con trayectorias continuas que solucionan la ecuación (3.0.1).

Demostración. Análogamente a las estimativas anteriores, se tiene que

$$\begin{aligned}
\sup_{0\leq t\leq T} \|x_t^{(k+1)} - x_t^{(k)}\|^2 \leq 2TC^2 \int_0^t \|x_s^{(k)} - x_s^{(k-1)}\|^2 ds + \\
2\sup_{0\leq t\leq T}\left\| \int_0^t b(s, x_s^{(k)}) - b(s, x_s^{(k-1)})dW_s \right\|^2.
\end{aligned}$$

Recordando, que cada componente de la integral estocástica anterior, el segundo termino del lado izquierdo es una suma de integrales estocásticas unidimensionales, las cuales son procesos martigales de acuerdo con el Teorema 2.4.1. Por lo tanto la desigualdad L^p de Doob (Teorema 1.3.2(b)) con $p = 2$, aplicada a cada termino se tiene que,

$$\begin{aligned}
&\mathbf{E}\left[\sup_{0\leq t\leq T} \|x_t^{(k+1)} - x_t^{(k)}\|^2\right] \\
&\quad \leq 2TC^2 \int_0^t \mathbf{E}\left[\|x_s^{(k)} - x_s^{(k-1)}\|^2 ds\right] + 8\mathbf{E}\left[\left\|\int_0^T \sigma(s, x_s^{(k)}) - \sigma(s, x_s^{(k-1)}) dW_s\right\|^2\right] \\
&\quad \leq (2TC^2 + 8C^2) \int_0^T \mathbf{E}[\|x_s^{(k)} - x_s^{(k-1)}\|^2] ds.
\end{aligned}$$

Luego, por (3.1.8), dado $K = 2T^2C^2 + 8TC^2$, se tiene que

$$\mathbf{E}\left[\sup_{0\leq t\leq T} \|x_t^{(k+1)} - x_t^{(k)}\|^2\right] \leq K\frac{(MT)^k}{k!}$$

Consecuentemente, usando la desigualdad de Tchebychev, tenemos que

$$\begin{aligned}
\mathbf{P}\left(\sup_{0\leq t\leq T} \|x_t^{(k+1)} - x_t^{(k)}\| > \frac{1}{2^k}\right) &\leq 4^k \mathbf{E}\left[\sup_{0\leq t\leq T} \|x_t^{(k+1)} - x_t^{(k)}\|^2\right] \\
&\leq K\frac{(4MT)^k}{k!}.
\end{aligned}$$

Como $\sum_{k=0}^{\infty} \dfrac{(4MT)^k}{k!} < \infty$, luego por el Lema de Borel-Cantelli se tiene que

$$\mathbf{P}\left(\sup_{0\leq t\leq T} \|x_t^{(k+1)} - x_t^{(k)}\| > \frac{1}{2^k}, \ \text{ para infinitos valores de } k\right) = 0.$$

Entonces, para casi todo $\omega \in \Omega$ existe un $k_0(\omega)$ tal que

$$\sup_{0\leq t\leq T} \|x_t^{(k+1)}(\omega) - x_t^{(k)}(\omega)\| \leq \frac{1}{2^k}$$

para todo $k \geq k_0(\omega)$. Por lo tanto, con probabilidad 1, la sucesión

$$x_t^{(k)} = x_t^{(0)} + \sum_{i=0}^{k-1} (x_t^{(i+1)} - x_t^{(i)})$$

converge uniformemente en $t \in [0, T]$ a un proceso $\{x_t\}_{t\in[0,T]}$. Naturalmente, el proceso $\{x_t\}_{t\in[0,T]}$ tiene trayectorias continuas c.s., y esta adaptado a la filtración $\{\mathcal{F}_t\}_{t\in[0,T]}$, pues $\{x_t^{(k)}\}_{t\in[0,T]}$ tiene las propiedades para todo k. Ahora, mostremos que $\{x_t\}_{t\in[0,T]} \in \mathbb{M}_n^2(0, T)$.

Después, por un resultado análogo al Lema 3.1.2, se tiene que

$$\begin{aligned}
\mathbf{E}[\|x_t^{(k+1)}\|^2] &\leq 3\mathbf{E}[\|Y\|^2] + 3\mathbf{E}\left[\left\|\int_0^t b(s, x_s^{(k)}) ds\right\|^2\right] + 3\mathbf{E}\left[\left\|\int_0^t \sigma(s, x_s^{(k)}) dW_s\right\|^2\right] \\
&\leq K(1 + \mathbf{E}[\|Y\|^2]) + K\int_0^t \mathbf{E}[\|x_t^{(k)}\|^2] ds,
\end{aligned}$$

donde K es una constante que depende solamente de C y T. Entonces, por inducción,

$$\mathbf{E}[\|x_t^{(k+1)}\|^2] \leq \left(K + K^2 + \cdots + K^{k+2}\frac{t^{k+1}}{(k+1)!}\right)(1+\mathbf{E}[\|Y\|^2]).$$

Consecuentemente,

$$\mathbf{E}[\|x_t^{(k+1)}\|^2] \leq K(1+\mathbf{E}[\|Y\|^2])e^{Kt}.$$

Haciendo $k \to \infty$ y usando el Lema de Fatou, se concluye que

$$\begin{aligned}\mathbf{E}[\|x_t^{(k+1)}\|^2] &= \mathbf{E}\left[\lim_{k\to\infty} \|x_t^{(k+1)}\|^2\right] \\ &\leq \liminf_{k\to\infty} \mathbf{E}[\|x_t^{(k+1)}\|^2] \\ &\leq K(1+\mathbf{E}[\|Y\|^2])e^{Kt}.\end{aligned}$$

Este último resultado implica que

$$\mathbf{E}\left[\int_0^T \|x_t\|^2 dt\right] = \int_0^T \mathbf{E}\left[\|x_t\|^2\right] dt < \infty$$

y por lo tanto, $\{x_t\}_{t\in[0,T]} \in \mathbb{M}_n^2(0,T)$. Para concluir, debemos verificar que $\{x_t\}_{t\in[0,T]}$ es una solución de (3.1.7). De (3.1.8) tenemos que, para $j > i \geq 0$,

$$\begin{aligned}\mathbf{E}[\|x_t^{(j)} - x_t^{(i)}\|^2]^{1/2} &= \|x_t^{(j)} - x_t^{(i)}\|_{L^2(\Omega)} \\ &= \left\| \sum_{k=i+1}^{j} (x_t^{(j)} - x_t^{(i)}) \right\|_{L^2)(\Omega)} \\ &\leq \sum_{k=i+1}^{j} \|x_t^{(j)} - x_t^{(i)}\|_{L^2(\Omega)} \\ &\leq \sum_{k=i+1}^{\infty} \left(\frac{M^k t^k}{k!}\right)^{1/2} \to 0 \text{ cuando } i \to \infty. \end{aligned} \tag{3.1.9}$$

Por lo tanto, identificando las v.a. que son c.s. iguales, se tiene que, $x_t^{(k+1)} \to x_t$ cuando $k \to \infty$ en $L^2(\Omega)$. Como, para c.s. $\omega \in \Omega$, se tiene $x_t^{(k+1)}(\omega) \to x_t(\omega)$ cuando $k \to \infty$, uniformemente en $t \in [0,T]$, y por el Lema de Fatou y (3.1.9) se obtiene que

$$\mathbf{E}\left[\int_0^T \|x_t - x_t^{(i)}\|^2 dt\right] \leq \liminf_{j\to\infty} \mathbf{E}\left[\int_0^T \|x_t^{(j)} - x_t^{(i)}\|^2 dt\right] \to 0$$

cuando $i \to \infty$. Entonces, por la hipótesis (i) y por la isometría de Itô tenemos que

$$\mathbf{E}\left[\left\|\int_0^t \sigma(s,x_s) - \sigma(s,x_s^{(k)})\right\|^2\right] \to 0,$$

cuando $k \to \infty$. Usando la desigualdad de Cauchy-Schwarz se tiene que

$$\mathbf{E}\left[\left\|\int_0^t b(s,x_s) - b(s,x_s^{(k)})\right\|^2\right] \to 0.$$

Por lo tanto, al tomar limite en $L^2(\Omega)$ cuando $k \to \infty$ en (3.1.7), se obtiene,

$$x_t = Y + \int_0^t b(s, x_s)ds + \int_0^t \sigma(s, x_s)dW_s.$$

Osea, el proceso $\{x_t\}_{t\in[0,T]}$ satisface (3.1.7). De esta forma, el teorema esta completo. ■

3.2. Ejemplos

Ejemplo 3.2.1. *Considere la siguiente EDO en* $\mathbb{R}$

$$\begin{cases} \dfrac{dx_t}{dt} = a(t)x_t \\ x_{t_0} = x_0. \end{cases}$$

Esta ecuación sirve, por ejemplo, como un modelo simple de crecimiento poblacional. En este caso, x_t *representa el tamaño de la población, y el termino* $a(t)$ *se interpreta como la tasa de crecimiento poblacional en el instante* t*. Ahora veamos que sucede si adicionamos un ruido aleatorio a la función* $a(\cdot)$*. Suponga que*

$$a(\cdot) = r(\cdot) + \text{“}ruido\text{''}$$

en donde $r(\cdot)$ *es una función conocida. Un posible modelo para tratar esta situación es dado por la siguiente EDE*

$$\begin{cases} dx_t = r(t)x_t dt + \sigma x_t dW_t, \\ x_{t_0} = x_0. \end{cases}$$

en donde σ *es una constante. Simplificando un poco la situación, asumimos que* $r(t) = \mu$ *es una constante. Ahora veamos como es una solución de la siguiente ecuación*

$$\begin{cases} dx_t = \mu x_t dt + \sigma x_t dW_t, \\ x_{t_0} = x_0. \end{cases} \tag{3.2.10}$$

Economizando un poco las cuentas, podemos decir que, dividiendo por x_t *el modelo anterior* (3.2.10)*, tenemos* $\frac{dx_t}{x_t} = \mu dt + \sigma dW_t$*, luego usando la formula de Itô* (2.5.13) *aplicada a la función* $h(t,x) = \log(x)$*, tenemos que,*

$$\begin{aligned} d(\log(x_t)) &= \frac{\partial h}{\partial t}(t, x_t) + \frac{\partial h}{\partial x}(t, x_t)dx_t + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}(t, x_t)\sigma^2 x_t^2 dt, \\ &= \frac{1}{x_t}dx_t + \frac{1}{2}\left(-\frac{1}{x_t^2}\right)\sigma^2 x_t^2 dt \\ &= \left(\mu - \frac{\sigma^2}{2}\right)dt + \sigma dW_t \end{aligned}$$

en donde

$$x_t = x_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t}. \tag{3.2.11}$$

Este tipo de procesos son denominados de movimientos brownianos geométricos y son muy usados en el mundo de las finanzas para modelar la evolución de los precios de las acciones. En ese sentido, σ *es conocido como volatilidad de la acción.*

En seguida, calculamos $\mathbf{E}[x_t]$. *De* (3.2.10), *tenemos que*

$$x_t = x_0 + \int_0^t \mu x_s ds + \int_0^t \sigma x_s dW_s.$$

Por el teorema de Fubini y por el el item (b) del Lema (2.3.1),

$$\mathbf{E}[x_t] = \mathbf{E}[x_0] + \mathbf{E}\left[\int_0^t \mu x_s ds\right] + \mathbf{E}\left[\int_0^t \sigma x_s ds\right] \tag{3.2.12}$$

$$= \mathbf{E}[x_0] + \int_0^t \mu \mathbf{E}\left[x_s\right] ds + 0. \tag{3.2.13}$$

Por lo tanto $\varphi(t) = \mathbf{E}[x_t]$ *satisfaz la ecuación diferencial* $d\varphi(t) = \mu\varphi(t)dt$, *que corresponde a la ecuación* (3.2.10) *sin ruido* ($\sigma = 0$). *Luego* $\mathbf{E}[x_t] = x_0 e^{\mu t}$.

Para calcular la solución de la EDE en determinado momento consideramos $\log(x_t)$, *esto puede causar alguna confusión, pues* x_t *en realidad puede asumir valores negativos. Este problema se evita simplemente definiendo el proceso por la igualdad dada en* (3.2.11) *y verificando mediante la formula de Itô que* x_t, *satisfaz* (3.2.10).

El siguiente ejemplo calculamos y verificamos que la solución para el caso mas general, en donde, el Ejemplo 3.2.1 queda como un caso particular.

Ejemplo 3.2.2. *Suponga que* $n = m = 1$. *Sean* $a, b : [0, T] \to \mathbb{R}$ *funciones continuas, y* x_0 *es una constante. La ecuación*

$$\begin{cases} dx_t = a(t)x_t dt + b(t)x_t dW_t, \\ x_{t_0} = x_0, \end{cases} \tag{3.2.14}$$

tiene una única solución dada por

$$x_t = x_0 e^{\int_0^t \left(a(s) - \frac{1}{2}b(s)^2\right)ds + \int_0^t b(s)dW_s},$$

para $0 \leq t \leq T$. *Para verifica esto, definimos el proceso* $\{y_t\}_{t\geq 0}$ *de la siguiente forma*

$$y_t = \int_0^t \left(a(s) - \frac{1}{2}b(s)^2\right) ds + \int_0^t b(s)dW_s,$$

en donde

$$dy_t = \left(a(t) - \frac{1}{2}b(t)^2\right) dt + b(t)dW_t.$$

Aplicando la formula de Itô para $h(t, x) = x_0 e^x$, *obtenemos que*

$$\begin{aligned} dx_t &= \frac{\partial h}{\partial t}(t, y_t)dt + \frac{\partial h}{\partial x}(t, y_t)dy_t + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}(t, y_t)b(t)^2 dt \\ &= 0 + x_0 e^{y_t} dy_t + \frac{1}{2}x_0 e^{y_t} b(t)^2 dt \\ &= x_0 e^{y_t}\left(\left(a(t) - \frac{1}{2}b(t)^2\right) dt + b(t)dW_t + \frac{1}{2}b(t)^2 dt\right) \\ &= x_t\left(a(t)dt + b(t)dW_t\right) \\ &= a(t)x_t dt + b(t)x_t dW_t. \end{aligned}$$

Para la unicidad basta verificar que están satisfechas las hipótesis del Teorema 3.1.1.

Ejemplo 3.2.3. *La ecuación de Langevin*

$$\begin{cases} dx_t = \alpha x_t dt + \sigma x_t dW_t, \\ x_{t_0} = x_0. \end{cases} \tag{3.2.15}$$

Consideremos el proceso definido por $y_t = e^{-\alpha t} x_t$. *Note que* $y_0 = x_0$. *Usando la formula de Itô, aplicada a* $h(t,x) = e^{-\alpha t} x$, *tenemos que*

$$\begin{aligned} dy_t &= \frac{\partial h}{\partial t}(t, x_t)dt + \frac{\partial h}{\partial x}(t, x_t)dx_t + \frac{1}{2}\frac{\partial^2 h}{\partial x^2}(t, x_t)\sigma^2 dt \\ &= -\alpha e^{-\alpha t} x_t dt + e^{-\alpha t} dx_t + 0 \\ &= \sigma e^{-\alpha t} dW_t. \end{aligned}$$

entonces,

$$e^{-\alpha t} x_t = y_t = y_0 + \sigma \int_0^t e^{-\alpha s} dW_s,$$

tal que tenemos

$$x_t = e^{\alpha t} x_0 + \sigma e^{\alpha t} \int_0^t e^{-\alpha s} dW_s.$$

Ejemplo 3.2.4. *Considere la ecuación de Orntein-Uhlenbeck*

$$dx_t = -\alpha x_t dt + \sigma dW_t \tag{3.2.16}$$

en donde α *y* σ *son constante positivas. Esta ecuación fue propuesta por modela la variación de la velocidad en el movimiento difuso de una partículas para tiempos pequeños. La solución de* (3.2.16) *puede ser escrita de la siguiente forma*

$$x_t = x_0 e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-s)} dW_s \tag{3.2.17}$$

este proceso es conocido como el proceso de Ornstein-Uhlenbeck.

Para obtener una solución consideremos el siguiente proceso

$$y_t = a(t)\left(x_0 + \int_0^t b(s) dW_s\right)$$

en donde a *y* b *son funciones diferenciales. Derivando la expresión anterior y usando la formula de Itô tenemos que*

$$\begin{aligned} dy_t &= a'(t)\left(x_0 + \int_0^t b(s) dW_s\right) dt + a(t)b(t)dW_t \\ &= \frac{a'(t)}{a(t)} y_t dt + a(t)b(t)dW_t, \end{aligned}$$

luego haciendo una simple comparación con (3.2.16) *la funciones* a *y* b *satisfacen*

$$\frac{a'(t)}{a(t)} = -\alpha, \quad a(t)b(t) = \sigma.$$

Imponiendo $a(0) = 1$, *se obtiene que* $a(t) = e^{-\alpha t}$ *y* $b(t) = \sigma e^{\alpha t}$. *De esta forma se obtiene el proceso Ornstein-Uhlenbeck, véase [15].*

Ejemplo 3.2.5 (Modelo de Black-Scholes)**.** *En esta ejemplo, primero vamos a deducir la ecuación diferencial en derivadas parciales que desde su descubrimiento en* 1973 *ha sido denominada como el modelo de Black Scholes Merton. Supongamos que el valor de una acción, que se toma como activo subyacente, es S y satisface la siguiente ecuación diferencial estocástica*

$$dS = bSdt + \sigma dW_t$$

El valor de una opción sobre aquel activo subyacente, lo denotaremos por $V(t, S)$ y es una función del valor de ese activo S, y del tiempo t. Usando el lema de Itô (que es una conocida fórmula del cálculo estocástico) se tiene que

$$dV(t, S) = \left(\frac{\partial V}{\partial t} + bS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) + \sigma S \frac{\partial V}{\partial S} dW_t.$$

Como V y S son correlacionados, esto puede hacerse construyendo un portafolio que consiste de una opción y un número

$$-\frac{\partial V}{\partial S}$$

de acciones. El valor de este portafolio esta dado por

$$J = V - \frac{\partial V}{\partial S} S.$$

Por lo tanto el cambio del valor del portafolio es dado por

$$dJ = dV - \frac{\partial V}{\partial S} dS.$$

Que combinando con las expresiones dadas para dS y dV se convierte en

$$dJ = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt.$$

Además la ganancia de invertir S a una tasa sin riesgo r, durante un intervalo de tiempo dt, sería $rSdt$. Entonces asumiendo que no existe oportunidad de arbitraje y que no hay costos de transacción, se tendría que:

$$rJdt = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right) dt.$$

Sustituyendo $J = V - \frac{\partial V}{\partial S} S$, en la expresión anterior y dividiendo por t se obtiene la ecuación diferencial de Black-Scholes

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0.$$

La mayoría de las ecuaciones diferenciales de Black-Scholes tienen muchas soluciones, que dependen de las condiciones iniciales y de frontera, y que corresponden a la multitud de posibles instrumentos derivados financieros. En muchos casos prácticos, los procedimientos no permiten una solución analítica, y se hace necesario recurrir a métodos numéricos [9, 20, 21].

Capítulo 4

Conclusiones

De esta tesis podemos concluir, que el desarrollo matemático de la integral estocástica como una generalización de la integral de clásica de Riemann, permite tratar un conjunto mas amplio de problemas, los cuales son estudiados mediante ecuaciones diferenciales estocásticas. Con estas ecuaciones podemos tratar problemas con incerteza o riesgo diversas áreas de la matemática, las cuales no pueden ser abordadas por las ecuaciones diferenciales tradicionales. Este es un motivo de notable importancia del estudio de la integral estocástica para luego formular las ecuaciones diferenciales estocásticas. Ademas, cabe resaltar que un estudio riguroso de la integral estocástica junta muchas áreas de la matemática, como análisis funcional, cálculo estocástico y teoría de probabilidades.

En el interior de esta tesis, también se deja en claro la importancia de la integral de Riemann-Stieltjes y la integral de Lebesgue en la construcción de la integral estocástica. Representamos mediante ejemplos asociados a distintas áreas de las matemáticas e ingeniería para ilustrar la importancia de los procesos estocásticos de Itô.

Como parte de las consideraciones finales, notamos en el Capitulo 3 para las EDE, fue necesario juntar varias ramas de la matemática, que tradicionalmente se ensenãn de forma separada en la universidades peruanas. Es una necesidad actual en nuestras universidades el abordar temas multidisciplinarios que por lo general tienen como herramienta principal la matemática.

Una motivación a futuro, seria tratar estos mismos asuntos en espacios mas abstractos. Por ejemplo, ecuaciones diferenciales estocásticas en una variedad diferenciable i.e., procesos de Itô en variedades, en este sentido podemos revisar [13] para tener una visión primaria de este nuevo asunto.

Bibliografía

[1] A. Friedman: *Stochastic Differential Equations and Applications*, Academic Press, 1st. Ed, 1975.

[2] A. N. Kolmogorov: *Foundations of the Theory of Probability*, 2nd. Ed, trans. Nathan Morrison(1956), 1933.

[3] A. F. Karr: *Probability.* 1st. Ed, Springer-Verlag. New York. 1993.

[4] B. Oksendal: *Stochastic Differential Equations: An Introduction with Applications*, Springer, 6th. Ed, 2005.

[5] G. Grimmett and D. Stirzaker: *Probability and Random Processes*, 3rd. Ed, Oxford University Press, 2001.

[6] H. Brezis: *Functional Analysis, Sobolev Spaces and Partial Differential Equations.* 1st. Ed, Springer-Verlag, 2011.

[7] H.-H. Kuo: *Introduction to Stochastic Integration.* 1st. Ed, Springer-Verlag, New York, 2006.

[8] I. Karatzas and S. Shreve:*Brownian motion and stochastic calculus.* 2nd Ed. V. 113, Springer, New York. 1991.

[9] J. Hugger and S. Mashayekhi: *Feedback options in nonlinear numerical finance*, AIP Conference Proceedings, v. 1479, n. 1, p. 2266–2269, 2012.

[10] J. L. Doob: *Stochastic Processes.* 1st. Ed, John Wiley & Sons, New York. 1953.

[11] J. Lukkarinen and M. S. Pakkanen: *On the positivity of Riemann-Stieltjes Integrals.* Australian Mathematical Society. Bulletin 87, n. 3, p. 400-405, 2013.

[12] K. Itô: *Stochastic integral*, Proceedings of the Imperial Academy. Pag. 519–524, 1944.

[13] K. Itô: *Stochastic differential equations in a differentiable manifold*, J. Nagoya Math. p. 35–47, 1950.

[14] L. Hodgkin: *A History of Mathematics: From Mesopotamia to Modernity.* 1st Ed. Oxford University Press. New York. 2005.

[15] P. Kloeden and E. Platen : *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag Berlin Heidelberg, 1st. Ed, 1992.

[16] P. E. Protter: *Stochastic Integration and Differential Equations.* 2nd. Ed, Springer, 2005.

[17] R. G. Bartle: *The Elements of Integration and Lebesgue Measure*, Wiley Classics Library Ed. New York: Wiley, 1995.

[18] R. G. Bartle and D. R. Sherbert: *Introduction to Real Analysis*. 3rd. Ed. John Wiley & Sons Inc, 2000.

[19] R. J. Barry: *Probabilidade: um curso em nível intermediário*, Projeto Euclides - IMPA, 2nd. Ed, 1996.

[20] S. Mashayekhi and J. Hugger: *Finite difference schemes for a nonlinear Black-Scholes model with transaction cost and volatility risk*, Acta Mathematica Universitatis Comenianae (AMUC), v. 84, n. 2, p. 255–266, 2015.

[21] S. Mashayekhi and J. Hugger: *K_α-shifting, Rannacher Time Stepping and Mesh Grading in Crank-Nicolson FDM for Black-Scholes Option Pricing* Communications in Mathematical Finance, v. 5, n. 1, p. 1–31, 2016.

[22] S. Karlin: *A first course in stochastic processes.* Academic press. 2014.

[23] W. Rudin: *Principles of Mathematical Analysis.* 3rd. Ed. McGraw-Hill. New York. 1976.

[24] W. Feller: *An introduction to probability theory and its applications.* John Wiley & Sons, New York. 1971.