

¡Draft! Caminos rugosos y soluciones de
ecuaciones diferenciales.

David Alejandro Alquichire Rincón

21 de mayo de 2025

Idea: Estudiar ecuaciones diferenciales estocásticas por medio de caminos rugosos.

¿Hasta dónde? Peter Fritz... soluciones a PDE estocásticas... ¿Métodos Numéricos?

Propuesta capítulos:

1. Introducción y Preliminares

- a) Conceptos de Probabilidad y Teoría de la medida (LB, D. Cohn, Protter y el otro libro).
- b) Conceptos en Convergencia de Procesos Estocásticos.
- c) Conceptos de Procesos Estocásticos (Notas de Freddy, apoyo de Capinski)
- d) Integración de Riemann Stieltjes
- e) Teoría de la medida y la integral de Lebesgue
- f) Análisis Funcional
- g) Ecuaciones Diferenciales Ordinarias (Existencia y Unicidad)
- h) Ecuaciones Diferenciales Parciales

2. Construcción del Movimiento Browniano

- a) a

3. Construcción de la Integral de Itô

- a) a

4. Ecuaciones Diferenciales Estocásticas por Itô -¿Oksendal

- a) Integral de Itô, Cálculo Estocástico
- b) Ecuaciones Diferenciales Estocásticas, Solución clásica de Itô.
- c) Teoremas de Existencia y Unicidad.

5. Ecuaciones Diferenciales Estocásticas por caminos rugosos Pruebas de funciones α . Gráficas: Simulaciones de ecuaciones rugosas. α -Hölder. Caminos orden α , ¿Cómo luce?

6. EDP Estocásticas*
7. Métodos Numéricos y Aplicaciones*
8. Conclusiones
9. Bibliografía

Título

Camino rugoso y soluciones de ecuaciones diferenciales.

Title

Rough paths and solutions to differential equations.

Resumen:**Abstract:****Palabras clave:****Keywords:**

Índice general

1. Preliminares	7
1.1. Conceptos de Probabilidad.	7
1.1.1. Espacios de probabilidad.	7
1.1.2. Variables aleatorias.	10
1.1.3. Integración respecto a medida de probabilidad. Valor esperado.	12
1.1.4. Variables aleatorias independientes.	16
1.1.5. Distribuciones en \mathbb{R}^n y variables aleatorias Gaussianas.	21
1.1.6. Convergencia de variables aleatorias.	21
1.1.7. Esperanza Condicional.	21
1.2. Preliminares de Integración y Ecuaciones Diferenciales.	26
1.2.1. Ecuaciones Diferenciales Ordinarias Controladas e In- tegración.. . . .	27
1.2.2. Teoremas de Existencia y Unicidad.	29
2. Procesos Estocásticos y el Movimiento Browniano	35
2.1. Conceptos de Procesos Estocásticos	35
2.2. Movimiento Browniano	38
3. Ecuaciones diferenciales estocásticas: Enfoque de Itô.	43
3.1. La integral de Itô.	45
4. Caminos Rugosos e Integración Rugosa	47
4.1. Caminos α -Hölder	50
4.2. Caminos Rugosos y el movimiento Browniano.	50
4.3. Teoría de la Medida	55
4.3.1. Funciones L^p y \mathcal{L}^p	55

Capítulo 1

Preliminares

En este capítulo, nos dedicaremos a repasar conceptos de teoría de la probabilidad, teoría de integración y ecuaciones diferenciales.

1.1. Conceptos de Probabilidad.

En esta sección, daremos un breve repaso a conceptos esenciales en probabilidad, para poder entender mejor procesos estocásticos, y de igual forma, poder realizar la construcción de la integral de Itô. Para mayor información, puede consultar [6], del cuál se basará la gran parte de este capítulo.

1.1.1. Espacios de probabilidad.

Sea Ω un conjunto abstracto. Denotamos por 2^Ω el conjunto de partes de Ω .

Definimos a \mathcal{F} una σ -álgebra de Ω , como un subconjunto de 2^Ω que cumple las siguientes propiedades:

- $\emptyset, \Omega \in \mathcal{F}$
- Si $A \in \mathcal{A}$, luego $A^c \in \mathcal{A}$
- Dado $\{A_i\}_{i \in I}$ una sucesión de subconjuntos de Ω a lo más contable. Luego, si para todo $i \in I$, $A_i \in \mathcal{A}$, entonces $\cup_{i \in I} A_i \in \mathcal{A}$

El espacio (Ω, \mathcal{A}) se llama **espacio medible**.

Los elementos en \mathcal{A} se llamarán *eventos*.

Ejemplo:

- Para Ω un conjunto abstracto, $\mathcal{A} = \{\emptyset, \Omega\}$ es la σ -álgebra trivial.
- Sea $A \subset \Omega$, entonces $\sigma(A) = \{\emptyset, A, A^c, \Omega\}$ también es una σ -álgebra, llamada la **menor σ -álgebra** que contiene a A , que se genera mediante la intersección de todas las σ -álgebras que contienen a A .
- Para $\Omega = \mathbb{R}$, una σ -álgebra para este conjunto es la **σ -álgebra de Borel**, que se puede generar con intervalos de la forma $(-\infty, a]$ para todo $a \in \mathbb{Q}$. También, es la generada por todos los conjuntos abiertos (O cerrados, o semiabiertos...). Para más información consulte [2] Y [6]

□

Una **medida de probabilidad** definida en una σ -álgebra \mathcal{A} de Ω , es una función $P : \mathcal{A} \rightarrow [0, 1]$ que cumple:

- $P(\Omega) = 1$
- Para toda colección contable $\{A_n\}_{n \geq 1}$ de elementos en \mathcal{A} que son disyuntos par a par, se tiene:

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

Es decir, la función es *contablemente aditiva*. Se llama a $P(A)$ como la *probabilidad del evento* A .

La tripla (Ω, \mathcal{A}, P) se conoce como **espacio de probabilidad**.

De forma general, la medida de probabilidad, es un caso específico de una *función de medida*, en este caso, tendremos un *espacio de medida*. Vea COHN.

Note que, podemos ver una propiedad más débil que el axioma (2) en la anterior definición. Para toda colección $\{A_k\}_{k=1}^n$ *finita*, de disyuntos par a par, si tenemos:

$$P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n P(A_k)$$

entonces la función P es **aditiva (O finitamente aditiva)**.

Vamos a revisar algunas propiedades de las funciones de probabilidad, sin demostración. Para consultar los detalles, puede consultar PROTTER.

Teorema 1 Sea (Ω, \mathcal{A}) un espacio medible, y $P : \mathcal{A} \rightarrow [0, 1]$ una función finitamente aditiva y $P(\Omega) = 1$. Entonces, tenemos las siguientes equivalencias:

- La función es *contablemente aditiva*.
- Si $A_n \in \mathcal{A}$ y $A_n \downarrow \emptyset$, luego $P(A_n) \downarrow 0$.

- Si $A_n \in \mathcal{A}$ y $A_n \downarrow A$, luego $P(A_n) \downarrow P(A)$.
- Si $A_n \in \mathcal{A}$ y $A_n \uparrow \Omega$, luego $P(A_n) \uparrow 1$.
- Si $A_n \in \mathcal{A}$ y $A_n \uparrow A$, luego $P(A_n) \uparrow P(A)$.

Más aún, si P es una medida de probabilidad, y dado $\{A_n\}$ sucesión de eventos que converge a A . Entonces $A \in \mathcal{A}$ y $\lim_{n \rightarrow \infty} P(A_n) = P(A)$

Finalmente, damos el concepto de sub- σ -álgebra.

Sea (Ω, \mathcal{A}, P) un espacio de probabilidad. Decimos que \mathcal{T} es una **sub- σ -álgebra** si $\mathcal{T} \subseteq \mathcal{A}$ y \mathcal{T} es σ -álgebra de Ω

1.1.2. Variables aleatorias.

En esta sección, tomamos a (Ω, \mathcal{A}, P) un espacio abstracto, donde Ω no es necesariamente contable.

Sean (E, \mathcal{E}) y (F, \mathcal{F}) dos espacios medibles (No necesariamente tienen una medida de probabilidad). Una función $X : E \rightarrow F$ es una **función medible** si $X^{-1}(\Lambda) \in \mathcal{E}$ para todo $\Lambda \in \mathcal{F}$.

Si (E, \mathcal{E}, P) es un espacio de probabilidad, X posee el nombre de **variable aleatoria**.

Nuevamente, tenemos varias propiedades para las funciones medibles, que enunciaremos acá, sin la demostración respectiva. Para esto, consulte, PROTTER.

Corolario 1 Sea (E, \mathcal{E}) un espacio medible aleatorio, y $(\mathbb{R}, \mathcal{B})$. Sea $X, X_n : E \rightarrow \mathbb{R}$ funciones:

- X es medible si y sólo si $\{X \leq a\} = X^{-1}((-\infty, a]) \in \mathcal{E}$, para todo $a \in \mathbb{R}$.
- Si cada X_n es medible, luego $\sup X_n$, $\inf X_n$, $\limsup X_n$ y $\liminf X_n$ son medibles.
- Si cada X_n es medible, y $\{X_n\}$ converge puntualmente a X , luego X es medible.

Teorema 2 Sea X medible de (E, \mathcal{E}) en (F, \mathcal{F}) , y Y medible de (F, \mathcal{F}) en (G, \mathcal{G}) . Entonces, $Y \circ X$ es medible de (E, \mathcal{E}) en (G, \mathcal{G}) .

Teorema 3 Sean (E, \mathcal{U}) y (F, \mathcal{V}) espacios topológicos, y \mathcal{E}, \mathcal{F} sus σ -álgebras de Borel (generada por los abiertos), respectivamente. Entonces, cada función continua $X : E \rightarrow F$ es medible (O también llamada, función boreliana).

Recuerde que, la función indicadora, $f(x) = 1_A(x)$ se define como:

$$1_A(x) = \begin{cases} 0 & x \in A \\ 1 & x \notin A \end{cases}$$

Teorema 4 Sea $(F, \mathcal{F}) = (\mathbb{R}, \mathcal{B})$ y (E, \mathcal{E}) un espacio medible.

- Función indicadora 1_A en E es medible si y sólo si $A \in \mathcal{E}$
- Si X_1, \dots, X_n son funciones medibles de (E, \mathcal{E}) en $(\mathbb{R}, \mathcal{B})$, y si f es borel en \mathbb{R}^n , luego $f(X_1, \dots, X_n)$.
- Si X, Y son medibles, luego $X + Y, XY, \max(X, Y), \min X, Y$ y X/Y con $Y \neq 0$ son medibles.

Recordemos, para X una variable aleatoria, será una función entre los espacios medibles (Ω, \mathcal{A}) y (E, \mathcal{E}) . Si dotamos al primer espacio de una probabilidad, P , de forma canónica podemos dotar al segundo espacio, de una medida de probabilidad, según X .

Si X es una variable aleatoria entre (Ω, \mathcal{A}, P) , con valores en (E, \mathcal{E}) , la **distribución** (O **medida de distribución**) de X , está definida por:

$$P^X(B) = P(X^{-1}(B)) = P(\{\omega : X(\omega) \in B\}) = P(X \in B)$$

para todo $B \in \mathcal{E}$.

Como la inversa se comporta bien bajo uniones e intersecciones, no es muy difícil probar que:

Teorema 5 La distribución de X es una medida de probabilidad en (E, \mathcal{E})

Si X es una variable aleatoria en \mathbb{R} , P^X es una probabilidad en los reales, caracterizada por la función:

$$F_X(x) = P^X((-\infty, x]) = P(X \leq x)$$

por el hecho, que los elementos en los borelianos, \mathcal{B} , pueden ser generados por elementos de la forma $(-\infty, x]$. $F_X(x)$ se conoce como **función de distribución acumulativa**.

1.1.3. Integración respecto a medida de probabilidad. Valor esperado.

Dada una variable aleatoria, en un espacio de probabilidad (Ω, \mathcal{A}, P) , podríamos determinar un valor esperado, un promedio ponderado según la probabilidad, la imagen que se espera que tenga la variable aleatoria.

Para una variable aleatoria discreta, tenemos la definición:

Ahora, queremos hallar el valor esperado para variables aleatorias en general. Consideramos algunos casos especiales inicialmente:

Una variable aleatoria X es **simple** si su imagen es un conjunto finito, por ende, para una familia de conjuntos disyuntos medibles, $\{A_i\} \subset \mathcal{A}$, y constantes, $a_i \in \mathbb{R}$, para $1 \leq i \leq n$, veremos que la variable aleatoria tiene la forma:

$$X = \sum_{i=1}^n a_i 1_{A_i}$$

Para X variable aleatoria simple, podemos definir su **integral respecto a P o valor esperado** como:

$$\mathbb{E}[X] = \sum_{i=1}^n a_i P(A_i)$$

o también denotado por $\int X dP$.

Ahora, deseamos extender la definición para funciones más generales. Para esto, tendremos en cuenta los siguientes resultados:

Teorema 6 *Para cada variable aleatoria positiva X , existe una sucesión de variables aleatorias simples $\{A_n\}_{n \geq 1}$ tal que X_n tiende a X de forma creciente, para $n \rightarrow \infty$*

Demostración: Podemos tomar la sucesión:

$$X_n(\omega) = \begin{cases} k2^{-n} & \text{si } k2^{-n} \leq X(\omega) < (k+1)2^{-n} \text{ y } 0 \leq k \leq n2^n - 1 \\ n & \text{si } X(\omega) \geq n \end{cases}$$

AÑADIR UNA GRAFICA

□

Teorema 7 *Sea X una variable aleatoria positiva. Si $\{X_n\}$ es sucesión de variables aleatorias simples que tienden de forma creciente a X , entonces $\mathbb{E}[X_n]$ tiende a $\mathbb{E}[X]$*

Primero, podemos definir el valor esperado para variables aleatorias en positivas, esto es, que toma valores en $[0, \infty)$, como:

$$\int X dP = \mathbb{E}[X] = \sup \{ \mathbb{E}[Y] : Y \text{ es función simple con } 0 \leq Y \leq X \}$$

De este modo, podemos definir:

Sea $X^+ = \max(X, 0)$ y $X^- = -\min(X, 0)$. Una variable aleatoria X es **integrable** si $\mathbb{E}[X^+] < \infty$ y $\mathbb{E}[X^-] < \infty$. En este caso, el **valor esperado** de X se define como:

$$\int X dP = \mathbb{E}[X] = \int X^+ dP + \int X^- dP = \mathbb{E}[X^+] + \mathbb{E}[X^-]$$

Tenemos que \mathcal{L}^1 o $\mathcal{L}^1(\Omega, \mathcal{A}, P)$, es el conjunto de variables aleatorias que son integrables.

Ya estamos listos para enunciar varias propiedades importantes de las variables aleatorias.

Teorema 8 ■ \mathcal{L}^1 es espacio vectorial, donde \mathbb{E} es operador lineal, y para $X \in \mathcal{L}^1$ tal que $X \geq 0$, luego $\mathbb{E}[X] \geq 0$. Más aún, para $X \leq Y$, tenemos que $\mathbb{E}[X] \leq \mathbb{E}[Y]$.

- $X \in \mathcal{L}^1$ si y sólo si $|X| \in \mathcal{L}^1$.
- $|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$. Mas aún, cualquier variable aleatoria acotada es integrable.
- Si $X = Y$ casi siempre (Esto es, que $P(X = Y) = P(\{\omega : X(\omega) = Y(\omega)\}) = 1$), entonces, $\mathbb{E}[X] = \mathbb{E}[Y]$.

Teorema 9 (Teorema de la convergencia monótona.) Si las variables aleatorias X_n son positivas y tienden de forma creciente casi siempre a X , luego $\lim_{n \rightarrow \infty} \mathbb{E}[X_n] = \mathbb{E}[X]$.

En este caso, X_n tienden de forma creciente casi siempre a X , si $P(\{\omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}) = 1$.

Lema 1 (Lema de Fatou) Si las variables aleatorias X_n satisfacen $X_n \geq Y$ casi siempre ($P(X_n \leq Y) = P(\{\omega : X_n(\omega) \leq Y(\omega)\}) = 1$) con $Y \in \mathcal{L}^1$, para todo n , entonces:

$$\mathbb{E} \left[\liminf_{n \rightarrow \infty} X_n \right] \leq \liminf_{n \rightarrow \infty} \mathbb{E}[X_n]$$

o se puede escribir como:

$$\int_{\Omega} \liminf_{n \rightarrow \infty} X_n \leq \liminf_{n \rightarrow \infty} \int_{\Omega} X_n$$

En particular, si $X_n \leq 0$ casi siempre para todo n , entonces se cumple la desigualdad.

Teorema 10 (Teorema de la convergencia dominada de Lebesgue) Si las variables aleatorias X_n convergen casi siempre a X y si $|X_n| \leq Y$ casi siempre para $Y \in \mathcal{L}^1$, para todo n , entonces, $X_n, X \in \mathcal{L}^1$, y:

$$\lim_{n \rightarrow \infty} \mathbb{E}[X_n] = \mathbb{E}[X]$$

o

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

Para consultar las pruebas, sugerimos consultar [6]. Ahora, estos teoremas poderosos, van a traer una serie de consecuencias, que serán útiles en la práctica. Enunciamos sin demostración.

Teorema 11 *Sea X_n una sucesión de variables aleatorias.*

- Si para todo n , X_n es positiva, entonces:

$$\mathbb{E} \left[\sum_{n=1}^{\infty} X_n \right] = \sum_{n=1}^{\infty} \mathbb{E}[X_n]$$

o

$$\int_{\Omega} \sum_{n=1}^{\infty} X_n dP = \sum_{n=1}^{\infty} \int_{\Omega} X_n dP$$

- Si $\sum_{n=1}^{\infty} \mathbb{E}[|X_n|] < \infty$, luego $\sum_{n=1}^{\infty} X_n$ converge casi siempre y la suma de esta serie es integrable.

Antes de enunciar más propiedades, definimos los espacios \mathcal{L}^p .

Para $1 < p < \infty$, definimos \mathcal{L}^p el espacio de variables aleatorias tal que $|X|^p \in \mathcal{L}^1$.

Teorema 12 ▪ Si $X, Y \in \mathcal{L}^2$, entonces $XY \in \mathcal{L}^1$ y se cumple la desigualdad de Cauchy-Schwarz:

$$|\mathbb{E}[XY]| \leq \sqrt{\mathbb{E}[X^2]\mathbb{E}[Y^2]}$$

- $\mathcal{L}^2 \subset \mathcal{L}^1$. Si, $X \in \mathcal{L}^2$, luego $(\mathcal{E}[X])^2 \leq \mathbb{E}[X^2]$.
- \mathcal{L}^2 es un espacio vectorial.

El siguiente resultado, permite calcular el valor esperado de cualquier función medible de una variable aleatoria.

Teorema 13 (Regla del valor esperado.) *Sea X una variable aleatoria en (Ω, \mathcal{A}, P) , con valores en (E, \mathcal{E}) y distribución P^X . Sea $h : (E, \mathcal{E}) \rightarrow (\mathbb{R}, \mathcal{B})$ una función medible.*

- $h(X) \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$ si y sólo si $h \in \mathcal{L}^1(E, \mathcal{E}, P^X)$.
- Si, h es positiva, o se tienen las condiciones del inciso anterior, entonces:

$$\mathbb{E}[h(X)] = \int_{\Omega} h(x) P^X(dx)$$

Finalmente, definimos varianza y mostramos una desigualdad conocida y bastante útil.

Si $X \in \mathcal{L}^2$, la **varianza** de X , denotada por σ_X^2 , está dada por:

$$\sigma_X^2 = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

También llamado **segundo momento alrededor de la media** μ_2 .

Teorema 14 (Desigualdad de Chebyshev - Bienaymé)

$$P(|X| \geq a) \leq \frac{\mathbb{E}[X^2]}{\sigma^2}$$

1.1.4. Variables aleatorias independientes.

Al tener dos variables aleatorias independientes, nos dará varias propiedades, por ejemplo, al tener la esperanza del producto de esas dos variables. De igual forma, vamos a definir σ -álgebras en \mathbb{R}^n .

Dado (Ω, \mathcal{A}, P) un espacio de probabilidad. Dados eventos $A, B \in \mathcal{A}$, definimos la **probabilidad condicional** de B dado A , como:

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

siempre que $P(A) > 0$.

De manera intuitiva, podemos ver que si A y B son dos eventos, tal que la ocurrencia de uno, no afecta a otro, o dicho de otra forma, son *independientes*, entonces, sería razonable pensar que:

$$P(B|A) = P(B)$$

porque el hecho que el evento A ocurra, no afectará en nada a B . Formalizando, damos la siguiente definición:

Dado (Ω, \mathcal{A}, P) y $\{A_i\}_{i \in I}$ una colección (a lo más contable) de conjuntos medibles, que pertenecen a \mathcal{A} (También llamados *eventos*). Se dice que la colección es **colección independiente**, o **mutuamente independiente**, si dado $J \subset I$ conjunto finito de índices, se cumple que:

$$P(\cap_{j \in J} A_j) = \prod_{j \in J} P(A_j)$$

La colección es **independiente de a parejas**, si para todo $i, j \in I$, se tiene que $P(A_j \cap A_i) = P(A_j)P(A_i)$.

Tenga en cuenta, que si la colección $\{A_i\}_{i \in I}$ es mutuamente independiente, entonces es independiente de a parejas. Sin embargo, la recíproca no se tiene.

Damos algunas propiedades y teoremas importantes acerca de eventos independientes.

Teorema 15 *Si A, B son independientes, entonces:*

- A, B^c
- A^c, B
- A^c, B^c

son también independientes.

Teorema 16 (Ley de la probabilidad total.) *Dado (Ω, \mathcal{A}, P) un espacio de probabilidad. Sea $\{E_n\}_{n \geq 1}$ una partición a lo más contable de Ω . Luego, para un evento $A \in \mathcal{A}$, se cumple:*

$$P(A) = \sum_n P(A|E_n)P(E_n)$$

Teorema 17 (Teorema de Bayes.) *Dado (Ω, \mathcal{A}, P) un espacio de probabilidad. Sea $\{E_n\}_{n \geq 1}$ una partición a lo más contable de Ω , y sea $P(A) > 0$. Entonces:*

$$P(E_n|A) = \frac{P(A|E_n)P(E_n)}{\sum_m P(A|E_m)P(E_m)}$$

Ahora, podemos generalizar el concepto de independencia a σ -álgebras, e inclusive, a variables aleatorias.

Sea (Ω, \mathcal{A}, P) un espacio de probabilidad.

- Dadas sub- σ -álgebras $\{A_i\}_{i \in I}$ de \mathcal{A} . Decimos que esta colección es **independiente** si para todo $J \subseteq I$ con J finito, y todo $A_i \in \mathcal{A}_i$, se cumple que:

$$P(\cap_{j \in J} A_j) = \prod_{j \in J} P(A_j)$$

En este caso, hablamos de **independencia de σ -álgebras**.

- Sea $\{X_i\}_{i \in I}$ un conjunto de variables aleatorias en el espacio de probabilidad dado, tal que la imagen de X_i es (E_i, \mathcal{E}_i) . Decimos que las variables aleatorias son **independientes** si $\sigma(X_i) = X_i^{-1}(\mathcal{E}_i)$ (Esto es, las σ -álgebras generadas por X_i son independientes).

Enunciamos algunas propiedades para X y Y variables aleatorias independientes. De forma canónica, se puede extender el enunciado a un número a lo más contable de variables aleatorias independientes, $\{X_i\}_{i \in I}$.

Teorema 18 *Sea X, Y variables aleatorias cuya imagen son los espacios (E, \mathcal{E}) y (F, \mathcal{F}) , respectivamente. X y Y son independientes si y sólo sí, se tiene algunas de las siguientes condiciones:*

- $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$ para todo $A \in \mathcal{E}$ y $B \in \mathcal{F}$.
- $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$, para todo $A \in \mathcal{C}$, $B \in \mathcal{D}$ donde \mathcal{C}, \mathcal{D} son clases de conjuntos cerrados bajo intersecciones finitas, tal que $\sigma(\mathcal{C}) = \mathcal{E}$ y $\sigma(\mathcal{D}) = \mathcal{F}$

- Para f, g funciones medibles, $f(X)$ y $g(X)$ son independientes.
- Para f, g funciones medibles positivas, o medibles acotadas, $\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$.
- Sean E, F espacios métricos, y \mathcal{E}, \mathcal{F} sus σ -álgebras de Borel. Entonces, $\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$ para todas f, g funciones acotadas y continuas.

En este punto, vemos que estamos comenzando a tomar conjuntos de dos o más variables aleatorias. Sería deseable hablar de una noción de *conjuntamente medible*. Sean (E, \mathcal{E}) y (F, \mathcal{F}) espacios medibles. En general, $\mathcal{E} \times \mathcal{F} = \{A \subseteq E \times F \mid A = \Lambda \times \Gamma, \Lambda \in \mathcal{E}, \Gamma \in \mathcal{F}\}$ no será una σ -álgebra, por ejemplo, si tomamos el producto $\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R})$ (Producto de los conjuntos de Borel en \mathbb{R}), es tentador pensar que tal producto es σ -álgebra de \mathbb{R}^2 , sin embargo, note que el elemento $[0, 1] \times [0, 1] \cup [-1, -1/2] \times [0, 1]$ debe estar en la σ -álgebra de \mathbb{R}^2 , pero no está en $\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R})$, falla la estabilidad bajo uniones.

Por tanto, denotaremos a:

$$\mathcal{E} \otimes \mathcal{F} = \sigma(\mathcal{E} \times \mathcal{F})$$

como la menor σ -álgebra que contiene a $\mathcal{E} \times \mathcal{F}$. Tal cuál como las anteriores definiciones, enunciamos algunas propiedades y teoremas que serán de utilidad.

Teorema 19 Sea $f : (E \times F, \mathcal{E} \otimes \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{R})$ función medible. Entonces, las secciones $y \rightarrow f(x, y)$ (Para todo $x \in E$) y $x \rightarrow f(x, y)$ (Para todo $y \in F$) son, respectivamente, \mathcal{F} -medible y \mathcal{E} -medible.

Teorema 20 (Tonelli-Fubini) Sea (E, \mathcal{E}, P) y $(F, \mathcal{F}), Q$ espacios de probabilidad.

- Sea $R(A \times B) = P(A)Q(B)$, para $A \in \mathcal{E}$ y $B \in \mathcal{F}$. Entonces, R se extiende de forma unívoca a una probabilidad en $(E \times F, \mathcal{E} \otimes \mathcal{F})$, denotada por $P \otimes Q$.

- Cada función f que es $\mathcal{E} \otimes \mathcal{F}$ -medible, positiva o integrable, respecto a $P \otimes Q$, tenemos que $x \rightarrow \int f(x, y)Q(dy)$ es \mathcal{E} -medible, $y \rightarrow \int f(x, y)P(dx)$ es \mathcal{F} -medible. Además:

$$\int f dP \otimes Q = \int \left\{ \int f(x, y)Q(dy) \right\} P(dx) = \int \left\{ \int f(x, y)P(dx) \right\} Q(dy)$$

Antes de acabar esta sección, vamos a ver dos teoremas importantes en probabilidad. Primero damos una definiciones:

- Sea A_n una sucesión de eventos en \mathcal{A} . Definimos:

$$\limsup_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} (\bigcup_{m \geq n} A_m) = \lim_{n \rightarrow \infty} (\bigcup_{m \geq n} A_m)$$

De manera probabilística, podemos interpretar este evento como:

$$\limsup_{n \rightarrow \infty} A_n = A_n \text{ ocurre } \mathbf{infinitamente seguido}$$

o del inglés, *infinitely often (i.o)*. Esto es, que el evento A_n ocurre para un número infinito de n . Podemos escribir:

$$\limsup_{n \rightarrow \infty} A_n = \{A_n \text{ i.o.}\}$$

- Sean X_n variables aleatorias definidas en (Ω, \mathcal{A}, P) . Defina las σ -álgebras:

$$\begin{aligned} B_n &= \sigma(X_n) \\ C_n &= \sigma(\bigcup_{p \geq n} B_p) \\ C_\infty &= \bigcap_{n=1}^{\infty} C_n \end{aligned}$$

C_∞ es la σ -álgebra cola.

Teorema 21 (Lema de Borel-Cantelli.) Sea A_n una sucesión de eventos en (Ω, \mathcal{A}, P) .

- Si $\sum_{n=1}^{\infty} P(A_n) < \infty$, luego $P(A_i.o) = 0$.
- Si $P(A_n.i.o) = 0$ y si los eventos A_n son mutuamente independientes, entonces $\sum_{n=1}^{\infty} P(A_n) < \infty$.

Teorema 22 (Ley cero-uno de Kolmogorov.) Sea X_n una sucesión de variables aleatorias independientes, definidas en $(\Omega, \mathcal{A}, \mathcal{P})$, y sea C_∞ la σ -álgebra cola correspondiente. Si $C \in C_\infty$, entonces $P(C) = 0$ o $P(C) = 1$.

1.1.5. Distribuciones en \mathbb{R}^n y variables aleatorias Gaussianas.

1.1.6. Convergencia de variables aleatorias.

1.1.7. Esperanza Condicional.

Para X, Y variables aleatorias (Con Y tomando valores de \mathbb{R} y X toma valores a lo sumo contables), ya sabemos calcular el valor de $P(Y|X = i)$, esto es, la probabilidad condicional de Y dado $X = i$. Deseamos extender este concepto, para poder calcular el valor esperado de la variable aleatoria Y , dado que $X = i$, esto es, una *esperanza condicional*.

Dado X que toma valores en $\{x_1, \dots, x_n, \dots\}$ (Conjunto a lo sumo contable), y Y una variable aleatoria. Si $P(X = x_j) > 0$, entonces la **esperanza condicional de Y dado $\{X = x_j\}$** está definida por:

$$\mathbb{E}[Y|X = x_j] = \mathbb{E}_Q[Y] = \int Y dQ$$

tal que $Q(\Lambda) = P(\Lambda|X = x_j)$, y dado que $\mathbb{E}_Q[|Y|] < \infty$.

Una forma clásica de calcular esta esperanza condicional, con Y tomando valores de un conjunto a lo más contable, está enunciada en este teorema, que es consecuencia de la definición:

Teorema 23 Asuma las condiciones de la definición anterior. Además, si Y tiene valores contables $\{y_1, \dots, y_n, \dots\}$ y si $P(X = x_j) > 0$, entonces:

$$\mathbb{E}[Y|X = x_j] = \sum_{k=1}^{\infty} y_k P(Y = y_k|X = x_j)$$

dado que la serie converge absolutamente.

Sin embargo, no todas las variables aleatorias tomarán valores de un conjunto a lo más contable, deseamos extender la definición a variables aleatorias más generales. Primero, podemos definir la esperanza condicional, entre dos variables aleatorias.

Sea:

$$f(x) = \begin{cases} \mathbb{E}[Y|X = x] & \text{si } P(X = x) > 0 \\ \text{Otro valor} & \text{si } P(X = x) = 0 \end{cases}$$

Dado X con valores contables, y Y una variable aleatoria con valores en \mathbb{R} . La **esperanza condicional de Y dado X** está definida por:

$$\mathbb{E}[Y|X] = f(X)$$

Note que la definición está dada por eventos que ocurren *casi siempre* (*a.s.*), es decir, salvo en elementos, donde la probabilidad sea 0. Ahora, si X toma valores reales, en general los eventos $\{X = x_j\}$ tendrán probabilidad 0 y el enfoque dado no funciona.

Sin embargo, lo que se realizó fue hallar una función f auxiliar, para hallar la esperanza condicional, y este será el enfoque que se tomará para el caso en general. Para cumplir el objetivo, se usará el siguiente teorema.

Teorema 24 (Lema de Doob-Dynkin.) Sea X una variable aleatoria con valores en \mathbb{R}^n , y Y una variable aleatoria con valores en \mathbb{R} . Y es medible con respecto a $\sigma(X)$ (La σ -álgebra generada por X) si y sólo si existe una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ que sea Borel-medible tal que $Y = f(X)$.

Sea (Ω, \mathcal{A}, P) un espacio de probabilidad, y $X : \Omega \rightarrow \mathbb{R}^n$ una variable aleatoria. Recuerde que, \mathcal{L}^2 es el espacio de los *cuadrado-integrables*, esto es, que para $Y \in \mathcal{L}^2$, $\mathbb{E}[Y^2] < \infty$ (O $\int_{\Omega} Y^2 dP < \infty$). Como $\mathcal{L}^2(\Omega, \mathcal{A}, P)$ es un *espacio de Hilbert*, entonces se puede definir un producto interno en este espacio:

$$\langle Y, Z \rangle = \int_{\Omega} YZ dP = \mathbb{E}[YZ]$$

Como $\sigma(X) \subset \mathcal{A}$, entonces $(\Omega, \sigma(X), P)$ es también un subespacio de Hilbert. Dado estos conceptos, se procede a generalizar la esperanza condicional.

Sea $Y \in \mathcal{L}^2(\Omega, \mathcal{A}, P)$. La **esperanza condicional de Y dado X** es el único $\hat{Y} \in \mathcal{L}^2(\Omega, \sigma(X), P)$ tal que:

$$\mathbb{E}[\hat{Y}Z] = \mathbb{E}[YZ]$$

para todo $Z \in \mathcal{L}^2(\Omega, \sigma(X), P)$, se denota como $\mathbb{E}[Y|X] = \hat{Y}$.

Note que la esperanza condicional de Y dado X , será la proyección de Y en el subespacio de Hilbert $\mathcal{L}^2(\Omega, \sigma(X), P)$. Para más detalles de análisis funcional, consulte el apéndice.

Como \hat{Y} es $\sigma(X)$ -medible, por el lema de Doob-Dynkin, existe una función f Borel medible tal que:

$$\mathbb{E}[Y|X] = f(X)$$

Entonces, se tiene una expresión alternativa para el teorema de Doob-Dynkin; para toda función g Borel-medible tal que $g(X) \in \mathcal{L}^2$:

$$\mathbb{E}[f(X)g(X)] = \mathbb{E}[Yg(X)]$$

Ahora, se puede definir la esperanza condicional respecto a una σ -álgebra.

Sea $Y \in \mathcal{L}^2(\Omega, \mathcal{A}, P)$ y \mathcal{G} una sub- σ -álgebra de \mathcal{A} . Definimos la **esperanza condicional de Y respecto a \mathcal{G}** como el único elemento $\mathbb{E}[Y|\mathcal{G}] \in \mathcal{L}^2(\Omega, \mathcal{G}, P)$ tal que:

$$\mathbb{E}[YZ] = \mathbb{E}[\mathbb{E}[Y|\mathcal{G}]Z]$$

para todo $Z \in \mathcal{L}^2(\Omega, \mathcal{G}, P)$. También podemos escribir como:

$$\mathbb{E}[Y|\mathcal{G}] = \int_{\Lambda} Y dP$$

para todo $\Lambda \in \mathcal{G}$.

Note que, $\mathbb{E}[Y|\mathcal{G}]$ se puede ver como la proyección de Y sobre el subespacio de Hilbert (Ω, \mathcal{G}, P) .

Como dato adicional, es importante recalcar que la esperanza condicional es un elemento de \mathcal{L}^2 , y es una *clase de equivalencia* de alguna variable aleatoria. Por ende, relaciones del tipo $\mathbb{E}[Y|\mathcal{G}] \geq 0$ o $\mathbb{E}[Y|\mathcal{G}] = Z$ se sobreentienden como relaciones que se cumplen en *casi siempre* o en *casi todo sitio* (Esto es, se preserva la igualdad salvo en conjuntos nulos o de medida cero).

Al igual que el operador \mathbb{E} cumple varias propiedades, también $\mathbb{E}[\cdot|\cdot]$ cumple propiedades similares.

Teorema 25 (Propiedades de la esperanza condicional.) Sea $Y \in \mathcal{L}^2(\Omega, \mathcal{A}, P)$ y \mathcal{G} una sub- σ -álgebra de \mathcal{A} . Entonces,

- Si $Y \geq 0$, luego $\mathbb{E}[Y|\mathcal{G}] \geq 0$.
- Si $\mathcal{G} = \sigma(X)$ para cierta variable aleatoria X , entonces existe una función Borel medible f tal que $\mathbb{E}[Y|\mathcal{G}] = f(X)$.
- $\mathbb{E}[\mathbb{E}[Y|\mathcal{G}]] = \mathbb{E}[Y]$.
- La función $Y \rightarrow \mathbb{E}[Y|\mathcal{G}]$ es lineal.

Teorema 26 Sea $Y \in \mathcal{L}^+(\Omega, \mathcal{A}, P)$ (El espacio de variables aleatorias no negativas), y sea \mathcal{G} una sub- σ -álgebra de \mathcal{A} . Entonces existe un único elemento $\mathbb{E}[Y|\mathcal{G}]$ de $\mathcal{L}^+(\Omega, \mathcal{G}, P)$ tal que:

$$\mathbb{E}[YX] = \mathbb{E}[\mathbb{E}[Y|\mathcal{G}]X]$$

para todo $X \in \mathcal{L}^+(\Omega, \mathcal{G}, P)$. Más aún, si $0 \leq Y \leq Y'$, luego:

$$\mathbb{E}[Y|\mathcal{G}] \leq \mathbb{E}[Y'|\mathcal{G}],$$

es decir, la esperanza condicional es monótona creciente.

Teorema 27 Sea $Y \in \mathcal{L}^+(\Omega, \mathcal{A}, P)$ y \mathcal{G} una sub- σ -álgebra de \mathcal{A} . Entonces, existe un único elemento $\mathbb{E}[Y|\mathcal{G}] \in \mathcal{L}^1(\Omega, \mathcal{G}, P)$ tal que:

$$\mathbb{E}[YX] = \mathbb{E}[\mathbb{E}[Y|\mathcal{G}]X]$$

para toda X que es acotada y \mathcal{G} -medible. Además, se cumple:

- Si $Y \geq 0$, luego $\mathbb{E}[Y|\mathcal{G}] \geq 0$.

- $Y \rightarrow \mathbb{E}[Y|\mathcal{G}]$ es un mapa lineal.

Note que, se han visto algunas definiciones equivalentes a la esperanza condicional.

Ahora, se puede caracterizar la esperanza condicional para diversos casos, como cuando una variable aleatoria es medible respecto a una sub- σ -álgebra.

Teorema 28 Sea Y una variable aleatoria positiva o integrable en (Ω, \mathcal{A}, P) , y sea \mathcal{G} una sub- σ -álgebra. Entonces, $\mathbb{E}[Y|\mathcal{G}] = Y$ si y sólo si Y es \mathcal{G} -medible.

Al tomar dos variables aleatorias, y colocar condiciones, se pueden obtener resultados interesantes.

Teorema 29 Sea $Y \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$, y sea X, Y independientes, luego:

$$\mathbb{E}[Y|X] = \mathbb{E}[Y]$$

Teorema 30 Sea X, Y variables aleatorias en (Ω, \mathcal{A}, P) , sea \mathcal{G} una sub- σ -álgebra de \mathcal{A} , y suponga que X es \mathcal{G} -medible. Si X, Y y XY son integrables, o si X, Y son positivos, tendremos que: $\mathbb{E}[XY|\mathcal{G}] = X\mathbb{E}[Y|\mathcal{G}]$

En la esperanza condicional, también se tendrán los teoremas importantes de convergencia, que se tenían para la esperanza usual.

Teorema 31 (Teoremas de convergencia para la esperanza condicional.)

Sea $\{Y_n\}_{n \geq 1}$ una sucesión de variables aleatorias en (Ω, \mathcal{A}, P) , y sea \mathcal{G} una sub- σ -álgebra de \mathcal{A} .

- **(Convergencia monótona).** Si $Y_n \geq 0$, $n \geq 1$ y $Y_n \rightarrow Y$ de forma creciente, luego, casi siempre se tiene:

$$\lim_{n \rightarrow \infty} \mathbb{E}[Y_n|\mathcal{G}] = \mathbb{E}[Y|\mathcal{G}]$$

- **(Lema de Fatou).** Si $Y_n \geq 0$, $n \geq 1$, entonces:

$$\mathbb{E}[\liminf_{n \rightarrow \infty} Y_n|\mathcal{G}] \leq \liminf_{n \rightarrow \infty} \mathbb{E}[Y_n|\mathcal{G}]$$

- **(Teorema de la convergencia dominada de Lebesgue.)** Si $\lim_{n \rightarrow \infty} Y_n = Y$ casi siempre, y $|Y_n| \geq Z$, con $n \geq 1$ para algún $Z \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$, entonces, casi siempre:

$$\lim_{n \rightarrow \infty} \mathbb{E}[Y_n | \mathcal{G}] = \mathbb{E}[Y | \mathcal{G}]$$

Se finaliza la sección, con algunas desigualdades importantes que involucran a la esperanza condicional.

Teorema 32 (Desigualdad de Jensen.) Dada $\phi : \mathbb{R} \rightarrow \mathbb{R}$ una función convexa, y sea $X, \phi(X)$ variables aleatorias. Entonces, para toda sub- σ -álgebra, se tiene que:

$$\phi(\mathbb{E}[X | \mathcal{G}]) \leq \mathbb{E}[\phi(X) | \mathcal{G}]$$

Como consecuencia de la desigualdad de Jensen, se tiene:

Teorema 33 (Desigualdad de Hölder.) Dados X, Y variables aleatorias, tal que $\mathbb{E}[|X|^p] \leq \infty$, $\mathbb{E}[|Y|^q] \leq \infty$ con $p > 1$, y $\frac{1}{p} + \frac{1}{q} = 1$. Luego:

$$\mathbb{E}[|XY|] \leq \mathbb{E}[|X|^p]^{\frac{1}{p}} \mathbb{E}[|Y|^q]^{\frac{1}{q}}$$

Y de igual forma, se tiene una consecuencia de la desigualdad de Hölder:

Teorema 34 (Desigualdad de Minkowski.) Dada X, Y variables aleatorias, $1 \leq p < \infty$ con $\mathbb{E}[|X|^p] < \infty$ y $\mathbb{E}[|Y|^p] < \infty$. Luego:

$$\mathbb{E}[|X + Y|^p]^{\frac{1}{p}} \leq \mathbb{E}[|X|^p]^{\frac{1}{p}} + \mathbb{E}[|Y|^p]^{\frac{1}{p}}$$

1.2. Preliminares de Integración y Ecuaciones Diferenciales.

En esta sección, se hablará un poco acerca de temas de Integración, como la integral de Riemann-Stieltjes, la integración de Young y algunos teoremas importantes para poder estudiar caminos rugosos.

De igual forma, se verá una breve introducción a las ecuaciones diferenciales ordinarias, enunciando algunos teoremas de existencia y unicidad. Este será útil, al demostrar el teorema de existencia y unicidad en ecuaciones diferenciales estocásticas.

1.2.1. Ecuaciones Diferenciales Ordinarias Controladas e Integración..

En esta sección, se hablarán algunos aspectos breves acerca de las ecuaciones diferenciales controladas, y ciertos problemas que estas pueden poseer. También se hablará brevemente *integral de Riemann-Stieltjes* (Para un tratamiento más extenso, consulte [1]) y también sobre la *integral de Young*, y finalmente, se observará que estas integrales en ciertas ocasiones no son suficientes al tener caminos poco regulares.

Sea $\Omega \subset \mathbb{R} \times \mathbb{R}^n$. Definimos una *norma* en $\mathbb{R} \times \mathbb{R}^n$ como:

$$|(t, x)| = \max\{|t|, \|x\|\}.$$

Sea $f : \Omega \rightarrow \mathbb{R}^n$ una función continua, e I un intervalo conexo de \mathbb{R} no reducido a un punto. Entonces, se dice que una función diferencial $\phi : I \rightarrow \mathbb{R}^n$ es una **solución de la ecuación diferencial**

$$\frac{dx}{dt} = f(t, x)$$

en el intervalo I si se cumplen las condiciones:

- $\{(t, \phi(t)) | t \in I\} \subset \Omega$.
- $\frac{d\phi}{dt}(t) = f(t, \phi(t))$ para todo $t \in I$.

La ecuación diferencial mostrada, es una **ecuación diferencial ordinaria de primer orden**.

En general, se observa que las ecuaciones diferenciales ordinarias, son un tipo de **sistema controlado de ecuaciones diferenciales**:

$$dY_t = f(Y_t)dX_t$$

donde f se puede pensar como un campo vectorial, $Y : [0, T] \rightarrow \mathbb{R}^n$ es la solución y $X : [0, T] \rightarrow \mathbb{R}^d$ es una *señal de entrada*. Si tenemos que X_t es absolutamente continua, se tendrá una EDO. Pero, ¿Qué pasa si X_t toma

otra forma? De forma integral, la expresión de arriba se puede reescribir como:

$$Y_t = Y_0 + \int_0^t f(Y_s) dX_s$$

y el problema es sobre cómo definir la integral $\int f dX_s$, por ejemplo, si X_t es una función con baja regularidad, como un movimiento de ruido blanco, u otro proceso estocástico. Una solución a medias, bajo una topología sutil se podría aproximar por caminos suaves $X_n \rightarrow X$ y $Y_n \rightarrow Y$, como:

$$\int_0^t f(Y_s) dX_s = \lim_{n \rightarrow \infty} \int_0^t f(Y_s^n) dX_s^n$$

Por ejemplo, al tener funciones continuas, se podría tomar la norma del supremo, que es la topología de la convergencia uniforme. Comenzando por el caso más simple, se puede definir la integral de Riemann-Stieljes:

Sea $\pi = \{0 = t_0, t_1, \dots, t_N = T\}$ una partición de $[0, T]$. Sea $\pi_n = \{0 = t_0^n, t_1^n, \dots, t_N^n = T\}$ una sucesión de particiones tal que $|\pi_n| \rightarrow 0$ para $n \rightarrow \infty$, con $u_i^n \in [t_i^n, t_{i+1}^n]$. Para Y camino continuo, y X camino de variación acotada, definimos la **integral de Riemann-Stieljes** como:

$$\int_0^T Y_s dX_s = \lim_{n \rightarrow \infty} \sum_{i=0}^{N_n-1} Y_{u_i^n} (X_{t_{i+1}^n} - X_{t_i^n})$$

Gracias a las condiciones de los caminos Y y X , la integral no dependerá del punto u_i^n seleccionado en el intervalo de la partición, lo que es una gran ventaja. Sin embargo, este lujo se pierde al hacer los caminos menos regulares.

Se puede debilitar la condición de continuidad de los caminos, por un concepto más débil.

Para $\alpha \in (0, 1]$, un camino $X : [0, T] \rightarrow \mathbb{R}^d$ es **α -Hölder continuo** si existe C constante tal que:

$$|X_t - X_s| \leq C|t - s|^\alpha$$

para todo $t, s \in [0, T]$ con $s < t$.

1.2. PRELIMINARES DE INTEGRACIÓN Y ECUACIONES DIFERENCIALES.29

De esta forma, definimos la **integración de Young**.

Teorema 35 (Integración de Young.) Sea $\alpha, \beta \in (0, 1]$ tal que:

$$\alpha + \beta > 1$$

Sea X una función α -Hölder continua, y Y β -Hölder continua. Sea $\pi_n = \{t_0^n, \dots, t_N^n\}$ sucesión de particiones con $|\pi_n| \rightarrow 0$ para $n \rightarrow \infty$. Para todo $n \geq 1$, e $i = 0, 1, \dots, N-1$, tomando $u_i^n \in [t_i^n, t_{i+1}^n]$. Luego el límite

$$\int_0^T Y_s dX_s = \lim_{n \rightarrow \infty} \sum_{i=0}^{N_n-1} Y_{u_i^n} (X_{t_{i+1}^n} - X_{t_i^n})$$

llamada **integral de Young**, existe, y no depende de la sucesión $\{\pi_n\}$, ni de la elección del punto u_i^n .

Al trabajar en ecuaciones diferenciales estocásticas, usualmente el integrador es $X_s = B_s$, un movimiento Browniano, que tiene muy baja regularidad. Por el teorema de continuidad de Kolmogorov (Capítulo 2), las trayectorias del Browniano son *casi siempre* α -Hölder continuas para $\alpha < \frac{1}{2}$. En este caso, no se pueden cumplir siempre las condiciones de la integral de Young, por ende, es mejor cambiar el enfoque.

Más adelante, en el presente trabajo, se hablará acerca del movimiento Browniano, y la construcción de la integral de Itô, como una solución para atacar la baja regularidad.

1.2.2. Teoremas de Existencia y Unicidad.

De forma preliminar, las ecuaciones diferenciales se han usado a lo largo de los años para modelar distintos fenómenos naturales, económicos, etc... y es natural preguntarse si es posible solucionar cierta ecuación diferencial, obtener una solución y determinar un modelo matemático a cierto fenómeno. En este apartado, se formalizará estos conceptos, y se mostrarán los teoremas de existencia y unicidad, con su demostración.

Ahora, consideramos el *problema de Cauchy*, o específicamente, *problema de valor inicial*:

$$\frac{dx}{dt} = f(t, x), \text{ sujeto a } x(t_0) = x_0$$

Bajo algunas condiciones sutiles, esta ecuación diferencial tiene solución y es única. Primero, se muestran algunas definiciones y un lema importante, antes de pasar a trabajar la existencia y unicidad de la solución.

Una función $f : \Omega \subset \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ es una función **Lipschitz-continua** en Ω (En este caso, respecto a la segunda variable), si existe $K \in \mathbb{R}$ constante tal que:

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

para toda $(t, x), (t, y) \in \Omega$.

Si para una vecindad de $(t_0, x_0) \in \Omega$, $V_{\mathbb{R}}(t_0, x_0; \epsilon)$, la función es Lipschitz-continua, entonces se dice **localmente Lipschitz-continua**.

Note que, si f tiene derivada parcial respecto a la segunda variable, D_2f , que está acotada, $\|D_2f\| \leq K$ y Ω es un conjunto convexo, entonces, por el teorema del valor medio, tendremos:

$$|f(t, x) - f(t, y)| \leq \sup_{0 \leq \theta \leq 1} \|D_2f(t, \theta x + (1 - \theta)y)\| |x - y| \leq K|x - y|$$

Lema 2 (Lema de la contracción (Punto fijo de Banach).) *Dado (X, d) un espacio métrico completo, $F : X \rightarrow X$ una contracción, esto es, $d(F(x), F(y)) \leq Kd(x, y)$ para $0 < K < 1$. Entonces, F tiene un único punto fijo p . Más aún, $F^n(x) = F(F^{n-1}(x)) \rightarrow p$ cuando $n \rightarrow \infty$ para todo $x \in X$. En este caso, p se conoce como **atractor** de F .*

Corolario 2 *Dado (X, d) un espacio métrico completo. Si $F : X \rightarrow X$ es continua, y para algún m , F^m es contracción, luego existe un único punto fijo para F . Más aún, ese punto fijo p es atractor de F .*

Ya con estos resultados, comenzamos a trabajar los teoremas de existencia y unicidad de soluciones en ecuaciones diferenciales ordinarias.

Teorema 36 (Teorema de Picard.) *Sea f una función continua, y de Lipschitz en $\Omega = I_a \times B_b$, donde $I_a = \{t | |t - t_0| \leq a\}$ y $B_b = \{x | \|x - x_0\| \leq b\}$. Sea $|f| \leq M$, esto es, la función está acotada. Entonces, el problema:*

$$\frac{dx}{dt} = f(t, x), \quad x(t_0) = x_0$$

tiene solución única en I_α con $\alpha = \min \{a, b/M\}$

Demostración. Sea $X = \mathcal{C}(I_a, B_b)$ el espacio métrico completo de funciones $\phi : I_a \rightarrow B_b$ continuas, equipado de la métrica uniforme;

$$d(\phi_1, \phi_2) = \sup_{t \in I_\alpha} |\phi_1(t) - \phi_2(t)|$$

Para $\phi \in X$, sea $F(\phi) : I_\alpha \rightarrow \mathbb{R}^n$ definida como:

$$F(\phi)(t) = x_0 + \int_{t_0}^t f(s, \phi(s)) ds$$

con $t \in I_\alpha$. Podemos verificar lo siguiente:

- $F(X) \subseteq X$.

Tomando $t \in I_\alpha$, tenemos:

$$\begin{aligned} |F(\phi)(t) - x_0| &= \left| \int_{t_0}^t f(s, \phi(s)) ds \right|, \text{ definición de } F \\ &\leq M\alpha \quad , \text{ } f \text{ está acotado y } |t - t_0| \leq \alpha \\ &\leq b \quad \text{ porque } M\alpha = \min \{Ma, b\} \end{aligned}$$

y así, $F(\phi)(t) \in B_b$, y como $F(\phi)$ es continua, tenemos que $F(X) \subset X$.

- Para n suficientemente grande, $F^n(X)$ es contracción.

Tomando $\phi_1, \phi_2 \in X$, y $n \geq 0$. La idea es probar que:

$$|F^{(n)}(\phi_1)(t) - F^{(n)}(\phi_2)(t)| \leq \frac{K^n(t - t_0)^n}{n!} d(\phi_1, \phi_2)$$

con $t \in I_\alpha$. La prueba es por inducción sobre n . El caso $n = 0$ es trivial. Suponer que se tiene lo deseado para $n = k$, entonces, se prueba para $k + 1$;

$$\begin{aligned}
|F^{(k+1)}(\phi_1)(t) - F^{(k+1)}(\phi_2)(t)| &= |F(F^{(k)}(\phi_1))(t) - F(F^{(k)}(\phi_2))(t)| \\
&\leq \left| \int_{t_0}^t |f(s, F^{(k)}(\phi_1)(s)) - f(s, F^{(k)}(\phi_2)(s))| ds \right| \\
&\leq \left| \int_{t_0}^t K |F^{(k)}(\phi_1)(s) - F^{(k)}(\phi_2)(s)| ds \right| \\
&\leq K \left| \int_{t_0}^t \frac{K^k (t_0 - s)^k}{k!} d(\phi_1, \phi_2) ds \right| \\
&= \frac{K^{k+1} |t - t_0|^{k+1}}{(k+1)!} d(\phi_1, \phi_2)
\end{aligned}$$

donde este último paso se obtiene al realizar la integral. Entonces, ya queda probado que:

$$d(F^{(n)}(\phi_1), F^{(n)}(\phi_2)) \leq \frac{K^n \alpha^n}{n!} d(\phi_1, \phi_2)$$

para n suficientemente grande. Además, también para un n grande, se tendrá que $\frac{K^n \alpha^n}{n!} \leq 1$, porque el término de esta serie debe converger a $e^{K\alpha}$, si fuera mayor a 1, la serie no converge.

Así, queda probado que F^n es contracción de X .

Entonces, por el corolario consecuencia del lema de contracción, existe una función $\phi \in X$ tal que $F(\phi) = \phi$, y de esta forma, queda demostrado lo deseado.

□

En el teorema de Picard, se puede relajar las hipótesis, y se puede retirar la condición de Lipschitz, para obtener otra versión del teorema más general. Para ello, se presentan primero dos teoremas necesarios (Sin demostración) usados en la prueba del teorema general.

Teorema 37 (Teorema de Ascoli-Arzelá.) *Sea (X, d) un espacio métrico compacto. Sea F una familia equicontinua de funciones $\phi : X \rightarrow \mathbb{R}$ (Esto es, para todo $\epsilon > 0$, existe un $\delta > 0$ tal que si $d(x, y) < \delta$, luego $|\phi(x) - \phi(y)| < \epsilon$, para toda $\phi \in F$). Sea F uniformemente acotada (Existe $M > 0$ tal que para todo $\phi \in F$, $|\phi| < M$), entonces, toda sucesión $\{\phi_n\}$ en F , tiene subsucesión uniformemente convergente en X .*

1.2. PRELIMINARES DE INTEGRACIÓN Y ECUACIONES DIFERENCIALES.33

Teorema 38 (Teorema de la aproximación de Weierstrass.) *Para cualquier $\epsilon > 0$ y para cualquier función f continua en un intervalo $[a, b] \subset \mathbb{R}$, existe un polinomio de coeficientes reales p tal que*

$$\sup_{x \in [a, b]} |f(x) - p(x)| < \epsilon$$

Esto es, se puede aproximar una función por una sucesión de polinomios con coeficientes reales, tal que converjan uniformemente.

Teorema 39 (Teorema de Peano.) *Sea F continua en $\Omega = I_a \times B_b$, definidos como en el teorema de Picard. Si $|f| < M$ en Ω , luego la ecuación diferencial:*

$$\frac{dx}{dt} = f(t, x), \quad x(t_0) = x_0$$

tiene al menos una solución en I_α con $\alpha = \min \{a, b/M\}$

Demostración: PENDIENTE.

□

Para finalizar esta sección, se hablará un poco de ecuaciones diferenciales controladas, que también, son parte importante al hablar de caminos rugosos...

Capítulo 2

Procesos Estocásticos y el Movimiento Browniano

En este capítulo, inicialmente se hablará acerca de los procesos estocásticos y se darán algunos conceptos básicos. Luego, se estudiará un proceso estocástico a tiempo continuo, que es muy importante en la teoría de ecuaciones diferenciales estocásticas, que es el *movimiento Browniano*. Primero, se hablará sobre la construcción de este proceso (Usando los teoremas de Kolmogorov), luego se darán algunas propiedades bastante importantes.

2.1. Conceptos de Procesos Estocásticos

En esta sección, se introducirán los conceptos básicos de procesos estocásticos en general, tanto de tiempo discreto como de tiempo continuo.

Un **proceso estocástico** corresponde a la colección $\{X_t : \Omega \rightarrow S\}_{t \in T}$ de variables aleatorias definidas en un mismo espacio de probabilidad (Ω, \mathcal{F}, P) .

A S se le conoce como **espacio de estados** y T se le conoce como el tiempo.

Si T es un conjunto finito o contable, entonces el proceso corresponde a un *proceso a tiempo discreto*. De otro modo, se dice que se tiene un *proceso a tiempo continuo*.

Ejemplo: El *movimiento Browniano*, posee un espacio de estados continuo, $S = \mathbb{R}^n$, y también tiempo continuo, $T = [0, \infty)$. Otro ejemplo es el *paso aleatorio simple*, tal que su espacio de estados es discreto, $S = \{s_1, s_2, \dots\}$ (Se puede interpretar los vértices en un grafo que se pueden visitar), y además, también tiene tiempos discretos, por ejemplo, $T = \{t_1, t_2, \dots\}$.

□

Mayoritariamente, en procesos estocásticos, uno se interesa más en las distribuciones conjuntas de las variables aleatorias. Esto motiva la siguiente definición:

Las distribuciones conjuntas de $(X_{t_1}, X_{t_2}, \dots)$ son llamadas **distribuciones finito-dimensionales** del proceso $\{X_t\}_{t \in T}$.

Dado $\{X_t\}_{t \in T}$ un proceso con espacio de estados S definido en (Ω, \mathcal{F}, P) . Para cada $w \in \Omega$, se define como la **trayectoria**, a la función:

$$\begin{aligned} X(w) : T &\rightarrow S \\ t &\mapsto X_t(w) \end{aligned}$$

Note que, hay una equivalencia, entre hablar una probabilidad μ_X sobre el conjunto de las trayectorias $X : T \rightarrow S$, y una distribución conjunta de todos los tiempos $t \in T$ para el proceso $\{X_t\}_{t \in T}$.

Un proceso estocástico $\{X_t\}_{t \in T}$ en (Ω, \mathcal{F}) se llama **conjuntamente medible** si:

$$\begin{aligned} X : T \times \Omega &\rightarrow \mathbb{R} \\ (t, w) &\mapsto X_t(w) \end{aligned}$$

es medible respecto a la σ -álgebra producto $\mathcal{B}(T) \otimes \mathcal{F}$.

Dados $\{X_t\}_{t \in T}$ y $\{Y_t\}_{t \in T}$ en un mismo espacio de probabilidad (Ω, \mathcal{F}, P) .

- Se dice que $\{Y_t\}$ es **versión** de $\{X_t\}$ si:

$$P(X_t = Y_t) = P(\{w \in \Omega | X_t(w) = Y_t(w)\}) = 1$$

para todo $t \in T$, esto es para todo tiempo fijo.

- Se dice que estos dos procesos son **indistinguibles** si:

$$P[X_t = Y_t \text{ para todo } t \in T] = P[\{w \in \Omega | X_t(w) = Y_t(w) \text{ para todo } t \in T\}] = 1$$

Note que la segunda definición es mucho más fuerte que la primera definición, porque dos procesos *indistinguibles* tendrás trayectorias iguales (Salvo en un conjunto de puntos de medida cero, como un conjunto de puntos discreto).

Una familia $\{\mathcal{F}_t\}$ de sub- σ -álgebras de \mathcal{F} se dice **filtración** en (Ω, \mathcal{F}) si, para $t_1 \leq t_2$, $\mathcal{F}_{t_1} \subset \mathcal{F}_{t_2}$, esto es, una familia creciente de sub- σ -álgebras.

Dado $\{X_t\}_{t \in T}$ un proceso estocástico, se le llamará **proceso \mathcal{F}_\square -adaptado** si para todo $t \in T$, X_t es medible respecto a \mathcal{F}_t , para todo t .

Si se toma como $\mathcal{F}_t^X = \sigma(\{X_s\}_{s < t})$, esta es la menor filtración tal que $\{X_t\}$ es un proceso adaptado a esta filtración, conocida como **filtración natural de $\{X_t\}_{t \in T}$** .

Una variable aleatoria τ en (Ω, \mathcal{F}, P) se le conoce como **tiempo de parada** respecto a la filtración $\{\mathcal{F}_t\}_{t \in T}$ si:

- $P[\tau < \infty] = 1$
- $\{\tau \leq t\} = \{w \in \Omega | \tau(w) \leq t\} \in \mathcal{F}_t$

2.2. Movimiento Browniano

Primero, se da la definición de un movimiento Browniano, y luego, se hará la construcción. Esta, se puede hacer por tres métodos distintas (Consulte [3]):

1. Teoremas de existencia y continuidad de Kolmogorov.
2. Construcción propuesta por Levý, Wiener y Ciesielski, que usa fuertemente conceptos de espacios de Hilbert, aprovechando que el movimiento Browniano es Gaussiano. Usando funciones de Haar.
3. Principio de invarianza de Donsker, donde se busca una convergencia débil a una medida de Wiener.

En el presente trabajo, se enunciarán los teoremas de Kolmogorov, y en el apéndice se discutirán brevemente las demás construcciones.

Dado $\{W_t\}$ un proceso estocástico, en el espacio de probabilidad (Ω, \mathcal{F}, P) . El proceso $\{W_t\}$ es un **movimiento Browniano** en una dimensión, si se cumplen las siguientes condiciones:

- Para casi todo ω , los caminos $W_t(\omega)$ son continuos (En el sentido de la probabilidad).
- $\{W_t\}$ es un proceso Gaussiano, es decir, para $k \geq 1$, y todo $0 \leq t_1 \leq \dots \leq t_k$, el vector aleatorio, $Z = (W_{t_1}, \dots, W_{t_k}) \in \mathbb{R}^n$ tiene distribución multinormal (O Gaussiana), con media el vector 0, y la matriz de covarianza como $B(t_i, t_j) = \mathbb{E}[W_{t_i} W_{t_j}] = \min(t_i, t_j)$.

Más general, podemos considerar el movimiento Browniano respecto a una filtración.

Un proceso W_t en un espacio de probabilidad (Ω, \mathcal{F}, P) adaptado a una filtración $(\mathcal{F}_t)_{t \geq 0}$ es un **movimiento Browniano relativo a la filtración \mathcal{F}_t** , si:

- Los caminos $W_t(\omega)$ son trayectorias continuas de t para casi todo ω .
- $W_0(\omega) = 0$ para casi todo ω .
- Para $0 \leq s \leq t$, los incrementos $W_t - W_s$ son variables aleatorias Gaussianas con media 0 y varianza $t - s$.
- Para $0 \leq s \leq t$, los incrementos $W_t - W_s$ son independientes de las σ -álgebras \mathcal{F}_s .

Una pregunta que natural, es acerca de la existencia y de la unicidad de este proceso. Kolmogorov propone dos teoremas para verificar que el proceso existe y es único (En el sentido de distribuciones). Antes, se define el concepto de familia *consistente*.

Sea T un conjunto de sucesiones finitas $\tilde{t} = (t_1, \dots, t_n)$ de números positivos distintos. Suponga que para cada \tilde{t} de longitud n , existe una medida de probabilidad $Q_{\tilde{t}}$ en $(\mathbb{R}, \mathcal{B}(\mathbb{R}^n))$. Luego, $\{Q_{\tilde{t}}\}_{\tilde{t} \in T}$ es una **familia de distribuciones finito-dimensionales**. La familia se dice **consistente** si cumple las siguientes condiciones:

- Si $\tilde{s} = (t_{i_1}, \dots, t_{i_n})$ es una permutación de $\tilde{t} = (t_1, \dots, t_n)$, luego para todo $A_i \in \mathcal{B}(\mathbb{R})$ con $i = 1, \dots, n$ tenemos:

$$Q_{\tilde{t}}(A_1 \times \dots \times A_n) = Q_{\tilde{s}}(A_{i_1} \times \dots \times A_{i_n})$$

esto es, es invariante bajo permutaciones.

- Si $\tilde{t} = (t_1, \dots, t_n)$ con $n \geq 1$, y $\tilde{s} = (t_1, \dots, t_{n-1})$, con $A \in \mathcal{B}(\mathbb{R}^{n-1})$, luego:

$$Q_{\tilde{t}}(A \times \mathbb{R}) = Q_{\tilde{s}}(A)$$

Note que, al tener un espacio de probabilidad, $(\mathbb{R}^{[0,\infty)}, \mathcal{B}(\mathbb{R}^{[0,\infty)}), P)$, se puede definir una familia de distribuciones finito-dimensionales:

$$Q_{\tilde{t}}(A) = P[\{w \in \mathbb{R}^{[0,\infty)} | (w(t_1), \dots, w(t_n)) \in A\}]$$

Para construir el movimiento Browniano, se usará el hecho, de tener las distribuciones finito-dimensionales, para construir una medida de probabilidad en el espacio, como primer paso. Para ello, se usa el siguiente teorema:

Teorema 40 (Teorema de consistencia de Kolmogorov y Daniell.) *Sea $\{Q_{\tilde{t}}\}$ una familia de distribuciones finito-dimensionales consistentes. Así, existe una medida de probabilidad P en $(\mathbb{R}^{[0,\infty)}, \mathcal{B}(\mathbb{R}^{[0,\infty)}))$ tal que:*

$$Q_{\tilde{t}}(A) = P[\{w \in \mathbb{R}^{[0,\infty)} | (w(t_1), w(t_2), \dots, w(t_n)) \in A\}]$$

se cumpla para todo $\tilde{t} = \{t_1, t_2, \dots, t_n\} \in T$.

Como consecuencia, se tiene el siguiente resultado:

Corolario 3 *Existe una medida de probabilidad P en $(\mathbb{R}^{[0,\infty)}, \mathcal{B}(\mathbb{R}^{[0,\infty)}))$ tal que el proceso:*

$$B_t(w) = w(t), \quad w \in \mathbb{R}^{[0,\infty)}, t \geq 0$$

tiene incrementos estacionarios e independientes. Además, los incrementos $B_t - B_s$, con $0 \leq s \leq t$, es normalmente distribuido con media 0 y varianza $t - s$,

Ya se tiene una medida de probabilidad. Ahora, se debe construir el proceso en todo el espacio $\mathbb{R}^{[0,\infty)}$. Sin embargo, hay que notar que $C[0, \infty) \notin \mathcal{B}(\mathbb{R}^{[0,\infty)})$, y por ende, se debe construir una modificación continua del proceso. Se tiene entonces, el segundo teorema de Kolmogorov.

Teorema 41 (Teorema de continuidad de Kolmogorov y Chenstov.) *Suponga que un proceso $X = \{X_t | 0 \leq t \leq T\}$ en un espacio de probabilidad (Ω, \mathcal{F}, P) cumple la condición:*

$$\mathbb{E}[|X_t - X_s|^\alpha] \leq C|t - s|^{1+\beta}$$

con $0 \leq s, t \leq T$, para constantes positivas α , β y C . Entonces, existe una modificación $\tilde{X} = \{\tilde{X}_t | 0 \leq t \leq T\}$ de X , que es localmente Hölder-continua, con exponente γ , para todo $\gamma \in (0, \beta/\alpha)$, esto es:

$$P \left[w \mid \sup_{0 < t-s < h(w), s, t \in [0, T]} \frac{|\tilde{X}_t(w) - \tilde{X}_s(w)|}{|t-s|^\gamma} \leq \delta \right] = 1$$

donde $h(w)$ es una variable aleatoria positiva casi siempre y $\delta > 0$ una constante apropiada.

Como resultado final, tendremos:

Corolario 4 *Hay una medida de probabilidad P en $(\mathbb{R}^{[0, \infty)}, \mathcal{B}(\mathbb{R}^{[0, \infty)}))$ y un proceso estocástico $W = \{W_t, \mathcal{F}_t^W \mid t \leq 0\}$ en el mismo espacio, tal que W es un movimiento Browniano respecto a P .*

Como consecuencia, cada trayectoria del movimiento Browniano $\{W_t(w) \mid 0 \leq t < \infty\}$ es localmente Hölder-continua con exponente γ , para $\gamma \in (0, \frac{1}{2})$.

Además de eso, el movimiento Browniano tiene varias características y propiedades:

Capítulo 3

Ecuaciones diferenciales estocásticas: Enfoque de Itô.

En este capítulo, se comenzará la construcción de la integral de Itô. Para cumplir este objetivo, usaremos fuertemente los hechos vistos en el capítulo anterior del movimiento Browniano. Pero antes, se verán algunos ejemplos de ecuaciones diferenciales estocásticas, que poseen cierto ruido.

1. Se puede considerar un modelo de crecimiento poblacional, sujeto a un ruido:

$$\frac{dN}{dt} = (r(t) + \text{Ruido})N(t)$$

con la condición inicial $N(0) = N_0$. El término $(r(t) + \text{Ruido})$ corresponde a la tasa de crecimiento, esta puede variar en función de una distribución de probabilidad.

2. Otro problema conocido, es el **problema de filtrado**. Suponga que se hacen observaciones $Z(s)$ de, por ejemplo, la carga eléctrica en cierto medio, cuya función es denotada por $Q(s)$, para $s \leq t$. Se puede ver la relación entre ambas funciones como:

$$Z(s) = Q(s) + \text{Ruido}$$

donde el ruido proviene de errores al realizar las mediciones. Realmente, uno obtiene una versión perturbada de $Q(s)$. El problema consiste

en determinar: ¿Cuál es la mejor estimación de $Q(t)$ que cumple la relación dada, basado en las observaciones $Z(s)$?. En otras palabras, se desea filtrar el ruido, y obtener la medición más precisa. Este problema se puede plantear como una ecuación diferencial lineal con ruido.

3. En matemáticas financiera, es muy común el uso de estas ecuaciones diferenciales para modelar diferentes fenómenos financieros. Por ejemplo, suponga que una persona tiene dos posibles inversiones:

- a) Una inversión riesgosa, donde el precio $p_1(t)$ satisface la ecuación diferencial estocástica (Modelo poblacional estocástico):

$$\frac{dp_1}{dt} = (a + \alpha \cdot \text{Ruido})p_1$$

donde $a > 0$ y $\alpha \in \mathbb{R}$.

- b) Una inversión segura, donde el precio $p_2(t)$ crece exponencialmente:

$$\frac{dp_2}{dt} = bp_2$$

donde $0 < b < a$.

La persona, en cada instante t desea saber qué porcentaje u_t de su fortuna, sea X_t , colocar en la inversión riesgosa. Así, coloca $(1 - u_t)X_t$ de su fortuna en la inversión segura. Dada una *función de utilidad* U y un tiempo final T , se desea hallar la *cartera* óptima $u_t \in [0, 1]$ que maximizar la utilidad de la fortuna terminal, esto es:

$$\max_{0 \leq u_t \leq 1} \left\{ \mathbb{E}[U(X_T^{(n)})] \right\}$$

Si se supone que la inversión riesgosa, es una *opción de llamada Europea*, se deriva la famosa ecuación de **Black-Scholes**.

Con estos ejemplos, ya se sabe el por qué es necesario, en algunos casos, recurrir a estas ecuaciones diferenciales con ruido. En este capítulo, siguiendo el tratamiento de [5], se verá un enfoque *probabilístico* para solucionar estas ecuaciones, y en el próximo capítulo, se verán algunas desventajas de usar este método, proponiendo un método más analítico.

3.1. La integral de Itô.

Ahora, ¿Por qué es necesario construir una nueva integral? Veamos el objetivo inicial, solucionar una ecuación diferencial que tiene cierto ruido:

$$\frac{dX}{dt} = b(t, X_t) + \sigma(t, X_t) \cdot W_t$$

Note que el ruido se puede representar como el proceso estocástico W_t . Bajo experimentación, se interponen las siguientes condiciones sobre el ruido:

- Dos variables del proceso W_{t_1} y W_{t_2} con $t_1 \neq t_2$ son independientes.
- $\{W_t\}$ es un proceso estacionario.
- $\mathbb{E}[W_t] = 0$ para todo t .

No hay algún proceso estocástico tradicional que cumpla las condiciones dadas. Por ende, lo podemos ver como un proceso estocástico generalizado, un **proceso de ruido blanco**, esto es, un proceso que se puede construir como medida de probabilidad en cierto espacio sutil de funcionales $\mathcal{C}[0, \infty)$. Por ende, se nos sugiere que el proceso $\{W_t\}$ será el movimiento Browniano.

Ahora, la ecuación diferencial estocástica se puede representar forma integral, usando el hecho que W_t será el movimiento Browniano, como:

$$X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) \cdot B_s$$

Sin embargo, el inconveniente se presenta con la segunda integral de la derecha, cuyo integrador no es de variación acotada, y por ende, no existe la *integral de Riemann-Stieljes* en este caso:

$$\int_0^t f(s, w) dB_s(w)$$

Por ende, se debe realizar una construcción nueva. La idea será análoga a construir la integral de Lebesgue (Consulte [2]). Esto es, definir la integral para un conjunto de funciones simples, luego tomar sucesión de funciones simples uniformemente convergente a una función positiva (Donde ya se conoce que existe esta sucesión), y luego ver el caso para funciones en general.

Sea f de la forma:

$$\phi(t, w) = \sum_{j \geq 0} e_j(w) \chi_{[j \cdot 2^{-n}, (j+1) \cdot 2^{-n}]}(t)$$

para $n \in \mathbb{N}$. Entonces, se define la integral en este caso como:

$$\int_S^T \phi(t, w) dB_t(w) = \sum_{j \geq 0} e_j(w) [B_{t_{j+1}} - B_{t_j}](t)$$

donde

$$t_k = t_k^{(n)} = \begin{cases} k \cdot 2^{-n} & \text{Si } S \leq k \cdot 2^{-n} \leq T \\ S & \text{Si } k \cdot 2^{-n} < S \\ T & \text{Si } T < k \cdot 2^{-n} \end{cases}$$

pero, si no hay condiciones adicionales sobre e_j , habrán ciertos problemas. Suponga, que tenemos:

$$\phi_1(t, w) = \sum_{j \geq 0} B_{j \cdot 2^{-n}}(w) \cdot \chi_{[j \cdot 2^{-n}, (j+1) \cdot 2^{-n}]}(t)$$

y

$$\phi_2(t, w) = \sum_{j \geq 0} B_{(j+1) \cdot 2^{-n}}(w) \cdot \chi_{[j \cdot 2^{-n}, (j+1) \cdot 2^{-n}]}(t)$$

Capítulo 4

Camino Rugoso e Integración Rugosa

La teoría de caminos rugosos (Vea, por ejemplo [4]) permite extender una teoría de ecuaciones diferenciales controladas, para poder trabajar el caso al tener una señal de entrada que sea *ruidosa*, tal como una *semimartingala*, como lo es el movimiento Browniano. Sea una ecuación diferencial controlada:

$$dY_t = f(Y_t)dX_t, \quad Y_0 = \zeta$$

Si la ecuación determinística admite solución única, entonces se denota como $Y = I_f(X, \zeta)$, y se conoce a I_f como **mapeo de Itô asociado a f** . Podemos expresar, de forma *integral* la solución de esta ecuación diferencial como:

$$Y_t = Y_0 + \int_0^t f(Y_s)dX_s$$

Y, el problema de solucionar la ecuación diferencial, implica en definir la integral:

$$\int_0^t f(Y_s)dX_s$$

Recordando el Capítulo 1, se sabe que el integrador X_s debe, en principio, cumplir algunas condiciones de regularidad (Que sea α -Hölder, con $\alpha \geq \frac{1}{2}$) para permitir que exista su integral de Young (Consulte teorema 35).

Sin embargo, suponga que X_t es un camino mucho menos regular, como las trayectorias de un movimiento Browniano, que son α -Hölder con $\alpha \in (\frac{1}{3}, \frac{1}{2}]$. Se puede usar el enfoque de integración estocástica propuesto por Itô, pero en este caso, el mapa de Itô, carecerá de continuidad. Además, la integral de Young, no estará bien definida en este caso, por lo que dependerá de la elección de la partición, y más aún, el límite puede no existir. También, otro problema de la integral de Itô, es acerca de la elección de puntos, que puede afectar el valor de la integral, puesto que, las trayectorias tienen una variación muy rápida, y una integral como la de Riemann-Stieljes no es capaz de capturar esta información y más en intervalos de tiempos pequeños.

Se puede dar otro enfoque al problema. Sea $f : \mathbb{R}^d \rightarrow \mathbb{R}$ una función suave, $X : [0, T] \rightarrow \mathbb{R}^d$ un camino α -Hölder continuo. Suponga que se desea darle un significado a

$$\int_0^T f(X_r) dX_r$$

Para eso, se usará la expansión de Taylor. Tomando $[s, t] \subset [0, T]$ un intervalo de tiempo lo suficientemente pequeño y $r \in [s, t]$, tenemos:

$$f(X_r) = f(X_s) + \nabla f(X_s)(X_r - X_s) + \dots$$

Integrando respecto al camino X , obtenemos:

$$\int_s^t f(X_r) dX_r = f(X_s)(X_t - X_s) + \nabla f(X_s) \int_s^t (X_r - X_s) \otimes dX_r + \dots$$

Véase el apéndice XX para una introducción a los tensores. Dado que $\alpha > \frac{1}{3}$, se pueden omitir los términos de grado superior en la expansión. Ahora, suponiendo que $\alpha > \frac{1}{2}$, por la condición de Young, teorema 35, se puede probar que:

$$\lim_{|\pi| \rightarrow 0} \sum_{[s,t] \in \pi} \int_s^t (X_r - X_s) \otimes dX_r = 0$$

Sin embargo, al tomar $\alpha \leq \frac{1}{2}$, el término de segundo orden, no necesariamente se anula, por lo que quedará:

$$\begin{aligned}
\int_0^T f(X_r) dX_r &= \lim_{[s,t] \in \pi} \int_s^t f(X_r) dX_r \\
&= \lim_{[s,t] \in \pi} \left(f(X_s)(X_t - X_s) + \nabla f(X_s) \int_s^t (X_r - X_s) \otimes dX_r \right)
\end{aligned}$$

Note que, se le debe dar un significado a la integral, al término de segundo orden. Este se conocerá como **levantamiento** de X , que será un tipo de candidato para el valor de la integral. Este levantamiento se expresa como:

$$\int_s^t (X_r - X_s) \otimes dX_r := \mathbb{X}_{s,t}$$

Y vea que la integral se define como el valor propuesto para el levantamiento (No al contrario, como se podría pensar). Mientras se define qué es un camino rugoso, puede pensar a \mathbb{X} como mayor información codificada por X .

Con esto en mente, la idea es, obtener un camino rugoso para X , de tal manera, que se puede definir la integral $\int f(X) \otimes X$ como una *integral rugosa*, integrando respecto al camino rugoso (X, \mathbb{X}) . Con ello, el mapeo solución $(X, \mathbb{X}) \mapsto Y$ será continuo en una topología sutil.

Entonces, resolver ecuaciones diferenciales rugosas, implicará hallar dos funciones:

$$X \mapsto (X, \mathbb{X}) \mapsto Y$$

donde el primer mapeo consiste en agregar más información a X , y el segundo mapeo, conocido como **mapa de Itô-Lyons**, va a la solución del problema. Dado el levantamiento, este mapeo será continuo, e inclusive, en algunos casos, será localmente Lipschitz.

Ya con esta idea acerca de caminos rugosos, se puede ver algunas ventajas al trabajar con este enfoque. En el capítulo, antes de pasar a la definición formal de caminos rugosos, se hablará de caminos α -Hölder y algunas de sus propiedades. Con esto, ya se hablará de caminos rugosos, algunas propiedades, y también se verá algunos ejemplos, como lo es el *movimiento Browniano*.

4.1. Caminos α -Hölder

Sea $\alpha \in (0, 1]$. Recuerde que una trayectoria $X : [0, T] \rightarrow \mathbb{R}^d$ es **α -Hölder continua**, si se cumple que:

$$|X_t - X_s| \leq C|t - s|^\alpha$$

para $s, t \in [0, T]$ con $s < t$.

Para $\alpha \in (0, 1]$, defina una **seminorma α -Hölder** de X , como:

$$\|X\|_\alpha = \sup_{0 \leq s < t \leq T} \frac{|X_{s,t}|}{|t - s|^\alpha}$$

Si $\|X\|_\alpha < \infty$, el camino se denomina **α -Hölder continuo**. El espacio de los caminos α -Hölder continuo se denota por $\mathcal{C}^\alpha = \mathcal{C}^\alpha([0, T]; \mathbb{R}^d)$

4.2. Caminos Rugosos y el movimiento Browniano.

Recuerde, que la integral de Itô está definida como:

$$\int_s^t B_{s,r} dB_r$$

esto es, tomando el punto izquierdo. En esta sección, se va a ver que el movimiento Browniano, en conjunto con esta integral, conforman un camino rugoso para $\alpha \in [\frac{1}{3}, \frac{1}{2}]$. Con esto, se puede aplicar la teoría de caminos rugosos e integración rugosa, a la integración estocástica, lo que abre un gran mundo de posibilidades.

En primer lugar, se va a demostrar el teorema fuerte de esta sección, que corresponde al *criterio de Kolmogorov de caminos rugosos*, el cuál, parafraseando, mostrará la existencia de una modificación para ciertos procesos estocásticos medibles, cuya modificación será un camino rugoso. Más específicamente, el enunciado dice:

Teorema 42 (Criterio de Kolmogorov para caminos rugosos) Sea $(X, \mathbb{X}) : \Omega \times [0, T] \rightarrow \mathbb{R}^d \times \mathbb{R}^{d \times d}$ un proceso estocástico medible (Respecto a la σ -álgebra producto $\mathcal{F} \otimes \mathcal{B}[0, T]$) que, para casi todo $\omega \in \Omega$, satisface la relación de

Chen. Sea $q \geq 2$ y $\beta > \frac{1}{q}$. Suponga que existe una constante $C > 0$, tal que, para todo $(s, t) \in \Delta_{[0, T]}$,

$$\|X_{s,t}\|_{L^q} \leq C|t-s|^\beta, \quad \|\mathbb{X}_{s,t}\|_{L^{q/2}} \leq C|t-s|^{2\beta}$$

(Esta elección que constantes q y β se hace para concuerde con la definición de camino rugoso, que X sea β -Hölder y el levantamiento \mathbb{X} sea 2β -Hölder).

Entonces, para todo $\alpha \in [0, \beta - \frac{1}{q}]$, existe una modificación $(\tilde{X}, \tilde{\mathbb{X}})$ de (X, \mathbb{X}) y también existen variables aleatorias $K_\alpha \in L^q$, $\mathbb{K}_\alpha \in L^{q/2}$ tal que, para todo $(s, t) \in \Delta_{[0, T]}$,

$$|\tilde{X}_{s,t}| \leq K_\alpha |t-s|^\alpha, \quad |\tilde{\mathbb{X}}_{s,t}| \leq \mathbb{K}_\alpha |t-s|^{2\alpha}$$

(Esta relación es para casi toda $\omega \in \Omega$ ¿?). En particular, si $\beta - \frac{1}{q} > \frac{1}{3}$, entonces, para todo $\alpha \in (\frac{1}{3}, \beta - \frac{1}{q})$, tenemos que $(\tilde{X}, \tilde{\mathbb{X}})$

En primer lugar, se puede ver el proceso estocástico como la siguiente función:

$$Y_t(\omega) = (X_{s,t}(\omega), \mathbb{X}_{s,t}(\omega))$$

donde $\mathbb{X}_{s,t}(\omega) = \int_s^t X_{s,r}(\omega) \otimes dX_r$, tal que para casi todo $\omega \in \Omega$, se cumple que:

$$\mathbb{X}_{s,t}(\omega) = \mathbb{X}_{s,u}(\omega) + \mathbb{X}_{u,t}(\omega) + X_{s,u}(\omega) \otimes X_{u,t}(\omega)$$

Esto es, se cumple para casi todo ω la relación del Chen.

Demostración: En este caso, se probará primero el enunciado, para el conjunto de particiones diádicas, $\mathcal{D}_n = \{ \frac{k}{2^n} | k = 0, 1, \dots, 2^n - 1 \}$. Además, se puede, sin pérdida de generalidad, tomar $T = 1$. Defina, así:

$$K_n = \max_{t \in \mathcal{D}_n} |X_{t,t+2^{-n}}| \quad \mathbb{K}_n = \max_{t \in \mathcal{D}_n} |\mathbb{X}_{t,t+2^{-n}}|$$

(Este corresponde a la modificación propuesta!)

Echando cuentas, tenemos:

$$\begin{aligned}
\mathbb{E}[K_n^q] &= \mathbb{E} \left[\max_{t \in \mathcal{D}_n} |X_{t,t+2^{-n}}|^q \right] && \text{Sumar sobre los otros términos positivos de } \mathcal{D}_n \\
&\leq \mathbb{E} \left[\sum_{t \in \mathcal{D}_n} |X_{t,t+2^{-n}}|^q \right] && \text{Aplicar ¿DT?} \\
&= \sum_{t \in \mathcal{D}_n} \mathbb{E}[|X_{t,t+2^{-n}}|^q] && \text{Recuerde } \|X_{s,t}\|_{L^q} = \mathbb{E}[|X_{s,t}|^q]^{1/q} \\
&= \sum_{t \in \mathcal{D}_n} \|X_{t,t+2^{-n}}\|^q && \text{Use las hipótesis acerca de la norma} \\
&\leq \sum_{t \in \mathcal{D}_n} C^q |t - (t - 2^{-n})|^{q\beta} && \text{Hay } 2^n \text{ elementos en la partición} \\
&= C^q |2^{-n}|^{q\beta} \cdot 2^n \\
&= C^q 2^{-n\beta q + n} = C^q 2^{n(-\beta q + 1)}
\end{aligned}$$

Puede repetir esta cuenta con el levantamiento \mathbb{K}_n , ¡Vamos a ello!:

$$\begin{aligned}
\mathbb{E}[\mathbb{K}_n^{q/2}] &= \mathbb{E}[\max_{t \in \mathcal{D}_n} |\mathbb{X}_{t,t+2^{-n}}|^{q/2}] && \text{Sumar sobre los otros términos } \mathcal{D}_n \\
&\leq \mathbb{E} \left[\sum_{t \in \mathcal{D}_n} |\mathbb{X}_{t,t+2^{-n}}|^{q/2} \right] \\
&= \sum_{t \in \mathcal{D}_n} \mathbb{E}[|\mathbb{X}_{t,t+2^{-n}}|^{q/2}] && \text{Definición de norma} \\
&= \sum_{t \in \mathcal{D}_n} \|\mathbb{X}_{t,t+2^{-n}}\|_{L^{q/2}}^{q/2} && \text{Aplicar hipótesis, cota para } \mathbb{X} \\
&\leq \sum_{t \in \mathcal{D}_n} C^{q/2} 2^{-n \cdot q/2 \cdot 2\beta} \\
&= C^{q/2} 2^{-nq\beta + n} = C^{q/2} 2^{-n(q\beta - 1)}
\end{aligned}$$

Ya tenemos cotas para los valores esperados de K_n^q y \mathbb{K}_n^q ! ¿Y ahora? La idea es, con estas cotas, mostrar que $K_\alpha \in L^p$ y $\mathbb{K}_\alpha \in L^{(p/2)}$ (**VEA, EN EL APÉNDICE, ESPACIOS L^p**).

Dados $s < t$, elementos tomados de $\cup_{n \leq 0} \mathcal{D}_n$ (La unión de todos los elementos en la partición diádica, esto es, $t = \frac{k}{2^n}$ para algún k, n). Seleccione algún m , de tal forma que

$$2^{-(m+1)} < t - s < 2^{-m}$$

Para cada $n \geq m + 1$, tome $u_n, v_n \in \mathcal{D}_n$, tal que $u_n < v_n$, $[u_n, v_n] \subset [s, t]$ y $\cup [u_n, v_n] = [s, t]$, esto es, una partición. Note que, no es posible que existan 3 subintervalos con el mismo tamaño, ¿Por qué? Graficando esta situación tenemos:

En este caso, se aplicó un argumento de descomposición multiescala: Argumento de encadenamiento. Luego, tenemos

$$|X_{s,t}| \leq \max_{0 \leq i < N} |X_{s,u_{i+1}}| \leq \sum_{i=0}^{N-1} |X_{u_i, u_{i+1}}| \leq 2 \sum_{n=m+1}^{\infty} K_n$$

Y de manera análoga,

$$\begin{aligned} |\mathbb{X}_{s,t}| &= \left| \sum_{i=0}^{N-1} (\mathbb{X}_{u_i, u_{i+1}}) + X_{s, u_i} \otimes X_{u_i, u_{i+1}} \right| && \text{Aplique desigualdad triangular dos veces} \\ &\leq \sum_{i=0}^{N-1} |\mathbb{X}_{u_i, u_{i+1}}| + |X_{s, u_i}| |X_{u_i, u_{i+1}}| && \text{Saque el máximo} \\ &\leq \sum_{i=0}^{N-1} |\mathbb{X}_{u_i, u_{i+1}}| + \left(\max_{0 \leq i < N} |X_{s, u_i}| \right) \sum_{i=0}^{N-1} |X_{u_i, u_{i+1}}| && \text{Use las consecuencias de lo anterior} \\ &= 2 \sum_{i=m+1}^{\infty} K_n + 2 \sum_{i=m+1}^{\infty} K_n \cdot 2 \sum_{i=m+1}^{\infty} K_n \\ &= 2 \sum_{i=m+1}^{\infty} K_n + \left(2 \sum_{i=m+1}^{\infty} K_n \right)^2 \end{aligned}$$

Con esto, obtener otra cota para la norma α -Hólder de X :

$$\begin{aligned}
\frac{|X_{s,t}|}{|t-s|^\alpha} &\leq 2 \sum_{n=m+1}^{\infty} \frac{K_n}{2^{-(m+1)\alpha}} \quad |t-s| \geq 2^{-(m+1)} \\
&\leq 2 \sum_{n=m+1}^{\infty} \frac{K_n}{2^{-n\alpha}} \quad \text{Porque } 2^{-n} < 2^{-(m+1)} \\
&\leq 2 \sum_{n=0}^{\infty} \frac{K_n}{2^{-n\alpha}} =: K_\alpha \quad \text{Porque } 2^{-n} < 2^{-(m+1)}
\end{aligned}$$

Esta es la variable aleatoria deseada. Observe, finalmente que:

$$\begin{aligned}
\|K_\alpha\|_{L^q} &= \left\| 2 \sum_{n=0}^{\infty} \frac{K_n}{2^{-n\alpha}} \right\| \quad \text{Aplicar desigualdad triangular} \\
&\leq 2 \sum_{n=0}^{\infty} \frac{\|K_n\|_{L^q}}{2^{-n\alpha}} \quad \text{Definición de la norma} \\
&= 2 \sum_{n=0}^{\infty} \frac{\mathbb{E}[K_n^q]^{1/q}}{2^{-n\alpha}} \\
&\leq 2 \sum_{n=0}^{\infty} \frac{(C^q 2^{-n(\beta q - 1)})^{1/q}}{2^{-n\alpha}} \quad \text{Por el resultado anterior. Hagamos más cuentas!} \\
&= 2 \sum_{n=0}^{\infty} \frac{C \cdot 2^{-n(\beta - \frac{1}{q})}}{2^{-n\alpha}} \\
&= 2C \sum_{n=0}^{\infty} 2^{-n(\beta - \frac{1}{q} - \alpha)} < \infty
\end{aligned}$$

Este último es, debido a que tendremos una serie geométrica. Note que, por hipótesis, $\beta > \frac{1}{q}$ o $\beta - \frac{1}{q} > 0$. Como $\alpha \in \left[0, \beta - \frac{1}{q}\right)$, $\beta - \frac{1}{q} > \alpha$ o $\beta - \frac{1}{q} - \alpha > 0$, por ende, la serie es convergente.

Con esto, queda probado que $\|K_\alpha\|_{L^q}$. Ahora, queda probar el resultado análogo para la variable aleatoria \mathbb{K}_α .

Apéndice.

Este apéndice de forma temporal, tendrá todos los apuntes de temas relacionados pero no tan relacionados acerca de la tesis (Teoría de la medida, Análisis Funcional...) conceptos que se usen en cierta demostración o en alguna definición muy específica, pero que no sea necesaria realmente para continuar con el desarrollo de la tesis.

Finalmente, se hará un apéndice "limpio", con los elementos realmente necesarios :).

4.3. Teoría de la Medida

Algunas notas acerca de teoría de la medida, usadas para el presente trabajo de grado.

4.3.1. Funciones L^p y \mathcal{L}^p

Sea (X, \mathcal{A}, μ) un espacio de medida, con $1 < p < \infty$ (p no necesariamente entero). Denotará a:

$$\mathcal{L}^p(X, \mathcal{A}, \mu, \mathbb{R})$$

el conjunto de funciones medibles respecto a \mathcal{A} , tal que $f : X \rightarrow \mathbb{R}$, donde $|f|^p$ es integrable, esto es:

$$\int |f|^p d\mu < \infty$$

Ese conjunto conforma un espacio vectorial sobre \mathbb{R} (¿Por qué?). Más aún, $\mathcal{L}^p(X, \mathcal{A}, \mu, \mathbb{C})$ también es un espacio vectorial sobre \mathbb{C} .

Si $p = \infty$, $\mathcal{L}^p(X, \mathcal{A}, \mu, \mathbb{R})$ es el conjunto de funciones acotadas y medibles respecto a \mathcal{A} (O *esencialmente acotada*).

Podemos definir una **seminorma** en \mathcal{L}^p como:

$$\|f\|_p = \left(\int |f|^p d\mu \right)^{1/p}$$

Falla al ser norma al pedir que $\|x\| = 0 \Leftrightarrow x = 0$. Podemos construir a partir de este espacio, un espacio Banach normado, $L^p(X, \mathcal{A}, \mu)$. Sea $\mathcal{N}^p(X, \mathcal{A}, \mu) \subset \mathcal{L}^p(X, \mathcal{A}, \mu)$ las funciones que cumplen:

$$f \in \mathcal{L}^p(X, \mathcal{A}, \mu) \text{ y } \|f\|_p = 0$$

Esto es, funciones que en casi todo punto se anulan para $1 < p < \infty$. Si $p = \infty$, son funciones acotadas medibles respecto a \mathcal{A} en X , tal que se anulan en casi todo punto localmente. \mathcal{N}^p es subespacio lineal de \mathcal{L}^p . Defina:

$$L^p(X, \mathcal{A}, \mu) = \mathcal{L}^p(X, \mathcal{A}, \mu) / \mathcal{N}^p(X, \mathcal{A}, \mu)$$

Esto es, una colección de cosets, que están definidos por la relación de equivalencia definida por:

$$f \tilde{g} \Leftrightarrow f - g \in \mathcal{N}^p$$

Esto es, si las funciones son iguales en casi todo punto. (Para $p = \infty$, corresponde a las funciones iguales en casi todo punto localmente). En este caso, $\|\cdot\|_p$ es una norma en $L^p(X, \mathcal{A}, \mu)$.

Consulte: [https://en.wikipedia.org/wiki/Quotient_space_\(linear_algebra\)](https://en.wikipedia.org/wiki/Quotient_space_(linear_algebra)).

Espacios de Hilbert.

Bibliografía

- [1] Tom M. Apostol, *Análisis matemática*, Editorial Reverté, 1976.
- [2] Donald L. Cohn, *Measure theory*, Springer-Birkhäuser, 2013.
- [3] Ioannis Karatzas and Steven Shreve, *Brownian motion and stochastic calculus*, Springer, 2005.
- [4] Terry Lyons and Michael Caruana, *Differential equations driven by rough paths*, Springer, 2007.
- [5] Bernt Oksendal, *Stochastic differential equations*, Springer, 2000.
- [6] Protter Philip and Jean Jacob, *Probability essentials*, Springer-Verlag, 2004.