

¿QUÉ ES EL CÁLCULO ESTOCÁSTICO?

Miguel Ángel García Álvarez

1. INTRODUCCIÓN

El Cálculo Estocástico es en realidad únicamente Cálculo Integral Estocástico, por lo menos en lo que se refiere a la teoría que quedó prácticamente concluída hacia finales de los años 70's del siglo pasado y que en lo que sigue denominaré Cálculo Estocático Clásico. Después esa teoría se fue ampliando y se desarrolló en varias direcciones. Ahora bien, al igual que ocurrió con el Cálculo Diferencial e Integral que conocemos, la teoría de Integración Estocástica se desarrolló para poderla aplicar en la solución de cierto tipo de problemas de interés. De manera específica, el objetivo es poder resolver las llamadas ecuaciones diferenciales estocásticas, las cuales también son en realidad ecuaciones integrales estocásticas. Como veremos más adelante, en el Cálculo Estocástico Clásico no hay derivadas; es decir, no hay un Cálculo Diferencial Estocástico ya que, en general, se trabaja con funciones que no son diferenciables. A su vez, el resolver ecuaciones diferenciales estocásticas tiene por objetivo el poder construir procesos estocásticos que modelen fenómenos de interés en diversas áreas. Una de las aplicaciones más desarrolladas es la referente a las Finanzas, pero actualmente las aplicaciones están más diversificadas. Por otra parte, si bien, como lo mencioné antes, la teoría del Cálculo Estocástico Clásico quedó bien establecida en los 70's del siglo pasado, no fue así en lo que se refiere a las ecuaciones diferenciales estocásticas, tema que actualmente aún sigue en desarrollo.

Con estas notas pretendo dar una idea de lo que es el Cálculo Estocástico Clásico y un poco lo que son las ecuaciones diferenciales estocásticas y cómo se aplica la teoría de integración estocástica para resolverlas.

Para entender el concepto de integral estocástica es muy útil tener una idea clara de cómo se fue desarrollando la teoría de integración tal como se enseña actualmente en los cursos de Cálculo (integral de Riemann) y en los de Análisis Matemático (integral de Lebesgue). Esto no únicamente por curiosidad, sino porque las investigaciones alrededor del concepto de integral y sus propiedades condujeron a lo que ahora se llama Teoría de la Medida, la cual es una de las herramientas fundamentales que se utilizan en el Cálculo Estocástico. Para entender a fondo este último es necesario un conocimiento amplio y detallado de la Teoría de la Medida.

El Cálculo Diferencial e Integral fue inventado por Isaac Newton y Gottfried Wilhelm Leibniz a finales del siglo XVII. En su trabajo definieron el concepto de derivada de una función y geoméricamente la interpretaban como la pendiente de la tangente a su gráfica. La integral de una función la vieron como la operación inversa de la derivada y geoméricamente la interpretaban como el área de la región delimitada, en el intervalo de integración, por la gráfica de la función y el eje horizontal.

A medida que la teoría se fue desarrollando se plantearon problemas cada vez más complejos, los cuales hicieron ver la necesidad de definir los conceptos con mayor precisión y de demostrar resultados con métodos analíticos, en lugar de algunos métodos geométricos que se utilizaban.

En particular, la manera en que se trataba con la integral de una función llevó a cuestionamientos acerca de la validez de algunas propiedades que se asumían como válidas. De particular importancia fue el trabajo de Jean-Baptiste Joseph Fourier, publicado en el año

1822 bajo el título *Théorie analytique de la chaleur*. Afirmó ahí que una función arbitraria f (hay que aclarar aquí que en esa época aún no se había llegado al concepto de función que tenemos actualmente) definida y acotada en el intervalo $[-L, L]$, puede representarse mediante una serie trigonométrica de la forma:

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} \left[a_n \cos \frac{n\pi x}{L} + b_n \operatorname{sen} \frac{n\pi x}{L} \right]$$

La demostración de Fourier de esta afirmación consiste básicamente en tratar el desarrollo anterior como una ecuación para la cual tendrían que encontrarse los coeficientes a_0 , a_n y b_n ($n \in \mathbb{N}$) que la hacen válida. Para esto, integrando entre $-L$ y L ambos lados de la expresión, se obtiene:

$$\int_{-L}^L f(x) dx = 2a_0L$$

$$\text{Así que } a_0 = \frac{1}{2L} \int_{-L}^L f(x) dx$$

Ahora, multiplicando por $\cos \frac{n\pi x}{L}$ ambos lados de la expresión e integrando, se obtiene:

$$\int_{-L}^L f(x) \cos \frac{n\pi x}{L} dx = a_nL$$

$$\text{Así que } a_n = \frac{1}{L} \int_{-L}^L f(x) \cos \frac{n\pi x}{L} dx$$

Finalmente, multiplicando por $\operatorname{sen} \frac{n\pi x}{L}$ ambos lados de la expresión e integrando, se obtiene:

$$\int_{-L}^L f(x) \operatorname{sen} \frac{n\pi x}{L} dx = b_nL$$

$$\text{Así que } b_n = \frac{1}{L} \int_{-L}^L f(x) \operatorname{sen} \frac{n\pi x}{L} dx$$

Fourier argumentaba que las integrales que definen los coeficientes a_0 , a_n y b_n están bien definidas pues cada una puede obtenerse mediante el cálculo del área bajo la gráfica de la función correspondiente.

Además de la necesidad de clarificar el concepto de función, la demostración de Fourier planteaba los siguientes tres problemas:

1. Definiendo los coeficientes a_n y b_n como lo hacía Fourier,

¿la serie $\frac{1}{2}a_0 + \sum_{n=1}^{\infty} \left[a_n \cos \frac{n\pi x}{L} + b_n \operatorname{sen} \frac{n\pi x}{L} \right]$ converge a $f(x)$ para cualquier $x \in [-L, L]$?

2. ¿Para qué funciones f , las integrales que definen los coeficientes a_n y b_n ($n \in \mathbb{N}$) están definidas?

3. ¿Se puede integrar término a término una serie de funciones?

En el año 1823 se publicó el libro de Augustin-Louis Cauchy titulado *Résumé des leçons données à l'École Royale Polytechnique sur le calcul infinitésimal*, en el cual trató el problema de la definición de la integral, primero para las funciones continuas y después para funciones con discontinuidades.

En ese trabajo, Cauchy definió el concepto de continuidad básicamente como se conoce actualmente y formuló la definición analítica de la integral de una función continua, demostrando su existencia:

Sea f una función continua en el intervalo $[a, b]$, entonces las sumas:

$$S = \sum_{k=1}^n f(x_{k-1}) (x_k - x_{k-1}),$$

correspondientes a particiones $P = \{a = x_0 < \dots < x_n = b\}$ tienden a un límite cuando los elementos $x_k - x_{k-1}$ se hacen infinitamente pequeños; a ese límite se le llama la integral definida de f y se le denota por $\int_a^b f(x)dx$. Se obtiene el mismo límite si se consideran sumas de la forma $S = \sum_{k=1}^n f[x_{k-1} + \theta_k(x_k - x_{k-1})] (x_k - x_{k-1})$, donde $\theta_k \in [0, 1]$.

Demostró además que si f es una función continua y $F(x) = \int_a^x f(y)dy$, entonces $F'(x_0) = f(x_0)$ para cualquier $x_0 \in (a, b)$.

La integral así definida es conocida actualmente como la integral de Riemann y no como la integral de Cauchy. La razón de esto parece justa pues fue el trabajo de Riemann, publicado en el año 1867, el que dió la pauta para desarrollar una Teoría de Integración, la cual a su vez llevaría más tarde a una Teoría del Contenido y finalmente a la moderna Teoría de la Medida.

2. LA INTEGRAL DE RIEMANN

Georg Friedrich Bernhard Riemann, en un artículo titulado *Sur la possibilité de représenter une fonction par une série trigonométrique*, el cual fue elaborado en 1854 pero publicado en 1867, cambió el enfoque para atacar el problema de la integración de funciones. Cauchy y quienes le siguieron buscaban extender la definición de la integral a funciones tan discontinuas como fuera posible, pero no partiendo de una definición general, sino dando una definición distinta dependiendo del tipo de funciones que se querían integrar. En cambio, Riemann planteó una definición general de la integral para cualquier función y se abocó al problema de caracterizar a las funciones para las cuales esa integral está definida.

Planteaba Riemann: ¿Qué se debe entender por $\int_a^b f(x)dx$?

Consideremos una partición x_0, x_1, \dots, x_n del intervalo $[a, b]$ y definamos $\delta_k = x_k - x_{k-1}$. Si, independientemente de como se elijan las cantidades $\varepsilon_k \in [0, 1]$, las sumas $\sum_{k=1}^n \delta_k f(x_{k-1} + \varepsilon_k \delta_k)$ tienden a un límite cuando todas las cantidades δ_k tienden a cero, a ese límite se le llama el valor de la integral definida $\int_a^b f(x)dx$.

Decía Riemann: “Busquemos ahora la extensión y el límite de la definición precedente y hagámonos esta pregunta: ¿En qué casos una función es susceptible de integración?, ¿en qué casos no lo es?”

Estableció dos criterios, ambos basados en el concepto de oscilación de una función en un intervalo, mediante los cuales aclaró algunas confusiones que había en ese momento con relación a la integrabilidad de una función; en particular, exhibió una función integrable, definida en un intervalo cerrado, para la cual el conjunto de sus discontinuidades es denso

en el intervalo. El ejemplo de Riemann echó abajo una idea que se tenía según la cual para que una función sea integrable se requiere que el conjunto de sus discontinuidades sea topológicamente pequeño.

Definición 1. Sea $f : [a, b] \rightarrow \mathbb{R}$ una función acotada. La diferencia:

$$\sup \{f(x) : x \in [a, b]\} - \inf \{f(x) : x \in [a, b]\}$$

es llamada la *oscilación de f en el intervalo $[a, b]$* .

Definición 2. Sea $f : [a, b] \rightarrow \mathbb{R}$ una función acotada, $x \in (a, b)$ e $(I_n)_{n \in \mathbb{N}}$ una sucesión de intervalos cerrados encajados que contengan a x como punto interior y tales que $\bigcap_{n=1}^{\infty} I_n = \{x\}$; denotemos por O_n a la oscilación de f en el intervalo I_n ; entonces el límite $\lim_{n \rightarrow \infty} O_n$ existe y es independiente de la sucesión particular de intervalos encajados con las propiedades dadas antes. A ese límite se le llama la **oscilación de la función f en el punto x** .

3. TEORÍA DEL CONTENIDO

Después de la publicación del artículo de Riemann, las investigaciones se orientaron hacia la búsqueda de una caracterización de las funciones integrables, basada en el tipo de conjunto que forman sus discontinuidades. Fue así que se llegó a un nuevo concepto, el de contenido cero.

Definición 3. Se dice que un conjunto tiene contenido cero si, para cualquier $\varepsilon > 0$, existe una familia finita de intervalos abiertos cuya unión cubre al conjunto y tales que la suma de sus longitudes es menor que ε .

Una vez definido este concepto se pudo establecer con toda claridad la condición para que una función sea integrable. Axel Harnack demostró en 1881 que **una función es integrable si y sólo si, para cualquier $\sigma > 0$, el conjunto de puntos donde la oscilación de la función es mayor que σ tiene contenido cero.**

El concepto de contenido cero se convertiría desde ese momento en uno clave para la teoría de la integración. Surgiría más adelante en conexión con la teoría de integrales dobles sobre una región E del plano, cuya frontera requiere tener contenido cero para que la integral pueda ser definida.

En ese momento se tuvieron entonces las bases para desarrollar una teoría del contenido, lo cual fue llevado a cabo, sobre todo, por Camille Jordan durante el periodo que va de 1883 a 1892:

Las definiciones y propiedades se establecieron en ese periodo tanto para el caso de subconjuntos de los reales como para subconjuntos del plano, siendo similares en los dos casos. También surgieron en este periodo los conceptos de integral superior e inferior de una función.

Sea A un conjunto acotado de números reales y $[a, b]$ un intervalo que lo contenga. Para cada partición P del intervalo $[a, b]$ sea $\bar{S}(P, A)$ la suma de las longitudes de los subintervalos de P que contienen puntos de A y $\underline{S}(P, A)$ la suma de las longitudes de los subintervalos de

P contenidos en A . Se define entonces el **contenido exterior** de A , $c_e(A)$, y el **contenido interior** de A , $c_i(A)$, mediante las relaciones:

$$c_e(A) = \inf \{ \bar{S}(P, A) : P \text{ es partición del intervalo } [a, b] \}$$

$$c_i(A) = \inf \{ \underline{S}(P, A) : P \text{ es partición del intervalo } [a, b] \}$$

Se dice entonces que A es **Jordan-medible** si $c_e(A) = c_i(A)$ y, en este caso, a esta cantidad común se le llama el **contenido** de A y se le denota por $c(A)$.

Evidentemente todo conjunto de contenido cero es Jordan-medible. También todo intervalo acotado es Jordan-medible y su contenido es igual a su longitud.

Consideremos un intervalo $[a, b]$, entonces la familia de subconjuntos de $[a, b]$ que son Jordan-medibles es cerrada bajo complementos y uniones finitas. Además, si A_1, A_2, \dots, A_n es una familia finita de subconjuntos de $[a, b]$ que son Jordan-medibles y ajenos por parejas, entonces:

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

Se observó también durante ese periodo que la teoría del contenido está íntimamente relacionada con la teoría de integración de Riemann, no únicamente porque la caracterización de la integrabilidad de una función se establece con base en el concepto de contenido cero o porque para integrar sobre una región del plano se requiere que la frontera de ésta tenga contenido cero. La relación resulta bastante más profunda, a tal grado que puede decirse que constituyen en realidad la misma teoría, formulada por un lado para los conjuntos y por el otro para las funciones. Por ejemplo, se tiene el siguiente resultado:

Sea $f : [a, b] \rightarrow \mathbb{R}$ una función acotada no negativa y E la región en \mathbb{R}^2 acotada por el eje x y la gráfica de f entre a y b , entonces f es Riemann integrable si y sólo si E es Jordan medible. Además, en ese caso, se tiene $\int_a^b f(x)dx = c(E)$.

4. TEORÍA DE LA MEDIDA DE BOREL

En 1894-1895, Émile Borel dio las bases para un nuevo avance al introducir el concepto de medida cero:

Definición 4. *Se dice que un conjunto tiene medida cero si para cualquier $\varepsilon > 0$ existe una colección numerable de intervalos abiertos $\{I_n\}$ cuya unión cubre al conjunto y tales que la suma de sus longitudes es menor que ε .*

Además de introducir el concepto de medida cero al resolver el problema que lo motivó para dar esa definición, la demostración de Borel contiene un resultado que sería clave para que más adelante Lebesgue pudiera definir el concepto de medida. Ese resultado se puede enunciar de la siguiente manera:

Sea I un intervalo cerrado y acotado e $(I_j)_{j \in \mathbb{N}}$ una sucesión de intervalos abiertos tales que $I \subset \bigcup_{j=1}^{\infty} I_j$, entonces:

$$l(I) \leq \sum_{j=1}^{\infty} l(I_j)$$

El resultado parece trivial ya que al evaluar la suma de las longitudes de los intervalos I_j , si dos de ellos se traslapan, podría haber una parte de I cuya longitud se está sumando dos veces; si no se traslapan, al sumar las longitudes de los dos intervalos, esa suma es por lo menos igual a la suma de las longitudes de las partes de I que se encuentran dentro de esos intervalos. Sin embargo, al tratar de formalizar esta idea se llega nuevamente al problema inicial.

Si el conjunto de intervalos abiertos cuya unión cubre I fuera finito, el resultado se puede demostrar fácilmente.

Para demostrar el resultado enunciado, Borel demostró que existe una colección finita de los intervalos I_j cuya unión también contiene a I , resultado que más adelante llevó al concepto de compacidad.

Más adelante, en un libro publicado en 1898, Borel retomó el concepto de conjunto de medida cero para desarrollar una teoría de la medida. Para Borel la idea fundamental consistía en definir los elementos nuevos que se introducen con ayuda de sus propiedades esenciales, es decir, aquellas que son estrictamente indispensables para los razonamientos que siguen. Las propiedades esenciales que planteó Borel son las siguientes:

1. La medida de la unión de una colección numerable de conjuntos ajenos es igual a la suma de sus medidas.
2. La medida de la diferencia de dos conjuntos de medida finita A y B , con $A \subset B$, es igual a la diferencia de sus medidas $m(B) - m(A)$.
3. La medida de un conjunto nunca es negativa.

5. TEORÍA DE LA MEDIDA DE LEBESGUE

El paso siguiente en el desarrollo de la Teoría de la Medida, así como el último paso hacia la caracterización de las funciones Riemann-integrables lo dió Henri Lebesgue en 1902.

Una función acotada $f : [a, b] \mapsto \mathbb{R}$ es Riemann integrable si y sólo si el conjunto de puntos donde la función es discontinua tiene medida cero.

Lebesgue desarrolló su teoría de la integral en su tesis doctoral titulada *Integrale, longueur, aire*. Más tarde la expuso en su libro *Leçons sur l'intégration et la recherche des fonctions primitives*.

Mostró Lebesgue que para dar una definición constructiva de la integral de cualquier función acotada, basta con hacerlo para las funciones que únicamente toman como valores 0 y 1 y, para una función f de este tipo, el problema de integración se traduce en asignar un número al conjunto de números reales $x \in (a, b)$ tales que $f(x) = 1$; de manera que entonces se planteó Lebesgue el problema de la medida de conjuntos, el cual consiste en asignar a cada conjunto acotado de números reales E , un número no negativo, $m(E)$, al cual llamará la medida de E , debiéndose satisfacer las siguientes propiedades:

1. Si E es un conjunto acotado y $a \in \mathbb{R}$, entonces $m(E + a) = m(E)$.
2. Si E_1, E_2, \dots es una familia finita o infinita numerable de conjuntos, ajenos por parejas y contenidos en un conjunto acotado, entonces $m(\bigcup_n E_n) = \sum_n m(E_n)$.
3. $m([0, 1]) = 1$.

Las condiciones sobre la medida implican que si $x \in \mathbb{R}$, entonces $m(\{x\}) = 0$.

También, la medida de un intervalo acotado $[a, b]$ debe de ser igual a su longitud.

Para definir la medida de cualquier conjunto acotado, Lebesgue hizo el siguiente razonamiento:

Si E es un conjunto acotado e I_1, I_2, \dots es una colección finita o infinita numerable de intervalos, ajenos por parejas, tales que $E \subset \bigcup_n I_n$, entonces se debe de tener $m(E) \leq \sum_n l(I_n)$; definió entonces la **medida exterior** de E , $m_e(E)$, como el ínfimo de esas sumas, es decir:

$$m_e(E) = \inf \left\{ \sum_n l(I_n) : I_1, I_2, \dots \text{ son intervalos ajenos por parejas y } E \subset \bigcup_n I_n \right\}$$

Ahora bien, como se tiene que $m([a, b]) = m(E) + m([a, b] - E)$, entonces:

$$m(E) = m([a, b]) - m([a, b] - E) \geq m([a, b]) - m_e([a, b] - E) = l([a, b]) - m_e([a, b] - E)$$

Se sigue que la cantidad $l([a, b]) - m_e([a, b] - E)$ es una cota inferior para la medida de E , la cual define como la **medida interior** de E y la denota por $m_i(E)$.

Como ya lo mencionamos, Lebesgue hizo lo anterior asumiendo que es posible asignarle una medida a todo conjunto acotado E , sin embargo las definiciones de medida exterior e interior son independientes de esta consideración y pueden darse para cualquier conjunto. Mostró entonces que se tienen las siguientes relaciones para cualquier conjunto acotado E :

$$c_i(E) \leq m_i(E) \leq m_e(E) \leq c_e(E)$$

Además, como se mostró arriba, de ser posible asignar una medida $m(E)$ al conjunto E , se debe de tener $m_i(E) \leq m(E) \leq m_e(E)$. Por lo tanto, la medida asignada a E será única cuando sus medidas interior y exterior coincidan. De aquí que Lebesgue estableció la siguiente definición:

Definición 5. *Se dice que un conjunto acotado E es medible si $m_i(E) = m_e(E)$.*

Finalmente observó Lebesgue que, debido a la relación:

$$c_i(E) \leq m_i(E) \leq m_e(E) \leq c_e(E)$$

cualquier conjunto Jordan medible es también Lebesgue medible y, dado que los intervalos son medibles y la familia de conjuntos medibles tiene las propiedades enunciadas arriba, todo conjunto medible, de acuerdo a la definición de Borel, es también Lebesgue medible. De esta

forma la teoría de la medida de Lebesgue resulta más general tanto que la de Jordan como de la de Borel y las engloba a ambas.

Lo mismo ocurre con la teoría de integración que desarrolla Lebesgue a partir de su concepto de medida. La integral de Lebesgue resulta ser más general que la de Riemann: toda función Riemann-integrable es Lebesgue-integrable, pero hay muchísimas funciones que son Lebesgue-integrables, pero no Riemann-integrables. Además, la integral definida por Lebesgue tiene propiedades que no las tiene la integral de Riemann y eso la hace mucho más manejable.

6. LA INTEGRAL DE STIELTJES

Previamente al trabajo de Lebesgue, Thomas Joannes Stieltjes había extendido el concepto de integral en una dirección distinta a la de Lebesgue. En el año 1894 publicó un artículo titulado *Recherches sur les fractions continues*, donde planteó el problema de determinar el límite, si existe, de una fracción continua de la forma:

$$\frac{1}{a_1 z + \frac{1}{a_2 z + \frac{1}{a_3 z + \frac{1}{\dots}}}}$$

donde $(a_n)_{n \in \mathbb{N}}$ es una sucesión de números reales positivos y z un número real o un número complejo.

En el desarrollo que realizó Stieltjes en su artículo, obtuvo una expresión que lo llevó a introducir el concepto de momento de una función monótona creciente y al problema de la determinación de esa función a partir de sus momentos. Para ello, decía que una distribución de masa sobre la parte positiva de una recta de origen O representa una función creciente de la distancia x al origen. Agregaba que, inversamente, una función creciente, definida sobre la parte positiva de la recta, se puede imaginar como representando una distribución de masa. Dada una función creciente Φ , definida en un intervalo $[a, b]$ sobre la parte positiva de una recta, consideraba una partición $\{x_0, x_1, \dots, x_n\}$ del intervalo $[a, b]$, tomaba un punto ξ_i en cada subintervalo $[x_{i-1}, x_i]$, consideraba la suma $\sum_{i=1}^n \xi_i (\Phi(x_i) - \Phi(x_{i-1}))$ y definía el momento de Φ como el límite de esa suma cuando las longitudes de los subintervalos de la partición tienden a cero (para $k \in \mathbb{N}$, el momento de orden k de Φ sería el límite de las sumas $\sum_{i=1}^n \xi_i^k (\Phi(x_i) - \Phi(x_{i-1}))$). Generalizando esta idea, consideró una función continua f , definida sobre el intervalo $[a, b]$, y definió la integral de f con respecto a Φ en el intervalo $[a, b]$, denotada por $\int_a^b f(u) d\Phi(u)$, como el límite de las sumas $\sum_{i=1}^n f(\xi_i) (\Phi(x_i) - \Phi(x_{i-1}))$. De esta forma surgió lo que ahora se conoce como la integral de Riemann-Stieltjes.

La integral de Stieltjes se vincula con otro concepto de importancia central, el de función de variación acotada. En su libro *Leçons sur l'intégration et la recherche des fonctions primitives* (1904), Lebesgue atribuye la invención de este concepto a Jordan. En ese libro, Lebesgue dedicó un capítulo al estudio de las funciones de variación acotada, aunque no motivado por el trabajo de Stieltjes, el cual ni siquiera menciona. La motivación de Lebesgue provenía del estudio que hacía de la rectificación de curvas, es decir de la medida de la longitud de

una curva, y también del hecho de que, si $f : [a, b] \rightarrow \mathbb{R}$, es integrable, entonces la función $x \rightarrow \int_a^x f(y) dy$, definida sobre $[a, b]$, es de variación acotada.

La relación entre la integral de Stieltjes y las funciones de variación acotada proviene de que una función $g : [a, b] \rightarrow \mathbb{R}$ es de variación acotada si y sólo si se puede expresar como la diferencia de dos funciones no decrecientes, resultado que Lebesgue demostró en el libro citado. De aquí que la integral de Stieltjes se pueda extender al caso en que Φ es de variación acotada.

Una propiedad muy importante de la integral de Stieltjes es la fórmula de integración por partes.

Teorema 1 (Fórmula de integración por partes). *Supongamos que la función acotada $f : [a, b] \rightarrow \mathbb{R}$ es integrable con respecto a la función acotada $g : [a, b] \rightarrow \mathbb{R}$, entonces g es integrable con respecto a f y se tiene:*

$$\int_a^b gdf = f(b)g(b) - f(a)g(a) - \int_a^b f dg$$

Por otra parte, para que una teoría de integración tenga alguna utilidad práctica, tal vez la mínima condición que se puede pedir es que todas las funciones continuas sean integrables y es aquí donde cobra gran importancia un concepto clave para la teoría desarrollada por Lebesgue y su generalización, el concepto de función de variación acotada, el cual Lebesgue ya lo había identificado como muy importante.

Se dice que una función $g : [a, b] \rightarrow \mathbb{R}$ es de variación acotada si:

$$\sup_P \sum_{k=1}^n |g(t_k) - g(t_{k-1})| < \infty$$

donde $P = \{a = t_0 < t_1 < \dots < t_n = b\}$ es una partición del intervalo $[a, b]$.

Se puede demostrar que una función $g : [a, b] \rightarrow \mathbb{R}$ es de variación acotada si y sólo si existen dos funciones monótonas no decrecientes $u_1 : [a, b] \rightarrow \mathbb{R}$ y $u_2 : [a, b] \rightarrow \mathbb{R}$ tales que $g = u_1 - u_2$.

Teorema 2. *Teorema Sea $g : [a, b] \rightarrow \mathbb{R}$ una función de variación acotada y $f : [a, b] \rightarrow \mathbb{R}$ una función continua, entonces f es integrable con respecto a g .*

Teorema 3. *Sea $g : [a, b] \rightarrow \mathbb{R}$ una función y supongamos que todas las funciones continuas $f : [a, b] \rightarrow \mathbb{R}$ son integrables con respecto a g , entonces g es de variación acotada.*

Este resultado es fundamental y se generalizaría, llegando a ser un concepto clave en la teoría general de la medida y la teoría general de integración. Podría decirse de la siguiente manera: para tener una "buena" teoría de integración de una función f con respecto a otra función g , se requiere que g sea de variación acotada.

7. TEORÍA GENERAL DE LA MEDIDA

Después de que Lebesgue desarrolló su teoría de integración en \mathbb{R} , se extendió al caso de \mathbb{R}^n sin mucha dificultad.

En 1913, Johann Karl August Radon mostró que se puede desarrollar una teoría general en la cual quedan incluidas la integral de Lebesgue en \mathbb{R}^n y la integral de Stieltjes.

Radon introdujo el concepto de funcional aditiva sobre una familia de subconjuntos de \mathbb{R}^n . Aún no se trataba de la definición general de una medida ya que Radon se restringió a familias de subconjuntos de \mathbb{R}^n que contienen a los conjuntos borelianos.

Finalmente, también siguiendo un procedimiento similar al de Lebesgue, Radon desarrolló una teoría de integración para las funcionales aditivas.

Con base en el trabajo de Radon, Maurice René Fréchet extendió la teoría de la medida de Lebesgue a espacios abstractos en un artículo de 1915 titulado *Sur l'intégrale d'une fonctionnelle étendue a un ensemble abstrait*.

Finalmente, Fréchet definió la integral con respecto a una función aditiva utilizando el método de Darboux, el cual consiste en definir la integral superior y la integral inferior.

En su artículo de 1923, al cual le siguió uno de 1924, Fréchet desarrolló aún más su teoría iniciada en 1915, quedando así ya establecido lo esencial de lo que posteriormente se llamaría la teoría de la medida y la teoría de integración con respecto a una medida.

Podría decirse que el ciclo de investigación alrededor de los conceptos de medida y de integral con respecto a una medida, así como de sus propiedades básicas, se cerró con el trabajo de Otton Nikodym de 1930.

Nikodym retomó el trabajo de Radon y obtuvo un resultado general, ahora conocido como el teorema de Radon-Nikodym.

Nikodym hacía referencia en su artículo a la formulación general que hace Fréchet de la teoría de la medida de Lebesgue, pero modificó un poco la terminología. Consideraba una familia no vacía \mathcal{H} de subconjuntos de un conjunto H , la cual es cerrada bajo uniones numerables y complementos (en particular H pertenece a la familia); es decir, lo que ahora se denomina σ -álgebra de subconjuntos de H . Una medida μ la definió entonces como una función (con valores reales), no negativa, definida sobre \mathcal{H} , la cual es "perfectamente aditiva", es decir, si E_1, E_2, \dots son elementos de la familia, ajenos por parejas, entonces $\mu(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu(E_n)$; es decir, μ es σ -aditiva, en la terminología moderna. En otras palabras, Nikodym trabajaba con medidas tal y como se les define actualmente.

Con este trabajo de Nikodym quedó formulada la teoría de la medida como se le conoce actualmente y quedaron establecidos los resultados básicos de la teoría de integración con respecto a una medida.

8. TEORÍA DE LA PROBABILIDAD

Por el lado del Cálculo de Probabilidades, una de las propiedades básicas con la que se contaba para calcular probabilidades es la que formula Henri Poincaré en un libro publicado en el año 1896: cuando un evento puede producirse de dos maneras diferentes, de tal forma que esas dos maneras no puedan ocurrir simultáneamente, la probabilidad de ocurrencia de

este evento es igual a la suma de la probabilidad de que se produzca de la primer manera y de la probabilidad de que se produzca de la segunda manera."De aquí se sigue inmediatamente que si un evento puede producirse de n maneras diferentes, de tal forma que cualesquiera dos de esas maneras no puedan ocurrir simultáneamente, la probabilidad de ocurrencia de este evento es igual a la suma $p_1 + p_2 + \dots + p_n$, donde, para $k \in \{1, 2, \dots, n\}$, p_k es la probabilidad de que se produzca de la k -ésima manera. ¡Es la aditividad finita!

¿Qué se podía decir (estamos ubicándonos a finales del siglo XIX) si un evento puede producirse de una infinidad numerable de maneras diferentes, de tal forma que cualesquiera dos de esas maneras no puedan ocurrir simultáneamente? Hacía muchos años que Bernoulli había dado una respuesta a esta pregunta, en su libro publicado en el año 1713. No lo hizo para el problema general que estamos planteando, sino para un problema particular donde se presentaba esta situación. La respuesta, para el problema particular que planteó Bernoulli, es simple: la probabilidad de ocurrencia de tal evento es la serie $\sum_{k=1}^{\infty} p_k$, donde para $k \in \mathbb{N}$, p_k es la probabilidad de que el evento se produzca de la k -ésima manera. ¡Es la σ -aditividad!

La historia no finalizó con lo que hizo Lebesgue ni el Cálculo de Probabilidades se estancó en el estado en que se encontraba a finales del siglo XIX. Hacia 1930 se tenía ya desarrollada una Teoría General de la medida, la cual incluía la definición y el estudio de medidas definidas sobre una familia de subconjuntos de un conjunto cualquiera y se contaba con un método para construir esas medidas. La propiedad básica de cualquier medida es la σ -aditividad.

Casi inmediatamente después de publicado el trabajo de Lebesgue, la naciente Teoría de la Medida comenzó a utilizarse en algunos problemas de probabilidad, por ejemplo para calcular un tipo de probabilidades llamadas geométricas, las cuales consisten en considerar la elección al azar de un punto en una determinada región del plano y calcular la probabilidad de que el punto seleccionada pertenezca a un subconjunto dado de esa región. La probabilidad buscada se calculaba simplemente dividiendo el área del subconjunto dado entre el área de la región (la cual obviamente tendría que ser positiva). Con la teoría de Lebesgue ese problema podía ser resuelto para una familia más grande de subconjuntos de la región donde se selecciona el punto. En ese caso la función de probabilidad resulta ser σ aditiva ya que está definida mediante la medida de Lebesgue en el plano. Pero, se trataba únicamente de un tipo de problemas de probabilidad.

Se fueron planteando problemas de probabilidad cada vez más complejos, en particular con sucesiones de variables aleatorias independientes. Un problema de gran importancia que se resolvió fue el de construir un modelo matemático (probabilístico) para el llamado movimiento browniano, el cual consiste en el movimiento de un grano de polen que se coloca sobre agua. Las posibles trayectorias que sigue el grano de polen sobre el agua son funciones continuas definidas en el intervalo de tiempo en que se observa el movimiento, el cual podríamos asumir que es el intervalo $[0, \infty)$ y podríamos imaginar que el recipiente de agua es infinito, de manera que cada posible trayectoria que sigue el grano de polen es una función continua $f : [0, \infty) \rightarrow \mathbb{R}^2$; un espacio de dimensión infinita. Fue Norbert Wiener quien, en el año 1923, construyó ese modelo utilizando la moderna teoría de la medida y asumiendo que la función de probabilidad es σ -aditiva. Teníamos nuevamente la σ -aditividad de la función de probabilidad, pero todavía para un caso particular, aunque esta vez de mucha importancia.

Sin embargo aún quedaban algunos obstáculos que salvar para una aceptación general de esta propiedad. El problema central era el siguiente:

A una variable aleatoria se le asocia una función no decreciente denominada su función de distribución y, desde el año 1913, August Radon había demostrado que a partir de una función de ese tipo se puede construir una medida utilizando el método de Lebesgue. A una familia finita formada por n variables aleatorias se le asocia también una función de n variables denominada su función de distribución conjunta y, utilizando el método que Constantin Carathéodory había publicado en el año 1914, también a partir de una función de distribución conjunta se podía construir una medida.

El problema que quedaba por resolver era como construir una medida asociada a una infinidad, numerable o no numerable, de variables aleatorias.

Este problema lo resolvió, utilizando el método de Carathéodory, Andrei Nikolayevich Kolmogorov en el año 1933. De esta manera la fusión del Cálculo de Probabilidades con la Teoría de la Medida quedaba consumada. El sistema matemático formal que surgió con las investigaciones entre los años 1900 y 1933, las cuales culminaron con la publicación del trabajo de Kolmogorov, es lo que propiamente podemos llamar ahora la Teoría de la Probabilidad.

9. PROCESOS ESTOCÁSTICOS

El estudio de los procesos estocásticos se inició a principios del siglo XX, en forma paralela a la conclusión del estudio de las sucesiones de

variables aleatorias independientes. Sus principales precursores fueron Andrei Andreyevich Markov, A. N. Kolmogorov, A. Ya. Khintchine, N. Wiener y P. Lévy.

Desde la publicación del teorema de Bernoulli, en 1713, el motor de desarrollo de la teoría de la probabilidad fue la búsqueda de resultados que permitieran mejorar y generalizar ese teorema. Ese proceso culminó en la década de los 30's del siglo XX con la formulación acabada de los teoremas límite para sucesiones de variables aleatorias independientes.

El teorema de Bernoulli establece que si \mathcal{E} es un experimento aleatorio y A un evento relativo a ese experimento, de probabilidad igual a p , y consideramos un nuevo experimento aleatorio consistente en la repetición indefinida del experimento \mathcal{E} , de tal manera que cada repetición es independiente de las otras, entonces, llamando Y_n al número de veces que ocurre el evento A en las primeras n repeticiones del experimento, la sucesión de variables aleatorias $\left(\frac{Y_n}{n}\right)_{n \in \mathbb{N}}$ converge a p en probabilidad. Formulado de otra manera, el resultado dice que si $(X_n)_{n \in \mathbb{N}}$ es una sucesión de variables aleatorias independientes, todas con distribución Bernoulli de parámetro p , entonces la sucesión $\left(\frac{X_1+X_2+\dots+X_n}{n}\right)_{n \in \mathbb{N}}$ converge a p en probabilidad.

A la forma general del resultado de Bernoulli se le conoce como ley débil de los grandes números. En una primera versión, el resultado general fue demostrado por Chebyshev, iniciador de la llamada .^{escuela rusa.}en la teoría de la probabilidad. En el año 1867 Chebyshev demostró (utilizando la que ahora se conoce como "desigualdad de Chebyshev") que si $(X_n)_{n \in \mathbb{N}}$ es una sucesión de variables aleatorias independientes, idénticamente distribuidas

y de varianza finita, entonces la sucesión $\left(\frac{X_1+X_2+\dots+X_n}{n}\right)_{n \in \mathbb{N}}$ converge en probabilidad a la esperanza común de las variables aleatorias X_n .

En el año 1928, el resultado de Chebyshev fue mejorado por Khintchine, quien logró demostrar la ley débil sin la condición de que las variables aleatorias X_n tengan varianza finita.

De igual manera, en este periodo se estableció la ley fuerte de los grandes números a partir del resultado de Borel del año 1909, siendo Kolmogorov quien obtuvo el resultado general en el año 1930.

El teorema del límite central, demostrado en el año 1733 por Abraham de Moivre para el caso particular de una sucesión de variables aleatorias independientes, todas con distribución Bernoulli de igual parámetro, fue demostrado de manera general por Lindeberg en el año 1922.

Dentro de este contexto del estudio de los teoremas límite, Markov introdujo en 1906 una nueva línea de investigación al demostrar la ley débil de los grandes números para una sucesión de variables aleatorias no independientes, sino, de acuerdo con el término que él utilizó, encadenadas. Surgió así un tipo de proceso estocástico que ahora es llamado de Markov. De manera general e informal, se dice que un proceso estocástico es de Markov si dado cualquier tiempo $t \in \Gamma$, las familias de variables aleatorias $\{X_s : s < t\}$ y $\{X_s : s > t\}$ son independientes, propiedad que en ocasiones se enuncia diciendo que, dado el presente, el pasado y el futuro son independientes. Por esta propiedad, también se dice que un proceso de este tipo es un proceso sin memoria.

Un gran avance en el estudio de los procesos estocásticos lo hizo Wiener entre 1921 y 1923 al construir un modelo matemático del movimiento browniano. En el año 1827, al estudiar el proceso de fertilización de las flores de varias plantas, Robert Brown observó que los granos de polen se movían: .Al examinar la forma de estas partículas inmersas en agua, observé muchas de ellas muy evidentemente en movimiento; éste consistía no solamente en un cambio de lugar en el fluido, manifestado por alteraciones en sus posiciones relativas, sino que también, con no poca frecuencia, por un cambio en la forma de la misma partícula misma ... Estos movimientos eran tales que me convencieron, después de observaciones repetidas frecuentemente, de que no surgían de corrientes en el fluido, ni de su gradual evaporación, sino que pertenecían a la partícula misma..^Este movimiento ya había sido observado por otros investigadores anteriormente a Brown, sin embargo fue él quien realizó un estudio más detallado, aunque no completo del mismo. En su honor se le conoce como movimiento browniano.

Mientras tanto, la escuela rusa, que surgió con los trabajos de P. L. Chebyshev, realizó avances significativos generalizando el trabajo de Markov. Entre 1931 y 1937, Kolmogorov realizó un estudio a fondo de las cadenas de Markov en tiempo discreto y en tiempo continuo; también estudio los procesos de difusión, los cuales son procesos de Markov con espacios de estados continuos. Por su parte, en el año 1934, Khintchine introdujo un tipo de procesos que pueden ser no markovianos: los procesos estacionarios.

Otro gran avance en el desarrollo de la teoría de los procesos estocásticos, lo realizó Paul Lévy en su libro *Théorie de l'addition des variables aléatoires*, cuya primera edición se publicó en

1937. Ahí planteó que una extensión natural de las sumas o series de variables aleatorias independientes es el siguiente: .En lugar de un índice entero n , introduciremos un parámetro t que varía de una manera continua; las sumas sucesivas S_n (se refiere a $S_n = \sum_{k=1}^n X_k$, donde X_1, X_2, \dots, X_n son variables aleatorias independientes) serán entonces reemplazadas por una función aleatoria $X(t)$ de ese parámetro, y la condición de independencia de los diferentes elementos cuya suma es $X(t)$ debe expresarse como sigue: si $t_0 < t_1$, el crecimiento $X(t_1) - X(t_0)$ es una variable aleatoria independiente de $X(t_0)$ y del conjunto de valores de $X(t)$ para $t \leq t_0$."(p. 158). Tomando $t \in \mathbb{R}^+$, lo anterior implica que si t_1, t_2, \dots, t_n son números reales no negativos tales que $t_1 < t_2 < \dots < t_n$, entonces las variables aleatorias $X(t_2) - X(t_1), X(t_3) - X(t_2), \dots, X(t_n) - X(t_{n-1})$ son independientes.

Lévy está definiendo ahí lo que se conoce ahora como un proceso con incrementos independientes. Aclara que algunos casos particulares de ese tipo de funciones aleatorias habían sido consideradas previamente por Bachelier, en 1913, Wiener, en 1923 y Kolmogorov, en 1932.

Obsérvese que si $(Y_n)_{n \in \mathbb{Z}^+}$ es un proceso con incrementos independientes, entonces, definiendo, para $k \in \mathbb{N}$, $X_k = Y_k - Y_{k-1}$, $(X_k)_{k \in \mathbb{N}}$ es una sucesión de variables aleatorias independientes y, para cualquier $n \in \mathbb{N}$, se tiene: $Y_n - Y_0 = \sum_{k=1}^n X_k$. De aquí que, para estudiar el proceso $(Y_n)_{n \in \mathbb{Z}^+}$, se puedan utilizar los resultados que se tienen para las sumas de variables aleatorias independientes.

La definición general de proceso estocástico es la siguiente:

Definición 6. Sean $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad, (E, \mathcal{E}) un espacio medible y Γ un conjunto cualquiera. Llamaremos proceso estocástico definido sobre $(\Omega, \mathfrak{F}, P)$, con conjunto de índices en Γ y con espacio de estados E , a toda familia $\{X_t\}_{t \in \Gamma}$ de funciones medibles definidas sobre Ω y con valores en E . Dada $\omega \in \Omega$, la función $t \mapsto X_t(\omega)$ es llamada una trayectoria del proceso. Para referirnos a un proceso estocástico $\{X_t\}_{t \in \Gamma}$ utilizaremos la notación $(X_t)_{t \in \Gamma}$.

En general, un proceso estocástico representa la dinámica de un fenómeno aleatorio, en cuyo caso el conjunto Γ de índices representa el tiempo durante el cual evoluciona el fenómeno en consideración, así que Γ es un conjunto de números reales. El caso más común consiste en el estudio de un fenómeno aleatorio que evoluciona a partir de un tiempo inicial $t_0 \geq 0$, de manera que Γ es un subconjunto de $[0, \infty)$; en muchos casos de interés, se trata del conjunto de números enteros no negativos, del conjunto de números enteros no positivos o del conjunto de números reales no negativos. En algunos casos será de interés considerar también una variable aleatoria X_∞ .

Denotaremos por \mathbb{Z}^+ al conjunto de números enteros no negativos y por \mathbb{R}^+ al conjunto de números reales no negativos. $\overline{\mathbb{Z}^+}$ denotará al conjunto $\mathbb{Z}^+ \cup \{\infty\}$ y $\overline{\mathbb{R}^+}$ al conjunto $\mathbb{R}^+ \cup \{\infty\}$.

Si $\Gamma = \mathbb{Z}^+$ o $\Gamma = \mathbb{Z}^-$ diremos que $(X_t)_{t \in \Gamma}$ es un proceso estocástico en tiempo discreto, mientras que si $\Gamma = \mathbb{R}^+$ diremos $(X_t)_{t \in \Gamma}$ es un proceso estocástico en tiempo continuo.

Únicamente consideraremos el caso de procesos estocásticos con valores reales, de manera que E será el conjunto de números reales \mathbb{R} y \mathcal{E} la σ -álgebra de los conjuntos borelianos en

\mathbb{R} , la cual denotaremos por $\mathcal{B}(\mathbb{R})$. Así que, para cada $t \in \Gamma$, X_t será una variable aleatoria real.

Para cada $t \in \Gamma$, X_t representa lo que llamaremos el estado del proceso en el tiempo t . Este estado no está determinado, puede ser uno cualquiera de los valores que toma X_t . Lo único conocido es la ley probabilística mediante la cual evoluciona el proceso, la cual, en particular, nos dice cuál es la distribución de X_t . Esta distribución no es más que una medida de probabilidad definida sobre la σ -álgebra generada por X_t , cuyos elementos son aquellos eventos que dependen únicamente del estado del proceso en el tiempo t , a saber, todos los de la forma $[X_t \in B]$, donde $B \in \mathcal{E}$.

De manera más general, resulta de interés considerar, para cada $t \in \Gamma$, la σ -álgebra formada por todos aquellos eventos que dependen únicamente del estado del proceso, desde su inicio, hasta el tiempo t , es decir, la σ -álgebra generada por la familia de variables aleatorias $\{X_s : s \in \Gamma \text{ y } s \leq t\}$. Esta familia es no decreciente en el sentido de que si $t_1, t_2 \in \Gamma$ y $t_1 < t_2$ entonces $\sigma\{X_s : s \in \Gamma \text{ y } s \leq t_1\} \subset \sigma\{X_s : s \in \Gamma \text{ y } s \leq t_2\}$. Una familia de este tipo será llamada una filtración.

Definición 7. Una *filtración* $\{\mathfrak{S}_t : t \in \Gamma\}$ de (Ω, \mathfrak{S}) es una familia de σ -álgebras tales que $\mathfrak{S}_t \subset \mathfrak{S}$ para cualquier $t \in \Gamma$ y si $s, t \in \Gamma$ son tales que $s < t$, entonces $\mathfrak{S}_s \subset \mathfrak{S}_t$. Denotaremos por $\mathfrak{S}_{\infty-}$ a la σ -álgebra generada por los eventos que pertenecen a \mathfrak{S}_t para alguna $t \in \Gamma$.

Definición 8. Diremos que una filtración $\{\mathfrak{S}_t : t \in \Gamma\}$ es completa si, para cualquier $t \in \Gamma$, todos los eventos de probabilidad cero pertenecen a \mathfrak{S}_t .

Definición 9. Diremos que un proceso estocástico $(X_t)_{t \in \Gamma}$ está adaptado a la filtración $\{\mathfrak{S}_t : t \in \Gamma\}$ si, para cualquier $t \in \Gamma$, la variable aleatoria X_t es \mathfrak{S}_t -medible.

Definición 10. Si $(X_t)_{t \in \Gamma}$ es un proceso estocástico y $\{t_1, t_2, \dots, t_n\}$ es un subconjunto finito de Γ , la distribución del vector aleatorio $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ será llamada una distribución finito dimensional del proceso.

Recordemos que la distribución de un vector aleatorio (X_1, X_2, \dots, X_n) , formado por variables aleatorias reales, es el conjunto de probabilidades de la forma $P[(X_1, X_2, \dots, X_n) \in B]$, donde B es un conjunto boreliano de \mathbb{R}^n . Estas probabilidades quedan determinadas por las probabilidades de la forma $P[X_1 \in B_1, X_2 \in B_2, \dots, X_n \in B_n]$, donde B_1, B_2, \dots, B_n son conjuntos borelianos de \mathbb{R} . Utilizando funciones en lugar de conjuntos, la distribución de (X_1, X_2, \dots, X_n) queda determinada por el conjunto de esperanzas de la forma $E[f_1(X_1) f_2(X_2) \cdots f_n(X_n)]$, donde f_1, f_2, \dots, f_n son funciones medibles y acotadas de \mathbb{R} en \mathbb{R} .

El estudio que hizo Lévy de los procesos con incrementos independientes resultó ser de gran importancia en la teoría de los procesos estocásticos. Para enunciar algunos de sus resultados, requerimos de las siguientes definiciones:

Definición 11. Diremos que un proceso $(X_t)_{t \in \mathbb{R}^+}$ es fuertemente continuo si existe un conjunto $C \in \mathfrak{S}$ de probabilidad 1, tal que, para cualquier $\omega \in C$, la función $t \rightarrow X_t(\omega)$, definida sobre \mathbb{R}^+ , es continua.

Definición 12. Diremos que un proceso $(X_t)_{t \in \mathbb{R}^+}$ es débilmente continuo si:

i) Para cada $t_0 \in \mathbb{R}^+$, existe un conjunto $C_{t_0} \in \mathfrak{S}$, de probabilidad 1, tal que, para cualquier $\omega \in C_{t_0}$, la función $t \rightarrow X_t(\omega)$, definida sobre \mathbb{R}^+ , es continua en $t = t_0$.

ii) Existe un conjunto $C \in \mathfrak{S}$, de probabilidad 1, tal que, para cualquier $\omega \in C$, la función $t \rightarrow X_t(\omega)$, definida sobre \mathbb{R}^+ , tiene límites por la derecha y por la izquierda en cualquier punto.

El siguiente resultado fue demostrado por Lévy y resultó ser de gran importancia.

Teorema 4. Si $(X_t)_{t \in \mathbb{R}^+}$ es un proceso con incrementos independientes, entonces $(X_t)_{t \in \mathbb{R}^+}$ es fuertemente continuo si y sólo si, para cualesquiera $s, t \in \mathbb{R}^+$ tales que $s < t$, la variable aleatoria $X_t - X_s$ tiene distribución normal.

Los ejemplos básicos, de procesos con incrementos independientes, que menciona Lévy en su libro de 1937, son el proceso de Wiener, el cual es fuertemente continuo, y el proceso de Poisson, el cual es débilmente continuo, sin discontinuidades fijas:

Muestra que las propiedades que caracterizan al proceso de Wiener, $(W_t)_{t \in \mathbb{R}^+}$, son las siguientes:

i) $W_0 = 0$.

ii) Si $0 < t_1 < \dots < t_n$, entonces las variables aleatorias $W_{t_1}, W_{t_2} - W_{t_1}, \dots, W_{t_n} - W_{t_{n-1}}$ son independientes.

iii) Si $0 \leq s < t$, entonces la variable aleatoria $W_t - W_s$ tiene distribución normal de parámetros $\mu = 0$ y $\sigma^2 = t - s$.

iv) $(W_t)_{t \in \mathbb{R}^+}$ es un proceso fuertemente continuo.

Los resultados de la teoría de los procesos estocásticos obtenidos durante la primera mitad del siglo XX fueron recogidos, sistematizados y complementados con nuevos resultados, en dos libros que marcaron la pauta para continuar desarrollando esta teoría por lo menos durante los 30 años que les siguieron: El libro de Paul Lévy *Processus stochastiques et mouvement brownien*, publicado en 1948 y el libro de Joseph Leo Doob *Stochastic processes*, publicado en 1953.

Lévy comienza su libro dando una nueva construcción del proceso de Wiener y pasa después a considerar el problema general de la construcción de procesos estocásticos. En los capítulos siguientes expone la teoría de los procesos de Markov y de los procesos estacionarios. Después expone una versión más completa de su teoría de los procesos con incrementos independientes, llamados ahí procesos aditivos. Tal vez la parte más importante de su libro es el estudio a fondo que hace del movimiento browniano, tanto en la recta real, como en el plano y en otros espacios.

El libro de J. L. Doob fue fuente de inspiración para muchos matemáticos ya que ahí plantea varias líneas de investigación para diferentes problemas. Doob sigue lo que podríamos llamar el enfoque moderno: se parte de los axiomas de Kolmogorov y la construcción de procesos se realiza vía el teorema de extensión de Kolmogorov. Comienza su libro estudiando

propiedades generales de los procesos estocásticos, en particular, la existencia de versiones de un proceso con trayectorias que son regulares en algún sentido. Realiza un estudio a fondo de los procesos de Markov en tiempo discreto y en tiempo continuo. Trata también el estudio de los procesos con incrementos independientes y de los procesos estacionarios. Tal vez las principales aportaciones de Doob en ese libro son, por una parte, el estudio que hace de los procesos conocidos como martingalas, los cuales adquirirían una importancia central dentro de la teoría de los procesos estocásticos, y, por otra parte, la invención del concepto de tiempo opcional, el cual es un tiempo aleatorio. Meyer, en la segunda edición de su libro dice al respecto (p. 184): "Il a sans doute fallu autant de génie aux créateurs du calcul infinitésimal pour expliciter la notion si simple de dérivée, qu'à leurs successeurs pour faire toute le reste. L'invention des temps d'arrêt par Doob est tout a fait comparable".

Es interesante la diferencia que existió entre Paul Lévy y J. L. Doob en cuanto a la forma en que pensaban un proceso estocástico. El libro de Doob se publicó en 1954 y en 1955 se publicó una segunda edición del libro de Lévy *Théorie de l'addition des variables aléatoires*, de manera que Lévy pudo leer el libro de Doob antes de que se imprimiera esta segunda edición y, después de hacerlo, decidió agregar una nota al final de su libro. En ella, dice: "Un proceso estocástico es en principio un fenómeno en cuya evolución el azar interviene a cada instante. Para Doob, un proceso estocástico es simplemente una función aleatoria $X(t)$ de una variable t la cual puede uno imaginar que representa el tiempo. Cualquiera que sea el conjunto \mathfrak{E} de determinaciones posibles de $X(t)$, se puede asociar a cada elemento de \mathfrak{E} un valor de una variable simbólica ω ; eso le permite a Doob considerar $X(t)$ como una función cierta $x_t(\omega)$ de t y de una variable aleatoria ω cuya elección resume todas las intervenciones sucesivas del azar. Es una notación a menudo cómoda, aunque dé como un todo, nacido en un instante, lo que, para mí, es esencialmente un perpetuo devenir."

La teoría que más se desarrolló en los años posteriores a la publicación de los libros de Lévy y de Doob fue la de los procesos de Markov, en particular, la de los procesos de difusión. Uno de los problemas centrales que se atacó fue el de la construcción de procesos con trayectorias regulares en algún sentido.

La escuela rusa, con Kolmogorov a la cabeza, hizo las más importantes contribuciones, desarrollando ampliamente la teoría de los procesos de Markov. Uno de los más destacados en este periodo fue E. B. Dynkin, quien vinculó la teoría de los semigrupos de operadores con la de los procesos de Markov.

No se puede dejar de mencionar a G. A. Hunt por sus trabajos en la teoría de los procesos de Markov y su relación con la teoría del potencial y a K. L. Chung por sus aportaciones a la teoría de los procesos de Markov y por su incansable labor de difusión de la teoría de la probabilidad y de los procesos estocásticos en varios libros.

En los 60's comenzó a desarrollarse la escuela francesa, con Paul Meyer a la cabeza. En el año 1966 se publicó el primero de sus libros, titulado *Probabilités et Potentiel*, en el cual inició el desarrollo de lo que se llamaría después la teoría general de los procesos estocásticos. En 1967 se publicó otro libro suyo, titulado *Processus de Markov*, en el cual presenta algunos de los capítulos que incluiría la segunda parte de su primer libro. Sin embargo, el rumbo de las investigaciones cambió y no hubo tal segunda parte; en su lugar, en la segunda edición de su primer libro, publicada en 1975 y teniendo como coautor a Claude Dellacherie, hubo

una amplia reformulación, a tal grado que más bien esa segunda edición puede verse como un libro distinto. De esa segunda edición sí hubo una segunda parte, titulada *Théorie des martingales*.

Son interesantes los comentarios de Meyer y Dellacherie en el prefacio de la segunda edición de *Probabilités et Potentiel*. Para empezar, hacen la aclaración de que, a pesar del título, el libro contiene poco de probabilidad y nada de teoría del potencial. La razón para ir en otra dirección fue, como ellos lo dicen, la evolución tan rápida de la teoría, con lo cual se crearon nuevos conceptos y algunas cosas que parecían muy importantes antes, ya no lo eran tanto, mientras que aspectos que parecían sin interés, se volvieron importantes. Como enseñanza, dicen los autores: "Hemos tratado en particular de liberarnos de la actitud consistente en creer que un texto matemático presenta, para nuestra contemplación, verdades eternas, salidas del mundo de las ideas, y cuyo valor no puede sufrir alguna inflación... Pero las verdades existen no por el hecho que se impriman sobre buen papel, sino por el interés que se les da. Muchas verdades inmutables de 1966, están muertas, mientras que pequeños comentarios de 1966 aclaran ahora grandes partes de la teoría."

Esto es un ejemplo de cómo la matemática se va creando como producto social y no mediante "descubrimientos" de "verdades" que se encuentran escondidas en alguna parte, las cuales hay que ir encontrando.

En su libro *Processus stochastiques et mouvement brownien* (1948), Lévy trató el problema de la construcción de procesos mediante la solución de ecuaciones diferenciales estocásticas. La idea que formuló es que el incremento infinitesimal de un proceso $(X_t)_{t \in \mathbb{R}^+}$ en un tiempo t , lo cual denota por $dX(t)$ podría estar dado por una función que depende de t , dt , $X(t)$ y una o varias variables aleatorias de tipo familiar, como, por ejemplo, variables aleatorias independientes con distribución normal. En el caso del proceso de Wiener $(W_t)_{t \in \mathbb{R}^+}$, como el incremento $W_t - W_s$, para $s, t \in \mathbb{R}^+$ tales que $s < t$, tiene distribución normal con esperanza cero y varianza $t - s$, entonces en un incremento dt del tiempo, el incremento $dW(t)$ tiene distribución normal con esperanza cero y varianza dt , de manera que, denotando por ξ a una variable aleatoria con distribución normal estándar, el proceso de Wiener puede definirse por la ecuación diferencial estocástica $dW(t) = \xi \sqrt{dt}$.

Analizó algunos casos particulares de ecuaciones diferenciales estocásticas y planteó el problema general de resolver una ecuación diferencial estocástica del tipo siguiente:

$$dX(t) = f(t, X, Y, Z) dY(t) + g(t, X, Y, Z) dZ(t)$$

donde $(Y_t)_{t \in \mathbb{R}^+}$ y $(Z_t)_{t \in \mathbb{R}^+}$ son procesos estocásticos definidos previamente.

Hacía énfasis en que hay una dificultad para resolver una ecuación de este tipo cuando $Y(t)$ y $Z(t)$ no son funciones de variación acotada, por ejemplo cuando Y y Z representan las coordenadas de un movimiento browniano en el plano. Mostró entonces que una integral del tipo $\int_{t_0}^{\tau} f(t, X, Y, Z) dY(t)$, a la cual denominaba integral estocástica, puede evaluarse no siguiendo el método clásico de la integral de Riemann, sino considerando números τ_1, τ_2, \dots , elegidos al azar en el intervalo (t_0, τ) y tomando después, para cada $n \in \mathbb{N}$, los primeros n de esos números, ordenados del menor al mayor, para formar con ellos una suma de tipo Riemann. Consideró el caso del cálculo de áreas estocásticas, es decir, áreas de curvas

definidas por un proceso estocástico y mostró que las integrales estocásticas que resultan quedan bien definidas con el método mencionado.

Siguiendo un trabajo de S. Bernstein, consideró los procesos de Markov que son solución a ecuaciones diferenciales estocásticas de la forma:

$$dX(t) = A[t, X(t)] dt + B[t, X(t)] \xi \sqrt{dt}$$

donde ξ es una variable aleatoria auxiliar que cumple con determinadas condiciones.

Obsérvese que en el caso en que ξ es una variable aleatoria con distribución normal estándar, la ecuación anterior puede escribirse en la siguiente forma:

$$dX(t) = A[t, X(t)] dt + B[t, X(t)] dW(t)$$

Planteó entonces el problema general de la determinación de todos los procesos de Markov como soluciones de ecuaciones diferenciales estocásticas. Señalaba que ese problema no estaba resuelto de manera general, pero dio la solución para el caso de los procesos con incrementos independientes.

10. ESPERANZA CONDICIONAL

El concepto de Esperanza Condicional es básico en la Teoría de la Probabilidad Moderna. Su definición fue formulada por Kolmogorov en su libro de 1933 (*Foundations of the Theory of Probability*), en el cual estableció la formulación de la Teoría de la Probabilidad que prevalece hasta nuestros días. La definición general de la esperanza condicional está basada en el teorema de Radon Nikodym, publicado en el año 1930. Ese teorema fue la conclusión de la investigación que inició Lebesgue acerca de la condición para que una función sea una integral indefinida, la cual consiste en que esa función tiene que ser absolutamente continua. Radon continuó esa investigación y estableció el ahora llamado teorema de Radon-Nikodym para el caso de medidas en \mathbb{R}^n . El resultado de Nikodym fue mucho más general ya que lo formuló habiéndose desarrollado ya una teoría general de la medida. Únicamente pasaron 3 años para que Kolmogorov hiciera ver la importancia del resultado de Nikodym en la teoría de la probabilidad.

Ahora el concepto de esperanza condicional es una de las principales bases para el estudio e incluso la definición misma de los procesos estocásticos, ya que precisamente se constituyó en la herramienta central para tratar con variables aleatorias dependientes.

Sea (Ω, \mathcal{F}, P) un espacio de probabilidad y X una variable aleatoria de esperanza finita. Como primer paso para la definición de la esperanza condicional, queremos definir la esperanza condicional de X dada la ocurrencia de un evento A de probabilidad positiva. Para esto observemos que la función $P_A : \mathcal{F} \rightarrow \mathbf{R}$ definida por $P_A(B) = P(B | A)$ es una medida de probabilidad, lo cual motiva la siguiente definición:

Definición 13. *Sea A un evento de probabilidad positiva y X una variable aleatoria de esperanza finita. Se define la esperanza condicional de X dada la ocurrencia del evento A , $E[X | A]$, mediante la relación $E[X | A] = \int_{\Omega} X dP_A$.*

Obsérvese que $\int_{\Omega} X dP_A = \frac{1}{P(A)} \int_A X dP$, de manera que $E[X | A]$ puede interpretarse como el promedio de los valores de X en el conjunto A .

Con esta definición, $E[X | A]$ tiene como valor un número real; es simplemente una extensión natural de la probabilidad condicional. El concepto que queremos definir es mucho más general; podríamos continuar con una definición de $E[X | Y]$ donde Y es una variable aleatoria. Esto ya tiene su dificultad para el caso de una variable aleatoria continua ya que, en ese caso, tendríamos $P[Y = y] = 0$ para cualquier $y \in \mathbb{R}$, de manera que no podríamos utilizar la definición de probabilidad condicional para definir $E[X | Y = y]$. Vamos a proceder extendiendo la definición anterior al caso de una familia finita de eventos mutuamente excluyentes y de probabilidad positiva (esto equivale a definir $E[X | Y]$ para el caso en que Y es una variable aleatoria discreta cuyo conjunto de posibles valores es finito). Veremos entonces que la esperanza condicional queda caracterizada por determinadas propiedades, las cuales permitirán definir el concepto general.

Proposición 1. *Sea X una variable aleatoria de esperanza finita y A_1, A_2, \dots, A_n eventos de probabilidad positiva, mutuamente excluyentes y tales que $\bigcup_{k=1}^n A_k = \Omega$. Definamos la función $Z : \Omega \rightarrow \mathbb{R}$ mediante la relación $Z(\omega) = E[X | A_k]$ si $\omega \in A_k$. Entonces,*

1. Z es una variable aleatoria medible con respecto a la σ -álgebra generada por los eventos A_1, A_2, \dots, A_n .
2. Z tiene esperanza finita.
3. $\int_B Z dP = \int_B X dP$ para cualquier evento $B \in \sigma(A_1, A_2, \dots, A_n)$
4. Z es la única variable aleatoria con las propiedades i, ii y iii.

El teorema de Radon-Nikodym permite extender la idea anterior, mediante el siguiente resultado:

Teorema 5. *Sea X una variable aleatoria de esperanza finita y \mathcal{G} una sub σ -álgebra de \mathcal{F} , existe entonces una variable aleatoria Z , medible con respecto a \mathcal{G} , de esperanza finita y tal que $\int_B Z dP = \int_B X dP$ para cualquier conjunto $B \in \mathcal{G}$. Además, si Y es otra variable aleatoria con las mismas propiedades que Z , entonces $Y = Z$ c.s.*

Definición 14. *Sea X es una variable aleatoria de esperanza finita y \mathcal{G} una sub σ -álgebra de \mathcal{F} . Se dice que Z es una versión de la **esperanza condicional de X con respecto a \mathcal{G}** y se escribe $E[X | \mathcal{G}] = Z$ c.s. si Z una variable aleatoria medible con respecto a \mathcal{G} , de esperanza finita y tal que $\int_B Z dP = \int_B X dP$ para cualquier evento $B \in \mathcal{G}$.*

La siguiente proposición muestra que la esperanza condicional tiene propiedades similares a las de la esperanza no condicional. Se muestra también que tiene las propiedades que podrían esperarse con una buena definición, por corresponder a la idea intuitiva del concepto. Finalmente, se muestran otras propiedades específicas de la esperanza condicional, las cuales no resultan evidentes a partir de la idea intuitiva.

Proposición 2. *Sean X y Y dos variables aleatorias de esperanza finita, \mathcal{G} y \mathcal{H} dos sub-álgebras de \mathcal{F} y c una constante. Se tienen entonces las siguientes propiedades:*

1. $E[c | \mathcal{G}] = c$.
2. $E[cX | \mathcal{G}] = cE[X | \mathcal{G}]$.

3. $E[X + Y | \mathcal{G}] = E[X | \mathcal{G}] + E[Y | \mathcal{G}]$.
4. Si X es \mathcal{G} -medible, entonces $E[X | \mathcal{G}] = X$.
5. $E[E[X | \mathcal{G}]] = E[X]$.
6. Si $\mathcal{G} \subset \mathcal{H}$, entonces $E[E[X | \mathcal{H}] | \mathcal{G}] = E[X | \mathcal{G}]$.
7. Si X es independiente de \mathcal{G} , entonces $E[X | \mathcal{G}] = E[X]$.
8. Si $X \leq Y$, entonces $E[X | \mathcal{G}] \leq E[Y | \mathcal{G}]$.
9. $|E[X | \mathcal{G}]| \leq E[|X| | \mathcal{G}]$.

Si \mathcal{G} es una σ -álgebra de subconjuntos Ω y $\mathcal{G} \subset \mathfrak{F}$ diremos que \mathcal{G} es una **sub σ -álgebra** de \mathfrak{F} .

Sea \mathcal{G} una sub σ -álgebra de \mathfrak{F} . Se dice que una función $X : \Omega \rightarrow \mathbb{R}$ es una **\mathcal{G} -variable aleatoria** si $[X \leq x] \equiv \{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{G}$ para cualquier $x \in \mathbb{R}$.

Para cada $t \geq 0$ consideremos una σ -álgebra $\mathfrak{F}_t \subset \mathfrak{F}$, la cual representa la historia hasta el tiempo t . En particular se tiene $\mathfrak{F}_s \subset \mathfrak{F}_t$ para $s < t$. A la familia $(\mathfrak{F}_t)_{t \geq 0}$ se le llama una **filtración**.

Se dice que un proceso estocástico $(X_t)_{t \geq 0}$ está **adaptado a la filtración** $(\mathfrak{F}_t)_{t \geq 0}$ si, para cada $t \geq 0$, la función $X_t : \Omega \rightarrow \mathbb{R}$ es una \mathfrak{F}_t -variable aleatoria.

Un **\mathfrak{F}_t -proceso de Wiener** o \mathfrak{F}_t -movimiento browniano estándar es un proceso $(W_t)_{t \geq 0}$, el cual satisface las siguientes propiedades:

1. $(W_t)_{t \geq 0}$ está adaptado a la filtración $(\mathfrak{F}_t)_{t \geq 0}$.
2. $W_0 = 0$ con probabilidad 1.
3. Con probabilidad 1, las trayectorias $t \rightarrow W_t$ son funciones continuas.
4. Dados $0 \leq s < t$, $W_t - W_s \sim N(0, t - s)$.
5. Dados $0 \leq s < t$, $W_t - W_s$ es independiente de \mathfrak{F}_s .

La existencia de un espacio de probabilidad $(\Omega, \mathfrak{F}, P)$, una filtración $(\mathfrak{F}_t)_{t \geq 0}$ y un \mathfrak{F}_t -proceso de Wiener estándar definido sobre ese espacio, fué, como ya lo mencionamos, demostrada básicamente por N. Wiener en el año 1924, aunque actualmente no se utiliza la metodología utilizada por él y existen diferentes métodos para probar tal existencia.

11. MARTINGALAS

Otro concepto central en la teoría de los procesos estocásticos es el de Martingala.

Se dice que un proceso $(M_t)_{t \geq 0}$ es una **martingala** si se satisfacen las siguientes tres propiedades:

1. El proceso $(M_t)_{t \geq 0}$ está adaptado a la filtración $(\mathfrak{F}_t)_{t \geq 0}$.
2. Para cada $t \geq 0$, M_t es una variable aleatoria de esperanza finita.
3. Dados $s < t$, se tiene $E[M_t | \mathfrak{F}_s] = M_s$.

Una propiedad muy importante de cualquier martingala, es la siguiente:

Sea $(M_t)_{t \geq 0}$ una martingala y sean $s < t$, entonces:

$$\begin{aligned} E[(M_t - M_s)^2 | \mathfrak{F}_s] &= E[M_t^2 - 2M_s M_t + M_s^2 | \mathfrak{F}_s] \\ &= E[M_t^2 + M_s^2 | \mathfrak{F}_s] - 2M_s E[M_t | \mathfrak{F}_s] \\ &= E[M_t^2 + M_s^2 | \mathfrak{F}_s] - 2M_s^2 = E[M_t^2 + M_s^2 | \mathfrak{F}_s] - E[2M_s^2 | \mathfrak{F}_s] \\ &= E[M_t^2 - M_s^2 | \mathfrak{F}_s] \end{aligned}$$

Se dice que un proceso $(M_t)_{t \geq 0}$ es una **martingala local** si existe una sucesión no decreciente de tiempos aleatorios $(T_n)_{n \in \mathbb{N}}$ tales que:

1. $\lim_{n \rightarrow \infty} T_n = \infty$
2. El proceso $(M_t^{T_n})_{t \geq 0}$, definido por:

$$M_t^{T_n} = \begin{cases} M_t & \text{si } t < T_n \\ M_{T_n} & \text{si } t \geq T_n \end{cases}$$

es una martingala.

12. CONVERGENCIA DE VARIABLES ALEATORIAS

Sea $(X_n)_{n \geq 1}$ una sucesión de variables aleatorias.

1. Se dice que la sucesión X_n **converge en probabilidad** si existe una variable aleatoria X tal que $\lim_{n \rightarrow \infty} P[|X_n - X| > \varepsilon] = 0$ para cualquier $\varepsilon > 0$. En ese caso se escribe $\lim_{n \rightarrow \infty} (P)X_n = X$.
2. Se dice que la sucesión X_n **converge con probabilidad 1**, o casi seguramente, si existe una variable aleatoria X y $A \in \mathfrak{F}$ tales que $P(A) = 1$ y $\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)$ para cualquier $\omega \in A$.
3. Si las variables aleatorias X_n tienen esperanza y varianzas finitas, se dice que la sucesión X_n **converge en L^2** , o en media cuadrática, si existe una variable aleatoria X , de esperanza y varianzas finitas, tal que $\lim_{n \rightarrow \infty} E[(X_n - X)^2] = 0$.

Proposición 3. *Sea X_n una sucesión de variables aleatorias que converge a la variable aleatoria X con probabilidad 1 o en media cuadrática, entonces X_n converge a X en probabilidad.*

13. PROPIEDADES DEL PROCESO DE WIENER

Proposición 4. *Si $(W_t)_{t \in \mathbb{R}^+}$ es un \mathfrak{F}_t -proceso de Wiener estándar, entonces los procesos $(W_t)_{t \in \mathbb{R}^+}$ y $(W_t^2 - t)_{t \in \mathbb{R}^+}$ y $(e^{W_t - \frac{1}{2}t})_{t \in \mathbb{R}^+}$ son martingalas.*

Demostración

Para $s < t$, se tiene:

$$E[W_t - W_s | \mathfrak{F}_s] = E[W_t - W_s] = 0$$

$$E [W_t^2 - W_s^2 \mid \mathfrak{F}_s] = E [(W_t - W_s)^2 \mid \mathfrak{F}_s] = E [(W_t - W_s)^2] = t - s$$

Así que, $E [W_t \mid \mathfrak{F}_s] = W_s$ y $E [W_t^2 - t \mid \mathfrak{F}_s] = W_s^2 - s$.

$$\begin{aligned} E \left[e^{W_t - \frac{1}{2}t} \mid \mathfrak{F}_t \right] &= E \left[e^{W_t - W_s + W_s - \frac{1}{2}t} \mid \mathfrak{F}_t \right] = e^{W_s - \frac{1}{2}t} E \left[e^{W_t - W_s} \right] \\ &= e^{W_s - \frac{1}{2}t} \int_{-\infty}^{\infty} e^y \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{1}{2(t-s)}y^2} dy = e^{W_s - \frac{1}{2}t} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{1}{2(t-s)}(y^2 - 2(t-s)y)} dy \\ &= e^{W_s - \frac{1}{2}t} e^{\frac{1}{2}(t-s)} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{1}{2(t-s)}(y-(t-s))^2} dy = e^{W_s - \frac{1}{2}s} \end{aligned}$$

Teorema 6. $\lim_{\|P\| \rightarrow 0} \sum_{k=1}^n (W_{t_k} - W_{t_{k-1}})^2 = b - a$ en L^2

donde $P = \{a = t_0 < t_1 < \dots < t_n = b\}$ es una partición del intervalo $[a, b]$.

Demostración

Recuérdese que si X es una variable aleatoria con distribución normal de esperanza 0 y varianza σ^2 , entonces $E [X^{2m}] = 1 \cdot 3 \cdot 5 \dots (2m - 1)\sigma^{2m}$ para cualquier $m \in \mathbb{N}$.

Sea $\Delta_k = (W_{t_k} - W_{t_{k-1}})^2$. Se tiene entonces,

$$E [\Delta_k] = t_k - t_{k-1}$$

$$\begin{aligned} V [\Delta_k] &= E [(\Delta_k)^2] - (E [\Delta_k])^2 = E [(W_{t_k} - W_{t_{k-1}})^4] - (t_k - t_{k-1})^2 \\ &= 3(t_k - t_{k-1})^2 - (t_k - t_{k-1})^2 = 2(t_k - t_{k-1})^2 \end{aligned}$$

Por lo tanto,

$$E \left[\sum_{k=1}^n (W_{t_k} - W_{t_{k-1}})^2 \right] = E \left[\sum_{k=1}^n \Delta_k \right] = \sum_{k=1}^n (t_k - t_{k-1}) = b - a$$

Además, como $\Delta_1, \Delta_2, \dots, \Delta_n$ son variables aleatorias independientes,

$$\begin{aligned} E \left[\left(\sum_{k=1}^n (W_{t_k} - W_{t_{k-1}})^2 - (b - a) \right)^2 \right] &= V \left[\sum_{k=1}^n \Delta_k \right] = \sum_{k=1}^n V [\Delta_k] \\ &= 2 \sum_{k=1}^n (t_k - t_{k-1})^2 \leq 2 \|P\| \sum_{k=1}^n (t_k - t_{k-1}) = 2(b - a) \|P\| \end{aligned}$$

Por lo tanto,

$$\lim_{\|P\| \rightarrow 0} E \left[\left(\sum_{k=1}^n (W_{t_k} - W_{t_{k-1}})^2 - (b - a) \right)^2 \right] = 0$$

Al límite $\lim_{\|P\| \rightarrow 0} \sum_{k=1}^n (W_{t_k} - W_{t_{k-1}})^2$ se le llama la **variación cuadrática de (W_t)** en el intervalo $[a, b]$.

Corolario 1. Las trayectorias del proceso de Wiener no son de variación acotada.

El hecho de que las trayectorias del proceso de Wiener no sean de variación acotada se refleja en problemas del siguiente tipo:

Proposición.

1. $\lim_{\|P\| \rightarrow 0} (P) \sum_{k=1}^n W_{t_{k-1}} [W_{t_k} - W_{t_{k-1}}] = \frac{1}{2} W_t^2 - \frac{1}{2} t$
2. $\lim_{\|P\| \rightarrow 0} (P) \sum_{k=1}^n W_{t_k} [W_{t_k} - W_{t_{k-1}}] = \frac{1}{2} W_t^2 + \frac{1}{2} t$

donde $P = \{0 = t_0 < t_1 < \dots < t_n = t\}$ es una partición del intervalo $[0, t]$.

Demostración

$$\begin{aligned}
\sum_{k=1}^n W_{t_{k-1}} [W_{t_k} - W_{t_{k-1}}] &= \sum_{k=1}^n [W_{t_{k-1}} W_{t_k} - W_{t_{k-1}}^2] \\
&= \frac{1}{2} \left[\sum_{k=1}^n 2W_{t_{k-1}} W_{t_k} - \sum_{k=1}^n W_{t_{k-1}}^2 - \sum_{k=1}^n W_{t_{k-1}}^2 \right] \\
&= \frac{1}{2} \left[\sum_{k=1}^n 2W_{t_{k-1}} W_{t_k} - \sum_{k=1}^n W_{t_{k-1}}^2 - \sum_{k=1}^n W_{t_k}^2 + W_t^2 \right] \\
&= \frac{1}{2} \left[W_t^2 - \sum_{k=1}^n (W_{t_k} - W_{t_{k-1}})^2 \right] = \frac{1}{2} W_t^2 - \frac{1}{2} \sum_{k=1}^n (W_{t_k} - W_{t_{k-1}})^2
\end{aligned}$$

Por lo tanto,

$$\begin{aligned}
&\lim_{\|P\| \rightarrow 0} (P) \sum_{k=1}^n W_{t_{k-1}} [W_{t_k} - W_{t_{k-1}}] \\
&= \frac{1}{2} W_t^2 - \frac{1}{2} \lim_{\|P\| \rightarrow 0} (P) \sum_{k=1}^n (W_{t_k} - W_{t_{k-1}})^2 = \frac{1}{2} W_t^2 - \frac{1}{2} t
\end{aligned}$$

$$\begin{aligned}
\sum_{k=1}^n W_{t_k} [W_{t_k} - W_{t_{k-1}}] &= \sum_{k=1}^n [W_{t_k}^2 - W_{t_k} W_{t_{k-1}}] \\
&= \frac{1}{2} \left[-\sum_{k=1}^n 2W_{t_{k-1}} W_{t_k} + \sum_{k=1}^n W_{t_k}^2 + \sum_{k=1}^n W_{t_k}^2 \right] \\
&= \frac{1}{2} \left[-\sum_{k=1}^n 2W_{t_{k-1}} W_{t_k} + \sum_{k=1}^n W_{t_k}^2 + \sum_{k=1}^n W_{t_{k-1}}^2 + W_t^2 \right] \\
&= \frac{1}{2} \left[W_t^2 + \sum_{k=1}^n (W_{t_k} - W_{t_{k-1}})^2 \right] = \frac{1}{2} W_t^2 + \frac{1}{2} \sum_{k=1}^n (W_{t_k} - W_{t_{k-1}})^2
\end{aligned}$$

Por lo tanto,

$$\begin{aligned}
&\lim_{\|P\| \rightarrow 0} (P) \sum_{k=1}^n W_{t_k} [W_{t_k} - W_{t_{k-1}}] \\
&= \frac{1}{2} W_t^2 + \frac{1}{2} \lim_{\|P\| \rightarrow 0} (P) \sum_{k=1}^n (W_{t_k} - W_{t_{k-1}})^2 = \frac{1}{2} W_t^2 + \frac{1}{2} t
\end{aligned}$$

Como puede verse, los dos límites de la proposición anterior son distintos por el hecho de que la variación cuadrática de (W_t) es distinta de cero.

14. ECUACIONES INTEGRALES ESTOCÁSTICAS

Una ecuación integral estocástica es una relación de la forma:

$$X_t = X_0 + \int_0^t \mu(X_t, t)dt + \int_0^t \sigma(X_t, t)dW_t$$

donde $(X_t)_{t \geq 0}$ es un proceso estocástico, μ y σ son dos funciones reales definidas en \mathbb{R}^2 y $(W_t)_{t \geq 0}$ es un proceso de Wiener. Ecuaciones de este tipo se presentan en una diversidad de situaciones. La idea general es que si un sistema dinámico es modelado por una ecuación diferencial ordinaria $\frac{dX}{dt} = \mu(t, X)$ y dicho sistema es perturbado por la presencia de un ruido aleatorio, entonces la modelación estaría dada por una ecuación de la forma $\frac{dX}{dt} = \mu(t, X) + \sigma(t, X)\xi(t)$, en donde ξ representa la perturbación aleatoria. El ruido ξ se modela usualmente como la "derivada" de un proceso de Wiener, de manera que la última ecuación se escribiría en la forma:

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

μ es llamada el **arrastre** (drift) y σ la **volatilidad**.

Se sabe que con probabilidad 1, **las funciones $t \rightarrow W_t$ no son diferenciables**, así que la ecuación anterior es en realidad una ecuación integral:

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

También sabemos que, con probabilidad 1, las funciones $t \rightarrow W_t$ no son de variación acotada en ningún intervalo finito. No es posible entonces definir la integral estocástica $\int_0^t \sigma(X_s, s)dW_s$ como una integral de Stieltjes.

15. INTEGRALES ESTOCÁSTICAS

A pesar de las dificultades planteadas para definir la integral con respecto al proceso de Wiener, K. Ito (Stochastic Integral, 1944) logró darle un sentido preciso a una integral de la forma $\int_0^t Z_s dW_s$ para una familia amplia de procesos $(Z_t)_{t \geq 0}$, entre los que se incluyen a todos los procesos adaptados y continuos.

Son dos las propiedades básicas del proceso de Wiener que permitieron a Ito definir la integral estocástica $\int_0^t Z_s dW_s$.

1. El proceso de Wiener $(W_t)_{t \geq 0}$ es una martingala.
Para $s < t$, se tiene, $E[W_t - W_s | \mathfrak{F}_s] = E[W_t - W_s] = 0$. Por lo tanto, $E[W_t | \mathfrak{F}_s] = W_s$.
2. El proceso $W_t^2 - t$ es una martingala.

15.1. Definición de la integral estocástica con respecto al proceso de Wiener.

Si $H_t(\varpi) = I_{(a,b] \times F}(t, \varpi)$ donde $F \in \mathfrak{F}_t$, se define:

$$N_t = \int_0^t H_s dW_s = \begin{cases} 0 & \text{si } t < a \\ (W_t - W_a)I_F & \text{si } a \leq t \leq b \\ (W_b - W_a)I_F & \text{si } t \geq b \end{cases}$$

Propiedades:

1. $(N_t)_{t \geq 0}$ es un proceso continuo.
2. $(N_t)_{t \geq 0}$ es una martingala.
3. $N_t^2 - \int_0^t H_s^2 ds$ es una martingala.

En particular, se tiene:

$$E [N_t^2] = E \left[\int_0^t H_s^2 ds \right]$$

Esto define una isometría entre dos espacios L^2 , la cual permite extender la integral estocástica a todos los procesos que se puedan generar mediante procesos del tipo de los procesos H_t definidos arriba. La integral estocástica queda así definida para una familia bastante amplia de procesos, la cual incluye a todos los procesos adaptados y continuos $(Z_t)_{t \geq 0}$ tales que $E \left[\int_0^t Z_s^2 ds \right] < \infty$.

Sea $(Z_t)_{t \geq 0}$ un proceso adaptado y continuo tal que $E \left[\int_0^t Z_s^2 ds \right] < \infty$, la integral estocástica $M_t = \int_0^t Z_s dW_s$ está entonces bien definida y tiene las siguientes propiedades:

1. $M_0 = 0$
2. $(M_t)_{t \geq 0}$ es un proceso continuo.
3. $(M_t)_{t \geq 0}$ es una martingala.
4. $M_t^2 - \int_0^t Z_s^2 ds$ es una martingala.
5. $\int_0^t Z_s dW_s = \lim(P) \sum Z_{t_{k-1}} (W_{t_k} - W_{t_{k-1}})$
6. $\int_0^t (aZ_s + bY_s) dW_s = a \int_0^t Z_s dW_s + b \int_0^t Y_s dW_s$

El proceso $\int_0^t Z_s^2 ds$ es llamado el **compensador de la martingala** M_t y se le denota por $\langle M, M \rangle_t$. Obsérvese que se trata de un proceso adaptado y no decreciente.

Proposición 5. *El proceso $A_t = \int_0^t Z_s^2 ds$ es el único proceso adaptado, no decreciente, nulo en cero y continuo tal que $M_t^2 - A_t$ es una martingala.*

Sea $(Z_t)_{t \geq 0}$ un proceso adaptado y continuo y $M_t = \int_0^t Z_s dW_s$. Si $(Y_t)_{t \geq 0}$ es otro proceso adaptado y continuo, se define:

$$\int_0^t Y_s dM_s = \int_0^t Y_s Z_s dW_s$$

El proceso $N_t = \int_0^t Y_s dM_s$ está entonces bien definido y tiene las siguientes propiedades:

1. $N_0 = 0$
2. $(N_t)_{t \geq 0}$ es un proceso continuo.

3. $(N_t)_{t \geq 0}$ es una martingala local.
4. Existe un único proceso $(A_t)_{t \geq 0}$ adaptado, no decreciente, nulo en cero y continuo tal que $M_t^2 - A_t$ es una martingala local.
5. $\int_0^t Y_s dM_s = \lim(P) \sum Y_{t_{k-1}} (M_{t_k} - M_{t_{k-1}})$
6. $\int_0^t (aX_s + bY_s) dM_s = a \int_0^t X_s dM_s + b \int_0^t Y_s dM_s$
7. Si $(Z'_t)_{t \geq 0}$ es otro proceso adaptado y continuo y $M'_t = \int_0^t Z'_s dW_s$, entonces:

$$\int_0^t Y_s d(M_s + M'_s) = \int_0^t Y_s dM_s + \int_0^t Y_s dM'_s$$

El siguiente resultado es clave para poder calcular integrales estocásticas.

Teorema 7. Sea $(Z_t)_{t \geq 0}$ un proceso adaptado y continuo y sea $M_t = \int_0^t Z_s dW$, entonces:

$$\lim_{\|P\| \rightarrow 0} (P) \sum_{k=1}^n (M_{t_k} - M_{t_{k-1}})^2 = M_t^2 - 2 \int_0^t M_s dM_s$$

en donde $P = \{0 = t_0 < t_1 < \dots < t_n = t\}$ es una partición del intervalo $[0, t]$.

Demostración

$$\begin{aligned} \sum_{k=1}^n (M_{t_k} - M_{t_{k-1}})^2 &= \sum_{k=1}^n (M_{t_k}^2 + M_{t_{k-1}}^2 - 2M_{t_{k-1}}M_{t_k}) \\ &= \sum_{k=1}^n (M_{t_k}^2 - M_{t_{k-1}}^2 + 2M_{t_{k-1}}^2 - 2M_{t_{k-1}}M_{t_k}) \\ &= \sum_{k=1}^n [M_{t_k}^2 - M_{t_{k-1}}^2 - 2M_{t_{k-1}}(M_{t_k} - M_{t_{k-1}})] \\ &= M_t^2 - 2 \sum_{k=1}^n M_{t_{k-1}} (M_{t_k} - M_{t_{k-1}}) \end{aligned}$$

Por lo tanto,

$$\begin{aligned} \lim_{\|P\| \rightarrow 0} (P) \sum_{k=1}^n (M_{t_k} - M_{t_{k-1}})^2 &= M_t^2 - 2 \lim_{\|P\| \rightarrow 0} (P) \sum_{k=1}^n M_{t_{k-1}} (M_{t_k} - M_{t_{k-1}}) \\ &= M_t^2 - 2 \int_0^t M_s dM_s \end{aligned}$$

■

Corolario 2. El proceso $\langle M, M \rangle_t = M_t^2 - 2 \int_0^t M_s dM_s$ es adaptado, no decreciente, nulo en cero y continuo.

Corolario 3. $\int_0^t M_s dM_s = \frac{1}{2} M_t^2 - \frac{1}{2} \langle M, M \rangle_t$

Corolario 4. $\lim_{\|P\| \rightarrow 0} (P) \sum_{k=1}^n (M_{t_k} - M_{t_{k-1}})^2 = \langle M, M \rangle_t$

donde $P = \{0 = t_0 < t_1 < \dots < t_n = t\}$ es una partición del intervalo $[0, t]$.

Al límite $\lim_{\|P\| \rightarrow 0} (P) \sum_{k=1}^n (M_{t_k} - M_{t_{k-1}})^2$ se le llama la **variación cuadrática de (M_t)** .

Al igual que en el caso del proceso de Wiener, los límites $\lim_{\|P\| \rightarrow 0} (P) \sum_{k=1}^n M_{t_{k-1}} [M_{t_k} - M_{t_{k-1}}]$

y $\lim_{\|P\| \rightarrow 0} (P) \sum_{k=1}^n M_{t_k} [M_{t_k} - M_{t_{k-1}}]$ no coinciden. En efecto:

$$\begin{aligned}
\sum_{k=1}^n M_{t_k} [M_{t_k} - M_{t_{k-1}}] &= \sum_{k=1}^n [M_{t_k}^2 - M_{t_k} M_{t_{k-1}}] \\
&= \frac{1}{2} \left[-\sum_{k=1}^n 2M_{t_{k-1}} M_{t_k} + \sum_{k=1}^n M_{t_k}^2 + \sum_{k=1}^n M_{t_k}^2 \right] \\
&= \frac{1}{2} \left[-\sum_{k=1}^n 2M_{t_{k-1}} M_{t_k} + \sum_{k=1}^n M_{t_k}^2 + \sum_{k=1}^n M_{t_{k-1}}^2 + M_t^2 \right] \\
&= \frac{1}{2} \left[M_t^2 + \sum_{k=1}^n (M_{t_k} - M_{t_{k-1}})^2 \right] = \frac{1}{2} M_t^2 + \frac{1}{2} \sum_{k=1}^n (M_{t_k} - M_{t_{k-1}})^2
\end{aligned}$$

Por lo tanto:

$$\begin{aligned}
\lim_{\|P\| \rightarrow 0} (P) \sum_{k=1}^n M_{t_k} [M_{t_k} - M_{t_{k-1}}] &= \frac{1}{2} M_t^2 + \frac{1}{2} \lim_{\|P\| \rightarrow 0} (P) \sum_{k=1}^n (M_{t_k} - M_{t_{k-1}})^2 \\
&= \frac{1}{2} W_t^2 + \frac{1}{2} \langle M, M \rangle_t
\end{aligned}$$

La diferencia de estos límites es consecuencia de que la variación cuadrática de (M_t) es distinta de cero.

16. FÓRMULA DE INTEGRACIÓN POR PARTES

Sea $(Z_t)_{t \geq 0}$ un proceso adaptado y continuo $M_t = \int_0^t Z_s dW_s$.

La fórmula $\int_0^t M_s dM_s = \frac{1}{2} M_t^2 - \frac{1}{2} \langle M, M \rangle_t$ puede escribirse de la siguiente manera:

$$M_t^2 = 2 \int_0^t M_s dM_s + \langle M, M \rangle_t$$

Sea $(Y_t)_{t \geq 0}$ otro proceso adaptado y continuo y sea $N_t = \int_0^t Y_s dW$, entonces:

$$M_t + N_t = \int_0^t (Z_s + Y_s) dW_s$$

Así que:

$$\begin{aligned}
\langle M + N, M + N \rangle_t &= \int_0^t (Z_s + Y_s)^2 ds = \int_0^t Z_s^2 ds + 2 \int_0^t Z_s Y_s ds + \int_0^t Y_s^2 ds \\
&= \langle M, M \rangle_t + \langle N, N \rangle_t + 2 \int_0^t Z_s Y_s ds
\end{aligned}$$

Además:

$$\begin{aligned}
M_t^2 + 2M_t N_t + N_t^2 &= (M_t + N_t)^2 \\
&= 2 \int_0^t (M_s + N_s) d(M_s + N_s) + \langle M + N, M + N \rangle_t \\
&= 2 \int_0^t M_s dM_s + 2 \int_0^t M_s dN_s + 2 \int_0^t N_s dM_s + 2 \int_0^t N_s dN_s + \langle M + N, M + N \rangle_t \\
&= M_t^2 - \langle M, M \rangle_t + N_t^2 - \langle N, N \rangle_t + 2 \int_0^t M_s dN_s + 2 \int_0^t N_s dM_s + \langle M + N, M + N \rangle_t \\
&= M_t^2 + N_t^2 + 2 \int_0^t M_s dN_s + 2 \int_0^t N_s dM_s + 2 \int_0^t Z_s Y_s ds
\end{aligned}$$

Así que:

$$M_t N_t = \int_0^t M_s dN_s + \int_0^t N_s dM_s + \int_0^t Z_s Y_s ds$$

En particular, se tiene que $M_t N_t - \int_0^t Z_s Y_s ds$ es una martingala (local).

Tiene entonces sentido definir:

$$\langle M, N \rangle_t = \int_0^t Z_s Y_s ds$$

Obsérvese que así definido, el proceso $B_t = \langle M, N \rangle_t$ es el único proceso adaptado, de variación acotada, nulo en cero y continuo tal que $M_t N_t - B_t$ es una martingala (local).

De esta manera, se puede escribir la **fórmula de integración por partes para martingalas (locales)**:

$$M_t N_t = \int_0^t M_s dN_s + \int_0^t N_s dM_s + \langle M, N \rangle_t$$

17. SEMIMARTINGALAS

Sea $(Z_t)_{t \geq 0}$ un proceso adaptado y continuo y $M_t = \int_0^t Z_s dW_s$. Sea, además, $(A_t)_{t \geq 0}$ un proceso adaptado, de variación acotada, nulo en cero y continuo. Entonces, con probabilidad 1, las integrales $\int_0^t M_s dA_s$ existen. Por lo tanto, por la fórmula de integración por partes para las integrales de Riemann-Stieltjes, se tiene que, con probabilidad 1, las integrales $\int_0^t A_s dM_s$ existen como integrales de Riemann-Stieltjes y se tiene:

$$\int_0^t A_s dM_s = A_t M_t - \int_0^t M_s dA_s$$

Las integrales $\int_0^t A_s dM_s$ también están bien definidas como integrales de Ito.

Teorema 8. *La integral $\int_0^t A_s dM_s$ definida como integral de Riemann-Stieltjes coincide con su definición como integral de Ito.*

Demostración

Ambas integrales se obtienen como el límite $\lim(P) \sum A_{t_{k-1}} (M_{t_k} - M_{t_{k-1}})$. ■

Sea $(X_t)_{t \geq 0}$ un proceso de la forma $X_t = A_t + M_t$, en donde $(A_t)_{t \geq 0}$ un proceso adaptado, de variación acotada, nulo en cero y continuo y $M_t = \int_0^t Z_s dW_s$, en donde $(Z_t)_{t \geq 0}$ es un proceso adaptado y continuo. Un proceso de este tipo es llamado una **semimartingala**.

Si $(Z_t)_{t \geq 0}$ es un proceso adaptado y continuo, se define:

$$\int_0^t Z_s dX_s = \int_0^t Z_s dA_s + \int_0^t Z_s dM_s$$

Se tiene entonces:

$$\begin{aligned}
X_t^2 &= A_t^2 + M_t^2 + 2A_tM_t = 2 \int_0^t A_s dA_s + 2 \int_0^t M_s dM_s + \langle M, M \rangle_t + 2 \int_0^t A_s dM_s + 2 \int_0^t M_s dA_s \\
&= 2 \int_0^t (A_s + M_s) dA_s + 2 \int_0^t (A_s + M_s) dM_s + \langle M, M \rangle_t \\
&= 2 \int_0^t X_s dA_s + 2 \int_0^t X_s dM_s + \langle M, M \rangle_t = 2 \int_0^t X_s dX_s + \langle M, M \rangle_t
\end{aligned}$$

De manera que si definimos $\langle X, X \rangle_t = \langle M, M \rangle_t$, se tiene:

$$X_t^2 = 2 \int_0^t X_s dX_s + \langle X, X \rangle_t$$

Ahora, si $Y_s = B_t + N_t$, en donde $(B_t)_{t \geq 0}$ un proceso adaptado, de variación acotada, nulo en cero y continuo y $N_t = \int_0^t Z'_s dW_s$, en donde $(Z'_t)_{t \geq 0}$ es un proceso adaptado y continuo, se tiene:

$$X_t + Y_t = (A_t + B_t) + (M_t + N_t)$$

Así que:

$$\begin{aligned}
\langle X + Y, X + Y \rangle_t &= \langle M + N, M + N \rangle_t = \langle M, M \rangle_t + \langle N, N \rangle_t + 2 \langle M, N \rangle_t \\
&= \langle X, X \rangle_t + \langle Y, Y \rangle_t + 2 \langle M, N \rangle_t
\end{aligned}$$

Además:

$$\begin{aligned}
X_t^2 + 2X_tY_t + Y_t^2 &= (X_t + Y_t)^2 = 2 \int_0^t (X_s + Y_s) d(X_s + Y_s) + \langle X + Y, X + Y \rangle_t \\
&= 2 \int_0^t X_s dX_s + 2 \int_0^t X_s dY_s + 2 \int_0^t Y_s dX_s + 2 \int_0^t Y_s dY_s + \langle X + Y, X + Y \rangle_t \\
&= X_t^2 - \langle X, X \rangle_t + Y_t^2 - \langle Y, Y \rangle_t + 2 \int_0^t X_s dY_s + 2 \int_0^t Y_s dX_s + \langle X, X \rangle_t + \langle Y, Y \rangle_t + 2 \langle M, N \rangle_t \\
&= X_t^2 + Y_t^2 + 2 \int_0^t X_s dY_s + 2 \int_0^t Y_s dX_s + 2 \langle M, N \rangle_t
\end{aligned}$$

Así que:

$$X_tY_t = \int_0^t X_s dY_s + \int_0^t Y_s dX_s + 2 \langle M, N \rangle_t$$

De manera que si definimos $\langle X, Y \rangle_t = \langle M, N \rangle_t$, se tiene la **fórmula de integración por partes para semimartingalas**:

$$X_tY_t = \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle X, Y \rangle_t$$

18. FÓRMULA DE ITO

La fórmula de integración por partes para semimartingalas permite calcular:

$$X_t^3 = 3 \int_0^t X_s^2 dX_s + 3 \int_0^t X_s d \langle X, X \rangle_s$$

$$X_t^4 = 4 \int_0^t X_s^3 dX_s + 6 \int_0^t X_s^2 d \langle X, X \rangle_s$$

⋮

$$X_t^n = n \int_0^t X_s^{n-1} dX_s + \frac{1}{2} n(n-1) \int_0^t X_s^{n-2} d\langle X, X \rangle_s$$

Gracias a la linealidad de la integral, se obtiene, para un polinomio p :

$$p(X_t) = p(X_0) + \int_0^t p'(X_s) dX_s + \frac{1}{2} \int_0^t p''(X_s) d\langle X, X \rangle_s$$

Mediante un proceso de límite, esta fórmula se extiende a todas las funciones $F : \mathbb{R} \rightarrow \mathbb{R}$ de clase C^2 :

$$F(X_t) = F(X_0) + \int_0^t F'(X_s) dX_s + \frac{1}{2} \int_0^t F''(X_s) d\langle X, X \rangle_s$$

Este resultado es conocido como la **fórmula de Ito**.

Se puede extender la fórmula de Ito al caso multidimensional. Para el caso de dos semimartingalas, (X_t) y (Y_t) y una función $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ de clase C^2 , la fórmula es como sigue:

$$\begin{aligned} F(X_t, Y_t) &= F(X_0, Y_0) + \int_0^t F_x(X_s, Y_s) dX_s + \int_0^t F_y(X_s, Y_s) dY_s \\ &+ \frac{1}{2} \int_0^t F_{xx}(X_s, Y_s) d\langle X, X \rangle_s + \frac{1}{2} \int_0^t F_{yy}(X_s, Y_s) d\langle Y, Y \rangle_s + \int_0^t F_{xy}(X_s, Y_s) d\langle X, Y \rangle_s \end{aligned}$$

La fórmula de Ito unidimensional puede escribirse en la forma siguiente:

$$\int_0^t F'(X_s) dX_s = F(X_t) - F(X_0) - \frac{1}{2} \int_0^t F''(X_s) d\langle X, X \rangle_s$$

En particular:

$$\int_0^t F'(W_s) dW_s = F(W_t) - F(W_0) - \frac{1}{2} \int_0^t F''(W_s) ds$$

También, como caso particular, se tiene:

$$F(t, W_t) = F(0, W_0) + \int_0^t F_x(s, W_s) ds + \int_0^t F_y(s, W_s) dW_s + \frac{1}{2} \int_0^t F_{yy}(s, W_s) ds$$

Así que:

$$\int_0^t F_y(s, W_s) dW_s = F(t, W_t) - F(0, W_0) - \int_0^t F_x(s, W_s) ds - \frac{1}{2} \int_0^t F_{yy}(s, W_s) ds$$

Observemos que las funciones $t \rightarrow F''(W_t)$, $t \rightarrow F_x(t, W_t)$ y $t \rightarrow F_{yy}(t, W_t)$ son continuas, así que las integrales $\int_0^t F''(W_s) ds$, $\int_0^t F_x(s, W_s) ds$ y $\int_0^t F_{yy}(s, W_s) ds$ son integrales de Lebesgue.

19. CÁLCULO DE INTEGRALES ESTOCÁSTICAS

Las fórmulas anteriores nos permite realizar el cálculo de algunas integrales estocásticas, expresándolas en términos de integrales de Lebesgue. En particular, la fórmula del caso unidimensional nos permite calcular integrales estocásticas a partir del cálculo de la integral no estocástica indefinida correspondiente ya que si g es una función de clase C^2 y denotamos por f a su derivada, entonces se tiene $\int f(x) dx = g(x)$; aplicando entonces la fórmula de Ito, se obtiene:

$$\int_0^t f(W_s) dW_s = g(W_t) - g(W_0) - \frac{1}{2} \int_0^t f'(W_s) ds$$

El procedimiento se puede mecanizar: si queremos obtener la integral estocástica $\int_0^t f(W_s) dW_s$, evaluamos la integral indefinida $\int f(x) dx$ y, con la derivada de f , obtenemos inmediatamente la integral estocástica en términos de una integral de Lebesgue.

Ejemplos

1. Como $\int x^n dx = \frac{1}{n+1} x^{n+1}$ y la derivada de $f(x) = x^n$ es $f'(x) = nx^{n-1}$, entonces:

$$\int_0^t W_s^n dW_s = \frac{1}{n+1} W_t^{n+1} - \frac{n}{2} \int_0^t W_s^{n-1} ds$$

2. Como $\int e^x dx = e^x$ y la derivada de $f(x) = e^x$ es $f'(x) = e^x$, entonces:

$$\int_0^t e^{W_s} dW_s = e^{W_t} - 1 - \frac{1}{2} \int_0^t e^{W_s} ds$$

3. Como $\int x^n e^x dx = e^x \sum_{k=0}^n (-1)^{n-k} \frac{n!}{k!} x^k$ y la derivada de $f(x) = x^n e^x$ es $f'(x) = x^n e^x + nx^{n-1} e^x$, entonces:

$$\int_0^t W_s^n e^{W_s} dW_s = e^{W_t} \sum_{k=0}^n (-1)^{n-k} \frac{n!}{k!} W_t^k + (-1)^{n+1} - \frac{1}{2} \int_0^t (W_s^n e^{W_s} + nW_s^{n-1} e^{W_s}) ds$$

4. Como $\int x \operatorname{sen} x dx = \operatorname{sen} x - x \cos x$ y la derivada de $f(x) = x \operatorname{sen} x$ es $f'(x) = \operatorname{sen} x + x \cos x$, entonces:

$$\int_0^t W_s \operatorname{sen} W_s dW_s = \operatorname{sen} W_t - W_t \cos W_t - \frac{1}{2} \int_0^t (\operatorname{sen} W_s + W_s \cos W_s) ds$$

5. Como $\int e^x \operatorname{sen} x dx = \frac{1}{2} e^x (\operatorname{sen} x - \cos x)$ y la derivada de $f(x) = e^x \operatorname{sen} x$ es $f'(x) = e^x (\operatorname{sen} x + \cos x)$, entonces:

$$\int_0^t e^{W_s} \operatorname{sen} W_s dW_s = \frac{1}{2} e^{W_t} (\operatorname{sen} W_t - \cos W_t) + \frac{1}{2} - \frac{1}{2} \int_0^t e^{W_s} (\operatorname{sen} W_s + \cos W_s) ds$$

6. Como $\int x^2 (1-x)^n dx = -\frac{1}{n+3} (1-x)^{n+3} + 2\frac{1}{n+2} (1-x)^{n+2} - \frac{1}{n+1} (1-x)^{n+1}$ y la derivada de $f(x) = x^2 (1-x)^n$ es $f'(x) = 2x(1-x)^n - nx^2(1-x)^{n-1}$, entonces:

$$\begin{aligned} \int_0^t W_s^2 (1-W_s)^n dW_s &= -\frac{1}{n+3} (1-W_t)^{n+3} + 2\frac{1}{n+2} (1-W_t)^{n+2} - \frac{1}{n+1} (1-W_t)^{n+1} \\ &\quad - \frac{1}{n+3} - 2\frac{1}{n+2} + \frac{1}{n+1} - \frac{1}{2} \int_0^t [2W_s (1-W_s)^n - nW_s^2 (1-W_s)^{n-1}] ds \end{aligned}$$

Cuando el proceso a integrar depende tanto de t como de W_t , conviene utilizar la fórmula de Ito bidimensional.

7. Tomando $F(x, y) = x^n e^y$, se obtiene:

$$t^n e^{W_t} = F(t, W_t) = n \int_0^t s^{n-1} e^{W_s} ds + \int_0^t s^n e^{W_s} dW_s + \frac{1}{2} \int_0^t s^n e^{W_s} ds$$

Así que:

$$\int_0^t s^n e^{W_s} dW_s = t^n e^{W_t} - n \int_0^t s^{n-1} e^{W_s} ds - \frac{1}{2} \int_0^t s^n e^{W_s} ds$$

8. Tomando $F(x, y) = e^{xy}$, se obtiene:

$$e^{tW_t} = F(t, W_t) = 1 + \int_0^t W_s e^{sW_s} ds + \int_0^t s e^{sW_s} dW_s + \frac{1}{2} \int_0^t s^2 e^{sW_s} ds$$

Así que:

$$\int_0^t s e^{sW_s} dW_s = e^{tW_t} - 1 - \int_0^t W_s e^{sW_s} ds - \frac{1}{2} \int_0^t s^2 e^{sW_s} ds$$

9. Tomando $F(x, y) = \frac{1}{m+1} x^n y^{m+1}$, se obtiene:

$$\frac{1}{m+1} t^n W_t^{m+1} = F(t, W_t) = \frac{n}{m+1} \int_0^t s^{n-1} W_s^{m+1} ds + \int_0^t s^n W_s^m dW_s + \frac{m}{2} \int_0^t s^n W_s^{m-1} ds$$

Así que:

$$\int_0^t s^n W_s^m dW_s = \frac{1}{m+1} t^n W_t^{m+1} - \frac{n}{m+1} \int_0^t s^{n-1} W_s^{m+1} ds - \frac{m}{2} \int_0^t s^n W_s^{m-1} ds$$

En ocasiones es más simple utilizar la fórmula de integración por partes.

Por ejemplo, para calcular $\int_0^t s^n dW_s$ podemos tomar $X_t = t^n$ y $Y_t = W_t$ para obtener:

$$t^n W_t = n \int_0^t s^{n-1} W_s ds + \int_0^t s^n dW_s$$

Así que:

$$\int_0^t s^n dW_s = t^n W_t - n \int_0^t s^{n-1} W_s ds$$

20. SOLUCIÓN DE ECUACIONES INTEGRALES ESTOCÁSTICAS

La fórmula de Ito permite resolver **ecuaciones diferenciales estocásticas** del tipo:

$$dX_t = a(X_t, t)dt + b(X_t, t)dW_t$$

Ejemplos

1. Queremos resolver:

$$dX_t = X_t dW_t \quad X_0 = 1$$

Es decir:

$$X_t = 1 + \int_0^t X_s dW_s$$

Sea $F(x) = \ln x$, entonces:

$$\begin{aligned} Y_t = \ln X_t &= F(X_t) = F(X_0) + \int_0^t F'(X_s) dX_s + \frac{1}{2} \int_0^t F''(X_s) d\langle X, X \rangle_s \\ &= \ln(X_0) + \int_0^t \frac{1}{X_s} dX_s - \frac{1}{2} \int_0^t \frac{1}{X_s^2} d\langle X, X \rangle_s \end{aligned}$$

Se tiene:

$$\begin{aligned} dX_s &= X_s dW_s \\ \langle X, X \rangle_t &= \int_0^t X_s^2 ds \end{aligned}$$

Así que:

$$Y_t = \int_0^t \frac{1}{X_s} X_s dW_s - \frac{1}{2} \int_0^t \frac{1}{X_s^2} X_s^2 ds = \int_0^t dW_s - \frac{1}{2} \int_0^t ds = W_t - \frac{1}{2}t$$

Por lo tanto:

$$X_t = \exp \left\{ W_t - \frac{1}{2}t \right\}$$

Verificación:

$$\text{Definamos } G(x) = \exp\{x\} \text{ y } U_t = W_t - \frac{1}{2}t.$$

Se tiene:

$$\langle U, U \rangle_t = t$$

Así que, aplicando la fórmula de Ito:

$$\begin{aligned} X_t = \exp \left\{ W_t - \frac{1}{2}t \right\} &= G(U_t) = G(U_0) + \int_0^t G'(U_s) dU_s + \frac{1}{2} \int_0^t G''(U_s) d\langle U, U \rangle_s \\ &= 1 + \int_0^t X_s dW_s - \frac{1}{2} \int_0^t X_s ds + \frac{1}{2} \int_0^t X_s ds = 1 + \int_0^t X_s dW_s \end{aligned}$$

2. Haciendo el cambio de variable $Y_t = \ln X_t$, se pueden resolver los siguientes problemas con valores iniciales:

- a) $dX_t = tX_t dW_t$ $X_0 = 1$
- b) $dX_t = X_t dt + tX_t dW_t$ $X_0 = 1$
- c) $dX_t = tX_t dt + X_t dW_t$ $X_0 = 1$
- d) $dX_t = 2t^2 X_t dt + 2tX_t dW_t$ $X_0 = 1$

Solución

Sea $F(x) = \ln x$, entonces:

$$Y_t = \ln X_t = F(X_t) = F(X_0) + \int_0^t F'(X_s) dX_s + \frac{1}{2} \int_0^t F''(X_s) d\langle X, X \rangle_s$$

$$= \ln(X_0) + \int_0^t \frac{1}{X_s} dX_s - \frac{1}{2} \int_0^t \frac{1}{X_s^2} d\langle X, X \rangle_s$$

a) Queremos resolver la ecuación integral:

$$X_t = 1 + \int_0^t sX_s dW_s$$

Se tiene:

$$\langle X, X \rangle_t = \int_0^t s^2 X_s^2 ds$$

Así que:

$$Y_t = \int_0^t \frac{1}{X_s} sX_s dW_s - \frac{1}{2} \int_0^t \frac{1}{X_s^2} s^2 X_s^2 ds = \int_0^t s dW_s - \frac{1}{2} \int_0^t s^2 ds = \int_0^t s dW_s - \frac{1}{6} t^3$$

Por lo tanto:

$$X_t = \exp \left\{ \int_0^t s dW_s - \frac{1}{6} t^3 \right\}$$

Verificación:

$$\text{Definamos } G(x) = \exp\{x\} \text{ y } U_t = \int_0^t s dW_s - \frac{1}{6} t^3$$

Se tiene:

$$\langle U, U \rangle_t = \int_0^t s^2 ds = \frac{1}{3} t^3$$

Así que, aplicando la fórmula de Ito:

$$\begin{aligned} X_t &= \exp \left\{ \int_0^t s dW_s - \frac{1}{6} t^3 \right\} = G(U_t) = G(U_0) + \int_0^t G'(U_s) dU_s + \frac{1}{2} \int_0^t G''(U_s) d\langle U, U \rangle_s \\ &= 1 + \int_0^t sX_s dW_s - \frac{1}{2} \int_0^t s^2 X_s ds + \frac{1}{2} \int_0^t s^2 X_s ds = 1 + \int_0^t sX_s dW_s \end{aligned}$$

Por la fórmula de integración por partes, se tiene:

$$\int_0^t s dW_s = tW_t - \int_0^t W_s ds$$

Así que:

$$X_t = \exp \left\{ tW_t - \frac{1}{6} t^3 - \int_0^t W_s ds \right\}$$

Otro método:

$$\text{Definamos } G(x, y) = \exp \left\{ y - \frac{1}{6} x^3 \right\}, U_t = t \text{ y } V_t = \int_0^t s dW_s.$$

Se tiene:

$$\langle U, U \rangle_t = 0$$

$$\langle V, V \rangle_t = \int_0^t s^2 ds$$

$$\langle U, V \rangle_t = 0$$

$$G_x(x, y) = -\frac{1}{2}x^2 \exp\left\{y - \frac{1}{6}x^3\right\}$$

$$G_y(x, y) = \exp\left\{y - \frac{1}{6}x^3\right\}$$

$$G_{yy}(x, y) = \exp\left\{y - \frac{1}{6}x^3\right\}$$

Aplicando la fórmula de Ito, se tiene:

$$\begin{aligned} X_t &= \exp\left\{\int_0^t s dW_s - \frac{1}{6}t^3\right\} = G(U_t, V_t) \\ &= G(U_0, V_0) + \int_0^t G_x(U_s, V_s) dU_s + \int_0^t G_y(U_s, V_s) dV_s \\ &\quad + \frac{1}{2} \int_0^t G_{xx}(U_s, V_s) d\langle U, U \rangle_s + \frac{1}{2} \int_0^t G_{yy}(U_s, V_s) d\langle V, V \rangle_s + \int_0^t G_{xy}(U_s, V_s) d\langle U, V \rangle_s \\ &= 1 + \int_0^t -\frac{1}{2}s^2 X_s ds + \int_0^t s X_s dW_s + \frac{1}{2} \int_0^t s^2 X_s ds = 1 + \int_0^t s X_s dW_s \end{aligned}$$

b) Queremos resolver la ecuación integral:

$$X_t = 1 + \int_0^t X_s ds + \int_0^t s X_s dW_s$$

Se tiene:

$$\langle X, X \rangle_t = \int_0^t s^2 X_s^2 ds$$

Así que:

$$\begin{aligned} Y_t &= \int_0^t \frac{1}{X_s} X_s ds + \int_0^t \frac{1}{X_s} s X_s dW_s - \frac{1}{2} \int_0^t \frac{1}{X_s^2} s^2 X_s^2 ds = t + \int_0^t s dW_s - \frac{1}{6}t^3 \\ &= t - \frac{1}{6}t^3 + tW_t - \int_0^t W_s ds \end{aligned}$$

Por lo tanto:

$$X_t = \exp\left\{t - \frac{1}{6}t^3 + tW_t - \int_0^t W_s ds\right\}$$

$$c) X_t = \exp\left\{W_t + \frac{1}{2}t^2 - \frac{1}{2}t\right\}$$

$$d) X_t = \exp\left\{\int_0^t 2s dW_s\right\} = \exp\left\{2tW_t - 2 \int_0^t W_s ds\right\}$$

$$3. dX_t = \sec X_t \left[\frac{1}{2}(1 + \sec X_t \tan X_t) dt + dW_t\right]$$

Solución

Consideremos el cambio de variable $Y_t = \sec X_t$.

$$X_t = X_0 + \int_0^t \frac{1}{2} \sec X_s (1 + \sec X_s \tan X_s) ds + \int_0^t \sec X_s dW_s$$

$$\sec X_t = \sec X_0 + \int_0^t \cos X_s dX_s - \frac{1}{2} \int_0^t \sec X_s d\langle X, X \rangle_s$$

$$\begin{aligned}
&= \operatorname{sen} X_0 + \int_0^t \frac{1}{2} \cos X_s \sec X_s (1 + \sec X_s \tan X_s) ds + \int_0^t \cos X_s \sec X_s dW_s - \frac{1}{2} \int_0^t \operatorname{sen} X_s \sec^2 X_s ds \\
&= \operatorname{sen} X_0 + \int_0^t \frac{1}{2} (1 + \sec X_s \tan X_s) ds + \int_0^t dW_s - \frac{1}{2} \int_0^t \sec X_s \tan X_s ds \\
&= \operatorname{sen} X_0 + \frac{1}{2} t + W_t
\end{aligned}$$

Así que:

$$X_t = \operatorname{arcsen} \left(\operatorname{sen} X_0 + \frac{1}{2} t + W_t \right)$$

21. VERIFICACIÓN DE SOLUCIONES DE ECUACIONES DIFERENCIALES ESTOCÁSTICAS

Ejemplos

1. Verifiquemos que:

$$X_t = e^{\beta W_t - \frac{1}{2} \beta^2 t} \left(X_0 + \int_0^t \alpha e^{-(\beta W_s - \frac{1}{2} \beta^2 s)} ds \right)$$

es solución de la ecuación:

$$dX_t = \alpha dt + \beta X_t dW_t$$

donde α y β son constantes.

Definamos:

$$F(x, y) = xe^y$$

$$U_t = X_0 + \int_0^t \alpha e^{-(\beta W_s - \frac{1}{2} \beta^2 s)} ds$$

$$Z_t = \beta W_t - \frac{1}{2} \beta^2 t.$$

Se tiene:

$$\langle Z, Z \rangle_t = \beta^2 t$$

Aplicando la fórmula de Ito, se tiene:

$$\begin{aligned}
X_t &= F(U_t, Z_t) = X_0 + \alpha t + \int_0^t \beta X_s dW_s - \frac{1}{2} \beta^2 \int_0^t X_s ds + \frac{1}{2} \beta^2 \int_0^t X_s ds \\
&= X_0 + \alpha t + \int_0^t \beta X_s dW_s
\end{aligned}$$

2. Verifiquemos que:

$$S_t = S_0 e^{(\mu - \frac{1}{2} \sigma^2)t + \sigma W_t}$$

es solución de la ecuación:

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

En efecto, por la fórmula de Ito, con $F(x) = e^x$, se tiene

$$\begin{aligned} F\left(\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right) &= 1 + \int_0^t \left(\mu - \frac{1}{2}\sigma^2\right) F\left(\left(\mu - \frac{1}{2}\sigma^2\right)s + \sigma W_s\right) ds \\ &+ \int_0^t \sigma F\left(\left(\mu - \frac{1}{2}\sigma^2\right)s + \sigma W_s\right) dW_s + \frac{1}{2} \int_0^t \sigma^2 F\left(\left(\mu - \frac{1}{2}\sigma^2\right)s + \sigma W_s\right) ds \\ &= 1 + \int_0^t \mu F\left(\left(\mu - \frac{1}{2}\sigma^2\right)s + \sigma W_s\right) ds + \int_0^t \sigma F\left(\left(\mu - \frac{1}{2}\sigma^2\right)s + \sigma W_s\right) dW_s \end{aligned}$$

Por lo tanto,

$$\begin{aligned} S_t &= S_0 F\left(\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right) \\ &= S_0 + \mu \int_0^t S_0 F\left(\left(\mu - \frac{1}{2}\sigma^2\right)s + \sigma W_s\right) ds + \sigma \int_0^t S_0 F\left(\left(\mu - \frac{1}{2}\sigma^2\right)s + \sigma W_s\right) dW_s \\ &= S_0 + \mu \int_0^t S_s ds + \sigma \int_0^t S_s dW_s \end{aligned}$$

3. Verifiquemos que

$$X_t = X_0 \exp\left\{\int_0^t \left(\alpha_s - \frac{1}{2}\beta_s^2\right) ds + \int_0^t \beta_s dW_s\right\}$$

es solución de la ecuación:

$$dX_t = \alpha_t X_t dt + \beta_t X_t dW_t$$

donde $(\alpha_t)_{t \in \mathbb{R}^+}$ y $(\beta_t)_{t \in \mathbb{R}^+}$ son procesos estocásticos continuos e independientes de $(X_t)_{t \geq 0}$.

Solución

Definamos $F(x, y) = xe^y$, $U_t = X_0$ y $Z_t = \int_0^t (\alpha_s - \frac{1}{2}\beta_s^2) ds + \int_0^t \beta_s dW_s$.

Se tiene:

$$\langle Z, Z \rangle_t = \int_0^t \beta_s^2 ds$$

Aplicando la fórmula de Ito, se tiene:

$$\begin{aligned} X_t &= F(U_t, Z_t) = X_0 + \int_0^t X_s (\alpha_s - \frac{1}{2}\beta_s^2) ds + \int_0^t X_s \beta_s dW_s + \frac{1}{2} \int_0^t X_s \beta_s^2 ds \\ &= X_0 + \int_0^t \alpha_s X_s ds + \int_0^t \beta_s X_s dW_s \end{aligned}$$

4. Verifiquemos que

$$X_t = \frac{\exp\{W_t + \frac{1}{2}t\}}{X_0^{-1} + \int_0^t \exp\{W_s + \frac{1}{2}s\} ds}$$

es solución de la ecuación:

$$dX_t = (X_t - X_t^2) dt + X_t dW_t$$

Solución

Definamos $F(x, y) = \frac{e^y}{x}$, $U_t = X_0^{-1} + \int_0^t \exp\{W_s + \frac{1}{2}s\} ds$ y $Z_t = W_t + \frac{1}{2}t$.

Se tiene:

$$\langle Z, Z \rangle_t = t$$

Aplicando la fórmula de Ito, se tiene:

$$X_t = F(U_t, Z_t)$$

$$= X_0 + \int_0^t -\frac{\exp\{W_t + \frac{1}{2}t\}}{U_t^2} \exp\{W_s + \frac{1}{2}s\} ds + \int_0^t X_s dW_s + \frac{1}{2} \int_0^t X_s ds + \frac{1}{2} \int_0^t X_s ds$$

$$= X_0 - \int_0^t X_s^2 dW_s + \int_0^t X_s dW_s + \int_0^t X_s ds$$

$$= X_0 + \int_0^t (X_s - X_s^2) dW_s + \int_0^t X_s ds$$

$$V_t = \exp\{W_t + \frac{1}{2}t\} = 1 + \int_0^t V_s dW_s + \frac{1}{2} \int_0^t V_s ds + \frac{1}{2} \int_0^t V_s ds = 1 + \int_0^t V_s dW_s + \int_0^t V_s ds$$

$$dV_t = V_t dW_t + V_t dt$$

En general, si $V_t = F(X_t)$, entonces

$$F(X_t) = F(X_0) + \int_0^t F'(X_s) dX_s + \frac{1}{2} \int_0^t F''(X_s) d\langle X, X \rangle_s$$

Así que:

$$dV_t = F'(X_s) dX_s + \frac{1}{2} F''(X_s) d\langle X, X \rangle_s$$

5. Verifiquemos que

$$X_t = e^{W_t - \frac{1}{2}t} \left(X_0 + \int_0^t s e^{-(W_s - \frac{1}{2}s)} dW_s \right)$$

es solución de la ecuación:

$$dX_t = t dt + (X_t + t) dW_t$$

Solución

Definamos $F(x, y) = xe^y$, $U_t = X_0 + \int_0^t s e^{-(W_s - \frac{1}{2}s)} dW_s$ y $Z_t = W_t - \frac{1}{2}t$.

Se tiene:

$$\langle Z, Z \rangle_t = t$$

$$\langle U, Z \rangle_t = \int_0^t s e^{-(W_s - \frac{1}{2}s)} ds$$

Aplicando la fórmula de Ito, se tiene:

$$X_t = F(U_t, Z_t)$$

$$= X_0 + \int_0^t s dW_s + \int_0^t X_s dW_s - \frac{1}{2} \int_0^t X_s ds + \frac{1}{2} \int_0^t X_s ds + \int_0^t s ds$$

$$= X_0 + \int_0^t s dW_s + \int_0^t X_s dW_s + \int_0^t s ds$$

$$= X_0 + \int_0^t s ds + \int_0^t X_s dW_s + \int_0^t s dW_s$$

22. EJERCICIOS

1. Encuentra la solución general de la ecuación diferencial estocástica $dX_t = -\alpha X_t dt + \sigma dW_t$, donde α y σ son constantes.

$$X_t = \left(X_0 + \int_0^t \sigma e^{\alpha s} dW_s \right) e^{-\alpha t}$$

2. Haciendo el cambio de variable $Y_t = \ln X_t$, resuelve los siguientes problemas con valores iniciales:

a) $dX_t = tX_t dW_t$ $X_0 = 1$

b) $dX_t = X_t dt + tX_t dW_t$ $X_0 = 1$

c) $dX_t = tX_t dt + X_t dW_t$ $X_0 = 1$

d) $dX_t = 2t^2 X_t dt + 2tX_t dW_t$ $X_0 = 1$

3. Encuentra la solución general de la siguiente ecuación diferencial estocástica:

$$dX_t = \sec X_t \left[\frac{1}{2}(1 + \sec X_t \tan X_t) dt + dW_t \right]$$

Sugerencia: Considera el cambio de variable $Y_t = \operatorname{sen} X_t$.

4. Demuestra que cada una de las siguientes ecuaciones diferenciales estocásticas tiene la solución que se indica:

a) $dX_t = \alpha_t X_t dt + \beta_t X_t dW_t$, donde $(\alpha_t)_{t \geq 0}$ y $(\beta_t)_{t \geq 0}$ son procesos estocásticos continuos e independientes de $(X_t)_{t \geq 0}$.

$$X_t = X_0 \exp \left\{ \int_0^t (\alpha_s - \frac{1}{2} \beta_s^2) ds + \int_0^t \beta_s dW_s \right\}$$

b) $dX_t = (X_t - X_t^2) dt + X_t dW_t$

$$X_t = \frac{\exp\{W_t + \frac{1}{2}t\}}{X_0^{-1} + \int_0^t \exp\{W_s + \frac{1}{2}s\} ds}$$

c) $dX_t = t dt + (X_t + t) dW_t$

$$X_t = e^{W_t - \frac{1}{2}t} \left(X_0 + \int_0^t s e^{-(W_s - \frac{1}{2}s)} dW_s \right)$$