UNIVERSIDAD NACIONAL AUTÓNOMA DE MEXICO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MATEMÁTICAS Y DE LA ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA

ESTIMACIÓN DE LA DENSIDAD DE UN PROCESO DE DIFUSIÓN MEDIANTE EL CÁLCULO DE MALLIAVIN

TESIS
QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS

PRESENTA:
ROBERTO CARLOS MINERO PÉREZ GAVILÁN

DIRECTOR DE TESIS:
DRA. MARÍA EMILIA CABALLERO ACOSTA
INSTITUTO DE MATEMÁTICAS
UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

MÉXICO, D.F., JUNIO 2016

Agradecimientos

- A mis queridos padres, María de Lourdes y Roberto por su gran amor y apoyo invaluable.

- A mi directora de tesis, la Dra. María Emilia Caballero Acosta porque gracias a su infinita paciencia y junto con su gran apoyo y enseñanzas es que se pudo concretar este trabajo.

- A mis sinodales, la Dra. Reyla Areli Navarro Cruz, el Dr. Fernando Baltazar Larios, el Dr. José Luis Ángel Pérez Garmendia y el Dr. Carlos Gabriel Pacheco González por enriquecer a esta tesis con sus valiosas sugerencias y observaciones.

- A toda mi familia que siempre ha sido una piedra angular en mi vida.

- Al IIMAS, IMATE y a la Facultad de Ciencias de la UNAM por la formación y los conocimientos adquiridos durante la maestría.

- Al CONACYT por el valioso apoyo económico que me brindó para poder completar mis estudios de maestría.

- A los profesores de la maestría por fomentar mi pasión hacia las matemáticas.

- A todos y cada uno de mis amigos.

Índice

1. Introducción

La rama de la probabilidad conocida como procesos estocásticos constituye una bella teoría complementada con sus diversas aplicaciones en múltiples áreas de nuestro mundo actual. Especialmente, han evolucionado los procesos estocásticos a tiempo continuo cuya construcción teórica y aportación hacia las matemáticas forman un sólido conocimiento digno de estudiarse a profundidad.

El presente trabajo tiene como objetivo dar una exposición introductoria sobre los elementos del cálculo estocástico de variaciones; llamado también cálculo de Malliavin en honor a Paul Malliavin quien introdujo esta teoría en 1976.

Dicha teoría tiene sus consecuencias importantes tanto en la rama teórica como en las aplicaciones de la probabilidad y es por eso que resulta muy atractivo exponer sus elementos así como sus implicaciones en los procesos estocásticos.

En nuestra exposición nos enfocaremos en el estudio de algunos resultados elementales del cálculo de Malliavin y mediante el cual ilustramos las técnicas del cálculo diferencial sobre un espacio particular de dimensión infinita conocido como el espacio de Wiener.

Posteriormente estudiaremos la implementación del cálculo de Malliavin en la estimación de la densidad de un proceso de difusión, que de hecho constituye el corazón y la razón primordial del presente trabajo, el cual me fue motivado por la Dra. María Emilia mediante uno de sus artículos de investigación[1].

Cabe mencionar que el contenido del presente trabajo está dirigido a un lector con conocimientos de procesos estocásticos a tiempo continuo y del cálculo estocástico continuo. No obstante, la exposición pretende abordar las herramientas necesarias para el completo estudio de los resultados sin la pretensión de ser exhaustivo en la ardua tarea de demostrar todos los conocimientos preliminares en cuestión; puesto que se volvería una lectura muy árida y esto no compete con nuestro objetivo primordial.

Para complementar las necesidades del lector con curiosidad de los resultados preliminares no demostrados; exhibiremos a lo largo de la exposición, mediante referencias detalladas, aquellas fuentes donde se pueden consultar las demostraciones en su totalidad. El espíritu del presente trabajo consiste en exponer las bases necesarias del cálculo de Malliavin en virtud de su aplicación hacia los procesos de difusión.

El presente trabajo inicia en el capítulo 2 con el estudio de las derivadas de Gâteaux y de Fréchet que serán el concepto clave para comprender a fondo a la derivada de Malliavin, las cuales serán presentadas en la sección 2.1. Además expondremos los resultados elementales de ambas derivadas y después retomamos algunas propiedades del movimiento browniano y en seguida presentamos el espacio de Cameron-Martin e introducimos el concepto de derivada de Malliavin que fungirá como elemento primordial del cálculo de Malliavin, los cuales serán abordados en las secciones 2.2 y 2.3.

Posteriormente, en la sección 2.4, retomamos un capítulo dedicado a recapitular los procesos de difusión y mediante el cual solamente evocaremos los resultados principales pero cuya demostración puede ser consultada con las referencias detalladas en los pies de página. El objetivo es tener presente las propiedades importantes del proceso de difusión que será al que aplicaremos las técnicas del cálculo de Malliavin.

En la sección 2.5 estudiaremos a fondo la integral de Skorokhod y su relación con la derivada de Malliavin; definido como su operador adjunto. Es importante mencionar que en este capítulo se exponen las

[1] Véase [3] en la bibliografía.

prinicipales extensiones de los resultados previamente desarrollados y que a su vez serán primordiales en nuestro estudio en aras de sus implicaciones hacia la estimación de las densidades de una difusión.

A lo largo del capítulo 3 analizamos unas desigualdes para martingalas que enriquecerán nuestra presentación con diversos resultados interesantes y que además jugarán un papel esencial para la estimación deseada. Nuestra exposición cúspide la condensamos en la sección 3.3 con la implementación de los resultados expuestos a lo largo del presente trabajo y es ahí donde evidenciamos nuestro objetivo primordial de estudio.

2. La Derivada de Malliavin

2.1. Derivadas de Gâteaux y de Fréchet

Tomemos como punto de partida un concepto ampliamente estudiado, a saber, el de diferenciación. En este primer capítulo estudiaremos el concepto de derivadas para funciones definidas en espacios de Banach y específicamente nos enfocaremos en dos conceptos importantes: el de Gâteaux y el de Fréchet.

Estos conceptos guardan una relación entre sí y además nos serán esenciales para comprender, con mayor precisión, a la derivada especial que servirá como elemento clave de nuestro trabajo.

El objetivo central de nuestro estudio será la derivada de Malliavin, sin embargo, para poderla estudiar apropiadamente necesitaremos introducir estos conceptos de diferenciación y consecuentemente podremos apreciar la conexión entre la derivada de Malliavin con la de Gâteaux y la de Fréchet.

Teniendo en claro estos conceptos preliminares es que podremos ir estudiando con detalle a los elementos del cálculo de Malliavin. De manera que demos inicio con la idea de diferenciación de funciones definidas en espacios de Banach.

Sean $X = (X, \|\cdot\|_X)$ y $Y = (Y, \|\cdot\|_Y)$ espacios de Banach y sea $T : X \to Y$ una función lineal y continua. Sabemos que el espacio dual de X, que denotamos por X^*, está formado por el espacio de funciones $T : X \to Y$ lineales y continuas.

Si consideramos el espacio $X^* = (X^*, \|\cdot\|_{X^*})$ donde la norma está definida como

$$\|T\|_{X^*} := \sup_{\substack{x \in X \\ x \neq 0}} \frac{\|T(x)\|_Y}{\|x\|_X} \qquad \text{para cada } T \in X^*,$$

entonces X^* es un nuevo espacio de Banach[2].

Observemos que de la definición anterior se sigue que

$$\|T(x)\|_Y \leq \|T\|_{X^*} \|x\|_X \qquad \text{para toda } x \in X. \tag{1}$$

Definición 2.1. *Sean U un subconjunto abierto de X, x un elemento fijo de U, $\varepsilon \in \mathbb{R}$. Una función $f : U \to Y$ es **Gâteaux-diferenciable**, si para cada $y \in X$, existe el límite*

$$\lim_{\varepsilon \to 0} \frac{f(x + \varepsilon y) - f(x)}{\varepsilon}$$

y la función

$$y \longmapsto D_y f(x) := \lim_{\varepsilon \to 0} \frac{f(x + \varepsilon y) - f(x)}{\varepsilon}$$

*es lineal y continua. A $D_y f(x) \in Y$ se le conoce como **la derivada de Gâteaux o direccional** de f en el punto x en la dirección y.*

*Se dice que f es **Gâteaux-diferenciable** en U si lo es para todo elemento $x \in U$.*

[2] [5] Prop. 9.2. Pág. 155

Es importante notar que si x es un elemento fijo de U y el límite

$$\lim_{\varepsilon\to 0}\frac{f(x+\varepsilon y)-f(x)}{\varepsilon}$$

existe para toda $y\in X$, esto no garantiza que la función $y\mapsto D_y f(x)$ sea lineal.

Para ilustrar la aseveración anterior consideremos la función $f:\mathbb{R}^2\to\mathbb{R}$

$$f(x_1,x_2)=\begin{cases}\frac{x_1x_2^2}{x_1^2+x_2^2} & (x_1,x_2)\neq(0,0),\\ 0 & (x_1,x_2)=(0,0).\end{cases}$$

Sea $x=(x_1,x_2)$ fijo, entonces se puede verificar que para toda $y=(y_1,y_2)\in\mathbb{R}^2$ existe el límite

$$\lim_{\varepsilon\to 0}\frac{f(x_1+\varepsilon y_1,x_2+\varepsilon y_2)-f(x_1,x_2)}{\varepsilon}.$$

Pero, si en particular $x=(0,0)$, entonces

$$\begin{aligned} y\longmapsto D_y f(x) &= \lim_{\varepsilon\to 0}\frac{f(0+\varepsilon y_1,0+\varepsilon y_2)-f(0,0)}{\varepsilon}=\lim_{\varepsilon\to 0}\frac{(\varepsilon y_1)(\varepsilon y_2)^2-0}{\varepsilon[(\varepsilon y_1)^2+(\varepsilon y_2)^2]}\\ &=\lim_{\varepsilon\to 0}\frac{\varepsilon^3 y_1y_2^2}{\varepsilon^3[y_1^2+y_2^2]}=\frac{y_1y_2^2}{y_1^2+y_2^2}\end{aligned}$$

no es función lineal de $y=(y_1,y_2)$.

Observemos que si U es un subconjunto abierto de X y si tomamos $x\in U$ fijo y una función $f:U\to\mathbb{R}^m$, entonces, por la definición anterior, f tiene *derivada direccional* en $x\in U$ en la dirección $y\in X$ si existe el límite

$$D_y f(x):=\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}f(x+\varepsilon y)\right|_{\varepsilon=0}\in\mathbb{R}^m$$

y si además la función $y\mapsto D_y f(x)$ es lineal y continua. Ilustremos la derivada de Gâteaux con el siguiente ejemplo.

Ejemplo 2.1. *Consideremos la función* $f:\mathbb{R}^2\to\mathbb{R}$

$$f(x_1,x_2)=\begin{cases}\frac{x_1^4x_2}{x_1^6+x_2^3} & (x_1,x_2)\neq(0,0),\\ 0 & (x_1,x_2)=(0,0).\end{cases}$$

Entonces para $x=(0,0)$, $y=(y_1,y_2)\in\mathbb{R}^2$ *y con* $\varepsilon\in\mathbb{R}\setminus\{0\}$

$$\frac{f(0+\varepsilon y_1,0+\varepsilon y_2)-f(0,0)}{\varepsilon}=\frac{f(\varepsilon y_1,\varepsilon y_2)-0}{\varepsilon}=\begin{cases}\frac{\varepsilon y_1^4y_2}{\varepsilon^3y_1^6+y_2^3} & (y_1,y_2)\neq(0,0)\\ 0 & (y_1,y_2)=(0,0)\end{cases}$$

y esto conlleva a que

$$\lim_{\varepsilon\to 0}\frac{f(0+\varepsilon y_1,0+\varepsilon y_2)-f(0,0)}{\varepsilon}=0\qquad \textit{para toda } y\in\mathbb{R}^2.$$

Además, si $x = (0,0)$ lo anterior implica que

$$y \longmapsto D_y f(x) \;=\; \lim_{\varepsilon\to 0} \frac{f(x+\varepsilon y) - f(x)}{\varepsilon} \;=\; 0 \qquad \textit{para toda } y \in \mathbb{R}^2$$

y por lo tanto $D_y f(x)$ es lineal y continua, es decir, f es Gâteaux-diferenciable en $x = (0,0)$.

Un concepto más fuerte de diferenciabilidad, como probaremos más adelante, es el siguiente.

Definición 2.2. *Sea U un subconjunto abierto de X y sea x un elemento fijo de U. Una función $f : U \to Y$ es **Fréchet-diferenciable** en x, si existe una función lineal y continua $T : X \to Y$ tal que*

$$\lim_{\substack{y\to 0\\ y\in X}} \frac{\|f(x+y) - f(x) - T(y)\|_Y}{\|y\|_X} = 0. \tag{2}$$

*A T se le conoce como **la derivada de Fréchet** de f en x y se denota por $f'(x)$.*

La expresión (2) de la definición anterior refiere que para cada $\varepsilon > 0$ existe $\delta > 0$ tal que, para todo $y \in X$ con $\|y\|_X < \delta$, se cumple que

$$x + y \;\text{ pertenece a } U \;\text{ y que }\; \left\|f(x+y) - f(x) - f'(x)(y)\right\|_Y < \varepsilon \|y\|_X .$$

En virtud de esta observación veamos el siguiente resultado.

Teorema 2.1. *Si una función f es Fréchet-diferenciable en un punto x, entonces es continua en ese mismo punto.*

Demostración. Por hipótesis f es Fréchet-diferenciable en x y por lo tanto existe $\delta_1 > 0$ tal que si $\|y - x\|_X < \delta_1$, entonces se sigue de la observación anterior que

$$\left\|f(y) - f(x) - f'(x)(y-x)\right\|_Y < \|y - x\|_X$$

El hecho de que f sea Fréchet-diferenciable en x también implica que $f'(x)$ es una función lineal y continua, es decir, $f'(x) \in X^*$. En consecuencia, por la desigualdad (1) y junto con la desigualdad anterior se tiene que si $\|y - x\|_X < \delta_1$, entonces

$$\begin{aligned}
\|f(y) - f(x)\|_Y &= \left\|f(y) - f(x) - f'(x)(y-x) + f'(x)(y-x)\right\|_Y \\
&\le \left\|f(y) - f(x) - f'(x)(y-x)\right\|_Y + \left\|f'(x)(y-x)\right\|_Y \\
&< \|y-x\|_X + \left\|f'(x)\right\|_{X^*} \|y-x\|_X \\
&= \left(1 + \left\|f'(x)\right\|_{X^*}\right) \|y-x\|_X
\end{aligned}$$

Sea $\varepsilon > 0$ y fijemos $\delta := \min\left\{\delta_1, \dfrac{\varepsilon}{(1 + \|f'(x)\|_{X^*})}\right\}$, entonces por los cálculos anteriores

$$\text{si } \|y - x\|_X < \delta, \quad \text{entonces } \|f(y) - f(x)\|_Y < \varepsilon$$

y con esto demostramos que f es continua en x. □

Notemos la diferencia entre ambas derivadas. En el Ejemplo 2.1 vimos que f es Gâteaux-diferenciable en $(0,0)$, sin embargo, si tomamos $(x_1,x_2)\neq(0,0)$ y definimos $x_2 := x_1^2$ tenemos entonces que

$$f(x_1,x_1^2) \;=\; \frac{x_1^6}{x_1^6+x_1^6} \;=\; \frac{1}{2}$$

pero por definición $f(0,0)=0$ y esto implica que f no es continua en $(0,0)$.

Acabamos de probar que si f es Fréchet-diferenciable en un punto, entonces es continua en ese mismo punto, de manera que f no puede ser Fréchet-diferenciable en el $(0,0)$.

Con esto exhibimos un ejemplo de una función que es Gâteaux-diferenciable en el punto $(0,0)$ pero que no es Fréchet-diferenciable en ese mismo punto.

De manera conversa al ejemplo anterior, probaremos que si una función es Fréchet-diferenciable en un punto, entonces es Gâteaux-diferenciable en ese mismo punto; y en tal caso, ambas derivadas coinciden. Con esto verificamos que el concepto de Fréchet-diferenciable es más fuerte que Gâteaux-diferenciable.

Teorema 2.2. *Sea U un subconjunto abierto de X y sea x un elemento fijo de U. Si la función $f:U\to Y$ es Fréchet-diferenciable en x, entonces f es Gâteaux-diferenciable en x y además $Df(x)=f'(x)$.*

Demostración. Por hipótesis f es Fréchet-diferenciable en $x\in U$, entonces existe una función lineal y continua $T:X\to Y$ tal que

$$\lim_{\substack{y\to 0\\ y\in X}} \frac{\|f(x+y)-f(x)-T(y)\|_Y}{\|y\|_X}=0. \tag{3}$$

En particular fijemos cualquier $y\in X$ distinto de cero. Como T es una función lineal y junto con la expresión (3) obtenemos que

$$\begin{aligned}
\lim_{\varepsilon\to 0}\left\|\frac{f(x+\varepsilon y)-f(x)}{\varepsilon}-T(y)\right\|_Y &= \lim_{\varepsilon\to 0}\left\|\frac{f(x+\varepsilon y)-f(x)-\varepsilon T(y)}{\varepsilon}\right\|_Y \\
&= \lim_{\varepsilon\to 0}\frac{\|f(x+\varepsilon y)-f(x)-T(\varepsilon y)\|_Y}{|\varepsilon|\cdot\|y\|_X}\cdot\|y\|_X \\
&= \|y\|_X\cdot\lim_{\varepsilon\to 0}\frac{\|f(x+\varepsilon y)-f(x)-T(\varepsilon y)\|_Y}{\|\varepsilon y\|_X} \\
&= \|y\|_X\cdot 0 \\
&= 0
\end{aligned}$$

y esto implica que para cada $y\in X$ el $\lim\limits_{\varepsilon\to 0}\dfrac{f(x+\varepsilon y)-f(x)}{\varepsilon}$ existe y de hecho

$$\lim_{\varepsilon\to 0}\frac{f(x+\varepsilon y)-f(x)}{\varepsilon} \;=\; T(y) \qquad \text{para toda } y\in X,$$

más aún, como T es una función lineal y continua, entonces

$$y\longmapsto D_yf(x)=\lim_{\varepsilon\to 0}\frac{f(x+\varepsilon y)-f(x)}{\varepsilon}=T(y) \qquad \text{es lineal y continua,}$$

y por lo tanto T es la derivada Gâteaux en x, en donde $D_yf(x)=T(y)$ para toda $y\in X$.

Pero por hipótesis T es la derivada Fréchet en x lo que implica que $T(y) = f'(x)(y)$ para toda $y \in X$ y concluimos entonces que

$$Df(x) \ = \ f'(x).$$

$\square$

Observemos que si U es un subconjunto abierto de X y si tomamos $x \in U$ fijo y una función $f : U \to \mathbb{R}^m$, entonces f es *Fréchet-diferenciable* en $x \in U$ si existe una función lineal y continua

$$A : X \to \mathbb{R}^m, \quad \text{donde} \ \ A = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix} \ \text{con} \ A_i \in X^* \ \text{para} \ 1 \leq i \leq m,$$

tal que

$$\lim_{\substack{h \to 0 \\ h \in X}} \frac{\|f(x+h) - f(x) - A(h)\|_{\mathbb{R}^m}}{\|h\|_X} = 0.$$

Y en tal el caso, le llamaremos a A *la derivada de Fréchet* de f en x y escribimos

$$A = f'(x) = \begin{bmatrix} f'(x)_1 \\ \vdots \\ f'(x)_m \end{bmatrix} \in (X^*)^m.$$

Notemos la analogía con el cálculo vectorial clásico. Recordemos que si U es un subconjunto abierto de $\mathbb{R}^n$ y si $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en $x \in U$, entonces f es continua en x; como en el caso Fréchet-diferenciable del Teorema 2.1.

Recordemos también que si f es diferenciable en x, con diferencial $f'(x)$, entonces existe la derivada $Df(x)$ en todas las direcciones $y \in \mathbb{R}^n$ y además $Df(x) = f'(x)$ como en el Teorema 2.2. Más aún, recordemos también que en este caso se tiene que $D_y f(x) = \langle f'(x), y \rangle$ para toda $y \in \mathbb{R}^n$.

Veamos una aplicación de los teoremas anteriores, sobre un espacio de Hilbert, en el siguiente ejemplo.

Ejemplo 2.2. *Sea H un espacio de Hilbert y $f : H \to \mathbb{R}$. Si f es Fréchet-diferenciable en algún $x \in H$, entonces $f'(x) \in H^*$.*

Por el Teorema de Representación de Riesz[3] *existe un elemento $y \in H$ tal que*

$$f'(x)(z) = \langle z, y \rangle \qquad \textit{para toda } z \in H.$$

De manera que la derivada de Fréchet $f'(x)$ puede ser identificada con el elemento $y \in H$ que es conocido como el ***gradiente*** *de f en x, denotado por $\nabla f(x)$, tal que*

$$f'(x)(v) = \langle \nabla f(x), v \rangle \qquad \textit{para toda } v \in H.$$

Junto con esta observación tenemos que:

(i) *Si f es Fréchet-diferenciable en $x \in U \subset H$, entonces f tiene derivada direccional en el punto x*

[3][21] Teo. 3.7.1.

en todas las direcciones $y \in H$ con

$$D_y f(x) = \left\langle f'(x), y \right\rangle \in \mathbb{R}^m$$

en donde $\langle f'(x), y\rangle = (\langle f'(x)_1, y\rangle, \ldots, \langle f'(x)_m, y\rangle)$.

(ii) *Conversamente, si f tiene derivada direccional en todo elemento $x \in U$ en todas las direcciones $y \in X$ y la función*

$$y \longmapsto D_y f(x), \quad y \in X$$

es lineal y continua para toda $x \in U$, entonces existe un elemento $\nabla f(x) \in (X^*)^m$ tal que

$$D_y f(x) = \langle \nabla f(x), y \rangle$$

y si la función $x \longmapsto \nabla f(x) \in (X^*)^m$ es lineal y continua en U, entonces f es Fréchet-diferenciable y se tiene que

$$f'(x) = \nabla f(x).$$

Notemos que ***(i)*** es consecuencia del Teorema 2.2 y del Teorema de Representación de Riesz.

Por otro lado, vemos que para ***(ii)*** la existencia y la representación de $\nabla f(x)$ proviene de la identificación inicial que hicimos y junto con los supuestos iniciales podemos verificar que se cumple

$$\lim_{\substack{y \to 0 \\ y \in H}} \frac{\|f(x+y) - f(x) - \nabla f(x)(y)\|_{\mathbb{R}^m}}{\|y\|_H} = 0.$$

Al agregar la hipótesis de que la función $x \mapsto \nabla f(x)$ sea lineal y continua en U, y como se satisface el límite anterior, todo esto implica[4] que la derivada de Fréchet $f'(x)$ es igual a $\nabla f(x)$.

En el siguiente capítulo retomaremos al movimiento browniano junto con algunas de sus propiedades elementales en virtud de que este proceso será esencial en el desarrollo del presente trabajo.

2.2. Movimiento Browniano

Recordemos a este importante proceso originalmente introducido por el botánico inglés Robert Brown en 1827 y que además es una pieza clave en la teoría moderna de los procesos estocásticos. Especialmente será de suma importancia en la definición de los espacios que vamos a estudiar así como en los procesos de difusión. Retomemos entonces su definición y la filtración con la que trabajaremos.

Definición 2.3. *Dado $(\Omega, \mathscr{F}, \mathbb{P})$ un espacio de probabilidad completo y $B = (B_t)_{t\geq 0}$ un proceso estocástico, decimos que B es un* ***movimiento browniano*** *si cumple que*

(1) B inicia en cero, es decir, $B_0 = 0$.

(2) B tiene incrementos independientes y estacionarios.

(3) B_t se distribuye normal con media 0 y varianza t.

[4][25] Teo. 2. Pág. 116.

Para el espacio de probabilidad completo $(\Omega, \mathscr{F}, \mathbb{P})$, en el que está definido B, denotemos por $\mathscr{N}$ la clase de conjuntos $\mathbb{P}$-nulos y consideremos la filtración

$$\mathscr{F}_t \ = \ \sigma(B_s \, : \, s \leq t) \vee \mathscr{N}$$

entonces B es un movimiento browniano con respecto a $(\mathscr{F}_t)$ y esta filtración satisface las condiciones usuales, es decir, $(\mathscr{F}_t)$ es completa y continua por la derecha[5].

Recordemos que por el criterio de continuidad de Kolmogorov[6] existe una modificación de B_t con trayectorias continuas casi seguramente[7] a la que denotamos por W_t. De ahora en adelante nos enfocaremos en esta modificación continua y esto nos da pie hacia la definición del siguiente espacio.

Definición 2.4. *Sea $T > 0$ un número real fijo y definimos el espacio de funciones*

$$C_0([0,T]) \ := \ \{\omega : [0,T] \to \mathbb{R} \quad \textit{tal que} \ \ \omega(0) = 0 \ \ \textit{y} \ \ \omega \ \textit{es una función continua}\}.$$

Observemos que si dotamos a este espacio con la norma

$$\|\omega\|_\infty := \sup_{t \in [0,T]} |\omega(t)|$$

tenemos entonces que $(C_0([0,T]), \|\cdot\|_\infty)$ es un espacio de Banach[8] separable.

Notemos entonces que las trayectorias del movimiento browniano pertenecen a $C_0([0,T])$ salvo por un conjunto de medida cero de elementos de Ω.

Ahora bien, para identificar algunas propiedades clave del movimiento browniano retomemos una importante función que juega un papel central en las propiedades, tanto de consistencia como distribucionales, de un proceso estocástico a tiempo continuo.

Definimos la ***proyección*** $\pi_t : C_0([0,T]) \to \mathbb{R}$ como

$$\pi_t(\omega) = \omega(t) \qquad \text{para cada } \omega \in C_0([0,T]).$$

En consecuencia, tenemos que la σ-álgebra de Borel en $C_0([0,T])$ coincide[9] con la σ-álgebra generada por los cilindros finito-dimensionales, y además es la mínima σ-álgebra que hace medible a las proyecciones $\{\pi_t \, : \, t \in [0,T]\}$.

Por otro lado, retomemos también las propiedades distribucionales del movimiento browniano. Sean $0 < t_1 < t_2 < \cdots < t_k$, entonces las distribuciones finito-dimensionales de W_t

$$\mathbb{P} \, : \, \mathscr{F} \longrightarrow [0,1] \qquad \text{definidas para cada} \ \ [W_{t_1} \in A_1, \ldots, W_{t_k} \in A_k] \in \mathscr{F}_{t_k}$$

[5] [24] Prop. 2.2. Pág. 24.
[6] [13] Teo. 3.23. Pág. 57.
[7] [20] Teo. 1.9. Pág. 19.
[8] [22] Teo. 4.4. Pág. 69.
[9] [23] Lema 4.2. Pág. 30.

donde $A_i \in \mathcal{B}_{\mathbb{R}}$ para $i \in \{1, \ldots, k\}$ están dadas por[10]

$$\mathbb{P}[\, W_{t_1} \in A_1, \ldots, W_{t_k} \in A_k \,] \;=\; \int_{A_1} \cdots \int_{A_k} \rho(t_1, 0, x_1)\rho(t_2 - t_1, x_1, x_2) \cdots \rho(t_k - t_{k-1}, x_{k-1}, x_k)\, dx_1 \cdots dx_k \tag{4}$$

en donde

$$\rho(t, x, y) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x-y)^2}{2t} \right), \quad x, y \in \mathbb{R},\ t > 0.$$

Siguiendo con esta línea de razonamiento tenemos que es posible trasladar a un proceso estocástico tomando un nuevo punto incial de partida.

En nuestro caso, un movimiento browniano que empieza en x es el proceso trasladado $(x + W_t)_{t \geq 0}$. En este caso si $0 < t_1 < t_2 < \cdots < t_k$, entonces hablamos de la medida de probabilidad

$$\mathbb{P}^x \,:\, \mathscr{F} \longrightarrow [0,1] \qquad \text{definida para cada } \;[W_{t_1} \in A_1, \ldots, W_{t_k} \in A_k] \in \mathscr{F}_{t_k}$$

donde $A_i \in \mathcal{B}_{\mathbb{R}}$ para $i \in \{1, \ldots, k\}$, la cual está dada por

$$\mathbb{P}^x[\, W_{t_1} \in A_1, \ldots, W_{t_k} \in A_k \,] \;=\; \int_{A_1} \cdots \int_{A_k} \rho(t_1, x, x_1)\rho(t_2 - t_1, x_1, x_2) \cdots \rho(t_k - t_{k-1}, x_{k-1}, x_k)\, dx_1 \cdots dx_k$$

Finalmente, retomemos a la medida de probabilidad con la que sustentaremos nuestro estudio posterior. Esta medida nos permitirá establecer el espacio de probabilidad en el cual trabajaremos.

Sean $0 < t_1 < t_2 < \cdots < t_k$, entonces existe una única medida de probabilidad[11] definida en todos los borelianos de $C_0([0,T])$

$$\mu \,:\, \mathcal{B}_{C_0([0,T])} \longrightarrow [0,1]$$

donde $\mathcal{B}_{C_0([0,T])}$ denota a la σ-álgebra de Borel en $C_0([0,T])$ y cumple que

$$\mu(\{\omega \,:\, \omega(t_1) \in A_1, \ldots, \omega(t_k) \in A_k\}) = \mathbb{P}[\, W_{t_1} \in A_1, \ldots, W_{t_k} \in A_k \,]$$

en donde $A_i \in \mathcal{B}_{\mathbb{R}}$ para $i \in \{1, \ldots, k\}$ y $\mathbb{P}$ está determinada por la expresión (4). A la medida de probabilidad μ se le conoce como **la medida de Wiener**.

Una motivación sobre la importancia del espacio que estamos por definir se debe a que, de acuerdo con nuestra exposición de diferenciablidad de Gâteaux y de Fréchet, necesitaremos capturar la información probabilística de una derivada particular en una dirección determinada.

Es justamente el espacio de Cameron-Martin de donde tomaremos las direcciones en las cuales podremos diferenciar y estudiar probabilísticamente a dicha derivada y que además es caracterizada por la medida de Wiener[12].

[10] [23] Teo. 1.20. Pág. 140.
[11] [23] Teo. 2.1. Pág. 165.
[12] Para más detalles se puede consultar el trabajo original de Cameron y Martin en 1944 en la referencia [4].

2.3. Espacio de Cameron-Martin y el Operador Derivada

En virtud de la exposición anterior fijemos entonces $\Omega = C_0([0,T])$ con $\mathscr{F}$ la σ-álgebra de Borel en $C_0([0,T])$ y $\mathbb{P} = \mu$ de manera que $(\Omega, \mathscr{F}, \mu)$ es el espacio de probabilidad en donde vamos a trabajar.

Consideremos a continuación las siguientes definiciones preliminares.

Denotemos por λ la medida de Lebesgue en $([0,T], \mathcal{B}_{[0,T]})$ y definamos los siguientes espacios:

$$\begin{aligned}
L^2(\Omega) &:= \left\{ g : \Omega \to \mathbb{R} \quad | \quad g \text{ es } \mathscr{F}\text{-medible y } \int_\Omega g^2 \, d\mu < \infty \right\} \\
L^2([0,T] \times \Omega) &:= \left\{ g : [0,T] \times \Omega \to \mathbb{R} \quad | \quad g \text{ es } \mathcal{B}_{[0,T]} \otimes \mathscr{F}\text{-medible y } \int_{[0,T]\times\Omega} g^2(u,\omega) \, d\lambda d\mu < \infty \right\} \\
L^2([0,T]) &:= \left\{ g : [0,T] \to \mathbb{R} \quad | \quad g \;\; \text{es una función determinista con } \int_{[0,T]} g^2 \, d\lambda < \infty \right\}
\end{aligned}$$

Definición 2.5. *El espacio*

$$C = \left\{ H \in \Omega \quad | \quad \textit{existe } h \in L^2([0,T]) \;\; \textit{tal que} \;\; H(t) = \int_0^t h(s) \, ds \right\}$$

es llamado el espacio de ***Cameron-Martin.***

Es importante destacar que podemos dotar a C con el producto interno

$$\langle H, G \rangle = \int\limits_{[0,T]} h(t) g(t) \, dt \qquad \text{para toda} \quad H, G \in C,$$

y que además C es denso[13] en Ω con la métrica $d(f,g) = \|f - g\|_\infty$ para $f, g \in \Omega$.

Ahora bien, enfoquemos nuestra atención en la siguiente derivada que tanto hemos venido motivando.

Definición 2.6. *Sea $F : \Omega \to \mathbb{R}$ una variable aleatoria y supongamos que para cualquier $\gamma \in C$ de la forma $\gamma(t) = \int_0^t g(s) \, ds$ se tenga que*

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[F(\omega + \varepsilon\gamma) - F(\omega) \right]$$

existe en un sentido fuerte, es decir, el límite existe en $L^2(\Omega)$. Si este es el caso, entonces definimos el ***operador derivada***

$$\boldsymbol{D}_\gamma F(\omega) := \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[F(\omega + \varepsilon\gamma) - F(\omega) \right].$$

Adicionalmente, suponga que existe $\psi(t,\omega) \in L^2([0,T] \times \Omega)$ tal que

$$\boldsymbol{D}_\gamma F(\omega) = \int_0^T \psi(t,\omega) g(t) \, dt. \tag{5}$$

[13] [23] Prop. (a) Pág. 194.

Entonces decimos que F es ***diferenciable*** *y definimos* ***la derivada de*** *F como*

$$\boldsymbol{D}_t F(\omega) := \psi(t, \omega). \tag{6}$$

Donde $\boldsymbol{D}_t F(\omega) \in L^2(\,[0,T] \times \Omega\,)$ y al conjunto de todas las variables aleatorias diferenciables lo denotamos por $\mathcal{D}_{1,2}$.

Notemos que $\boldsymbol{D}_t$ es un operador lineal tal que

$$\boldsymbol{D}_t : L^2(\Omega) \longrightarrow L^2([0,T] \times \Omega).$$

Ilustremos la definición anterior con el siguiente ejemplo.

Ejemplo 2.3. *Definamos la variable aleatoria*

$$F : \Omega \longrightarrow \mathbb{R}$$
$$F(\omega) = \int_0^T f(s)\, dW(s) = \int_0^T f(s)\, d\omega(s),$$

donde $f : [0,T] \to \mathbb{R}$ es una función determinista en $L^2([0,T])$.

Sea γ un elemento de C tal que $\gamma(t) = \int_0^t g(s)\, ds$ y sea $\varepsilon > 0$, entonces directamente de la definición de F tenemos que

$$F(\omega + \varepsilon\gamma) = \int_0^T f(s) d\big(\,\omega(s) + \varepsilon\gamma(s)\,\big) = \int_0^T f(s) d\omega(s) + \int_0^T f(s) d\varepsilon\gamma(s)$$

y como $\gamma(t) = \int_0^t g(s)\, ds$, entonces $\int_0^T f(s) d\gamma(s)$ coincide con la integral de Lebesgue-Stieltjes y dado que $d\gamma(t) = g(t)dt$, entonces

$$\int_0^T f(s) d\varepsilon\gamma(s) = \varepsilon \int_0^T f(s) g(s)\, ds$$

de manera que en conjunto vemos que

$$\begin{aligned} F(\omega + \varepsilon\gamma) &= \int_0^T f(s) d\omega(s) + \int_0^T f(s) d\varepsilon\gamma(s) \\ &= \int_0^T f(s) d\omega(s) + \varepsilon \int_0^T f(s) g(s)\, ds \end{aligned} \tag{7}$$

y esto implica que

$$\frac{1}{\varepsilon}\left[F(\omega + \varepsilon\gamma) - F(\omega)\right] = \int_0^T f(s) g(s)\, ds \quad \text{para toda } \varepsilon > 0.$$

Si revisamos la definición anterior, en las expresiones (5) y (6), verificamos que $F \in \mathcal{D}_{1,2}$ y que

$$\boldsymbol{D}_t F(\omega) = f(t); \quad t \in [0,T],\ \omega \in \Omega. \tag{8}$$

Si en particular tomamos

$$f(t) = \mathbb{1}_{[0,t_1]}(t)$$

entonces

$$F(\omega) = \int_0^T \mathbb{1}_{[0,t_1]}(s)\, dW(s) = W(t_1, \omega)$$

y por lo tanto

$$\boldsymbol{D}_t(W(t_1,\omega)) = \mathbb{1}_{[0,t_1]}(t). \tag{9}$$

Por otro lado, denotemos por $\mathcal{P}$ a la familia de todas las variables aleatorias de la forma

$$F : \Omega \longrightarrow \mathbb{R}$$
$$F(\omega) = \varphi(\theta_1, \ldots, \theta_n)$$

tal que $\varphi(\theta_1, \ldots, \theta_n)$ es un polinomio de n variables en donde

$$\theta_i = \int_0^T f_i(t)\, dW(t), \quad \text{con } \ f_i \in L^2([0,T]) \ \text{ para } \ 1 \le i \le n.$$

A estas variables aleatorias se les conoce como **polinomios de Wiener**.

Notemos que $\mathcal{P}$ es denso[14] en el espacio de todas las funciones continuas en $[0,T]$ que inician en cero y que son cuadrado integrables para μ.

Si empleamos los cálculos del ejemplo anterior junto con la regla de la cadena usual obtenemos que $\mathcal{P} \subset \mathcal{D}_{1,2}$. A continuación lo demostramos exhibiendo explícitamente su derivada.

Proposición 2.1. *Sea $F(\omega) = \varphi(\theta_1, \ldots, \theta_n) \in \mathcal{P}$. Entonces $F \in \mathcal{D}_{1,2}$ y*

$$\boldsymbol{D}_t F(\omega) = \sum_{i=1}^{n} \frac{\partial \varphi}{\partial x_i}(\theta_1, \ldots, \theta_n) \cdot f_i(t) \tag{10}$$

donde $\theta_i = \int_0^T f_i(t)\, dW(t)$ y $f_i : [0,T] \to \mathbb{R}$ una función determinista en $L^2([0,T])$ para $1 \le i \le n$.

Demostración. Notemos primero que

$$\sup_{s \in [0,T]} \mathbb{E}[\, |W(s)|^n \,] < \infty \qquad \text{para toda } n \in \mathbb{N}$$

lo que nos permitirá tomar con libertad los límites respecto a ε.

Recordemos que la función $\gamma \in C$ es de la forma $\gamma(t) = \int_0^t g(s)\, ds$, $t \in [0,T]$ donde $g : [0,T] \to \mathbb{R}$ es una función determinista en $L^2([0,T])$ y que $\varphi : \mathbb{R}^n \to \mathbb{R}$ es un polinomio y donde $\theta_i = \int_0^T f_i(t)\, dW(t)$ para $1 \le i \le n$.

Entonces observemos que

$$F(\omega + \varepsilon\gamma) = \varphi(\, \theta_1(\omega + \varepsilon\gamma), \ldots, \theta_n(\omega + \varepsilon\gamma)\,).$$

[14] [13] Teo. 13.26. Pág. 266.

Por el cálculo del Ejemplo 2.3 en la expresión (7) podemos ver que

$$\theta_i(\omega + \varepsilon\gamma) = \theta_i + \varepsilon \int_0^T f_i(s)g(s)ds \qquad \text{para } \; i = 1, \dots, n$$

de manera que

$$\varphi(\, \theta_1(\omega + \varepsilon\gamma), \dots, \theta_n(\omega + \varepsilon\gamma)\,) = \varphi(\theta_1 + \varepsilon \langle f_1, g\rangle, \dots, \theta_n + \varepsilon \langle f_n, g\rangle).$$

Como φ es un polinomio, entonces es diferenciable y luego junto con la observación anterior y la regla de la cadena vemos que el siguiente límite existe y que de hecho

$$\frac{1}{\varepsilon}[\, F(\omega+\varepsilon\gamma) - F(\omega)\,] \xrightarrow[\varepsilon\to 0]{} \frac{\partial\varphi}{\partial x_1}(\theta_1, \dots, \theta_n) \cdot \int_0^T f_1(s)g(s)ds + \cdots + \frac{\partial\varphi}{\partial x_n}(\theta_1, \dots, \theta_n) \cdot \int_0^T f_n(s)g(s)ds.$$

En consecuencia junto con la observación que $\sup_{s\in[0,T]} \mathbb{E}[\,|W(s)|^n\,] < \infty$ para toda $n \in \mathbb{N}$ deducimos

$$\frac{1}{\varepsilon}[\, F(\omega+\varepsilon\gamma) - F(\omega)\,] \xrightarrow[\varepsilon\to 0]{L^2(\Omega)} \sum_{i=1}^n \frac{\partial\varphi}{\partial x_i}(\theta_1, \dots, \theta_n) \cdot \boldsymbol{D}_\gamma(\theta_i)$$

en donde $\boldsymbol{D}_\gamma(\theta_i) = \int_0^T f_i(t)g(t)\,dt$ y por la expresiones (7) y (8) del Ejemplo 2.3 tenemos que

$$\boldsymbol{D}_\gamma F(\omega) = \int_0^T \underbrace{\left[\sum_{i=1}^n \frac{\partial\varphi}{\partial x_i}(\theta_1, \dots, \theta_n) \cdot f_i(t) \right]}_{\psi(t,\omega)} g(t)\,dt$$

es decir, $\boldsymbol{D}_\gamma F(\omega) = \int_0^T \psi(t,\omega)g(t)\,dt$ donde $\psi(t,\omega) \in L^2([0,T]\times\Omega)$ y por lo tanto $F \in \mathcal{D}_{1,2}$ y

$$\boldsymbol{D}_t F(\omega) = \psi(t,\omega) = \sum_{i=1}^n \frac{\partial\varphi}{\partial x_i}(\theta_1, \dots, \theta_n) \cdot f_i(t).$$

□

Ahora definamos la norma $\|\cdot\|_{1,2}$ en el espacio $\mathcal{D}_{1,2}$ mediante

$$\|F\|_{1,2} := \Big[\mathbb{E}\big(\,|F|^2\,\big) \; + \; \mathbb{E}\Big(\,\|\boldsymbol{D}F\|^2_{L^2([0,T])}\Big)\Big]^{\frac{1}{2}} \quad \text{para cada } F \in \mathcal{D}_{1,2}$$

y consideremos la definición siguiente.

Definición 2.7. *Sea $\mathbb{D}_{1,2}$ la cerradura de la familia $\mathcal{P}$ respecto de la norma $\|\cdot\|_{1,2}$.*

En consecuencia, observemos que la familia $\mathbb{D}_{1,2}$ consiste de todas las $F \in L^2(\Omega)$ tal que existe $\{F_n\}_{n=1}^\infty \subset \mathcal{P}$ que cumple que

$$F_n \xrightarrow[n\to\infty]{L^2(\Omega)} F$$

y también

$$\{\boldsymbol{D}_t F_n\}_{n=1}^\infty \;\; \text{es convergente en } L^2([0,T]\times\Omega).$$

Veamos entonces si el operador $\boldsymbol{D}_t$ se puede extender a $\mathbb{D}_{1,2}$ de manera adecuada. Para dicho fin comencemos con un resultado preliminar conocido como *integración por partes.*

Lema 2.1. *Sea* $F \in \mathcal{D}_{1,2}$ *acotada y con derivada acotada para cada* $\omega \in \Omega$*. Sean* $\varphi \in \mathcal{D}_{1,2}$ *acotada y* $\gamma \in C$ *de la forma* $\gamma(t) = \int_0^t g(s)\, ds$ *con* $g \in L^2([0,T])$*. Entonces*

$$\mathbb{E}[\boldsymbol{D}_\gamma F \cdot \varphi] \;=\; \mathbb{E}[F \cdot \varphi \cdot \int_0^T g\, dW] \;-\; \mathbb{E}[F \cdot \boldsymbol{D}_\gamma \varphi]. \tag{11}$$

Demostración. Sea $\varepsilon > 0$ y como $g \in L^2([0,T])$, entonces

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T g^2 ds\right)\right] < \infty$$

y por el criterio de Novikov[15] implica que

$$\exp\left(-\varepsilon \int_0^t g\, dW - \frac{\varepsilon^2}{2}\int_0^t g^2 ds\right)$$

es una martingala. Entonces por el teorema de Girsanov y mediante un cambio de medida obtenemos

$$\mathbb{E}[F(\omega + \varepsilon\gamma) \cdot \varphi(\omega)] = \mathbb{E}\left[F(\omega)\varphi(\omega - \varepsilon\gamma) \cdot \exp\left(\varepsilon \int_0^T g\, dW - \frac{\varepsilon^2}{2}\int_0^T g^2 ds\right)\right]. \tag{12}$$

Por hipótesis F y su derivada son acotadas, φ es acotada y por Novikov $\exp(-\varepsilon \int_0^t g\, dW - \frac{\varepsilon^2}{2}\int_0^t g^2 ds)$ es una martingala, entonces podemos encontrar f_1 y f_2 que acoten

$$\begin{aligned} |[F(\omega + \varepsilon\gamma) - F(\omega)] \cdot \varphi(\omega)| \;&\leq\; f_1(\omega) \qquad \text{para toda } \omega \in \Omega \\ \left|F(\omega)\left[\varphi(\omega - \varepsilon\gamma)\exp\left(\varepsilon\int_0^T g\, dW - \frac{\varepsilon^2}{2}\int_0^T g^2 ds\right) - \varphi(\omega)\right]\right| \;&\leq\; f_2(\omega) \qquad \text{para toda } \omega \in \Omega \end{aligned}$$

y tales que $\mathbb{E}[f_1] < \infty$ y $\mathbb{E}[f_2] < \infty$.

Esto justifica el intercambio del límite con la esperanza[16] y junto con la expresión (12) obtenemos

$$\begin{aligned} \mathbb{E}[\boldsymbol{D}_\gamma F(\omega) \cdot \varphi(\omega)] &= \mathbb{E}\left[\lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left[F(\omega + \varepsilon\gamma) - F(\omega)\right] \cdot \varphi(\omega)\right] = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\mathbb{E}\left[F(\omega + \varepsilon\gamma)\varphi(\omega) - F(\omega)\varphi(\omega)\right] \\ &= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\mathbb{E}\left[F(\omega)\left[\varphi(\omega - \varepsilon\gamma)\exp\left(\varepsilon\int_0^T g\, dW - \frac{\varepsilon^2}{2}\int_0^T g^2 ds\right) - \varphi(\omega)\right]\right] \\ &= \mathbb{E}\left[F(\omega) \cdot \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left[\varphi(\omega - \varepsilon\gamma)\exp\left(\varepsilon\int_0^T g\, dW - \frac{\varepsilon^2}{2}\int_0^T g^2 ds\right) - \varphi(\omega)\right]\right] \\ &= \mathbb{E}\left[F(\omega) \cdot \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left[\varphi(\omega - \varepsilon\gamma)\exp\left(\varepsilon\int_0^T g\, dW - \frac{\varepsilon^2}{2}\int_0^T g^2 ds\right)\right]\Bigg|_{\varepsilon=0}\right] \\ &= \mathbb{E}\left[F(\omega) \cdot \varphi(\omega) \cdot \int_0^T g\, dW - F(\omega)\boldsymbol{D}_\gamma\varphi(\omega)\right] \\ &= \mathbb{E}[F \cdot \varphi \cdot \int_0^T g\, dW] \;-\; \mathbb{E}[F \cdot \boldsymbol{D}_\gamma \varphi]. \end{aligned}$$

□

En consecuencia veamos el siguiente resultado sobre el operador $\boldsymbol{D}_t$ que nos permitirá justificar la extensión que deseamos exponer.

[15] [20] Prop. 1.15 Pág. 332.
[16] [1] Cor. 5.7. Pág. 45.

Teorema 2.3. *Sea* $\{H_n\}_{n=1}^{\infty} \subset \mathcal{P}$ *tal que*

$$H_n \xrightarrow[n\to\infty]{L^2(\Omega)} 0$$

y también

$$\{\boldsymbol{D}_t H_n\}_{n=1}^{\infty} \quad \textit{es convergente en } L^2([0,T]\times\Omega).$$

Entonces

$$\lim_{n\to\infty} \boldsymbol{D}_t H_n = 0.$$

Demostración. Por el lema anterior en la expresión (11) tenemos que

$$\mathbb{E}[\boldsymbol{D}_\gamma H_n \cdot \varphi] \;=\; \mathbb{E}[H_n\varphi \cdot \int_0^T g\,dW] - \mathbb{E}[H_n \cdot \boldsymbol{D}_\gamma \varphi] \quad \xrightarrow[n\to\infty]{} 0 \qquad \text{para toda } \varphi \in \mathcal{P}.$$

Las hipótesis implican que $\{\boldsymbol{D}_\gamma H_n\}_{n=1}^{\infty}$ es convergente en $L^2(\Omega)$ y como $\mathcal{P}$ es denso en $L^2(\Omega)$ se sigue entonces que

$$\boldsymbol{D}_\gamma H_n \xrightarrow[n\to\infty]{L^2(\Omega)} 0$$

y como esto es válido para toda $\gamma = \int_0^{\cdot} g\,ds \in C$ concluimos entonces que

$$\boldsymbol{D}_t H_n \xrightarrow[n\to\infty]{L^2([0,T]\times\Omega)} 0$$

□

El resultado anterior permite que el operador $\boldsymbol{D}_t$ se pueda extender a $\mathbb{D}_{1,2}$ mediante la siguiente definición.

Definición 2.8. *Sea* $F \in \mathbb{D}_{1,2}$ *tal que existe* $\{F_n\}_{n=1}^{\infty} \subset \mathcal{P}$ *tal que*

$$F_n \xrightarrow[n\to\infty]{L^2(\Omega)} F$$

y también

$$\{\boldsymbol{D}_t F_n\}_{n=1}^{\infty} \quad \textit{es convergente en } L^2([0,T]\times\Omega).$$

Entonces definimos

$$D_t F := \lim_{n\to\infty} \boldsymbol{D}_t F_n$$

y

$$D_\gamma F := \int_0^T D_t F \cdot g(t)\,dt \quad \textit{para toda } \gamma(t) = \int_0^t g(s)\,ds \textit{ elemento de } C.$$

*A $D_t F$ le llamamos **la derivada de Malliavin** de F.*

Notemos que la **derivada de Malliavin** de F es un operador lineal tal que

$$D_t : L^2(\Omega) \longrightarrow L^2([0,T]\times\Omega).$$

Además $\lim\limits_{n\to\infty} \boldsymbol{D}_t F_n$ está bien definido, ya que si existiera otra $\{G_n\}_{n=1}^{\infty} \subset \mathcal{P}$ tal que

$$G_n \xrightarrow[n\to\infty]{L^2(\Omega)} F$$

y que también $\{\boldsymbol{D}_t G_n\}_{n=1}^{\infty}$ es convergente en $L^2([0,T]\times\Omega)$.

Entonces definimos $H_n := F_n - G_n$ y por el Teorema 2.3 se sigue que

$$\lim_{n\to\infty} \boldsymbol{D}_t F_n \;=\; \lim_{n\to\infty} \boldsymbol{D}_t G_n.$$

Observemos que por el momento contamos con dos *operadores derivada*:

a) Si $F \in \mathcal{D}_{1,2}$ tenemos el **operador derivada** $\boldsymbol{D}_t F$ de acuerdo con la Definición 2.6.

b) Si $F \in \mathbb{D}_{1,2}$ tenemos la **derivada de Malliavin** $D_t F$ de acuerdo con la Definición 2.8.

Pero si $F \in \mathcal{D}_{1,2} \cap \mathbb{D}_{1,2}$, entonces ambas derivadas coinciden como probaremos a continuación.

Teorema 2.4. *Sea $F \in \mathcal{D}_{1,2} \cap \mathbb{D}_{1,2}$ y suponga que $\{F_n\}_{n=1}^{\infty} \subset \mathcal{P}$ tal que*

$$F_n \xrightarrow[n\to\infty]{L^2(\Omega)} F \quad \textit{y que} \quad \{\boldsymbol{D}_t F_n\}_{n=1}^{\infty} \quad \textit{es convergente en } L^2([0,T]\times\Omega). \tag{13}$$

Entonces

$$\boldsymbol{D}_t F \;=\; \lim_{n\to\infty} \boldsymbol{D}_t F_n$$

y luego

$$D_t F \;=\; \boldsymbol{D}_t F \qquad \textit{para} \quad F \in \mathcal{D}_{1,2} \cap \mathbb{D}_{1,2}.$$

Demostración. Las hipótesis en la expresión (13) implican que $\{\boldsymbol{D}_\gamma F_n\}_{n=1}^{\infty}$ es convergente en $L^2(\Omega)$ para cada $\gamma = \int_0^{\cdot} g(s)\,ds \in C$.

Por el Lemma 2.1 en la expresión (11) y por las hipótesis en la expresión (13) tenemos que

$$\mathbb{E}[\,(\boldsymbol{D}_\gamma F_n - \boldsymbol{D}_\gamma F)\cdot\varphi\,] = \mathbb{E}[(F_n - F)\cdot\varphi\cdot\int_0^T g\,dW] \;-\; \mathbb{E}[(F_n - F)\cdot \boldsymbol{D}_\gamma\varphi] \;\longrightarrow 0 \qquad \text{para toda } \varphi \in \mathcal{P}.$$

Y como $\mathcal{P}$ es denso en $L^2(\Omega)$ esto implica que

$$\boldsymbol{D}_\gamma F_n \xrightarrow[n\to\infty]{L^2(\Omega)} \boldsymbol{D}_\gamma F$$

lo que conlleva a que $\boldsymbol{D}_t F = \lim\limits_{n\to\infty} \boldsymbol{D}_t F_n$ y por lo tanto $D_t F \;=\; \boldsymbol{D}_t F$ para $F \in \mathcal{D}_{1,2} \cap \mathbb{D}_{1,2}$. $\square$

Ilustremos un par de ejemplos de derivadas de Malliavin haciendo uso del resultado anterior junto con las herramientas anteriormente expuestas.

Ejemplo 2.4. *Sea*

$$\begin{aligned} F &: \Omega \longrightarrow \mathbb{R} \\ F(\omega) &= W(T) \end{aligned}$$

Si definimos

$$\begin{aligned} f &: [0,T] \longrightarrow \mathbb{R} \\ f(t) &= \mathbb{1}_{[0,T]}(t). \end{aligned}$$

entonces $f \in L^2([0,T])$ *y si tomamos*

$$\theta = \int_0^T f(t)\, dW(t) \;=\; \int_0^T \mathbb{1}_{[0,T]}(t)\, dW(t) \;=\; W(T)$$

con el polinomio de primer grado

$$\varphi(\theta) \;=\; \theta$$

se sigue entonces que $F \in \mathcal{P}$ *y además* $F \in \mathcal{D}_{1,2} \cap \mathbb{D}_{1,2}$.

Retomemos el Ejemplo 2.3 en la expresión (9) *donde probamos que*

$$\boldsymbol{D}_t(W(t_1,\omega)) = \mathbb{1}_{[0,t_1]}(t).$$

y como ya exhibimos que $F \in \mathcal{D}_{1,2} \cap \mathbb{D}_{1,2}$ *entonces por el Teorema 2.4*

$$D_t(W(T,\omega)) \;=\; \mathbb{1}_{[0,T]}(t) \;=\; 1 \quad \textit{para} \;\; t \in [0,T].$$

Consideremos ahora

$$\begin{aligned} F &: \Omega \longrightarrow \mathbb{R} \\ F(\omega) &= \int_0^T t^2\, dW(t) \end{aligned}$$

Si definimos

$$\begin{aligned} f &: [0,T] \longrightarrow \mathbb{R} \\ f(t) &= t^2 \end{aligned}$$

entonces $f \in L^2([0,T])$ *y si tomamos*

$$\theta = \int_0^T f(t)\, dW(t) \;=\; \int_0^T t^2\, dW(t)$$

con el polinomio de primer grado

$$\varphi(\theta) = \theta$$

se sigue que $F \in \mathcal{P}$ *y además* $F \in \mathcal{D}_{1,2} \cap \mathbb{D}_{1,2}$. *En consecuencia, por la Proposición 2.1 y junto con el Teorema 2.4 obtenemos que*

$$D_t F(\omega) \;=\; \frac{\mathrm{d}\varphi}{\mathrm{d}x}(\theta) \cdot f(t) \;=\; 1 \cdot t^2 \;=\; t^2.$$

En línea con la derivada de Malliavin anterioremente expuesta, nuestro objetivo será calcular la derivada de Malliavin de un proceso de difusión la cual nos conllevará hacia una de las aplicaciones esenciales de nuestro trabajo y que a su vez retomaremos más adelante con mayor profundidad. Por esta razón es que retomaremos, en el siguiente capítulo, a dicho proceso junto con sus propiedades elementales siguiendo los resultados del libro de Oksendal[17].

2.4. Procesos de Difusión

Esta sección tiene como objetivo retomar algunas propiedades importantes de los procesos de difusión, puesto que dichos procesos recibirán la principal aplicación de las técnicas expuestas en el presente trabajo. La demostración de cada resultado mencionado puede ser consultada en su correspondiente referencia detallada en los pies de página.

Comencemos este apartado enunciando el *teorema de existencia y unicidad*[18] para una ecuación diferencial estocástica.

Teorema 2.5. *Sean $T > 0$ y $b(\cdot,\cdot) : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$, $\sigma(\cdot,\cdot) : [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n\times m}$ funciones medibles tales que*

$$|b(t,x)| + |\sigma(t,x)| \leq N(1 + |x|); \qquad \textit{para} \;\; x \in \mathbb{R}^n, \;\; t \in [0,T]$$

con N una constante, $|\sigma|^2 = \sum |\sigma_{ij}|^2$ y que además cumplen que

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq M|x - y|; \qquad \textit{para} \;\; x, y \in \mathbb{R}^n, \;\; t \in [0,T]$$

con M una constante.

Sea Z una variable aleatoria independiente de la σ-álgebra generada por $\{W_s : s \geq 0\}$ y denotada por $\mathscr{F}_\infty^{(m)}$; donde W_t es un movimiento browniano m-dimensional, y que además $\mathbb{E}[|Z|^2] < \infty$.

Entonces la ecuación diferencial estocástica

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t\,, \qquad \textit{con} \;\; 0 \leq t \leq T, \;\; X_0 = Z$$

tiene una única solución X_t con trayectorias continuas tal que si $\mathscr{F}_t^Z$ es la σ-álgebra generada por Z y por $\{W_s : s \leq t\}$, entonces

$$X_t(\omega) \;\; \textit{es} \;\; \mathscr{F}_t^Z\textit{-adaptado} \quad \textit{y} \quad \mathbb{E}\left[\int_0^T |X_t|^2 dt\right] < \infty.$$

Ahora definamos y recordemos algunas propiedades importantes de los procesos de difusión.

Definición 2.9. *Decimos que el proceso estocástico*

$$X_t(\omega) \;=\; X(t,\omega) \,:\, [0,\infty) \times \Omega \longrightarrow \mathbb{R}^n$$

[17] [18] Capítulo 7.
[18] [18] Teo. 5.2.1. Pág. 66.

es un ***proceso de difusión*** *si satisface la ecuación diferencial estocástica*

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t \,, \qquad \textit{con} \;\; t \geq s\,;\; X_s = x \tag{14}$$

donde W_t *es un movimiento browniano* m*-dimensional y las funciones medibles* $b : \mathbb{R}^n \to \mathbb{R}^n$ *y* $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ *satisfacen las hipótesis del teorema de existencia y unicidad; que en este caso*

$$|b(x) - b(y)| + |\sigma(x) - \sigma(y)| \leq M|x - y|; \qquad \textit{para} \;\; x, y \in \mathbb{R}^n$$

con M *una constante y* $|\sigma|^2 = \sum |\sigma_{ij}|^2$.

Denotemos a la única solución X_t de la ecuación diferencial estocástica (14) por

$$X_t^{s,x} \qquad \text{para} \;\; t \geq s$$

y si $s = 0$ entonces la denotamos por

$$X_t^{0,x} \;=\; X_t^x.$$

Observemos que si $\tilde{W}$ es el movimiento browniano $\tilde{W}_v := W_{s+v} - W_s$; $v \geq 0$, entonces

$$X_{s+h}^{s,x} \;=\; x + \int_s^{s+h} b(X_u^{s,x})\, du + \int_s^{s+h} \sigma(X_u^{s,x})\, dW_u \;=\; x + \int_0^h b(X_{s+v}^{s,x})\, dv + \int_0^h \sigma(X_{s+v}^{s,x})\, d\tilde{W}_v$$

en donde $u = s + v$ y además se tiene que

$$X_h^{0,x} \;=\; x + \int_0^h b(X_v^{0,x})\, dv + \int_0^h \sigma(X_v^{0,x})\, dW_v$$

y como $\{W_v\}_{v\geq 0}$ y $\{\tilde{W}_v\}_{v\geq 0}$ tienen las mismas distribuciones finito-dimensionales se sigue que por la unicidad débil de la solución de la ecuación diferencial estocástica

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t \,, \qquad X_0 = x$$

los procesos

$$\{X_{s+h}^{s,x}\}_{h\geq 0} \qquad \text{y} \qquad \{X_h^{0,x}\}_{h\geq 0}$$

tienen las mismas distribuciones finito-dimensionales, es decir, el proceso $\{X_t\}_{t\geq 0}$ es *homogéneo en el tiempo.*

Definición 2.10. *Sea* $\mathcal{M}_t$ *la* σ*-álgebra generada por* $\{X_s^y \,:\, s \leq t,\, y \in \mathbb{R}^n\}$ *y* $\mathcal{M} = \mathcal{M}_\infty$. *Definimos las* ***leyes de probabilidad*** $\mathbb{Q}^x$ *del proceso* $\{X_t\}_{t\geq 0}$

$$\mathbb{Q}^x \,:\, \mathcal{M} \longrightarrow [0, 1]$$

definidas en los elementos de $\mathcal{M}_{t_k}$ *por*

$$\mathbb{Q}^x[X_{t_1} \in E_1, \cdots, X_{t_k} \in E_k] \;=\; \mathbb{P}^0[X_{t_1}^x \in E_1, \cdots, X_{t_k}^x \in E_k]$$

donde $E_i \subset \mathbb{R}^n$ *conjuntos de Borel para* $1 \leq i \leq k$.

Notemos que si $\mathscr{F}_t^{(m)}$ es la σ-álgebra generada por $\{W_s : s \leq t\}$, entonces por el teorema de existencia y unicidad se tiene que X_t es medible con respecto a $\mathscr{F}_t^{(m)}$ y por lo tanto $\mathcal{M}_t \subset \mathscr{F}_t^{(m)}$.

En consecuencia recordemos dos propiedades importantes de los procesos de difusión.

Teorema 2.6 (**Propiedad de Markov**[19])**.** *Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función Borel-medible y acotada. Entonces, para $t, h \geq 0$ se tiene que*

$$\mathbb{E}^x[f(X_{t+h})|\mathscr{F}_t^{(m)}](\omega) \;=\; \mathbb{E}^{X_t(\omega)}[f(X_h)].$$

donde $\mathbb{E}^x$ denota a la esperanza con respecto a la medida de probabilidad $\mathbb{Q}^x$ y el lado derecho de la identidad denota a la función $y \mapsto \mathbb{E}^y[f(X_h)] = \mathbb{E}[f(X_h^y)]$, es decir, la esperanza con respecto a la medida de probabilidad $\mathbb{P}^0$ evaluada en $y = X_t(\omega)$.

Definición 2.11. *Sea $\mathcal{H}$ el espacio de funciones reales y medibles respecto a $\mathcal{M}_\infty$. Sean $s, t \geq 0$ y $\eta \in \mathcal{H}$ definimos **el operador shift***

$$\theta_t : \mathcal{H} \longrightarrow \mathcal{H}$$
$$(\theta_t \eta)_s \;=\; \eta_{s+t}$$

Ilustremos la definición anterior con el siguiente ejemplo. Sea $\eta \in \mathcal{H}$ de la forma

$$\eta \;=\; g_1(X_{t_1}) g_2(X_{t_2}) \cdots g_k(X_{t_k})$$

donde $g_i : \mathbb{R}^n \to \mathbb{R}$ es una función Borel-medible y con $t_i \geq 0$ para $1 \leq i \leq k$, entonces

$$\theta_t \eta \;=\; g_1(X_{t_1+t}) g_2(X_{t_2+t}) \cdots g_k(X_{t_k+t}).$$

Teorema 2.7 (**Propiedad Fuerte de Markov**[20])**.** *Sea τ un tiempo de paro con respecto a $\mathscr{F}_t^{(m)}$ tal que $\tau < \infty$ casi seguramente y sea η un elemento de $\mathcal{H}$ acotada, entonces*

$$\mathbb{E}^x[\theta_\tau \eta | \mathscr{F}_\tau^{(m)}] \;=\; \mathbb{E}^{X_\tau}[\eta].$$

Para identificar el operador diferencial asociado a un proceso de difusión veamos primero la siguiente definición.

Definición 2.12. *Sea $\{X_t\}_{t\geq 0}$ un proceso de difusión, definimos **el generador infinitesimal** A de X_t como*

$$Af(x) \;=\; \lim_{t \downarrow 0} \frac{\mathbb{E}^x[f(X_t)] - f(x)}{t} \qquad \text{con} \;\; x \in \mathbb{R}^n.$$

Denotamos por $\mathcal{D}_A(x)$ al conjunto de funciones $f : \mathbb{R}^n \to \mathbb{R}$ tales que el límite anterior existe y $\mathcal{D}_A$ al conjunto de funciones tales que el límite existe para toda $x \in \mathbb{R}^n$.

En consecuencia, para asociar al generador infinitesimal A con los coeficientes b y σ de la ecuación diferencial estocástica correspondiente al proceso de difusión recordemos el siguiente resultado[21].

[19] [18] Teo. 7.1.2. Pág. 109.
[20] [18] Teo. 7.2.4. Pág. 111-113.
[21] [18] Teo. 7.3.3. Pág. 117.

Teorema 2.8. *Sea X_t el proceso de difusión*

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t$$

Si $f \in C_0^2(\mathbb{R}^n)$, es decir, si $f : \mathbb{R}^n \to \mathbb{R}$ es continuamente diferenciable de orden dos y con soporte compacto, entonces $f \in \mathcal{D}_A$ y además

$$Af(x) = \sum_i b_i(x)\frac{\partial f}{\partial x_i} + \tfrac{1}{2}\sum_{i,j}(\sigma\sigma^T)_{i,j}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}. \tag{15}$$

Ilustremos el teorema anterior con el movimiento browniano n-dimensional. A este proceso lo podemos ver como la solución de la ecuación diferencial estocástica

$$dX_t = dW_t$$

donde $b = 0$ y $\sigma = I_n$ la matriz identidad n-dimensional. Entonces el generador infinitesimal de W_t está dado por

$$Af = \tfrac{1}{2}\sum_{i,j}\frac{\partial^2 f}{\partial x_i^2} \qquad \text{con } f \in C_0^2(\mathbb{R}^n).$$

es decir, $A = \frac{1}{2}\Delta$, donde Δ es el operador Laplaciano.

Al operador diferencial del lado derecho de la expresión (15) usualmente se le denota por L.

De manera más general tomemos el dominio D un subconjunto abierto y conexo de $\mathbb{R}^n$, el operador diferencial L definido en $C^2(\mathbb{R}^n)$ de la forma

$$L = \sum_{i=1}^n b_i(x)\frac{\partial}{\partial x_i} + \sum_{i,j=1}^n a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j}$$

donde $b_i(x)$ y $a_{ij}(x) = a_{ji}(x)$ son funciones continuas. Decimos que L es un **operador elíptico** si los valores propios de la matriz simétrica $a(x) = [a_{ij}(x)]_{i,j=1}^n$ *son positivos para toda x.*

Notemos que por el teorema anterior el operador diferencial asociado al proceso de difusión coincide con su generador infinitesimal en $C_0^2(\mathbb{R}^n)$ mediante

$$\frac{1}{2}\sigma(x)\sigma(x)^T = [a_{ij}(x)] \qquad b(x) = [b_i(x)]$$

donde las funciones σ y b satisfacen las condiciones del teorema de existencia y unicidad.

En consecuencia retomemos los siguientes resultados[22].

Teorema 2.9. *Si X_t es un proceso de difusión con generador infinitesimal A, entonces para toda $f \in C_0^2(\mathbb{R}^n)$ el proceso*

$$M_t = f(X_t) - \int_0^t Af(X_s)ds$$

es una martingala con respecto a $\mathcal{M}_t$.

[22] [18] Teo. 8.3.1. Pág. 139.

Teorema 2.10 (Fórmula de Dynkin[23]). *Sea $f \in C_0^2(\mathbb{R}^n)$ y suponga que τ es un tiempo de paro tal que $\mathbb{E}^x[\tau] < \infty$. Entonces*

$$\mathbb{E}^x[f(X_\tau)] \;=\; f(x) + \mathbb{E}^x\left[\int_0^\tau Af(X_s)\,ds\right].$$

Ahora definamos un operador que está ampliamente relacionado con el generador infinitesimal de una difusión pero que resulta más adecuado en otros contextos; por ejemplo el problema de Dirichlet.

Definición 2.13. *Sea $\{X_t\}_{t\geq 0}$ un proceso de difusión, definimos* ***el operador característico*** *$\mathcal{A}$ de X_t como*

$$\mathcal{A}f(x) \;=\; \lim_{U\downarrow x}\frac{\mathbb{E}^x[f(X_{\tau_U})] - f(x)}{\mathbb{E}^x[\tau_U]}$$

en donde los conjuntos U son abiertos y decrecen hacia el punto x en el sentido de que $U_k \supset U_{k+1}$ y $\bigcap_k U_k = \{x\}$ y en donde $\tau_U = \inf\{t > 0 \;:\; X_t \notin U\}$.

Denotamos por $\mathcal{D}_\mathcal{A}$ al conjunto de funciones $f : \mathbb{R}^n \to \mathbb{R}$ tales que el límite anterior existe para toda $x \in \mathbb{R}^n$ y para toda $\{U_k\}$. Si $\mathbb{E}^x[\tau_U] = \infty$ para todo abierto U tal que $x \in U$, entonces definimos $\mathcal{A}f(x) = 0$.

Cabe mencionar que siempre ocurre la contención[24] $\mathcal{D}_A \subset \mathcal{D}_\mathcal{A}$ y además $Af = \mathcal{A}f$ para toda $f \in \mathcal{D}_A$. Más aún contamos con el siguiente resultado[25].

Teorema 2.11. *Sea $f \in C^2$, es decir, una función continuamente diferenciable de orden dos. Entonces $f \in \mathcal{D}_\mathcal{A}$ y*

$$\mathcal{A}f \;=\; \sum_i b_i \frac{\partial f}{\partial x_i} \;+\; \tfrac{1}{2}\sum_{i,j}(\sigma\sigma^T)_{i,j}\frac{\partial^2 f}{\partial x_i \partial x_j}.$$

Por otro lado, observemos que si el proceso difusión $X_t = X_t^x$ de la forma

$$dX_t \;=\; b(X_t)dt + \sigma(X_t)dW_t \qquad t \geq 0,\; X_0 = x$$

donde $b : \mathbb{R}^n \to \mathbb{R}^n$ y $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ son funciones que satisfacen las condiciones del teorema de existencia y unicidad y W_t un movimiento browniano m-dimensional.

Entonces podemos definir un nuevo proceso de difusión $Y_t = Y_t^{(s,x)}$ en $\mathbb{R}^{n+1}$ mediante

$$Y_t \;=\; \begin{bmatrix} s+t \\ X_t^x \end{bmatrix} \qquad t \geq 0.$$

De manera que

$$dY_t \;=\; \begin{bmatrix} 1 \\ b(X_t) \end{bmatrix} dt \;+\; \begin{bmatrix} 0 \\ \sigma(X_t) \end{bmatrix} dW_t \;=\; \hat{b}(Y_t)dt + \hat{\sigma}(Y_t)W_t$$

en donde

$$\hat{b}(\eta) = \hat{b}(t,\xi) = \begin{bmatrix} 1 \\ b(\xi) \end{bmatrix} \in \mathbb{R}^{n+1}, \qquad \hat{\sigma}(\eta) = \hat{\sigma}(t,\xi) = \begin{bmatrix} 0\cdots 0 \\ \hline \sigma(\xi) \end{bmatrix} \in \mathbb{R}^{(n+1)\times m}$$

[23] [18] Teo. 7.4.1. Pág. 118.
[24] [10] Pág. 143.
[25] [18] Teo. 7.5.4. Pág. 121.

con $\eta = (t, \xi) \in \mathbb{R} \times \mathbb{R}^n$ y de esta manera Y_t es un proceso de difusión que comienza en $y = (s, x)$.

Es decir, podemos incorporar el parámetro temporal en los coeficientes de la ecuación diferencial estocástica del proceso de difusión y tal que su operador característico $\hat{\mathcal{A}}$ está dado por

$$\hat{\mathcal{A}}\phi(s,x) \;=\; \frac{\partial \phi}{\partial s}(s,x) + \mathcal{A}\phi(s,x), \qquad \phi \in C^2(\mathbb{R} \times \mathbb{R}^n)$$

en donde $\mathcal{A}$ denota al operador característico de X_t.

Para lograr nuestro objetivo de calcular la derivada de Malliavin de una difusión necesitaremos primero introducir un operador relacionado con la derivada de Malliavin, que de hecho será su operador adjunto.

Este operador es de suma importancia en el cálculo de Malliavin y se le conoce como la integral de Skorokhod.

2.5. Integral de Skorokhod

Para definir apropiadamente a la integral de Skorokhod notemos algunas extensiones que se pueden hacer sobre la derivada de Malliavin en virtud de lo que previamente hemos desarrollado.

Recordemos que estamos trabajando con la modificación continua del movimiento browniano $W = \{W_t : t \in [0,T]\}$ definido en el espacio de probabilidad canónico $(\Omega, \mathscr{F}, \mathbb{P})$, es decir, $\Omega = C_0([0,T])$, con $\mathscr{F}$ la σ-álgebra completada de Borel en Ω y con $\mathbb{P}$ la medida de Wiener.

También retomemos al espacio separable de Hilbert $L^2([0,T])$, que de ahora en adelante lo denotaremos por H, definido como

$$H = \left\{ g : [0,T] \to \mathbb{R} \quad | \quad g \;\text{ es una función determinista con} \int_{[0,T]} g^2 \, d\lambda < \infty \right\}$$

en donde λ es la medida de Lebesgue en $([0,T], \mathcal{B}_{[0,T]})$ y con producto interno

$$\langle h, g \rangle_H = \int_0^T h(t) g(t) \, dt \qquad \text{para cada} \quad h, g \in H.$$

Denotemos por $C_b^\infty(\mathbb{R}^n)$ al conjunto de todas las funciones infinitamente diferenciables $f : \mathbb{R}^n \to \mathbb{R}$ tales que f y todas sus derivadas parciales están acotadas.

Denotemos por $\mathcal{S}$ a la clase de variables aleatorias de la forma

$$F \;=\; f(W_{t_1}, \ldots, W_{t_n}) \tag{16}$$

donde f pertenece a $C_b^\infty(\mathbb{R}^n)$ y $t_1, \ldots, t_n \in [0,T]$.

En consecuencia, si $p \geq 1$ y la variable aleatoria $F : \Omega \to \mathbb{R}$ es como en la expresión (16), entonces la derivada de Malliavin

$$D\,:\,\mathcal{S}\subset L^p(\Omega)\longrightarrow L^p(\Omega;L^2([0,T]))$$

está dada por

$$D_tF\;=\;\sum_{i=1}^n\frac{\partial f}{\partial x_i}(W_{t_1},\ldots,W_{t_n})\mathbb{1}_{[0,t_i]}(t)\,,\qquad t\in[0,T].$$

Observemos que si $F=f(W_{t_1},\ldots,W_{t_n})\in\mathcal{S}$ con $0\leq t_1<\cdots<t_n\leq T$ y para $h\in H$, entonces

$$\begin{aligned}\langle DF,h\rangle_H&=\sum_{i=1}^n\frac{\partial f}{\partial x_i}(W_{t_1},\ldots,W_{t_n})\left\langle\mathbb{1}_{[0,t_i]},h\right\rangle_H\\&=\sum_{i=1}^n\frac{\partial f}{\partial x_i}(W_{t_1},\ldots,W_{t_n})\int_0^{t_i}h(t)dt\\&=\frac{\mathrm{d}}{\mathrm{d}\varepsilon}F\left(\omega+\varepsilon\int_0^{\cdot}h(t)dt\right)\Big|_{\varepsilon=0}\end{aligned}$$

de manera que $\langle DF,h\rangle_H$ coincide con la derivada direccional de F en el punto ω en la dirección $\int_0^{\cdot}h(t)dt$ que es un elemento del espacio de Cameron-Martin.

Para cada $p\geq 1$ fijo, la derivada de Malliavin es un operador lineal que puede extenderse[26] hacia la cerradura de $\mathcal{S}$ con respecto de la norma

$$\|F\|_{1,p}^p\;=\;\mathbb{E}(|F|^p)+\mathbb{E}\left(\left(\int_0^T|D_tF|^2dt\right)^{\frac{p}{2}}\right).$$

Denotemos por $\mathbb{D}^{1,p}$ a la completación de $\mathcal{S}$ con respecto a $\|\cdot\|_{1,p}$.

Para cualquier número natural k y una variable aleatoria F en $\mathbb{D}^{1,p}$ se define la derivada D^kF como k-veces la iteración del operador D de tal manera que $D^k_{t_1,\ldots,t_k}F$ es un proceso estocástico k-paramétrico resultado de iterar k-veces el operador D.

Por ejemplo para $F\in\mathbb{D}^{1,p}$ tenemos

$$D^2F\;=\;D(DF)\;=\;\sum_{i,j=1}^n\frac{\partial f^2}{\partial x_i\partial x_j}(W_{t_1},\ldots,W_{t_n})(\mathbb{1}_{[0,t_i]}\otimes\mathbb{1}_{[0,t_j]})$$

donde $\otimes$ denota el producto tensorial.

En consecuencia, para cada $p\geq 1$ y cualquier número natural k el operador D^k puede extenderse[27] hacia la cerradura de $\mathcal{S}$ con respecto de la norma

$$\|F\|_{k,p}^p\;=\;\mathbb{E}(|F|^p)+\sum_{j=1}^k\mathbb{E}\left(\left\|D^jF\right\|_{L^2([0,T]^j)}^p\right).$$

Denotemos por $\mathbb{D}^{k,p}$ a la completación de $\mathcal{S}$ con respecto a $\|\cdot\|_{k,p}$.

También se puede extender la definición de la derivada de Malliavin sobre una familia de variables

[26] [16] Prop. 1.2.1. Pág. 26.
[27] [16] Pág. 27.

aleatorias suaves con valores en un espacio separable de Hilbert en los reales; al cual denotamos por V. Tomemos la familia

$$\mathcal{S}_V \;=\; \left\{ u = \sum_{j=1}^{n} F_j v_j \;\middle|\; F_j \in \mathcal{S},\ v_j \in V,\ n \geq 1 \right\}$$

tal que para cada $k \geq 1$ definimos la derivada D^k de $u \in \mathcal{S}_V$ como

$$D^k u \;=\; \sum_{j=1}^{n} D^k F_j \otimes v_j .$$

En consecuencia tenemos que para cada $p \geq 1$ y cualquier número natural k el operador D^k puede extenderse[28] hacia la cerradura de $\mathcal{S}_V$ con respecto de la norma

$$\|u\|_{k,p,V}^p \;=\; \mathbb{E}(\|u\|_V^p) + \sum_{j=1}^{k} \mathbb{E}\left(\left\| D^j u \right\|_{H^{\otimes j} \otimes V}^p \right).$$

Denotamos por $\mathbb{D}^{k,p}(V)$ a la completación de $\mathcal{S}_V$ con respecto a $\|\cdot\|_{k,p,V}$.

Ahora sí podemos introducir a la integral de Skorokhod.

Definición 2.14. *Definimos* ***la integral de Skorokhod*** *como el operador adjunto de la derivada de Malliavin, denotado por δ, un operador lineal, no acotado y cerrado*

$$\delta \,:\, \mathrm{Dom}\,\delta \subset L^2(\Omega; L^2([0,T])) \longrightarrow L^2(\Omega)$$

tal que

(i) El dominio de δ, al que denotamos por $\mathrm{Dom}\,\delta$, consiste de aquellas variables aleatorias u que pertenecen a $L^2(\Omega; L^2([0,T]))$ tales que

$$\left| \mathbb{E}\left[\langle DF, u \rangle_H \right] \right| \leq c_u \|F\|_{L^2(\Omega)} \qquad \textit{para toda} \;\; F \in \mathbb{D}^{1,2}$$

en donde c_u es una constante que depende únicamente de u.

(ii) Si $u \in \mathrm{Dom}\,\delta$, entonces $\delta(u)$ es aquel elemento en $L^2(\Omega)$ caracterizado por satisfacer que

$$\mathbb{E}(F\delta(u)) \;=\; \mathbb{E}\left(\int_0^T D_t F u_t \, dt \right) \qquad \textit{para toda} \;\; F \in \mathbb{D}^{1,2}. \tag{17}$$

Enunciemos el siguiente resultado[29] del cálculo de Malliavin el cual nos permitirá identificar a los elementos del dominio de δ.

Teorema 2.12. *Para toda $p > 1$ y $k \geq 1$ el operador δ es continuo de $\mathbb{D}^{k,p}(H)$ en $\mathbb{D}^{k-1,p}$. Es decir, si $u \in \mathbb{D}^{k,p}(H)$, entonces existe una constante finita y positiva $c_{k,p}$ tal que*

$$\|\delta(u)\|_{k-1,p} \;\leq\; c_{k,p} \, \|u\|_{k,p,H}$$

y en particular $\mathbb{D}^{k,p}(H) \subset \mathrm{Dom}\,\delta$.

[28] [16] Pág. 31.
[29] [16] Prop. 1.5.7. Pág. 78.

El resultado anterior implica, en particular, que el operador δ es continuo de $\mathbb{D}^{1,p}(H)$ en $L^p(\Omega)$ y además existe una constante finita y positiva C_p tal que

$$\|\delta(u)\|_p \leq C_p \left(\|u\|_{L^p(\Omega;H)} + \|Du\|_{L^p(\Omega;H\otimes H)} \right)$$

en consecuencia el espacio $\mathbb{D}^{1,p}(H)$ pertenece al dominio de δ y tenemos la estimación

$$\mathbb{E}(|\delta(u)|^p) \leq C_p \left[\mathbb{E}\left(\left(\int_0^T |u_s|^2 ds \right)^{p/2} \right) + \mathbb{E}\left(\left(\int_0^T \int_0^T |D_s u_t|^2 ds dt \right)^{p/2} \right) \right]. \tag{18}$$

Es importante destacar que el operador δ es una extensión de la integral esocástica de Itô hacia procesos que no son adaptados. Para demostrar esto necesitaremos primero los siguientes resultados.

Lema 2.2. *Sean $p, q > 1$ fijos y tales que $1/p + 1/q = 1$. Sea $F \in \mathbb{D}^{1,p}$ y sea $u \in Dom\,\delta$ tal que $u \in L^q(\Omega; L^2([0,T]))$ y $\delta(u) \in L^q(\Omega)$. Entonces $Fu \in Dom\,\delta$ y*

$$\delta(Fu) \;=\; F\delta(u) - \langle DF, u \rangle_H \tag{19}$$

Demostración. Denotamos por A al miembro del lado derecho de la expresión (19) el cual pertenece a $L^1(\Omega)$ por hipótesis. Queremos probar que se cumple la identidad (17) de la definición, es decir,

$$\mathbb{E}(GA) \;=\; \mathbb{E}(\langle DG, Fu \rangle_H) \qquad \text{para toda } \; G \in \mathbb{D}^{1,2}. \tag{20}$$

Notemos que por la regla del producto

$$D(GF) = (DG)F + G(DF) \implies (DG)Fu = D(GF)u - G(DF)u$$

así que podemos reescribir el lado derecho de la expresión (20) como

$$\mathbb{E}(\langle DG, Fu \rangle_H) = \mathbb{E}(\langle D(GF), u \rangle_H) - \mathbb{E}(G \langle DF, u \rangle_H)$$

y entonces necesitamos probar que

$$\mathbb{E}(\langle D(GF), u \rangle_H) = \mathbb{E}(GF\delta(u)) \qquad \text{para toda } \; G \in \mathbb{D}^{1,2}.$$

Notemos que esto último es cierto, por definición de $\delta(u)$, si $F \in \mathcal{S}$ y $G \in \mathbb{D}^{1,2}$.

Para demostrar el caso en que $F \in \mathbb{D}^{1,p}$ es suficiente con tomar una sucesión de variables aleatorias suaves $\{F_n\}_{n\in\mathbb{N}}$ tal que $F_n \in \mathcal{S}$ y F_n converge a F en $\mathbb{D}^{1,p}$, entonces por la continuidad del producto interno y por el teorema de convergencia dominada obtenemos que

$$\mathbb{E}(\langle D(GF), u \rangle_H) = \lim_{n\to\infty} \mathbb{E}(\langle D(GF_n), u \rangle_H) = \lim_{n\to\infty} \mathbb{E}(GF_n\delta(u)) = \mathbb{E}(GF\delta(u))$$

y junto con el Teorema 2.12 concluimos que $Fu \in Dom\,\delta$ y cumple la identidad deseada. $\square$

Veamos el siguiente lema técnico. Tomemos un intervalo $[c,d] \subset [0,T]$ y sea $\mathscr{F}_{[c,d]}$ la σ-álgebra generada por $\{W_s \,:\, s \in [c,d]\}$.

Lema 2.3. *Sea F una variable aleatoria tal que $F \in \mathbb{D}^{1,2} \cap L^2(\Omega, \mathscr{F}_{[c,d]^c}, \mathbb{P})$. Entonces*

$$D_t F = 0 \quad \textit{para casi toda} \quad (t,\omega) \in [c,d] \times \Omega.$$

Demostración. Para un intervalo $[a,b] \subset [0,T]$ denotemos $W(\mathbb{1}_{[a,b]}) = \int_0^T \mathbb{1}_{[a,b]}(s) dW_s$. Tomemos $f \in C_b^\infty(\mathbb{R}^n)$ y sean $[a_i, b_i] \subset [c,d]^c$ para $i = 1, \ldots n$.

Si $F = f(W(\mathbb{1}_{[a_1,b_1]}), \ldots, W(\mathbb{1}_{[a_n,b_n]}))$ se tiene que $F \in \mathcal{S} \cap L^2(\Omega, \mathscr{F}_{[c,d]^c}, \mathbb{P})$ y entonces

$$D_t F \;=\; \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(W(\mathbb{1}_{[a_1,b_1]}), \ldots, W(\mathbb{1}_{[a_n,b_n]}))\mathbb{1}_{[a_i,b_i]}(t)$$

y como $\mathbb{1}_{[a_i,b_i]}(t) = 0$ para toda $t \in [c,d]$ pues $[a_i,b_i] \subset [c,d]^c$ para $i = 1, \ldots n$, entonces $D_t F = 0$ para toda $(t,\omega) \in [c,d] \times \Omega$.

Para el caso en que $F \in \mathbb{D}^{1,2} \cap L^2(\Omega, \mathscr{F}_{[c,d]^c}, \mathbb{P})$ existe una sucesión de variables aleatorias suaves $\{F_n\}_{n\in\mathbb{N}}$ tal que $F_n \in \mathcal{S}$ y F_n converge a F en $L^2(\Omega)$.

En particular DF_n converge a DF en $L^2(\Omega; H)$ y por el argumento anterior tenemos que $D_t F_n = 0$ para toda $(t,\omega) \in [c,d] \times \Omega$ implicando que $D_t F = 0$ para toda $(t,\omega) \in [c,d] \times \Omega$. □

Denotemos por $L_a^2([0,T] \times \Omega)$ al conjunto de procesos adaptados y cuadrado integrables.

Veamos que en efecto la integral de Skorokhod coincide con la de Itô para los elementos de $L_a^2([0,T] \times \Omega)$.

Teorema 2.13. *Si $u \in L_a^2([0,T] \times \Omega)$, entonces $u \in \mathit{Dom}\,\delta$ y $\delta(u)$ coincide con la integral estocástica de Itô, es decir,*

$$\delta(u) \;=\; \int_0^T u_s \, dW_s$$

Demostración. Sea $F_j \in L^2(\Omega, \mathscr{F}_{t_j}, \mathbb{P})$ donde $\mathscr{F}_{t_j} = \mathscr{F}_{[0,t_j]}$ y $0 \le t_1 < \cdots < t_{n+1} \le T$, entonces por el Lema 2.2 implica que $F_j \mathbb{1}_{(t_j,t_{j+1}]} \in \mathit{Dom}\,\delta$ y junto con el Lema 2.3 y la expresión (19) obtenemos

$$\delta(F_j \mathbb{1}_{(t_j,t_{j+1}]}) = F_j \delta(\mathbb{1}_{(t_j,t_{j+1}]}) - \int_0^T D_t F_j \mathbb{1}_{(t_j,t_{j+1}]}) dt = F_j W(\mathbb{1}_{(t_j,t_{j+1}]}) - 0 = F_j (W_{t_{j+1}} - W_{t_j})$$

Ahora tomemos el proceso elemental y adaptado

$$u_t \;=\; \sum_{j=1}^{n} F_j \mathbb{1}_{(t_j,t_{j+1}]}(t)$$

entonces por el argumento anterior y por el Lema 2.2 implica que $u \in \mathit{Dom}\,\delta$ y además

$$\delta(u) \;=\; \sum_{j=1}^{n} F_j (W_{t_{j+1}} - W_{t_j}) \;=\; \int_0^T u_s \, dW_s \tag{21}$$

En consecuencia, como todo proceso $u \in L_a^2([0,T] \times \Omega)$ puede ser aproximado por una sucesión $\{u_n\}_{n\in\mathbb{N}}$ de procesos elementales y adaptados se sigue que por la expresión (21) se tenga que $\delta(u_n)$ es la integral de Itô de u_n y converge en $L^2(\Omega)$ a la integral de Itô de u.

Como el operador δ es cerrado concluimos que $u \in Dom\,\delta$ y $\delta(u)$ es la integral de Itô de u. $\square$

Por otro lado, observemos que si $F, G \in \mathcal{S}$, $h \in H$ y si denotamos $W(h) = \int_0^T h(s)\,dW_s$, entonces la *fórmula de integración por partes* (Lema 2.1) nos dice que

$$\mathbb{E}\left[\langle DF, h\rangle_H G\right] \;=\; \mathbb{E}\left[FGW(h)\right] - \mathbb{E}\left[F\,\langle DG, h\rangle_H\right]. \tag{22}$$

Sea $u = \sum_{k=1}^n F_k h_k$ donde $F_k \in \mathcal{S}$ y $h_k \in H = L^2([0,T])$ para $k = 1, \ldots, n$, es decir, $u \in \mathcal{S}_H$ y denotemos por $W(h_k) = \int_0^T h_k(s)\,dW_s$, que son elementos de $\mathcal{S}$ para cada k. Entonces por el Teorema 2.12 implica que $u \in Dom\,\delta$ y por la expresión (22) conlleva a que

$$\begin{aligned}\mathbb{E}\left[\langle DF, u\rangle_H\right] = \sum_{k=1}^n \mathbb{E}\left[\langle DF, h_k\rangle_H F_k\right] &= \sum_{k=1}^n \mathbb{E}\left[FF_kW(h_k)\right] - \mathbb{E}\left[F\,\langle DF_k, h_k\rangle_H\right] \\ &= \mathbb{E}[(\underbrace{\sum_{k=1}^n (F_kW(h_k)) - \langle DF, h_k\rangle_H)}_{\delta(u)}F]\end{aligned}$$

y por lo tanto

$$\delta(u) = \sum_{k=1}^n (F_kW(h_k)) - \langle DF, h_k\rangle_H\,. \tag{23}$$

Ahora veamos, para este caso, la relación entre D y δ usando linealidad y la regla del producto

$$\begin{aligned}\langle D(\delta(u)), h\rangle_H &= \sum_{k=1}^n \left\{F_k\,\langle DW(h_k), h\rangle_H + \langle DF_k, h\rangle_H W(h_k) - \langle D\left(\langle DF_k, h\rangle_H\right), h_k\rangle_H\right\} \\ &= \underbrace{\sum_{k=1}^n F_k\,\langle h_k, h\rangle_H}_{\langle u,h\rangle_H} + \underbrace{\sum_{k=1}^n \langle DF_k, h\rangle_H W(h_k) - \langle D\left(\langle DF_k, h\rangle_H\right), h_k\rangle_H}_{\delta\left(\sum_{k=1}^n \langle DF_k,h\rangle_H h_k\right)}\end{aligned}$$

es decir δ y D se relacionan de la forma

$$\langle D(\delta(u)), h\rangle_H = \langle u, h\rangle_H + \delta\left(\sum_{k=1}^n \langle DF_k, h\rangle_H h_k\right). \tag{24}$$

Nuestro objetivo es establecer la relación en el caso que $u \in \mathbb{D}^{1,2}(H)$, junto con ciertas condiciones, lo que nos permitirá calcular la derivada de Malliavin de un proceso de difusión. En consecuencia necesitamos introducir la descomposición en caos de Wiener-Itô.

Tomemos la función $f : [0,T]^n \to \mathbb{R}$, entonces definimos la ***simetrización*** $\tilde{f}$ de f por

$$\tilde{f}(t_1, \ldots, t_n) = \frac{1}{n!}\sum_{\sigma} f(t_{\sigma(1)}, \ldots t_{\sigma(n)})$$

en donde la suma es tomada sobre todas las permutaciones σ de $(1, \ldots, n)$. Por ejemplo, si

$$f(t_1, t_2) = t_1^2 + t_2 \cos t_1$$

entonces

$$\tilde{f}(t_1,t_2) = \frac{1}{2}\left[t_1^2 + t_2\cos t_1 + t_2^2 + t_1\cos t_2\right].$$

Para $f \in L^2([0,T]^n)$ definimos

$$I_n(f) := n! \int_0^T \int_0^{t_n} \dots \int_0^{t_3} \left(\int_0^{t_2} f(t_1,\dots,t_n)\, dW(t_1) \right) dW(t_2) \cdots dW(t_{n-1}) dW(t_n)$$

y entonces contamos con la siguiente descomposición.

Teorema 2.14 (**Descomposición en caos de Wiener-Itô**[30]**).** *Cualquier variable aleatoria $F \in L^2(\Omega, \mathscr{F}, \mathbb{P})$ admite una descomposición de la forma*

$$F = \sum_{n=0}^{\infty} I_n(f_n),$$

donde $f_0 = \mathbb{E}[F]$, I_0 es el mapeo identidad sobre las constantes y las funciones $f_n \in L^2([0,T]^n)$ son simétricas y determinadas de manera única por F.

Ilustremos la descomposición con el siguiente ejemplo.

Ejemplo 2.5. *Sean $F = W(t)(W(T) - W(t))$ y $g(t_1,t_2) = \mathbb{1}_{\{t_1<t<t_2\}}(t_1,t_2)$, entonces notemos que*

$$\int_0^T \left(\int_0^{t_2} \mathbb{1}_{\{t_1<t<t_2\}}(t_1,t_2) dW(t_1) \right) dW(t_2) = \int_t^T W(t) dW(t_2) = W(t)(W(T) - W(t))$$

y de esta manera vemos que

$$F = I_2(f_2)$$

donde

$$f_2(t_1,t_2) = \tilde{g}(t_1,t_2) = \frac{1}{2}(\mathbb{1}_{\{t_1<t<t_2\}} + \mathbb{1}_{\{t_2<t<t_1\}}).$$

Veamos la relación entre la derivada de Malliavin y la descomposición en caos de Wiener-Itô.

Proposición 2.2. *Sea $F : \Omega \to \mathbb{R}$ una variable aleatoria cuadrado integrable con descomposición en caos de Wiener-Itô*

$$F = \sum_{n=0}^{\infty} I_n(f_n)$$

donde $f_n \in L^2([0,T]^n)$ son funciones simétricas. Entonces, F pertenece a $\mathbb{D}^{1,2}$ si y sólo si

$$\sum_{n=1}^{\infty} n n! \, \|f_n\|^2_{L^2([0,T]^n)} < \infty$$

[30] [16] Teo. 1.1.2. Pág. 13.

y en este caso

$$D_tF = \sum_{n=1}^{\infty} nI_{n-1}(f_n(\cdot, t)).$$

Demostración. Denotemos a la diagonal de $[0,T]^n$ por

$$D = \{(t_1, t_2, \ldots, t_n) \in [0,T]^n \quad | \quad \text{existe} \;\; i \neq j \text{ tal que } \;\; t_i = t_j\}.$$

Sea $\mathcal{E}_n$ el conjunto de funciones elementales en $[0,T]^n$ que se desvanecen en D de la forma

$$f(t_1, \ldots, t_n) = \sum_{i_1,\ldots,i_n=1}^{k} a_{i_1,\ldots,i_n} \mathbb{1}_{[s_{i_0}, s_{i_1}) \times \cdots \times [s_{i_{n-1}}, s_{i_n})}(t_1, \ldots, t_n)$$

donde $0 = s_0 < s_1 < s_2 < \cdots < s_k = T$, tal que

$$a_{i_1,\ldots,i_n} = 0 \quad \text{si} \;\; i_p = i_q \quad \text{para alguna} \;\; p \neq q.$$

Entonces el conjunto $\mathcal{E}_n$ es denso[31] en $L^2([0,T]^n)$ y vemos que

$$I_n(f) = \sum_{i_1,\ldots,i_n=1}^{k} a_{i_1,\ldots,i_n} \xi_{i_1} \cdots \xi_{i_n}$$

donde $\xi_{i_p} = W(s_{i_p}) - W(s_{i_{p-1}})$ para $p = 1, \ldots, n$.

Además se tiene la identidad[32]

$$\mathbb{E}(I_n(f)^2) = n! \left\| \tilde{f}_n \right\|^2_{L^2([0,T]^n)} \qquad \text{para toda} \;\; f \in L^2([0,T]^n). \tag{25}$$

Observemos que si $F = I_n(f_n)$ con $f_n \in \mathcal{E}_n$ funciones simétricas, entonces

$$D_tF = \sum_{j=1}^{n} \sum_{i_1,\ldots,i_n=1}^{k} a_{i_1,\ldots,i_n} \xi_{i_1} \cdots \mathbb{1}_{[s_{i_{j-1}}, s_{i_j})}(t) \cdots \xi_{i_n} = nI_{n-1}(f_n(\cdot, t))$$

y como $\mathcal{E}_n$ es denso en $L^2([0,T]^n)$ entonces podemos extender la identidad anterior para $I_n(f_n)$ en donde f_n son funciones simétricas en $L^2([0,T]^n)$ y se cumple que

$$D_t(I_n(f_n)) = nI_{n-1}(f_n(\cdot, t)) \tag{26}$$

Sea $F = \sum_{n=0}^{\infty} I_n(f_n)$ con $f_n \in L^2([0,T]^n)$ funciones simétricas y consideremos las sumas parciales $F_N = \sum_{n=0}^{N} I_n(f_n)$ para $N \geq 0$, entonces $F_N \in \mathbb{D}^{1,2}$ y F_N converge a F en $L^2(\Omega)$ conforme $N \to \infty$.

[31] [16] Pág. 10.
[32] [16] Pág. 9.

Sea $N > k$, entonces por la expresión (25) tenemos que

$$\begin{aligned}
\|D_tF_N - D_tF_k\|^2_{L^2(\Omega;H)} &= \left\| \sum_{n=k+1}^{N} nI_{n-1}(f_n(\cdot,t)) \right\|^2_{L^2(\Omega;H)} \\
&= \int_0^T \mathbb{E}\left[\left\{ \sum_{n=k+1}^{N} nI_{n-1}(f_n(\cdot,t)) \right\}^2 \right] dt \\
&= \int_0^T \sum_{n=k+1}^{N} n^2(n-1)!\, \|f_n(\cdot,t)\|^2_{L^2([0,T]^{n-1})}\, dt \\
&= \sum_{n=k+1}^{N} nn!\, \|f_n\|^2_{L^2([0,T]^n)}\, .
\end{aligned} \tag{27}$$

Por hipótesis $\sum_{n=1}^{\infty} nn!\, \|f_n\|^2_{L^2([0,T]^n)} < \infty$ y entonces por el cálculo anterior se tiene que DF_N converge a DF en $L^2(\Omega;H)$ conforme $N \to \infty$ y por lo tanto $F \in \mathbb{D}^{1,2}$.

Conversamente, sea $F \in \mathbb{D}^{1,2}$ con $F = \sum_{n=0}^{\infty} I_n(f_n)$ y $f_n \in L^2([0,T]^n)$ simétricas. Notemos que para cada $n \geq 1$, $I_n(f_n)$ es el límite en $L^2(\Omega)$ de una sucesión $\{I_n(f_n^k)\}_{k\in\mathbb{N}}$ donde $f_n^k \in \mathcal{E}_n$ y f_n^k converge a f_n en $L^2([0,T]^n)$ conforme $k \to \infty$.

En consecuencia, la sucesión $D(I_n(f_n^k))$ converge a $D(I_n(f_n))$ en $L^2(\Omega;H)$ conforme $k \to \infty$ y además $I_{n-1}(f_n^k(\cdot,\cdot))$ converge a $I_{n-1}(f_n(\cdot,\cdot))$ en $L^2(\Omega;H)$ conforme $k \to \infty$.

Se sigue entonces que D_tF_k converge a D_tF en $L^2(\Omega;H)$ conforme $k \to \infty$ y por la expresión (27) y por el lema de Fatou implica que

$$\begin{aligned}
\sum_{j=0}^{\infty} j\cdot j!\, \|f_j\|^2_{L^2([0,T]^j)} = \sum_{j=0}^{\infty} \lim_{k\to\infty} \left(j\cdot j! \left\| f_j^k \right\|^2_{L^2([0,T]^j)} \right) &\leq \liminf_{k\to\infty} \sum_{j=0}^{\infty} j\cdot j! \left\| f_j^k \right\|^2_{L^2([0,T]^j)} \\
&= \liminf_{k\to\infty} \|D_tF_k\|^2_{L^2(\Omega;H)} = \|D_tF\|^2_{L^2(\Omega;H)} < \infty
\end{aligned}$$

donde $f_j^k = 0$ para $j > k$ y por lo tanto $\sum_{n=1}^{\infty} nn!\, \|f_n\|^2_{L^2([0,T]^n)} < \infty$. □

Denotemos por $L^2([0,T] \times \Omega)$ al conjunto de procesos cuadrado integrables, entonces tomamos un elemento $u \in L^2([0,T] \times \Omega)$ con descomposición en caos de Wiener-Itô

$$u(t) = \sum_{n=0}^{\infty} I_n(f_n(\cdot,t)) \tag{28}$$

tal que para cada $n \geq 1$ las funciones $f_n \in L^2([0,T]^{n+1})$ son simétricas en sus primeras n entradas. Veamos la relación entre la integral de Skorokhod y la descomposición en caos de Wiener-Itô.

Proposición 2.3. *Sea $u \in L^2([0,T]\times\Omega)$ con descomposición en caos de Wiener-Itô como en la expresión* (28)*. Entonces $u \in \mathrm{Dom}\,\delta$ si y sólo si la serie*

$$\delta(u) = \sum_{n=0}^{\infty} I_{n+1}(\tilde{f}_n) \qquad \textit{converge en } \; L^2(\Omega) \tag{29}$$

en donde

$$\tilde{f}_n(t_1,\ldots,t_n,t) = \frac{1}{n+1}\left(f_n(t_1,\ldots,t_n,t) + \sum_{i=1}^{n} f_n(t_1,\ldots,t_{i-1},t,t_{i+1},\ldots,t_n,t_i)\right).$$

Demostración. Sea $H_n(x)$ el ***n*-ésimo polinomio de Hermite**, el cual está definido por

$$H_n(x) = \frac{(-1)^n}{n!} e^{\frac{x^2}{2}} \frac{\mathrm{d}^n}{\mathrm{d}x^n}(e^{-\frac{x^2}{2}}), \qquad n \geq 1,\ x \in \mathbb{R}, \qquad \text{con } H_0(x) = 1.$$

Para cada $n \geq 1$ denotamos por $\mathcal{H}_n$ al subespacio vectorial cerrado de $L^2(\Omega, \mathscr{F}, \mathbb{P})$ generado por la familia de variables aleatorias

$$\{H_n(W(h)),\ h \in H,\ \|h\|_H = 1\} \quad \text{donde } \ h \in H = L^2([0,T]) \text{ y } W(h) = \int_0^T h(s)\, dW_s.$$

Al espacio $\mathcal{H}_n$ le llamamos **caos de Wiener de orden *n*** y notemos que $\mathcal{H}_0$ corresponde al conjunto de constantes, $\mathcal{H}_1 = \{W(h),\ h \in H, \|h\|_H = 1\}$ y además $\mathcal{H}_n$ y $\mathcal{H}_m$ son ortogonales si $n \neq m$.

Notemos que el espacio $L^2(\Omega, \mathscr{F}, \mathbb{P})$ se puede descomponer como la suma ortogonal infinita[33]

$$L^2(\Omega, \mathscr{F}, \mathbb{P}) = \bigoplus_{n=0}^{\infty} \mathcal{H}_n.$$

Sean $n \geq 1$ y $g \in L^2([0,T]^n)$ una función simétrica y $u_t \in L^2([0,T] \times \Omega)$ con descomposición en caos de Wiener-Itô como en la expresión (28), entonces

$$\begin{aligned}
\mathbb{E}\left[\int_0^T u_t D_t(I_n(g))\, dt\right] &= \sum_{m=0}^{\infty} \int_0^T \mathbb{E}[I_m(f_m(\cdot,t)) n I_{n-1}(g(\cdot,t))] dt \\
&= n \int_0^T \mathbb{E}[I_{n-1}(f_{n-1}(\cdot,t)) I_{n-1}(g(\cdot,t))] dt \\
&= n(n-1)! \int_0^T \langle f_{n-1}(\cdot,t), g(\cdot,t)\rangle_{L^2([0,T]^{n-1})}\, dt \\
&= n!\, \langle f_{n-1}, g\rangle_{L^2([0,T]^n)} = n! \langle \tilde{f}_{n-1}, g\rangle_{L^2([0,T]^n)} = \mathbb{E}[I_n(\tilde{f}_{n-1}) I_n(g)].
\end{aligned} \tag{30}$$

Supongamos que $u \in Dom\, \delta$, entonces por la identidad (17) y la expresión (30)

$$\mathbb{E}[\delta(u) I_n(g)] = \mathbb{E}[\langle u, D(I_n(g))\rangle_H] = \mathbb{E}[I_n(\tilde{f}_{n-1}) I_n(g)]$$

para toda $I_n(g)$ con $n \geq 1$ y cualquier $g \in L^2([0,T]^n)$ simétrica.

La expresión anterior implica que $I_n(\tilde{f}_{n-1})$ es la proyección de $\delta(u)$ en $\mathcal{H}_n$ y por lo tanto $\sum_{n=0}^{\infty} I_{n+1}(\tilde{f}_n)$ converge en $L^2(\Omega)$ y su suma es igual a $\delta(u)$.

Conversamente, supongamos que $\sum_{n=0}^{\infty} I_{n+1}(\tilde{f}_n)$ converge en $L^2(\Omega)$ y denotemos por V a la suma. Tomemos las sumas parciales $F_N = \sum_{n=0}^{N} I_n(g_n)$ con $g_n \in L^2([0,T]^n)$ funciones simétricas y $N \geq 1$.

[33] [16] Teo. 1.1.1. Pág. 6.

Entonces por la expresión (30) obtenemos que para toda $N \geq 1$

$$\mathbb{E}\left[\int_0^T u_t D_t F_N \, dt\right] = \sum_{n=1}^{N} \mathbb{E}[I_n(\tilde{f}_{n-1}) I_n(g)]$$

y en particular

$$\left|\mathbb{E}\left[\int_0^T u_t D_t F_N \, dt\right]\right| \leq \|V\|_{L^2(\Omega)} \|F_N\|_{L^2(\Omega)} .$$

Sea $F \in \mathbb{D}^{1,2}$ con $F = \sum_{n=0}^{\infty} I_n(g_n)$ donde $g_n \in L^2([0,T]^n)$ son funciones simétricas. Entonces F_N converge a F en $L^2(\Omega)$ conforme $N \to \infty$ y también DF_N converge a DF en $L^2(\Omega; H)$ conforme $N \to \infty$ y se sigue que

$$\left|\mathbb{E}\left[\int_0^T u_t D_t F \, dt\right]\right| \leq \|V\|_{L^2(\Omega)} \|F\|_{L^2(\Omega)} .$$

lo que demuestra que $u \in Dom\, \delta$.

Por el teorema de convergencia dominada y la continuidad del producto interno conlleva a que

$$\mathbb{E}[\langle DF, u\rangle_H] = \lim_{N\to\infty} \mathbb{E}\left[\int_0^T u_t D_t F_N \, dt\right] = \lim_{N\to\infty} \mathbb{E}\left[\sum_{n=0}^{N} I_{n+1}(\tilde{f}_n) I_n(g)\right] = \mathbb{E}[\delta(u)F]$$

y con esto concluimos que $\delta(u) = \sum_{n=0}^{\infty} I_{n+1}(\tilde{f}_n)$. □

Notemos que por el teorema anterior y junto con el hecho de que[34]

$$\mathbb{E}[I_m(f) I_p(g)] = \begin{cases} 0 & \text{si } m \neq p, \\ m! \langle \tilde{f}, \tilde{g} \rangle_{L^2([0,T]^m)} & \text{si } m = p. \end{cases}$$

implica que el dominio de δ coincide con el subespacio de $L^2([0,T] \times \Omega)$ de aquellos procesos u que satisfacen

$$\mathbb{E}(\delta(u)^2) = \sum_{n=0}^{\infty} (n+1)! \left\|\tilde{f}_n\right\|^2_{L^2([0,T]^{n+1})} < \infty.$$

Notemos que el espacio $\mathbb{D}^{1,2}(H)$, donde $H = L^2([0,T])$, coincide con la clase de procesos u en $L^2([0,T] \times \Omega)$ tales que $u_t \in \mathbb{D}^{1,2}$ para casi toda $t \in [0,T]$ y que existe una modificación medible del proceso $D_s u_t$ tal que $\mathbb{E}(\int_0^T \int_0^T (D_s u_t)^2 ds dt) < \infty$.

Veamos un resultado que permite calcular la derivada de la integral de Skorokhod para el caso en que $u \in \mathbb{D}^{1,2}(H)$ y mediante el cual podremos deducir resultados interesantes.

Teorema 2.15. *Sea $u \in \mathbb{D}^{1,2}(H)$ tal que para casi toda $s \in [0,T]$ el proceso $\{D_s u_t,\, t \in [0,T]\}$ pertenece a Dom δ y $\mathbb{E}(\int_0^T |\delta(D_s u)|^2 ds) < \infty$. Entonces $\delta(u) \in \mathbb{D}^{1,2}$ y se tiene que*

$$D_s(\delta(u)) = u_s + \delta(D_s(u))$$

[34] [16] Pág. 9.

Demostración. Sea $n \geq 1$ y como $u \in \mathbb{D}^{1,2}(H)$, entonces por la descomposición en caos de Wiener-Itô tenemos que $u_s = \sum_{n=0}^{\infty} I_n(f_n(\cdot, s))$ en donde $f_n \in L^2([0,T]^{n+1})$ es una función simétrica en sus primeras n entradas.

En consecuencia, por las Proposiciones 2.2 y 2.3 obtenemos que

$$
\begin{aligned}
D_s(\delta(u)) = D_s\left(\sum_{n=0}^{\infty} I_{n+1}(\tilde{f}_n)\right) &= \sum_{n=0}^{\infty}(n+1)I_n(\tilde{f}_n(\cdot, s)) \\
&= u_s + \sum_{n=0}^{\infty} I_n\left(\sum_{i=1}^{n} f_n(s_1, \ldots, s_{i-1}, s, s_{i+1}, \ldots, s_n, s_i)\right) \\
&= u_s + \sum_{n=0}^{\infty} nI_n(\phi_n(\cdot, s, \cdot)),
\end{aligned}
$$

donde $\phi_n(\cdot, s, \cdot)$ es la simetrización de la función

$$(s_1, \ldots, s_n) \longmapsto f_n(s_1, \ldots, s_{n-1}, s, s_n).$$

Por otro lado, nuevamente por las Proposiciones 2.2 y 2.3 obtenemos que

$$\delta(D_s u) = \delta\left(\sum_{n=1}^{\infty} nI_{n-1}(f_n(\cdot, s))\right) = \sum_{n=0}^{\infty} nI_n(\phi_n(\cdot, s, \cdot))$$

lo que prueba que $\delta(u) \in \mathbb{D}^{1,2}$ así como la identidad deseada. □

Una primera consecuencia de este resultado es que para la expresión (18) en el caso $p = 2$ resulta ser una isometría que satisface la integral de Skorokhod.

Proposición 2.4 (Isometría de la integral de Skorokhod). *Sea u un proceso medible tal que $u_s \in \mathbb{D}^{1,2}$ para casi toda $s \in [0,T]$ y que cumple*

$$\mathbb{E}\left[\int_0^T u_t^2\, dt + \int_0^T \int_0^T |D_t u_s D_s u_t|\, dsdt\right] < \infty.$$

Entonces $u \in Dom\,\delta$ y

$$\mathbb{E}\left[\delta(u)^2\right] \;=\; \mathbb{E}\left[\int_0^T u_t^2 dt\right] \;+\; \mathbb{E}\left[\int_0^T \int_0^T D_t u_s D_s u_t\, dsdt\right].$$

Demostración. Si $u \in \mathbb{D}^{1,2}$, entonces por el Teorema 2.12 se sigue $u \in Dom\,\delta$.

Por el Teorema 2.15 y la identidad (17) tenemos que

$$
\begin{aligned}
\mathbb{E}\left[\delta(u)^2\right] &= \mathbb{E}\left[\delta(u)\delta(u)\right] = \mathbb{E}\left[\langle D(\delta(u)), u\rangle_H\right] \\
&= \mathbb{E}\left[\langle u + \delta(Du), u\rangle_H\right] \\
&= \mathbb{E}\left[\langle u, u\rangle_H\right] + \mathbb{E}\left[\langle \delta(Du), u\rangle_H\right].
\end{aligned}
$$

Si usamos nuevamente la identidad (17) y junto con las hipótesis vemos que

$$
\begin{aligned}
\mathbb{E}\left[\langle \delta(Du), u\rangle_H\right] &= \mathbb{E}\left[\int_0^T \delta(D_t u) u_t \, dt\right] = \int_0^T \mathbb{E}\left[\delta(D_t u) u_t\right] \, dt \\
&= \int_0^T \mathbb{E}\left[\langle D_t u, Du_t\rangle_H\right] \, dt = \int_0^T \mathbb{E}\left[\int_0^T D_t u_s D_s u_t \, ds\right] \, dt \\
&= \mathbb{E}\left[\int_0^T \int_0^T D_t u_s D_s u_t \, dsdt\right]
\end{aligned}
$$

y por los cálculos anteriores concluimos que

$$
\begin{aligned}
\mathbb{E}\left[\delta(u)^2\right] &= \mathbb{E}\left[\langle u, u\rangle_H\right] + \mathbb{E}\left[\langle \delta(Du), u\rangle_H\right] \\
&= \ \mathbb{E}\left[\int_0^T u_t^2 dt\right] \ + \ \mathbb{E}\left[\int_0^T \int_0^T D_t u_s D_s u_t \, dsdt\right].
\end{aligned}
$$

□

Notemos que si $\mathscr{F}_t = \mathscr{F}_{[0,t]}$ la σ-álgebra generada por $\{W_r : r \in [0,t]\}$ y si u es un proceso adaptado, entonces por el Lema 2.3 implica que

$$D_s u_t = 0, \qquad s > t$$

y por lo tanto

$$D_t u_s D_s u_t = 0 \qquad \text{para casi toda } \ (t,s) \in [0,T] \times [0,T]$$

de manera que la isometría anterior se reduce a

$$\mathbb{E}\left[\delta(u)^2\right] \ = \ \mathbb{E}\left[\int_0^T u_t^2 dt\right]$$

que de hecho es *la isometría de Itô* pues al ser adaptado el proceso habíamos probado que la integral de Skorokhod coincide con la de Itô.

Finalmente podemos calcular la derivada de Malliavin de un proceso de difusión.

Ejemplo 2.6. *Sean $\sigma : \mathbb{R} \to \mathbb{R}$ y $b : \mathbb{R} \to \mathbb{R}$ funciones acotadas y continuamente diferenciables tal que sus derivadas son acotadas también. Tomemos el proceso $X = \{X_t, \ t \in [0,T]\}$ que satisface la ecuación diferencial estocástica*

$$X_t = x_0 + \int_0^t \sigma(X_s) \, dW_s + \int_0^t b(X_s) \, ds.$$

Entonces, bajo estas condiciones tenemos que el proceso $\int_0^t \sigma(X_s) \, dW_s$ pertenece a $\mathbb{D}^{1,2}$ y por el Teorema 2.15 obtenemos que para $r \leq t$

$$D_r \left(\int_0^t \sigma(X_s) \, dW_s\right) = \sigma(X_r) + \int_r^t D_r(\sigma(X_s)) \, dW_s,$$

también bajo estas condiciones tenemos que $\int_0^t b(X_s)\,ds \in \mathbb{D}^{1,2}$ *y para* $r \leq t$ *se sigue que*

$$D_r\left(\int_0^t b(X_s)\,ds\right) = \int_r^t D_r(b(X_s))\,ds$$

y utilizando linealidad, la regla de la cadena y los cálculos anteriores deducimos que

$$D_rX_t = \int_r^t b'(X_s)D_rX_s\,ds + \int_r^t \sigma'(X_s)D_rX_s\,dW_s + \sigma(X_r)$$

Si denotamos por $Z_t = D_rX_t$ *para* $t \geq r$*, entonces obtenemos la ecuación diferencial estocástica*

$$\begin{cases} dZ_t = b'(X_t)Z_tdt + \sigma'(X_t)Z_tdW_t, & r \leq t \\ Z_r = \sigma(X_r). \end{cases}$$

Y mediante la fórmula de Itô y la exponencial estocástica deducimos la solución a la ecuación anterior la cual está dada para $r \leq t$ *por*

$$D_rX_t = \sigma(X_r)\exp\left(\int_r^t \sigma'(X_s)\,dW_s + \int_r^t \left[b'(X_s) - \frac{1}{2}(\sigma'(X_s))^2\right]ds\right). \tag{31}$$

Del ejemplo anterior podemos ver que si adicionalmente suponemos que $\sigma(x) \neq 0$ para toda $x \in \mathbb{R}$, entonces

$$D_rX_t = \frac{Y_t}{Y_r}\sigma(X_r)\mathbb{1}_{\{r\leq t\}}$$

donde Y_t es la solución a

$$dY_t = b'(X_t)Y_tdt + \sigma'(X_t)Y_tdW_t, \qquad Y_0 = 1.$$

A este proceso se le conoce como **la primera variación** y es de suma importancia en el cálculo de Malliavin.

Cabe resaltar que en los cálculos anteriores se les puede incorporar el parámetro temporal a los coeficientes b y σ, sin perder generalidad alguna, y obtenemos la misma solución simplemente agregando el parámetro temporal respectivamente.

Además podemos extender el análisis anterior al caso n-dimensional con las mismas justificaciones entrada por entrada.

Hemos llegado al momento cúspide de nuestra exposición. En donde ya contamos con las herramientas necesarias para poder aplicar las técnicas del cálculo de Malliavin para estimar la densidad de un proceso de difusión.

Cabe mencionar que para dicha estimación también necesitaremos desarrollar unas desigualdades para martingalas. Sin más preámbulo demos inicio con los capítulos de estimación.

3. Estimación de Densidades

3.1. Estimación explícita

Los resultados expuestos anteriormente nos permitirán estimar las densidades de variables aleatorias diferenciables así como de los procesos de difusión.

Comencemos la exposición con una estimación explícita para variables aleatorias que pertenecen a $\mathbb{D}^{1,1}$ bajo ciertas condiciones adicionales.

Para cualquier variable aleatoria $F \in \mathbb{D}^{1,1}$ y cualquier proceso $u \in L^1(\Omega; L^2([0,T]))$ denotamos

$$D_u F = \int_0^T D_t F u_t dt = \langle DF, u \rangle_H$$

Existe un resultado que afirma que si $F \in \mathbb{D}^{1,1}$ es tal que $\|DF\|_H > 0$ casi seguramente, entonces F posee una distribución absolutamente continua[35].

Si adicionamos supuestos al resultado anterior obtenemos una estimación explícita de la densidad de F como veremos a continuación.

Proposición 3.1. *Sea F una variable aleatoria en el espacio $\mathbb{D}^{1,1}$. Sea u un proceso en $L^1(\Omega; L^2([0,T]))$ tal que $D_u F \neq 0$ casi seguramente y que además $u/D_u F$ pertenece al dominio del operador δ. Entonces la ley de F tiene una densidad continua y acotada, denotada por $p(x)$, y dada por*

$$p(x) = \mathbb{E}\left(\mathbb{1}_{\{F>x\}}\, \delta\left(\frac{u}{D_u F}\right)\right) \tag{32}$$

Demostración. Sea ψ una función suave, no negativa y con soporte compacto. Sea $\varphi(y) = \int_{-\infty}^{y} \psi(z)dz$ con $y \in \mathbb{R}$ y bajo estas condiciones sabemos que $\varphi(F) \in \mathbb{D}^{1,1}$.

Usando la regla de la cadena y el teorema fundamental del cálculo vemos que

$$\langle D(\varphi(F)), u \rangle_H = \langle D(\int_{-\infty}^{F} \psi(z)dz), u \rangle_H = \langle \psi(F) DF, u \rangle_H = \psi(F) \langle DF, u \rangle_H = \psi(F) D_u F$$

y entonces

$$\psi(F) = \frac{1}{D_u F} \langle D(\varphi(F)), u \rangle_H = \left\langle D(\varphi(F)), \frac{u}{D_u F} \right\rangle_H$$

y por lo tanto

$$\mathbb{E}(\psi(F)) = \mathbb{E}\left(\left\langle D(\varphi(F)), \frac{u}{D_u F} \right\rangle_H\right).$$

Sea $\{F_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias suaves en $\mathcal{S}$ que converge a F en $\mathbb{D}^{1,1}$. Entonces usando la identidad (17), convergencia dominada y la continuidad del producto interno vemos que

$$\begin{aligned}\mathbb{E}\left(\left\langle D(\varphi(F)), \frac{u}{D_u F} \right\rangle_H\right) &= \lim_{n\to\infty} \mathbb{E}\left(\left\langle D(\varphi(F_n)), \frac{u}{D_u F} \right\rangle_H\right) \\ &= \lim_{n\to\infty} \mathbb{E}\left(\varphi(F_n)\, \delta\left(\frac{u}{D_u F}\right)\right) = \mathbb{E}\left(\varphi(F)\, \delta\left(\frac{u}{D_u F}\right)\right).\end{aligned}$$

[35] [16] Teo. 2.1.3. Pág. 98.

Y por lo tanto

$$\mathbb{E}(\psi(F)) = \mathbb{E}\left(\varphi(F)\,\delta\left(\frac{u}{D_u F}\right)\right) \qquad \text{para toda } F \in \mathbb{D}^{1,1}.$$

Dado que el espacio de funciones infinitamente diferenciables con soporte compacto, $C_0^\infty(\mathbb{R})$, es densa en $L^p(\mathbb{R})$ para $1 \leq p < \infty$ y dadas las hipótesis de ψ, entonces por el teorema de convergencia monótona se tiene que la identidad anterior se satisface para $\psi(y) = \mathbb{1}_{[a,b]}(y)$ con $y \in \mathbb{R}$ y $a < b$.

En consecuencia, si aplicamos el teorema de Fubini a lo anterior tenemos que

$$\begin{aligned}
\mathbb{P}(a \leq F \leq b) &= \mathbb{E}\left(\left(\int_{-\infty}^{F} \mathbb{1}_{[a,b]}(x)\,dx\right)\,\delta\left(\frac{u}{D_u F}\right)\right) \\
&= \int_a^b \mathbb{E}\left(\mathbb{1}_{\{F>x\}}\,\delta\left(\frac{u}{D_u F}\right)\right) dx
\end{aligned}$$

demostrando que $p(x)$ como en la expresión (32) es la densidad de F. □

Si a la proposición anterior tomamos $p, q \in [1, \infty]$ tal que $1/p + 1/q = 1$, entonces la desigualdad de Hölder implica que

$$\begin{aligned}
p(x) &\leq \left(\mathbb{E}[(\mathbb{1}_{\{F>x\}})^p]\right)^{1/p}\left(\mathbb{E}\left[\left|\delta\left(\frac{u}{D_u F}\right)\right|^q\right]\right)^{1/q} \\
&= (\mathbb{P}[F > x])^{1/p}\left(\mathbb{E}\left[\left|\delta\left(\frac{u}{D_u F}\right)\right|^q\right]\right)^{1/q}
\end{aligned}$$

y por lo tanto obtenemos la estimación

$$p(x) \leq \mathbb{E}\left(\left|\delta\left(\frac{u}{D_u F}\right)\right|\right). \tag{33}$$

Si aplicamos esta desigualdad nuevamente vemos que

$$\begin{aligned}
\int_{\mathbb{R}} p(x)^2 dx &\leq \int_{\mathbb{R}} p(x)\mathbb{E}\left(\left|\delta\left(\frac{u}{D_u F}\right)\right|\right) dx \\
&= \mathbb{E}\left(\left|\delta\left(\frac{u}{D_u F}\right)\right|\right)\int_{\mathbb{R}} p(x)dx \\
&= \mathbb{E}\left(\left|\delta\left(\frac{u}{D_u F}\right)\right|\right)\cdot 1
\end{aligned}$$

y de esta manera obtenemos la estimación

$$\int_{\mathbb{R}} p(x)^2 dx \leq \mathbb{E}\left(\left|\delta\left(\frac{u}{D_u F}\right)\right|\right). \tag{34}$$

Como consecuencia de las estimaciones anteriores junto con el Lema 2.2 deducimos la siguiente estimación.

Lema 3.1. *Sea F una variable aleatoria en el espacio $\mathbb{D}^{1,1}$. Sea u un proceso en $\operatorname{Dom}\delta$ tal que para p, q con $1/p + 1/q = 1$ satisface las siguientes propiedades:*

(i) u pertenece a $L^q(\Omega; L^2([0,T]))$.

(ii) $\delta(u)$ pertenece a $L^q(\Omega)$.

(iii) $(D_u F)^{-1}$ pertenece a $\mathbb{D}^{1,p}$.

Entonces F tiene una función de densidad, denotada por $p(x)$, tal que

$$p(x) \leq \mathbb{E}\left(\left|\frac{\delta(u)}{D_u F}\right|\right) + \mathbb{E}\left(\left|D_u\left(\frac{1}{D_u F}\right)\right|\right)$$

Demostración. Por la identidad (19) del Lema 2.2 vemos que

$$\begin{aligned}
\delta\left(\frac{u}{D_u F}\right) &= \delta\left(\frac{1}{D_u F}\cdot u\right) \\
&= \frac{1}{D_u F}\delta(u) - \left\langle D\left(\frac{1}{D_u F}\right), u\right\rangle_H \\
&= \frac{\delta(u)}{D_u F} - D_u\left(\frac{1}{D_u F}\right)
\end{aligned}$$

y esto implica que

$$\mathbb{E}\left(\left|\delta\left(\frac{u}{D_u F}\right)\right|\right) \leq \mathbb{E}\left(\left|\frac{\delta(u)}{D_u F}\right|\right) + \mathbb{E}\left(\left|D_u\left(\frac{1}{D_u F}\right)\right|\right)$$

y como consecuencia de la estimación (33) concluimos entonces que

$$\begin{aligned}
p(x) &\leq \mathbb{E}\left(\left|\delta\left(\frac{u}{D_u F}\right)\right|\right) \\
&\leq \mathbb{E}\left(\left|\frac{\delta(u)}{D_u F}\right|\right) + \mathbb{E}\left(\left|D_u\left(\frac{1}{D_u F}\right)\right|\right)
\end{aligned}$$

□

Podemos modificar los supuestos del Lema 3.1 y con el mismo razonamiento obtenemos lo siguiente[36].

Lema 3.2. *Sea F una variable aleatoria en el espacio $\mathbb{D}^{2,1}$. Sea u un proceso en $\operatorname{Dom}\delta$ tal que para p, q con $1/p + 1/q = 1$ satisface las siguientes propiedades:*

(i) u pertenece a $L^q(\Omega; L^2([0,T]))$.

(ii) $\delta(u)$ pertenece a $L^q(\Omega)$.

(iii) $(D_u F)^{-1}$ pertenece a $L^p(\Omega)$ y

$$\Phi_u := (D_u F)^{-2}\left(\left\|D^2 F\right\|_{H\otimes H}\|u\|_H + \|Du\|_{H\otimes H}\|DF\|_H\right) \in L^p(\Omega) \tag{35}$$

Entonces F tiene una función de densidad, denotada por $p(x)$, tal que

$$p(x) \leq \mathbb{E}\left(\left|\frac{\delta(u)}{D_u F}\right| + \Phi_u \|u\|_H\right) \tag{36}$$

[36] [3] Lem. 4. Pág. 836.

Observación 1[37]. Bajo los supuestos del Lema 3.2 podemos escribir la densidad de F como

$$p(x) = \mathbb{E}\left(\mathbb{1}_{\{F>x\}}\frac{\delta(u)}{D_uF} + (D_uF)^{-2}\left(\left\langle D^2F, u\otimes u\right\rangle_{H\otimes H} + \int_{[0,T]^2} u_tD_tu_sD_sF\,dsdt\right)\right)$$

Observación 2[38]. Cuando $u = DF$ la estimación (36) conlleva a

$$p(x) \le \mathbb{E}\left(\left|\frac{\delta(DF)}{\|DF\|_H^2}\right|\right) + 2\mathbb{E}\left(\frac{1}{\|DF\|_H^2}\left\|D^2F\right\|_{H\otimes H}\right) \tag{37}$$

Observación 3[39]. Sea $W = \{W_t,\, t\in[0,T]\}$ la versión continua del movimiento browniano y $u = \{u_t,\, t\in[0,T]\}$ un proceso adaptado que satisface las siguientes hipótesis:

(i) $\mathbb{E}\left(\int_0^T u_t^2\,dt\right) < \infty$, u_t pertenece al espacio $\mathbb{D}^{2,2}$ para cada $t\in[0,T]$ y

$$\lambda := \sup_{s,t\in[0,T]}\mathbb{E}(|D_su_t|^p) + \sup_{r,s\in[0,T]}\mathbb{E}\left(\left(\int_0^T |D_{r,s}^2u_t|^2dt\right)^{\frac{p}{2}}\right) < \infty$$

para alguna $p > 3$.

(ii) $0 < \rho \le |u_t|$ para toda $t\in[0,T]$ y con ρ una constante.

Sea $M_t = \int_0^t u_s\,dW_s$, y denotemos por $p_t(x)$ a la densidad de M_t. Entonces para cualquier $t > 0$

$$p_t(x) \le \frac{c}{\sqrt{t}}\mathbb{P}(|M_t| > |x|)^{\frac{1}{q}},$$

donde $q > \frac{p}{p-3}$ y la constante c depende únicamente de λ, ρ y p.

En particular, si tomamos $u = DF$ y junto con la estimación (37) obtenemos que

$$p(x) \le c\,\mathbb{P}(|F| > |x|)^{\frac{1}{q}}$$

Observación 4. Notemos que si u es un proceso adaptado y que si $D_tF = u_te^{N_t}$ (lo cual es cierto si F es el valor al tiempo t de un proceso de difusión como podemos ver en la identidad (31)), entonces

$$\left|\frac{\delta(u)}{D_uF}\right| \le \sup_{0\le t\le T} e^{-N_t}\frac{\left|\int_0^T u_t\,dW_t\right|}{\int_0^T u_t^2\,dt}$$

Observemos que hemos usado el hecho que por ser adaptado el proceso u, entonces

$$\delta(u) = \int_0^T u_t\,dW_t$$

y que además para la martingala browniana definida por $M_t = \int_0^T u_t\,dW_t$ tenemos que su variación

[37] Véase [3] en la bibliografía.
[38] Véase [3] en la bibliografía.
[39] [16] Prop. 2.1.3. Pág. 88.

cuadrática está dada por $\langle M\rangle_t = \int_0^T u_t^2\, dt$.

De manera que para obtener una estimación de la esperanza de este tipo de expresiones necesitamos estimar la norma p de $M_t/\langle M\rangle_t$ donde M_t es una martingala browniana lo que nos da pie al siguiente capítulo.

3.2. Desigualdades para Martingalas

En el siguiente teorema establecemos una estimación en la norma p de una martingala dividida por su variación cuadrática, en donde $1 \leq p < \infty$. Esta norma es acotada por una constante universal multiplicada por la norma q del inverso de la raíz cuadrada de la variación cuadrática.

Se le dice universal a la constante en el sentido de que puede ser tomada por igual para toda martingala local y en cualquier espacio de probabilidad.

Es importante mencionar que esta desigualdad es de suma importancia en el presente trabajo así como en otros trabajos relacionados como lo podemos apreciar en el artículo del Dr. Victor H. de la Peña[40] y que más adelante retomaremos en aras de enriquecer nuestra exposición del presente capítulo.

Teorema 3.1. *Sea $\{M_t,\, t \geq 0\}$ una martingala continua y nula en 0. Entonces para cada $1 \leq p < q = p + \varepsilon$, existe una constante universal $C := C(p,q)$ tal que*

$$\left\| \frac{M_t}{\langle M\rangle_t} \right\|_p \leq C \left\| \langle M\rangle_t^{-1/2} \right\|_q \tag{38}$$

Demostración. Como consecuencia de la fórmula de integración por partes tenemos que

$$\mathbb{E}\left[\left| \frac{M_t}{\langle M\rangle_t} \right|^p \right] = \int_0^\infty p x^{p-1} \mathbb{P}[\, |M_t| > x\langle M\rangle_t \,]\, dx.$$

Notemos que por subaditividad numerable

$$\mathbb{P}[\, |M_t| > x\langle M\rangle_t \,] \leq \mathbb{P}[\, M_t > x\langle M\rangle_t \,] + \mathbb{P}[\, -M_t > x\langle M\rangle_t \,]$$

De manera que es suficiente con estimar el término

$$\int_0^\infty x^{p-1} \mathbb{P}[\, M_t > x\langle M\rangle_t \,]\, dx.$$

Recordemos que si M es una martingala continua, entonces M es una martingala localmente continua y si definimos el tiempo de paro $\tau_n := \inf\{t \geq 0 \,:\, |M_t| \geq n\}$, entonces $M_{\tau_n \wedge t}$ es una martingala continua y acotada.

[40] Véase [7] en la bibliografía.

En consecuencia, podemos asumir sin pérdida de generalidad que M es acotada y junto con las desigualdades de Markov y de Hölder vemos que

$$\begin{aligned}
\mathbb{P}[\, M_t > x\langle M\rangle_t \,] &= \mathbb{P}[\, e^{\lambda M_t - \theta\langle M\rangle_t} > e^{(\lambda x-\theta)\langle M\rangle_t} \,] \\
&= \mathbb{P}[\, e^{\lambda M_t - \theta\langle M\rangle_t} e^{-(\lambda x-\theta)\langle M\rangle_t} > 1 \,] \\
&\leq 1 \cdot \mathbb{E}[\, e^{\lambda M_t - \theta\langle M\rangle_t} e^{-(\lambda x-\theta)\langle M\rangle_t} \,] \\
&\leq \left(\mathbb{E}[\, e^{\lambda\alpha M_t - \theta\alpha\langle M\rangle_t} \,]\right)^{1/\alpha} \left(\mathbb{E}[\, e^{-\beta(\lambda x-\theta)\langle M\rangle_t} \,]\right)^{1/\beta}
\end{aligned}$$

en donde $1/\beta + 1/\alpha = 1$. Por hipótesis M es una martingala continua, acotada y nula en cero y si elegimos λ, α, β, θ tal que $(\lambda\alpha)^2/2 = \theta\alpha$, es decir, $\theta = \lambda^2\alpha/2$, entonces

$$e^{(\lambda\alpha)M_t - (\theta\alpha)\langle M\rangle_t} = e^{(\lambda\alpha)M_t - \frac{(\lambda\alpha)^2}{2}\langle M\rangle_t} =: \mathcal{E}(M)_t$$

de manera que $\mathcal{E}(M)_t$ es una martingala local continua[41] y junto con la hipótesis de que M es nula en cero vemos que

$$\mathbb{E}[\, e^{(\lambda\alpha)M_t - \frac{(\lambda\alpha)^2}{2}\langle M\rangle_t} \,] = \mathbb{E}[\, e^{(\lambda\alpha)M_0 - \frac{(\lambda\alpha)^2}{2}\langle M\rangle_0} \,] = 1$$

En consecuencia con esta elección y junto con la desigualdad anterior obtenemos

$$\mathbb{P}[\, M_t > x\langle M\rangle_t \,] \leq \left(\mathbb{E}[\, e^{-\beta(\lambda x - \frac{\lambda^2\alpha}{2})\langle M\rangle_t} \,]\right)^{1/\beta}$$

Si optimizamos la función $e^{-\beta(\lambda x - \frac{\lambda^2\alpha}{2})\langle M\rangle_t}$ con respecto a λ obtenemos el óptimo en $\lambda = x/\alpha$ y notemos que

$$\frac{1}{\beta} + \frac{1}{\alpha} = 1 \iff 1 + \frac{\beta}{\alpha} = \beta \iff \frac{\beta}{\alpha} = \beta - 1$$

y junto con la desigualdad anterior se sigue que

$$\begin{aligned}
\mathbb{P}[\, M_t > x\langle M\rangle_t \,] &\leq \left(\mathbb{E}[\, e^{-\beta(\lambda x - \frac{\lambda^2\alpha}{2})\langle M\rangle_t} \,]\right)^{1/\beta} \\
&\leq \left(\mathbb{E}[\, e^{-\frac{\beta}{\alpha}\frac{x^2}{2}\langle M\rangle_t} \,]\right)^{1/\beta} \leq \left(\mathbb{E}[\, e^{-(\beta-1)\left(\frac{x^2}{2}\right)\langle M\rangle_t} \,]\right)^{1/\beta}
\end{aligned} \tag{39}$$

En consecuencia tenemos que para cada $\varepsilon > 0$ y $\delta > 1$

$$\begin{aligned}
\mathbb{E}\left[\left|\frac{M_t}{\langle M\rangle_t}\right|^p\right] &\leq 2\int_0^\infty px^{p-1}\left(\mathbb{E}[\, e^{-(\beta-1)\left(\frac{x^2}{2}\right)\langle M\rangle_t} \,]\right)^{1/\beta} dx \\
&\leq 2\varepsilon^p + 2\int_\varepsilon^\infty px^{p-1}\left(\mathbb{E}[\, e^{-(\beta-1)\left(\frac{x^2}{2}\right)\langle M\rangle_t} \,]\right)^{1/\beta} dx \\
&= 2\varepsilon^p + 2p\int_\varepsilon^\infty x^{-\delta}\left(\mathbb{E}[\, x^{(\delta+p-1)\beta}\; e^{-(\beta-1)\left(\frac{x^2}{2}\right)\langle M\rangle_t} \,]\right)^{1/\beta} dx \\
&\leq 2\varepsilon^p + 2p\left(\int_\varepsilon^\infty x^{-\delta}\, dx\right) \times \left(\mathbb{E}\left[\sup_{x\in\mathbb{R}}\left[x^{(\delta+p-1)\beta}\; e^{-(\beta-1)\left(\frac{x^2}{2}\right)\langle M\rangle_t}\right]\right]\right)^{1/\beta}
\end{aligned}$$

Sea

$$\phi(x) = x^{(\delta+p-1)\beta}\; e^{-(\beta-1)\left(\frac{x^2}{2}\right)\langle M\rangle_t}$$

[41] [23] Teo. 2.16. Pág. 332.

Entonces

$$\phi'(x) = \left[\beta(\delta+p-1)x^{(\delta+p-1)\beta-1} - x^{\beta(\delta+p-1)+1}(\beta-1)\langle M\rangle_t\right] \times e^{-(\beta-1)\left(\frac{x^2}{2}\right)\langle M\rangle_t}$$

La función ϕ alcanza su máximo en

$$x_0 = \sqrt{\frac{\beta(\delta+p-1)}{\beta-1}\langle M\rangle_t^{-1/2}}$$

En consecuencia

$$\phi(x_0) = \left(\frac{\beta(\delta+p-1)}{\beta-1}\right)^{\left(\frac{\delta+p-1}{2}\right)\beta} \langle M\rangle_t^{-\frac{(\delta+p-1)\beta}{2}} e^{-\left(\frac{\delta+p-1}{2}\right)\beta}$$

Y por lo tanto

$$\begin{aligned}\mathbb{E}\left[\left|\frac{M_t}{\langle M\rangle_t}\right|^p\right] &\leq 2\varepsilon^p + \frac{2p}{\delta-1}\varepsilon^{1-\delta}e^{-\frac{(\delta+p-1)}{2}}\left(\frac{\beta(\delta+p-1)}{\beta-1}\right)^{\frac{(\delta+p-1)}{2}}\left(\mathbb{E}\left[\langle M\rangle_t^{-\frac{(\delta+p-1)\beta}{2}}\right]\right)^{1/\beta} \\ &= 2\varepsilon^p + \varepsilon^{1-\delta}B\end{aligned}$$

en donde

$$B = \frac{2p}{\delta-1}e^{-\frac{(\delta+p-1)}{2}}\left(\frac{\beta(\delta+p-1)}{\beta-1}\right)^{\frac{(\delta+p-1)}{2}}\left(\mathbb{E}\left[\langle M\rangle_t^{-\frac{(\delta+p-1)\beta}{2}}\right]\right)^{1/\beta}$$

Ahora optimizamos con respecto a ε. Sea

$$A(\varepsilon) = 2\varepsilon^p + \varepsilon^{1-\delta}B$$

entonces

$$A'(\varepsilon) = 2p\varepsilon^{p-1} + (1-\delta)\varepsilon^{-\delta}B$$

y la única solución para $A'(\varepsilon) = 0$ está dada por

$$\varepsilon_0 = \left(\frac{\delta-1}{2p}\right)^{\frac{1}{(p+\delta-1)}} B^{1/(p+\delta-1)}$$

y luego

$$\begin{aligned}A(\varepsilon_0) &= 2\left(\frac{\delta-1}{2p}\right)^{\frac{p}{(p+\delta-1)}} B^{p/(p+\delta-1)} + \left(\frac{\delta-1}{2p}\right)^{\frac{(1-\delta)}{(p+\delta-1)}} B^{(1-\delta)/(p+\delta-1)+1} \\ &= B^{p/(p+\delta-1)}\left(2\left(\frac{\delta-1}{2p}\right)^{\frac{p}{(p+\delta-1)}} + \left(\frac{\delta-1}{2p}\right)^{\frac{(1-\delta)}{(p+\delta-1)}}\right) \\ &= \left(\frac{2p}{\delta-1}\right)^{\frac{p}{(p+\delta-1)}} e^{-\frac{p}{2}}\left(\frac{\beta(\delta+p-1)}{\beta-1}\right)^{\frac{p}{2}}\left(2\left(\frac{\delta-1}{2p}\right)^{\frac{p}{(p+\delta-1)}} + \left(\frac{\delta-1}{2p}\right)^{\frac{(1-\delta)}{(p+\delta-1)}}\right) \\ &\times\left(\mathbb{E}\left[\langle M\rangle_t^{-\frac{(\delta+p-1)\beta}{2}}\right]\right)^{p/(\beta(\delta+p-1))} \\ &= 2e^{-\frac{p}{2}}\left(\frac{\beta(\delta+p-1)}{\beta-1}\right)^{p/2}\left(1+\frac{p}{\delta-1}\right)\left(\mathbb{E}\left[\langle M\rangle_t^{-\frac{(\delta+p-1)\beta}{2}}\right]\right)^{p/(\beta(\delta+p-1))}\end{aligned}$$

En consecuencia

$$\left\|\frac{M_t}{\langle M\rangle_t}\right\|_p \leq 2^{1/p}e^{-1/2}\sqrt{\frac{\beta(\delta+p-1)}{\beta-1}}\left(1+\frac{p}{\delta-1}\right)^{1/p}\left\|\langle M\rangle_t^{-1/2}\right\|_{(\delta+p-1)\beta}$$

Tomemos $\delta = 2-(1/\beta)$ y pongamos $(\delta+p-1)\beta = p+\varepsilon$. Entonces $\varepsilon = (\beta-1)(p+1)$ y así

$$\sqrt{\frac{\beta(\delta+p-1)}{\beta-1}} = \left((p+1)\left(1+\frac{p}{\varepsilon}\right)\right)^{\frac{1}{2}}$$

y

$$\left(1+\frac{p}{\delta-1}\right)^{1/p} = \left((p+1)\left(1+\frac{p}{\varepsilon}\right)\right)^{\frac{1}{p}}$$

De esta manera definimos la constante

$$C_{p,\varepsilon} := \frac{2^{1/p}}{\sqrt{e}}\left((p+1)\left(1+\frac{p}{\varepsilon}\right)\right)^{\frac{1}{2}+\frac{1}{p}}$$

y finalmente obtenemos

$$\left\|\frac{M_t}{\langle M\rangle_t}\right\|_p \leq C_{p,\varepsilon}\left\|\langle M\rangle_t^{-1/2}\right\|_{p+\varepsilon}$$

lo que nos permite concluir con la demostración. $\square$

Al inicio del capítulo comentamos sobre trabajos relacionados, como el del Dr. de la Peña, en donde se menciona esta desigualdad y además se hacen extensiones relacionadas.

Para la martingala M_t, respecto de la filtración $\mathscr{F}_t$, obtuvimos una desigualdad en la expresión (39) de la demostración, para $\beta > 1$, dada por

$$\mathbb{P}[\,M_t > x\langle M\rangle_t\,] \leq \left(\mathbb{E}[\,e^{-(\beta-1)\left(\frac{x^2}{2}\right)\langle M\rangle_t}\,]\right)^{1/\beta}$$

A continuación veremos las extensiones de esta desigualdad expuestas en el trabajo del Dr. de la Peña. Comencemos con un lema que nos será de utilidad.

Lema 3.3. *Sean X, Y dos variables aleatorias tales que $X \geq 0$, $Y \geq 0$ y $X/Y \geq 0$ casi seguramente y además $\mathbb{E}(X/Y) \leq K$ para alguna constante K. Entonces*

$$\mathbb{E}(X^{1/q}) \leq K^{1/q}\,[\,\mathbb{E}(Y^{(p-1)})\,]^{1/p}$$

Demostración. Sean $p > 1$ y $1/p + 1/q = 1$, entonces por la desigualdad de Hölder y junto con la hipótesis tenemos que

$$\begin{aligned}
\mathbb{E}(X^{1/q}) &= \mathbb{E}\left(\left(\frac{X}{Y}\right)^{1/q} Y^{1/q}\right)\\
&\leq \left[\mathbb{E}\left(\frac{X}{Y}\right)\right]^{1/q}\left[\mathbb{E}(Y^{p/q})\right]^{1/p}\\
&\leq K^{1/q}\left[\mathbb{E}(Y^{p/q})\right]^{1/p}
\end{aligned}$$

Notemos que
$$\frac{1}{p}+\frac{1}{q}=1 \iff 1+\frac{p}{q}=p \iff \frac{p}{q}=p-1$$
Entonces por la desigualdad anterior concluimos que
$$\mathbb{E}(X^{1/q}) \leq K^{1/q}\left[\mathbb{E}(Y^{p/q})\right]^{1/p} = K^{1/q}\,[\,\mathbb{E}(Y^{(p-1)})\,]^{1/p}$$
$\square$

Veamos entonces una primer extensión de la desigualdad antes mencionada.

Teorema 3.2. *Sea M_t una martingala continua con respecto de la filtración $\mathscr{F}_t$ y tal que $M_0 = 0$. Sea A un conjunto medible con respecto de $\mathscr{F}_\infty$ y sean $x \geq 0$ y $p > 1$. Entonces*
$$\mathbb{P}[\,M_t \geq x\langle M\rangle_t, A\,] \leq \left(\mathbb{E}\left[\exp\left\{-\frac{(p-1)}{2}x^2\langle M\rangle_t\right\}\mathbb{1}\left(\frac{M_t}{\langle M\rangle_t}\geq x, A\right)\right]\right)^{1/p}$$

Demostración. Sea $\lambda \geq 0$ y q tal que $1/p + 1/q = 1$, entonces por la desigualdad de Markov
$$\begin{aligned}\mathbb{P}[\,M_t \geq x\langle M\rangle_t, A\,] &= \mathbb{P}\left[\exp\left\{\frac{\lambda}{q}M_t - \frac{\lambda}{q}x\langle M\rangle_t\right\}\mathbb{1}\,(M_t \geq x\langle M\rangle_t, A) \geq 1\right]\\ &\leq 1\cdot\mathbb{E}\left[\exp\left\{\frac{\lambda}{q}M_t - \frac{\lambda}{q}x\langle M\rangle_t\right\}\mathbb{1}\,(M_t \geq x\langle M\rangle_t, A)\right]\end{aligned}$$
Denotemos por $A_t = \{M_t \geq x\langle M\rangle_t\} \cap A = (M_t \geq x\langle M\rangle_t, A)$ y definamos las variables aleatorias no negativas
$$\begin{aligned}X &:= \exp\{\lambda M_t - \lambda x\langle M\rangle_t\}\,\mathbb{1}_{A_t}\\ Y &:= \exp\left\{\frac{\lambda^2}{2}\langle M\rangle_t - \lambda x\langle M\rangle_t\right\}\mathbb{1}_{A_t}\end{aligned}$$
y entonces vemos que
$$\frac{X}{Y} = \exp\left\{\lambda M_t - \frac{\lambda^2}{2}\langle M\rangle_t\right\}$$
es decir, X/Y es una martingala local continua[42] y como por hipótesis $M_0 = 0$, entonces
$$\mathbb{E}\left[\frac{X}{Y}\right] = \mathbb{E}\left[\exp\left\{\lambda M_t - \frac{\lambda^2}{2}\langle M\rangle_t\right\}\right] = \mathbb{E}\left[\exp\left\{\lambda M_0 - \frac{\lambda^2}{2}\langle M\rangle_0\right\}\right] = 1$$
En consecuencia, por la desigualdad del lema anterior
$$\mathbb{E}\left[\exp\left\{\frac{\lambda}{q}M_t - \frac{\lambda}{q}x\langle M\rangle_t\right\}\mathbb{1}_{A_t}\right] \leq 1^{1/q}\left(\mathbb{E}\left[\exp\left\{(p-1)\left[\frac{\lambda^2}{2}\langle M\rangle_t - \lambda x\langle M\rangle_t\right]\right\}\mathbb{1}_{A_t}\right]\right)^{1/p}$$
Si tomamos $\lambda = x$, entonces
$$\exp\left\{(p-1)\left[\frac{\lambda^2}{2}\langle M\rangle_t - \lambda x\langle M\rangle_t\right]\right\} = \exp\left\{-\frac{(p-1)}{2}x^2\langle M\rangle_t\right\}$$

[42] [23] Teo. 2.16. Pág. 332.

y de esta manera, por las desigualdades anteriores, concluimos que

$$\mathbb{P}[\, M_t \geq x\langle M\rangle_t, A\,] \leq \left(\mathbb{E}\left[\exp\left\{ -\frac{(p-1)}{2} x^2 \langle M\rangle_t \right\} \mathbb{1}\left(\frac{M_t}{\langle M\rangle_t} \geq x, A \right) \right] \right)^{1/p}$$

$\square$

Veamos ahora otra extensión de la desigualdad previamente mencionada.

Teorema 3.3. *Sea M_t una martingala continua con respecto de la filtración $\mathscr{F}_t$ con $M_0 = 0$ y tal que $\exp\{\lambda M_t - (\lambda^2 \langle M\rangle_t)/2\}$ es una supermartingala para toda $\lambda > 0$. Sea A un conjunto medible con respecto de la filtración $\mathscr{F}_\infty$.*

Entonces, para toda $0 < t < \infty$, $\beta > 0$ y $\alpha, x \geq 0$ se cumple que

$$\mathbb{P}[\, M_t \geq (\alpha + \beta\langle M\rangle_t)x, A\,] \leq \mathbb{E}\left[\exp\left\{ -x^2 \left[\frac{\beta^2}{2}\langle M\rangle_t + \alpha\beta \right] \right\} \middle| (M_t \geq (\alpha + \beta\langle M\rangle_t)x, A) \right], \tag{40}$$

$$\mathbb{P}\left[M_t \geq (\alpha + \beta\langle M\rangle_t)x, \frac{1}{\langle M\rangle_t} < y \quad \textit{para alguna} \;\; t < \infty \right] \leq \exp\left\{ -x^2 \left(\frac{\beta^2}{2y} + \alpha\beta \right) \right\} \tag{41}$$

Demostración. Por la desigualdad de Markov vemos que para toda $\lambda > 0$

$$\begin{aligned}
\mathbb{P}[\, M_t \geq (\alpha + \beta\langle M\rangle_t)x, A\,] &= \mathbb{P}\left[\exp\left\{ \frac{\lambda}{2} M_t - \left(\frac{\lambda\alpha x}{2} - \frac{\lambda\beta x}{2}\langle M\rangle_t \right) \right\} \mathbb{1}\left(\frac{M_t}{(\alpha + \beta\langle M\rangle_t)} \geq x, A \right) \geq 1 \right] \\
&\leq 1 \cdot \mathbb{E}\left[\exp\left\{ \frac{\lambda}{2} M_t - \left(\frac{\lambda\alpha x}{2} - \frac{\lambda\beta x}{2}\langle M\rangle_t \right) \right\} \mathbb{1}\left(\frac{M_t}{(\alpha + \beta\langle M\rangle_t)} \geq x, A \right) \right]
\end{aligned}$$

Denotemos por $A_t = \left\{ \frac{M_t}{(\alpha+\beta\langle M\rangle_t)} \geq x \right\} \cap A = \left(\frac{M_t}{(\alpha+\beta\langle M\rangle_t)} \geq x, A \right)$. Si agregamos un cero y usamos la desigualdad de Cauchy–Schwarz junto la hipótesis $\exp \lambda M_t - (\lambda^2\langle M\rangle_t)/2$ es supermartingala, que implica que $\mathbb{E}[\exp \lambda M_t - (\lambda^2\langle M\rangle_t)/2] \leq 1$, obtenemos entonces que

$$\begin{aligned}
&\mathbb{E}\left[\exp\left\{ \frac{\lambda}{2} M_t - \left(\frac{\lambda\alpha x}{2} - \frac{\lambda\beta x}{2}\langle M\rangle_t \right) \right\} \mathbb{1}_{A_t} \right] \\
&= \exp\left\{ -\frac{\lambda\alpha x}{2} \right\} \mathbb{E}\left[\exp\left\{ \frac{\lambda}{2} M_t - \frac{\lambda^2}{4}\langle M\rangle_t + \frac{\lambda^2}{4}\langle M\rangle_t - \frac{\lambda\beta x}{2}\langle M\rangle_t \right\} \mathbb{1}_{A_t} \right] \\
&\leq \exp\left\{ -\frac{\lambda\alpha x}{2} \right\} \sqrt{\mathbb{E}\left[\exp\left\{ \lambda M_t - \frac{\lambda^2}{2}\langle M\rangle_t \right\} \right]} \sqrt{\mathbb{E}\left[\exp\left\{ \left(\frac{\lambda^2}{2} - \lambda\beta x \right) \langle M\rangle_t \right\} \mathbb{1}_{A_t} \right]} \\
&\leq \exp\left\{ -\frac{\lambda\alpha x}{2} \right\} \sqrt{\mathbb{E}\left[\exp\left\{ \left(\frac{\lambda^2}{2} - \lambda\beta x \right) \langle M\rangle_t \right\} \mathbb{1}_{A_t} \right]}
\end{aligned}$$

Si tomamos $\lambda = \beta x$ se minimiza $\exp\left\{ \left(\frac{\lambda^2}{2} - \lambda\beta x \right) \langle M\rangle_t \right\}$ y por lo anterior obtenemos que

$$\mathbb{P}[\, M_t \geq (\alpha + \beta\langle M\rangle_t)x, A\,] \leq \exp\left\{ -\frac{\alpha\beta x^2}{2} \right\} \sqrt{\mathbb{E}\left[\exp\left\{ -\frac{\beta^2 x^2}{2}\langle M\rangle_t \right\} \mathbb{1}_{A_t} \right]}$$

Notemos que

$$\mathbb{E}\left[\exp\left\{-\frac{\beta^2x^2}{2}\langle M\rangle_t\right\}\mathbb{1}_{A_t}\right]=\mathbb{E}\left[\exp\left\{-\frac{\beta^2x^2}{2}\langle M\rangle_t\right\}\Big|\,A_t\right]\mathbb{P}\left[A_t\right]$$

De manera que si a la desigualdad anterior dividimos por $\sqrt{\mathbb{P}[M_t\geq(\alpha+\beta\langle M\rangle_t)x,A]}$ concluimos

$$\mathbb{P}[\,M_t\geq(\alpha+\beta\langle M\rangle_t)x,A\,]\leq\exp\{-\alpha\beta x^2\}\,\mathbb{E}\left[\exp\left\{-\frac{\beta^2x^2}{2}\langle M\rangle_t\right\}\Big|\left(\frac{M_t}{(\alpha+\beta\langle M\rangle_t)}\geq x,A\right)\right]$$

Demostrando la expresión (40). Ahora probemos la expresión (41). Definimos el tiempo de paro

$$\tau=\text{ínf}\left\{t>0\;:\;\frac{M_t}{(\alpha+\beta\langle M\rangle_t)}\geq x\right\}$$

y definimos

$$A=\left\{\frac{M_t}{(\alpha+\beta\langle M\rangle_t)}\geq x,\frac{1}{\langle M\rangle_t}<y\quad\text{para alguna}\;\;t<\infty\right\}$$

Vemos que $\tau<\infty$ en A y junto con la hipótesis de que $\exp\{\lambda M_t-(\lambda^2\langle M\rangle_t)/2\}$ es una supermartingala para toda $\lambda>0$ y por el teorema de muestreo opcional de Doob junto con el hecho de que $M_0=0$ implican que

$$\mathbb{E}\left[\exp\left\{\lambda M_\tau-\frac{\lambda^2}{2}\langle M\rangle_\tau\right\}\mathbb{1}(A)\right]\leq 1$$

Entonces utilizando la desigualdad anterior y nuevamente la misma línea de razonamiento de la prueba reemplazando a t por τ obtenemos

$$\mathbb{P}[\,M_\tau\geq(\alpha+\beta\langle M\rangle_\tau)x,A\,]\leq\exp\left\{-\frac{\alpha\beta x^2}{2}\right\}\mathbb{E}\left[\exp\left\{-\frac{\beta^2x^2}{2}\langle M\rangle_\tau\right\}\Big|\left(\frac{M_\tau}{(\alpha+\beta\langle M\rangle_\tau)}\geq x,A\right)\right]$$

Notemos que por definición de A y de la identidad $\mathbb{1}_{B\cap C\cap D}=\mathbb{1}_B\mathbb{1}_C\mathbb{1}_D$ se sigue que

$$\mathbb{1}(A)=\mathbb{1}(\{\tau<\infty\})\mathbb{1}(A)\mathbb{1}\left(\left\{\frac{M_\tau}{(\alpha+\beta\langle M\rangle_\tau)}\geq x\right\}\right)$$

En consecuencia junto con la desigualdad anterior que habíamos deducido

$$\begin{aligned}\mathbb{P}[A]&=\mathbb{P}\left[\frac{M_t}{(\alpha+\beta\langle M\rangle_t)}\geq x,\frac{1}{\langle M\rangle_t}<y\quad\text{para alguna}\;\;t<\infty\right]\leq\mathbb{P}\left[\frac{M_\tau}{(\alpha+\beta\langle M\rangle_\tau)}\geq x,A\right]\\&\leq\exp\left\{-\frac{\alpha\beta x^2}{2}\right\}\mathbb{E}\left[\exp\left\{-\frac{\beta^2x^2}{2}\langle M\rangle_\tau\right\}\Big|\left(\frac{M_\tau}{(\alpha+\beta\langle M\rangle_\tau)}\geq x,A\right)\right]\\&\leq\exp\left\{-x^2\left(\frac{\beta^2}{2y}+\alpha\beta\right)\right\}\end{aligned}$$

en donde la última desigualdad se debe a que $\langle M\rangle_\tau>(1/y)$ en A y con esto probamos (41). □

Para enriquecer nuestra exposición sobre las desigualdades para martingalas y en virtud del trabajo de la Dra. Nathalie Eisenbaum y del Dr. Victor H. de la Peña[43] expondremos a continuación una desigualdad aplicada al movimiento browniano.

Comencemos con retomar un resultado del Dr. Michael J. Klass[44] [45].

[43] Véase [6] en la bibliografía.
[44] Véase [14] en la bibliografía.
[45] Véase [15] en la bibliografía.

Teorema 3.4. *Sea* $(d_i)_{i\geq 1}$ *una sucesión de variables aleatorias independientes con valores en un espacio de Banach y sea* $(\tilde{d}_i)_{i\geq 1}$ *un copia independiente de esta sucesión. Sea* $S_n = \sum_{i=1}^n d_i$ *y similarmente* $\tilde{S_n} = \sum_{i=1}^n \tilde{d_i}$*. Entonces para cualquier* $p > 0$ *existen constantes estrictamente positivas* c_p *y* C_p*, tales que para cualquier tiempo de paro* T *con respecto a* S *se cumple que*

$$c_p\,\mathbb{E}\left(\sup_{n\leq T}|\tilde{S_n}|^p\right) \leq \mathbb{E}\left(\sup_{n\leq T}|S_n|^p\right) \leq C_p\,\mathbb{E}\left(\sup_{n\leq T}|\tilde{S_n}|^p\right)$$

Más aún, a la constante C_p *de la segunda desigualdad se puede tomar igual a* $20(18)^p$*.*

Veamos la extensión de este resultado a un proceso en tiempo continuo y con incrementos independientes.

Teorema 3.5. *Sea* $(X_t)_{t\geq 0}$ *un proceso en tiempo continuo con valores en un espacio de Banach tal que sus trayectorias son continuas y tiene incrementos independientes. Sea* $\tilde{X}$ *una copia independiente de* X*. Denotamos por*

$$X_t^* = \sup_{0\leq s\leq t}|X_s| \qquad y \qquad \tilde{X_t^*} = \sup_{0\leq s\leq t}|\tilde{X}_s|.$$

Sea $p > 0$ *y suponga que para cada* $t > 0$*,* $\mathbb{E}[(X_t^*)^p] < \infty$*. Entonces existen constantes positivas,* c_p *y* C_p*, independientes de* X*, tal que para cualquier tiempo de paro* T *con respecto a* X *se cumple*

$$c_p\,\mathbb{E}\left[(\tilde{X_T^*})^p\right] \leq \mathbb{E}\left[(X_T^*)^p\right] \leq C_p\,\mathbb{E}\left[(\tilde{X_T^*})^p\right] \tag{42}$$

Más aún, se puede elegir a C_p *igual a* $20(18)^p$*.*

Demostración. Para t fija consideremos una partición τ_n de $[0,t]$ tal que $\tau_n = (s_i^n)_{0\leq i\leq n}$ con

$$0 = s_0^n < s_1^n < \cdots < s_{n-1}^n < s_n^n = t$$

y $|\tau_n| \to 0$ conforme $n \to \infty$. Notemos que para $n > 0$

$$X_{s_j^n} = \sum_{i=1}^{j}(X_{s_i^n} - X_{s_{i-1}^n}) + X_{s_0^n}$$

De manera que la restricción de X a τ_n es un proceso a tiempo discreto con incrementos independientes.

Podemos elegir la sucesión de particiones $(\tau_n)_{n\geq 0}$ tal que $\tau_{n+1} \supset \tau_n$ y en consecuencia definimos

$$T_n(\omega) := s_{i+1}^n \qquad \text{en} \;\; s_i^n < T\wedge t \leq s_{i+1}^n$$

Y por lo tanto T_n es un tiempo de paro con respecto a $(\mathscr{F}_{s_i^n})_{i\geq 0}$ en donde $\mathscr{F}_s := \sigma(X_u : u \leq s)$.

Notemos que $(T_n)_{n\geq 0}$ es una sucesión decreciente que converge a $T\wedge t$. Por el Teorema 3.4 obtenemos

$$c_p\,\mathbb{E}[\sup_{\substack{s\in\tau_n\\ s\leq T_n}}|\tilde{X}_s|^p] \leq \mathbb{E}[\sup_{\substack{s\in\tau_n\\ s\leq T_n}}|X_s|^p] \leq C_p\,\mathbb{E}[\sup_{\substack{s\in\tau_n\\ s\leq T_n}}|\tilde{X}_s|^p]$$

y tal que se puede elegir a C_p igual a $20(18)^p$. Las hipótesis nos permiten emplear el Teorema de Convergencia Dominada de Lebesgue que conlleva a que

$$c_p\,\mathbb{E}[(\tilde{X}^*_{T\wedge t})^p] \leq \mathbb{E}[(X^*_{T\wedge t})^p] \leq C_p\,\mathbb{E}[(\tilde{X}^*_{T\wedge t})^p]$$

y nuevamente las hipótesis nos permiten emplear el Teorema de Convergencia Monótona para concluir con la demostración al hacer tender t hacia ∞. □

Retomemos la modificación continua del movimiento browniano $(W_t)_{t\geq 0}$. Por las desigualdades de Burkholder Davis Gundi sabemos que para toda $p > 0$ existen dos constantes estrictamente positivas c_p y C_p tales que para cualquier tiempo de paro T con respecto a W se cumple que

$$c_p\,\mathbb{E}(T^{p/2}) \leq \mathbb{E}\left(\sup_{s\leq T}|W_s|^p\right) \leq C_p\,\mathbb{E}(T^{p/2})$$

Sea $\tilde{W}$ una copia independiente de W, entonces se sigue inmediatamente que existen constantes c'_p y C'_p tal que para cualquier tiempo de paro T con respecto a W se tiene que

$$c'_p\,\mathbb{E}\left(\sup_{s\leq T}|\tilde{W}_s|^p\right) \leq \mathbb{E}\left(\sup_{s\leq T}|W_s|^p\right) \leq C'_p\,\mathbb{E}\left(\sup_{s\leq T}|\tilde{W}_s|^p\right) \tag{43}$$

Por la propiedad de escalamiento del movimiento browniano $\tilde{W}$ y por la independencia de $\tilde{W}$ y T

$$\mathbb{E}\left(\sup_{s\leq T}|\tilde{W}_s|^p\right) = \mathbb{E}(T^{p/2})\,\mathbb{E}\left(\sup_{s\leq 1}|W_s|^p\right).$$

Notemos que la expresión (43) es una fromulación equivalente a las desigualdades de Burkholder y Gundi, sin embargo, para el caso del movimiento browniano, son distintas esas desigualdades con respecto a la expresión (42).

La diferencia radica en las constantes, pues como vimos en la expresión (42) la constante C_p puede ser elegida igual a $20(18)^p$ y en la expresión (43) la constante C'_p puede ser elegida[46] por $(24\cdot p^2)^p$.

En consecuencia, veamos la siguiente extensión de las desigualdades del Teorema 3.5 a su versión exponencial.

Corolario 3.6. *Bajo la notación y supuestos del Teorema 3.5 y para cualquier $p > 0$ y cualquier tiempo de paro T con respecto a X y cualquier $\lambda > 0$ se cumple que*

$$\mathbb{E}\left(\exp\{\lambda(X^*_T)^p\}\right) \leq 20\cdot\mathbb{E}\left(\exp\left\{\lambda\,18^p(\tilde{X}^*_T)^p\right\}\right)$$
$$\mathbb{E}\left(\cosh\{\lambda X^*_T\} - 1\right) \leq 20\cdot\mathbb{E}\left(\cosh\left\{18\cdot\lambda\tilde{X}^*_T\right\} - 1\right)$$

Demostración. Sea $n \in \mathbb{N}$, entonces por el Teorema 3.5 tenemos que

$$\mathbb{E}\left[(X^*_T)^{np}\right] \leq 20\cdot\mathbb{E}\left[(18\cdot\tilde{X}^*_T)^{np}\right] \qquad \text{para toda } n\in\mathbb{N}$$

[46] Véase [8] en la bibliografía.

Tomemos las respresentaciones, para $x \in \mathbb{R}$, por

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \qquad \cosh(x) = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}$$

y si sumamos sobre n la desigualdad anterior junto con las representaciones mencionadas concluimos con la prueba. $\square$

Observemos que por la propiedad de escalamiento del movimiento browniano junto con el corolario anterior obtenemos para cualquier tiempo de paro T con respecto a W se cumple que

$$\mathbb{E}\left(\exp\left\{\lambda (W_T^*)^p\right\}\right) \leq 20 \cdot \mathbb{E}\left(\exp\left\{\lambda\, 18^p (\tilde{W}_1^*)^p\, T^{p/2}\right\}\right)$$
$$\mathbb{E}\left(\cosh\left\{\lambda W_T^*\right\} - 1\right) \leq 20 \cdot \mathbb{E}\left(\cosh\left\{18 \cdot \lambda \tilde{W}_1^*\, T^{1/2}\right\} - 1\right)$$

Dado que $\mathbb{P}(W_1^* > t) \leq 2\mathbb{P}(|W_1| > t)$ y directamente de la definición de esperanza y junto con la generadora de momentos de una normal estándar se sigue que

$$\mathbb{E}\left(\exp\{\lambda W_1^*\}\right) \leq 2\mathbb{E}\left(\exp\{\lambda |W_1|\}\right) \leq 4 \exp\left(\frac{\lambda^2}{2}\right)$$

y usando nuevamente la definición de esperanza y la generadora de momentos conlleva a que

$$\mathbb{E}(\cosh\{\lambda W_1^*\} - 1) \leq 2\mathbb{E}(\cosh\{\lambda |W_1|\} - 1) = 2\mathbb{E}(\exp\{\lambda W_1\} - 1) = 2\left(\exp\left(\frac{\lambda^2}{2}\right) - 1\right)$$

Finalmente, si seguimos una línea de razonamiento similar a la que tuvimos con las desigualdades anteriores podemos verificar que para cualquier tiempo de paro T con respecto a W se cumple[47]

$$\mathbb{E}\left(\exp\left\{\lambda W_T^*\right\}\right) \leq 80 \cdot \mathbb{E}\left(\exp\left\{18^2 \frac{\lambda^2}{2} T\right\}\right)$$
$$\mathbb{E}\left(\cosh\left\{\lambda W_T^*\right\} - 1\right) \leq 40 \cdot \mathbb{E}\left(\exp\left\{18^2 \frac{\lambda^2}{2} T\right\} - 1\right)$$

y notemos que estas desigualdades son muy parecidas a un caso particular de la desigualdad del lado derecho de Burkholder Davis Gundi pero en el caso exponencial del movimiento browniano.

Las desigualdades que hemos expuesto serán de gran ayuda para la estimación de la densidad de un proceso de difusión. Junto con las técnicas del cálculo de Malliavin que hemos desarrollado y junto con las estimaciones preliminares veamos en el siguiente capítulo el complemento de las estimaciones que deseamos exponer.

[47] [6] Teo. 2.2. Pág. 241.

3.3. Estimación de Densidades para un Proceso de Difusión

A continuación haremos uso de las estimaciones obtenidas en los capítulos anteriores así como de la derivada de Malliavin de un proceso de difusión, junto con las desigualdades de martingalas, con la finalidad de poder estimar su densidad.

Retomemos la modificación continua del movimiento browniano $W = \{W_t : t \in [0,T]\}$ definido en el espacio de probabilidad canónico $(\Omega, \mathscr{F}, \mathbb{P})$, es decir, $\Omega = C_0([0,T])$, con $\mathscr{F}$ la σ-álgebra completada de Borel en Ω y con $\mathbb{P}$ la medida de Wiener y consideremos el proceso de difusión $X = \{X_t, t \in [0,T]\}$ que satisface la ecuación diferencial estocástica:

$$X_t = x_0 + \int_0^t \sigma(s, X_s)\, dW_s + \int_0^t b(s, X_s)\, ds, \qquad t \in [0,T]$$

Para mejorar la presentación de las estimaciones usaremos la notación $\sigma' = \partial\sigma/\partial x$, $b' = \partial b/\partial x$, $\sigma'' = \partial^2\sigma/\partial x^2$ y $b'' = \partial^2 b/\partial x^2$ y tomemos las siguientes hipótesis sobre los coeficientes σ y b.

(H) Las funciones $\sigma(s,x)$ y $b(s,x)$ son de clase C^2 con respecto a x y

$$\begin{aligned} |\sigma(0,x)| &\leq K, & |b(0,x)| &\leq K \\ |\sigma'(s,x)| &\leq K, & |b'(s,x)| &\leq K \end{aligned}$$

para alguna constante $K > 0$ y $\sigma''(s,x)$ y $b''(s,x)$ tienen crecimiento polinomial en x uniformemente en t.

En lo que resta del presente trabajo denotaremos por C a una constante genérica que puede depender de $p > 1$, $T \geq 0$ y de los coeficientes σ y b.

Veamos la primera estimación de la densidad de un proceso de difusión.

Teorema 3.7. *Sean $\sigma(s,x)$ y $b(s,x)$ funciones que satisfacen las hipótesis **(H)**. Suponga que*

$$\mathbb{E}\left(\left| \int_0^t \sigma(s, X_s)^2 ds \right|^{-p_0/2} \right) < \infty$$

para alguna $p_0 > 2$ y para toda $t \in (0,T]$. Entonces para toda $t \in (0,T]$ la variable aleatoria X_t tiene una densidad continua, denotada por $p_t(x)$, tal que para toda $p > 1$

$$p_t(x) \leq C_p \left\| \left(\int_0^t \sigma(s, X_s)^2\, ds \right)^{-1/2} \right\|_p \tag{44}$$

para alguna constante $C_p > 0$.

Demostración. Sea $t \in (0,T]$ fija. Notemos que bajo las hipótesis **(H)** tenemos que $X_t \in \mathbb{D}^{2,p}$ para toda $p \geq 2$, y si $s \leq t$ tenemos que como consecuencia del Teorema 2.15 y junto con el cálculo en la expresión (31) más la observación sobre el parámetro temporal en los coeficientes conlleva a que

$$D_s X_t = \sigma(s, X_s) + \int_s^t \sigma'(r, X_r) D_s X_r\, dW_r + \int_s^t b'(r, X_r) D_s X_r\, dr$$

y de esta manera

$$D_sX_t = \sigma(s, X_s) \exp\left(\int_s^t \sigma'(r, X_r)\, dW_r + \int_s^t [b' - \tfrac{1}{2}(\sigma')^2](r, X_r)\, dr\right)$$

Ahora usemos la notación

$$M_{s,\,t} = \exp\left(\int_s^t \sigma'(r, X_r)\, dW_r + \int_s^t [b' - \tfrac{1}{2}(\sigma')^2](r, X_r)\, dr\right)$$

de manera que podemos reescribir D_sX_t como

$$D_sX_t = \sigma(s, X_s)\, M_{s,\,t}$$

para toda $s \leq t$.

Para $0 \leq s_2 < s_1 \leq t$ y si usamos la regla del producto, la regla de la cadena y junto con la expresión anterior obtenemos que

$$\begin{aligned}
D^2_{s_1,\,s_2}X_t = {} & (\sigma'(s_1, X_{s_1})D_{s_2}X_{s_1})M_{s_1,\,t} \\
& \quad + \sigma(s_1, X_{s_1})M_{s_1,\,t}\left(\int_{s_1}^t \sigma''(r, X_r)D_{s_2}X_r\, dW_r + \int_{s_1}^t [b'' - \sigma'\sigma''](r, X_r)D_{s_2}X_r\, dr\right) \\
= {} & \sigma'(s_1, X_{s_1})\sigma(s_2, X_{s_2})M_{s_1,\,t}M_{s_2,\,s_1} + \sigma(s_1, X_{s_1})\sigma(s_2, X_{s_2})M_{s_1,\,t} \\
& \quad \times \left(\int_{s_1}^t \sigma''(r, X_r)M_{s_2,\,r}\, dW_r + \int_{s_1}^t [b'' - \sigma'\sigma''](r, X_r)M_{s_2,\,r}\, dr\right)
\end{aligned}$$

El siguiente paso es aplicar el Lema 3.2 a los procesos X_t y $u_s := \sigma(s, X_s)\mathbb{1}_{[0,t]}(s)$.

Por el Lema 2.2 tenemos que u_s pertenece al dominio de δ y como u es un proceso adaptado, entonces δ coincide con la integral de Itô y por lo tanto

$$\delta\left(\sigma(s, X_s)\mathbb{1}_{[0,t]}(s)\right) = \int_0^t \sigma(s, X_s)\, dW_s$$

y entonces vemos que el proceso u cumple los supuestos (i) y (ii) del Lema 3.2 para toda $q > 1$.

Por otro lado tenemos que

$$\begin{aligned}
D_u(X_t) &= \int_0^T D_sX_tu_s\, ds \\
&= \int_0^T \sigma(s, X_s)\, M_{s,\,t}\, \sigma(s, X_s)\mathbb{1}_{[0,t]}(s)\, ds \\
&= \int_0^t \sigma(s, X_s)^2\, M_{s,\,t}\, ds
\end{aligned}$$

y consideremos las variables aleatorias

$$R_t = \int_0^t \sigma(s, X_s)^2 ds$$

y

$$S_t = \left(\sup_{s\in[0,t]} M_{0,\,s}^{-1} \right) \left(\sup_{s\in[0,t]} M_{0,\,s} \right)$$

Vemos que de la definición de $M_{s,\,t}$ se sigue que $\sup_{s\in[0,t]} M_{0,\,s} \geq 1$ y como $\exp(\cdot)$ es una función monótona no decreciente, entonces

$$\sup_{s\in[0,t]} M_{s,\,t}^{-1} \leq \left(\sup_{s\in[0,t]} M_{0,\,s}^{-1} \right) \left(\sup_{s\in[0,t]} M_{0,\,s} \right) = S_t \tag{45}$$

y junto con las hipótesis **(H)** se sigue que $\mathbb{E}(S_t^m) < \infty$ para toda $m \geq 2$.

Como consecuencia de las observaciones anteriores obtenemos la estimación

$$|D_u(X_t)|^{-1} = \left| \int_0^t \sigma(s, X_s)^2 \, M_{s,\,t} \, ds \right|^{-1} \leq (R_t)^{-1} S_t \tag{46}$$

Y con esto exhibimos que $(D_u(X_t))^{-1}$ pertenece a $L^{p'}(\Omega)$ para cualquier $1 < p' < p_0/2$, es decir, se cumple (iii) del Lema 3.2.

Para verificar que la variable Φ_u de la expresión (35) del Lema 3.2 pertenece a $L^{p'}(\Omega)$ para cualquier $1 < p' < p_0/2$ usaremos las siguientes estimaciones que son consecuencia inmediata de la estimación en (46) junto con la desigualdad en (45) y del cálculo de $D^2_{s_1,\,s_2} X_t$.

$$\begin{aligned} \left\| D^2 X_t \right\|_{H\otimes H} \leq (S_t)^2 \sqrt{R_t} \, \left\| \sigma' \right\|_\infty \\ + 2(S_t)^2 R_t \sup_{0\leq s\leq t} \left| \int_0^s \sigma''(r, X_r) M_{0,\,r} \, dW_r + \int_0^s [b'' - \sigma'\sigma''](r, X_r) M_{0,\,r} \, dr \right| \end{aligned}$$

$$\begin{aligned} \|u\|_H^2 &= R_t \\ \|Du\|_{H\otimes H} &\leq \left\| \sigma' \right\|_\infty \sqrt{R_t} \, S_t \\ \|DX_t\|_H &\leq \sqrt{R_t} \, S_t \end{aligned}$$

Entonces siguiendo la definición de Φ_u de la expresión (35) del Lema 3.2 y junto con las estimaciones anteriores tenemos que

$$\begin{aligned} \Phi_u &\leq (R_t)^{-1} \, (S_t)^4 \left(2 \left\| \sigma' \right\|_\infty + 2(R_t)^{1/2} \sup_{0\leq s\leq t} \left| \int_0^s \sigma''(r, X_r) M_{0,\,r} \, dW_r + \int_0^s [b'' - \sigma'\sigma''](r, X_r) M_{0,\,r} \, dr \right| \right) \\ &=: (R_t)^{-1} \Psi_u \end{aligned}$$

y se sigue entonces que Φ_u pertenece a $L^{p'}(\Omega)$ para cualquier $1 < p' < p_0/2$ puesto que $(R_t)^{-1/2}$ pertenece a $L^{p_0}(\Omega)$ y Ψ_u tiene momentos de cualquier orden como consecuencia de las hipótesis **(H)**.

Por la desigualdad en (46) y por el Teorema 3.1 tenemos que para cualquier $p > 1$

$$\begin{aligned} \mathbb{E}\left(\left| \frac{\delta(u)}{D_u X_t} \right| \right) &\leq \mathbb{E}\left(\frac{\left| \int_0^t \sigma(s, X_s) \, dW_s \right|}{\int_0^t \sigma(s, X_s)^2 \, ds} S_t \right) \\ &\leq C_p \left\| (R_t)^{-1/2} \right\|_p \end{aligned}$$

Y por la desigualdad de Hölder obtenemos que

$$\mathbb{E}(\Phi_u \|u\|_H) \leq \mathbb{E}((R_t)^{-1/2}\Psi_u) \leq C_p \left\|(R_t)^{-1/2}\right\|_p$$

y concluimos finalmente que por el Lema 3.2 se sigue que

$$p_t(x) \leq C_p \left\|(R_t)^{-1/2}\right\|_p$$

□

A continuación veremos una aplicación del Teorema 3.7 para el caso en que el coeficiente σ no depende del parámetro temporal.

Supondremos que $\sigma(0) = 0$ y estudiaremos el comportamiento de la densidad como una función de la condición inicial x_0, que en adelante la denotaremos por x, cuando ésta es cercana a 0. Cabe mencionar que el punto de degeneración es tomado igual a 0 para simplificar la exposición.

Teorema 3.8. *Suponga que $\sigma : \mathbb{R} \to \mathbb{R}$ es una función continuamente diferenciable de orden dos tal que σ' está acotada, $\sigma(x) > 0$ para toda $x > 0$ y $\sigma(0) = 0$. Sea $b : [0,T] \times \mathbb{R} \to \mathbb{R}$ una función que satisface las hipótesis **(H)**. Sea $x > 0$ fijo y X_t la solución a la ecuación diferencial estocástica*

$$X_t = x + \int_0^t \sigma(X_s)\, dW_s + \int_0^t b(s, X_s)\, ds, \qquad t \in [0,T] \tag{47}$$

Entonces la densidad $p_t(y)$ de X_t satisface que para cualquier $\varepsilon > 0$

$$p_t(y) \leq C_\varepsilon \left(\sigma(x)^{-1}\, t^{-1/2} + \sigma(x)^{-2-\varepsilon}\, t^{\varepsilon/2}\right) \tag{48}$$

Demostración. Definimos el tiempo de paro τ por

$$\tau = \inf\{s > 0 \ : \ \sigma(X_s) = \tfrac{1}{2}\sigma(x)\}.$$

Notemos que $\sigma(X_0) = \sigma(x) > \frac{1}{2}\sigma(x)$ y usando la hipótesis que $\sigma(x) > 0$ para toda $x > 0$ analicemos los siguientes casos tomando $p > 1$.

Si $t < \tau$, entonces

$$\int_0^t \sigma(X_s)^2\, ds \geq \int_0^t \left(\frac{\sigma(x)}{2}\right)^2 ds$$

y por lo tanto

$$\left(\int_0^t \sigma(X_s)^2\, ds\right)^{-p/2} \leq \left(\frac{\sigma(x)}{2}\right)^{-p} t^{-p/2}.$$

Si $t \geq \tau$, entonces

$$\int_0^t \sigma(X_s)^2\, ds \geq \int_0^\tau \sigma(X_s)^2\, ds$$

y como $0 \leq s \leq \tau$ y junto con el caso anterior vemos que

$$\left(\int_0^t \sigma(X_s)^2\, ds\right)^{-p/2} \leq \left(\int_0^\tau \sigma(X_s)^2\, ds\right)^{-p/2} \leq \left(\frac{\sigma(x)}{2}\right)^{-p} \tau^{-p/2}.$$

Sea $p > 1$, entonces por las cotas anteriores

$$
\begin{aligned}
\mathbb{E}\left(\left(\int_0^t \sigma(X_s)^2 ds\right)^{-p/2}\right) &= \mathbb{E}\left(\mathbb{1}_{\{\tau > t\}}\left(\int_0^t \sigma(X_s)^2 ds\right)^{-p/2}\right) + \mathbb{E}\left(\mathbb{1}_{\{\tau \leq t\}}\left(\int_0^t \sigma(X_s)^2 ds\right)^{-p/2}\right) \\
&\leq \left(\frac{\sigma(x)}{2}\right)^{-p} t^{-p/2} + \left(\frac{\sigma(x)}{2}\right)^{-p} \mathbb{E}\left(\mathbb{1}_{\{\tau \leq t\}}\, \tau^{-p/2}\right)
\end{aligned}
$$

Estimemos la esperanza del último término. Vemos que por la fórmula de integración por partes

$$
\mathbb{E}\left(\mathbb{1}_{\{\tau \leq t\}}\, \tau^{-p/2}\right) = \frac{p}{2}\int_{1/t}^{\infty} y^{(p/2)-1}\, \mathbb{P}\left(\tau < \frac{1}{y}\right) dy \tag{49}
$$

Y por otro lado

$$
\begin{aligned}
\mathbb{P}\left(\tau < \frac{1}{y}\right) &= \mathbb{P}\left(\inf_{0 \leq s \leq 1/y} \sigma(X_s) \leq \frac{1}{2}\sigma(x)\right) \\
&\leq \mathbb{P}\left(\sup_{0 \leq s \leq 1/y} |\sigma(X_s) - \sigma(x)| \geq \frac{1}{2}\sigma(x)\right)
\end{aligned}
$$

Por la desigualdad de Markov se sigue que

$$
\mathbb{P}\left(\sup_{0 \leq s \leq 1/y} |\sigma(X_s) - \sigma(x)| \geq \frac{1}{2}\sigma(x)\right) \leq \left(\frac{2}{\sigma(x)}\right)^q \mathbb{E}\left(\sup_{0 \leq s \leq 1/y} |\sigma(X_s) - \sigma(x)|^q\right)
$$

Usando la hipótesis de que σ es continuamente diferenciable de orden dos vemos que

$$
\left(\frac{2}{\sigma(x)}\right)^q \mathbb{E}\left(\sup_{0 \leq s \leq 1/y} |\sigma(X_s) - \sigma(x)|^q\right) \leq \left(\frac{2}{\sigma(x)}\right)^q \left\|\sigma'\right\|_\infty^q \mathbb{E}\left(\sup_{0 \leq s \leq 1/y} |X_s - x|^q\right)
$$

Las hipótesis **(H)** implican[48] que existe una constante $\tilde{C} > 0$ tal que

$$
\mathbb{E}\left(\sup_{0 \leq s \leq 1/y} |X_s - x|^q\right) \leq \tilde{C}\, y^{-q/2}
$$

De manera que por el análisis anterior, por la hipótesis de que σ' es acotada y tomando la constante genérica $C > 0$ obtenemos que

$$
\begin{aligned}
\mathbb{P}\left(\tau < \frac{1}{y}\right) &\leq \left(\frac{2}{\sigma(x)}\right)^q \left\|\sigma'\right\|_\infty^q \mathbb{E}\left(\sup_{0 \leq s \leq 1/y} |X_s - x|^q\right) \\
&\leq C\left(\frac{2}{\sigma(x)}\right)^q y^{-q/2}
\end{aligned}
$$

En consecuencia, si $p < q$ obtenemos de la identidad (49) y junto con el análisis anterior que

$$
\mathbb{E}\left(\mathbb{1}_{\{\tau \leq t\}}\, \tau^{-p/2}\right) \leq C\sigma(x)^{-q}\, t^{-(p-q)/2}
$$

[48] [23] Lema 2.1 Pág. 523.

De manera que junto con la estimación inicial deducimos que

$$\mathbb{E}\left(\left(\int_0^t \sigma(X_s)^2\, ds\right)^{-p/2}\right) \leq \left(\frac{\sigma(x)}{2}\right)^{-p} t^{-p/2} + \left(\frac{\sigma(x)}{2}\right)^{-p} \mathbb{E}\left(\mathbb{1}_{\{\tau\leq t\}}\, \tau^{-p/2}\right)$$
$$\leq C^p \left(\frac{1}{\sigma(x)\sqrt{t}} + \sigma(x)^{-1-(q/p)}\, t^{-(p-q)/2p}\right)^p < \infty$$

Finalmente, por el Teorema 3.7 y si fijamos $q = p(1+\varepsilon)$ concluimos que

$$p_t(y) \leq C_\varepsilon \left(\sigma(x)^{-1}\, t^{-1/2} + \sigma(x)^{-2-\varepsilon}\, t^{\varepsilon/2}\right)$$

□

Cuando al coeficiente de deriva es no negativo podemos obtener una estimación más fina en (48).

Teorema 3.9. *Suponga que $\sigma : \mathbb{R} \to \mathbb{R}$ es una función continuamente diferenciable de orden dos tal que $\sigma(0) = 0$, $\sigma(x) > 0$ y $\sigma'(x) > 0$ si $x > 0$ y además σ' está acotada. Sea $b : [0,T] \times \mathbb{R} \to \mathbb{R}$ una función que satisface las hipótesis **(H)** y tal que $b(s,x) \geq 0$.*

Entonces, la densidad $p_t(y)$ de la solución X_t de la ecuación diferencial estocástica (47) satisface

$$p_t(y) \leq \frac{C}{\sigma(x)\sqrt{t}} \tag{50}$$

para toda $x > 0$, $t \in (0,T]$.

Demostración. Notemos que por el teorema de comparación para ecuaciones diferenciales estocásticas[49] tenemos que $X_t \geq 0$ para toda $t \in [0,T]$.

Si aplicamos la fórmula de Itô vemos que

$$\sigma(X_s) = \sigma(x) + \int_0^s \sigma'(X_r)\, dX_r + \frac{1}{2}\int_0^s \sigma''(X_r)\, d\langle X\rangle_r$$
$$= \sigma(x) + \int_0^s (\sigma'\sigma)(X_r)\, dW_r + \int_0^s b(r, X_r)\,\sigma'(X_r)\, dr + \int_0^s \frac{(\sigma''\sigma^2)(X_r)}{2}\, dr$$

y se sigue entonces que

$$\frac{\sigma(X_s)}{\sigma(x)} = 1 + \int_0^s \frac{(\sigma'\sigma)(X_r)}{\sigma(x)}\, dW_r + \int_0^s \frac{(\sigma''\sigma^2)(X_r)}{2\sigma(x)}\, dr + \int_0^s \frac{b(r,X_r)\,\sigma'(X_r)}{\sigma(x)}\, dr$$

Sea $Z_s := \frac{\sigma(X_s)}{\sigma(x)}$, entonces podemos reescribir las cuentas anteriores con notación diferencial

$$dZ_s = \sigma'(X_s)\, Z_s\, dW_s + \frac{\sigma''\sigma}{2}(X_s)\, Z_s\, ds + \frac{b(s,X_s)\,\sigma'(X_s)}{\sigma(x)}\, ds$$
$$= \left[\sigma'(X_s)\, dW_s + \frac{\sigma''\sigma}{2}(X_s)\, ds + \frac{1}{Z_s}\frac{b(s,X_s)\,\sigma'(X_s)}{\sigma(x)}\, ds\right] Z_s$$

[49] [12] Teo. 1. Pág. 437.

y mediante la fórmula de Itô y la exponencial estocástica vemos que si

$$N_s = \int_0^s \sigma'(X_r)\, dW_r - \frac{1}{2}\int_0^s (\sigma'(X_r))^2 dr + \frac{1}{2}\int_0^s (\sigma''\sigma)(X_r)\, dr$$

entonces

$$\frac{\sigma(X_s)}{\sigma(x)} = e^{N_s}\left(1 + \int_0^s \frac{e^{-N_r}}{\sigma(x)} b(r, X_r)\,\sigma'(X_r)\, dr\right)$$

Por el Teorema 3.7 en la estimación (44) obtenemos que

$$\begin{aligned}
\sigma(x)\, p_t(y) &\leq C \left\| \left(\int_0^t \left(\frac{\sigma(X_s)}{\sigma(x)}\right)^2 ds \right)^{-1/2} \right\|_p \\
&\leq C \left\| \sup_{0\leq s\leq t} e^{-N_s} \left(\int_0^t \left(1 + \int_0^s \frac{e^{-N_r}}{\sigma(x)} b(r, X_r)\,\sigma'(X_r)\, dr\right)^2 ds \right)^{-1/2} \right\|_p \\
&\leq \frac{C}{\sqrt{t}}
\end{aligned}$$

es decir

$$p_t(y) \leq \frac{C}{\sigma(x)\sqrt{t}}$$

que era lo que deseábamos probar. □

Ilustremos la aplicación de los resultados anteriores con el siguiente ejemplo. Consideremos la ecuación diferencial estocástica

$$\begin{aligned}
dX_t &= \beta(\alpha - X_t)dt + \sigma dW_t \\
X_0 &= x_0
\end{aligned}$$

en donde $\beta > 0$, $\sigma > 0$, $\alpha \in \mathbb{R}$. Entonces notemos que las funciones

$$\begin{aligned}
b(s, x) &= \beta(\alpha - x) \\
\sigma(s, x) &= \sigma
\end{aligned}$$

satisfacen las hipótesis del teorema de existencia y unicidad así como las hipótesis **(H)**.

A la única solución adaptada a la filtración se le llama ***proceso de Ornstein-Uhlenbeck*** en donde α representa la media asintótica, β la velocidad de reversión a la media y σ la dispersión o tamaño del ruido.

Si aplicamos la fórmula de Itô a la expresión $X_t e^{\beta t}$ obtenemos que

$$X_t = \alpha + (x_0 - \alpha)e^{-\beta t} + \sigma \int_0^t e^{-\beta(t-s)}\, dW_s.$$

Sabemos que si $f : \mathbb{R} \to \mathbb{R}$ es una función determinista que pertenece a $L^2([0, t])$, entonces

$$\int_0^t f(s)\, dW_s \sim N\left(0, \int_0^t |f(s)|^2\, ds\right)$$

y luego junto con la Isometría de Itô vemos que

$$\begin{aligned}
\mathbb{E}(X_t|X_0 = x_0) &= \alpha + (x_0 - \alpha)e^{-\beta t} \\
\mathbb{V}(X_t|X_0 = x_0) &= \mathbb{E}(X_t^2|X_0 = x_0) - \mathbb{E}(X_t|X_0 = x_0)^2 = \frac{\sigma^2}{2\beta}(1 - e^{-2\beta t})
\end{aligned}$$

y como la media y la varianza determinan a la normal tenemos que

$$(X_t|X_0 = x_0) \sim N\left(\alpha + (x_0 - \alpha)e^{-\beta t}, \frac{\sigma^2}{2\beta}(1 - e^{-2\beta t})\right). \tag{51}$$

Por otro lado vemos que

$$\left\| \left(\int_0^t \sigma(s, X_s)^2 \, ds \right)^{-1/2} \right\|_p = \left\| \left(\int_0^t \sigma^2 \, ds \right)^{-1/2} \right\|_p = \left\| (\sigma^2 t)^{-1/2} \right\|_p = \frac{1}{\sigma \sqrt{t}}$$

Notemos que para cualquier $p_0 > 2$ se cumplen las hipótesis del Teorema 3.7 y entonces existe una constante $C > 0$ tal que si $p_t(x)$ denota a la densidad de X_t, entonces se cumple que

$$p_t(x) \leq C \left\| \left(\int_0^t \sigma(s, X_s)^2 \, ds \right)^{-1/2} \right\|_p = \frac{C}{\sigma \sqrt{t}} \qquad t \in (0, T], \; p > 1.$$

Aquí notemos también la conexión con la estimación del Teorema 3.9.

Para verificar la desigualdad anterior veamos primero que por la expresión (51) se tiene que

$$p_t(x) = \frac{1}{\sigma} \sqrt{\frac{\beta}{\pi}} \frac{1}{\sqrt{1 - e^{-2\beta t}}} \exp\left\{ -\frac{\beta}{\sigma^2} \frac{[x - \alpha - (x_0 - \alpha)e^{-\beta t}]^2}{1 - e^{-2\beta t}} \right\}$$

Ahora bien, por la regla de L'Hôpital podemos verificar que

$$\lim_{t \to 0^+} \frac{\sqrt{t}}{\sqrt{1 - e^{-2\beta t}}} = \frac{1}{\sqrt{2\beta}}$$

y entonces si definimos la función $g : \mathbb{R} \to \mathbb{R}$

$$g(t) = \begin{cases} \frac{1}{\sqrt{2\beta}} & t = 0, \\ \frac{\sqrt{t}}{\sqrt{1 - e^{-2\beta t}}} & t \in (0, T]. \end{cases}$$

tenemos que g es continua en el compacto $[0, T]$ y por lo tanto existe $K > 0$ tal que $|g(t)| \leq K$ para toda $t \in [0, T]$. Mediante el cálculo diferencial clásico podemos verificar que g es creciente en $[0, T]$ y por lo tanto

$$K = \frac{\sqrt{T}}{\sqrt{1 - e^{-2\beta T}}}$$

El argumento anterior conlleva a que

$$\frac{1}{\sqrt{1 - e^{-2\beta t}}} \leq \frac{K}{\sqrt{t}} \quad \text{para toda} \;\; t \in (0, T]. \tag{52}$$

Por otro lado notemos que $(1-e^{-2\beta t})>0$ y $[x-\alpha-(x_0-\alpha)e^{-\beta t}]^2\geq 0$ para toda $t\in(0,T]$. Y como por hipótesis $\beta>0$ y $\sigma>0$, entonces se sigue que

$$\left\{-\frac{\beta}{\sigma^2}\,\frac{[x-\alpha-(x_0-\alpha)e^{-\beta t}]^2}{1-e^{-2\beta t}}\right\}\leq 0 \qquad \text{para toda}\ \ t\in(0,T]$$

y como la función $\exp(\cdot)$ es monótona no decreciente concluimos que

$$\exp\left\{-\frac{\beta}{\sigma^2}\,\frac{[x-\alpha-(x_0-\alpha)e^{-\beta t}]^2}{1-e^{-2\beta t}}\right\}\leq 1 \tag{53}$$

Si definimos la constante $C>0$ como

$$C=\sqrt{\frac{\beta}{\pi}}\cdot\frac{\sqrt{T}}{\sqrt{1-e^{-2\beta T}}}$$

entonces por las expresiones (52) y (53) concluimos que

$$p_t(x)=\frac{1}{\sigma}\sqrt{\frac{\beta}{\pi}}\frac{1}{\sqrt{1-e^{-2\beta t}}}\exp\left\{-\frac{\beta}{\sigma^2}\,\frac{[x-\alpha-(x_0-\alpha)e^{-\beta t}]^2}{1-e^{-2\beta t}}\right\}\leq\frac{C}{\sigma\sqrt{t}}\cdot 1 \qquad \text{para toda}\ \ t\in(0,T]$$

como lo afirmaba desde un principio el Teorema 3.7.

Cabe mencionar que las estimaciones de los teoremas anteriores también se pueden aplicar en el caso en que la densidad es desconocida.

Para la ecuación diferencial estocástica del ejemplo anterior pudimos ilustrar la aplicación del Teorema 3.7 y su conexión con el Teorema 3.9 así como la obtención explícita de la constante de acotamiento.

Es importante resaltar la importancia de la desigualdad para martingalas del Teorema 3.1 para la obtención de las estimaciones que hemos expuesto.

Finalmente, con este ejemplo vimos la conexión entre los diversos resultados que hemos venido estudiando así como la manera en que cada una de sus componentes tiene su importancia individual.

4. Conclusiones

En el presente trabajo hemos atestiguado una interesante conjunción de dos áreas de las matemáticas; el análisis funcional y la teoría de probabilidad. Partimos con el concepto de derivación de funciones en espacios de Banach así como propiedades de densidad y espacios de Hilbert asociados con espacios de probabilidad.

En primera instancia ilustramos el concepto del operador derivada en un sentido débil y que jugó un papel crucial para nuestro estudio del cálculo estocástico de variaciones. Nuevamente retomamos conceptos como el operador adjunto de la derivada de Malliavin y presenciamos una simbiosis de elementos de funcional con los de probabilidad.

Una vez desarrolladas las herramientas clave del cálculo de Malliavin pudimos entonces aplicarlas sobre los procesos de difusión asociados a la solución de una ecuación diferencial estocástica que satisface ciertas condiciones sobre sus coeficientes.

No dejemos de lado el grandioso papel que jugaron las desigualdades para martingalas que estudiamos el Capítulo 3.2 y que pudimos enriquecer nuestra exposición hacia diversos resultados concernientes con artículos de investigación relacionados.

Finalmente podemos concluir la importancia de la teoría del cálculo de Malliavin tanto en su profundidad teórica, lo cual nos permitió consolidar los conocimientos del cálculo estocástico continuo así como en su implicaciones hacia los procesos estocásticos, en particular, los procesos de difusión.

En lo que respecta hacia las líneas de continuidad del presente trabajo tenemos que el cálculo de Malliavin se puede desarrollar hacia procesos más generales como los procesos de Lévy y cuya aplicación cotidiana ha sido empleada particularmente hacia las finanzas matemáticas.

Cabe mencionar que en este sentido también se buscarían resultados de estimación de densidades para ecuaciones diferenciales estocásticas con condiciones adicionales, como por ejemplo condiciones de frontera. Y también se buscaría como línea futura de estudio su implementación hacia el control estocástico y propiedades de regularidad para lo cual requiere un desarrollo más avanzado en lo que respecta al presente trabajo.

Bibliografía

[1] Bartle R. *The Elements of Integration and Lebesgue Measure.* Wiley Classics, 1995.

[2] Bell J. *Fréchet derivatives and Gâteaux derivatives.* Department of Mathematics, University of Toronto, 2014.

[3] Caballero, M.E., Fernández, B. and Nualart, D. *Estimation of Densities and Applications.* Journal of Theoretical Probability, Vol. 11, No. 3, 1998.

[4] Cameron, R.H. and Martin, W.T. *Transformations of weiner integrals under translations.* The Annals of Mathematics, 45(2):386–396, 1944.

[5] Clapp, M. *Introducción al Análisis Real.* Instituto de Matemáticas UNAM, 2010.

[6] De la Peña, V. H. and Eisenbaum, N. *Exponential Burkholder Davis Gundy Inequalities.* Bulletin of the London Mathematical Society (1997) 29 (2): 239-242.

[7] De la Peña V. H. *A General Class of Exponential Inequalities for Martingales and Ratios.* Annals of Probability, Volume 27, Number 1 (1999), 537-564.

[8] Dellacherie, C. and Meyer, A. *Probabilités et potentiel-Théorie des martingales.* Hermann Paris, 1980.

[9] Di Nunno, G., Oksendal, B. and Proske, F. *Malliavin Calculus for Lévy Processes with Applications to Finance.* Universitext. Springer-Verlag Berlin Heidelberg, 2009.

[10] Dynkin, E.B. *Markov Processes, Vol. I.* Springer, 1965.

[11] Grasselli, M. and Hurd, T. *Malliavin Calculus.* Department of Mathematics & Statistics McMaster University, 2005.

[12] Ikeda, N. and Watanabe, S. *Stochastic Differential Equations and Diffusion Processes.* Kodansha Scientific Books, 1989.

[13] Kallenberg, O. *Foundations of Modern Probability.* Probability and Its Applications. Springer, 2002.

[14] Klass M. J. *A best possible improvement of Wald's equation*, Annals of Probability, 16 (1988): 840-853.

[15] Klass M. J. *Uniform lower bounds for randomly stopped Banach space valued random sums*, Annals of Probability, 18 (1990): 790-809.

[16] Nualart, D. *The Malliavin Calculus and Related Topics.* Springer, 2006.

[17] Nualart, E. *Lectures on Malliavin calculus and its applications to finance.* University of Paris 13, 2009.

[18] Oksendal, B. *Stochastic Differential Equations.* Springer, 2000.

[19] Oksendal, B. *An Introduction to Malliavin Calculus with Applications to Economics.* Dept. of Mathematics, University of Oslo, 1997.

[20] Revuz, D. and Yor, M. *Continuous Martingales and Brownian Motion.* Grundlehren der mathematischen Wissenchaften. Springer Berlin Heidelberg, 2005.

[21] Siddiqi, A.H. *Applied Functional Analysis.* Marcel Dekker U.S.A., 2004.

[22] Taylor, A. and Lay, C. *Introduction to Functional Analysis*. Krieger Publishing, 1980.

[23] Tudor, C. *Procesos Estocásticos*. Sociedad Matemática Mexicana, 2002.

[24] Uribe, G. *Curso Avanzado de Probabilidad, Integración estocástica*. Instituto de Matemáticas UNAM, 2014-I.

[25] Ward, C. *Analysis for Applied Mathematics*. Graduate Texts in Mathematics. Springer-Verlag New York, 2001.