UNIVERSIDAD DE MURCIA

Posgrado de Matemáticas
Máster de Matemática Avanzada

TESIS DE MASTER

CONVERGENCIA DE MARTINGALAS Y EL TEOREMA DE RADON-NIKODYM

Kabuya Singya Ongo Muraïri

Curso 2006-7

D. Gabriel Vera Botí, Catedrático del Área de Análisis Matemático y Coordinador del Programa de Posgrado de Matemáticas,

INFORMA: Que la Tesis de Master titulada "Convergencia de Martingalas y el Teorema de Radon-Nikodým" ha sido realizada por D. Kabuya Singya Ongo Muraïri, dentro del Máster de Matemática Avanzada.

En Murcia, a 19 de Noviembre de 2007

Fdo: Gabriel Vera Botí

D. Gabriel Vera Botí, Catedrático del Área de Análisis Matemático del Departamento de Matemáticas de la Universidad de Murcia.

AUTORIZA: La presentación de la Tesis de Máster titulada "Convergencia de Martingalas y el Teorema de Radon-Nikodým", realizada por D. Kabuya Singya Ongo Muraïri, bajo mi inmediata dirección y supervisión, dentro del Máster de Matemática Avanzada.

En Murcia, a 19 de Noviembre de 2007

Fdo: Gabriel Vera Botí

Dedicatoria

There are several persons I would like to dedicate this paper to. First of all my parents, they are always present in my heart and I hope that they are proud of their son's academic achievements. Secondly Christelle, she has always been on my side while I was struggling studying and trying to understand deep mathematical concepts. However, I dedicate this paper to two persons in particular. The first is Bob Creagar, my mathematics instructor at Eastern Wyoming College. He was more than an instructor, he was a true friend and I want him to know how much I enjoyed learning at his side and how much I appreciated his friendship. The second is Dr. Michael Sterner from the University of Montevallo. Dr. Sterner is the one who encouraged me to start graduate studies in mathematics. More important, Dr. Sterner taught me a new way of doing mathematics, the passionate way.

Agradecimientos

Quiero darle las gracias a todos los profesores del Área de Análisis Matemático de la universidad de Murcia por su eseñanza de alta calidad. En particular he disfrutado mucho con las enseñanzas del Profesor José Orihuela y le estoy agradecido por sus múltiplos consejos y su franquesa. Tengo que agradacer especialmente al Profesor Gabriel Vera Botí por haber sido tremendamente paciente conmigo a la hora de acabar la Tesis de Máster.

Índice general

Introducción

El concepto de martingala fue introducido a las matemáticas por el probabilista Paul Pierre Lévy pero fue Joseph Leo Doob quién desarolló su teoría tal como se entiende hoy. Por tanto Doob es considerado el fundador de la teoría de martingalas. Además de ser una de las principales herramientas de la teoría de Procesos Estocásticos, la teoría de martingalas tiene importantes aplicaciones al Análisis Matemático. Esta teoría formaliza la idea intuitiva de un proceso que evoluciona de modo que se gana información conforme transcurre el tiempo. El teorema de convergencia de martingalas dice en qué sentido y cuanta información se puede recuperar mediante un proceso de paso al límite.

El significado intuitivo de la noción de martingala se describe frecuamente gracias al modelo de la evolución de la fortuna de un jugador que realiza una serie de apuestas en un juego de azar equitativo (por ejemplo cara y cruz con una moneda equilibrada). Con este modelo en mente, se pueden interpretar los tiempos de parada (una de las nociones básicas de la teoría) como las posibles estrategias que puede adoptar el jugador para decidir cuando terminan sus apuestas. Intuitivamente la gente estaba convencida de que no existiera ninguna estrategia de este tipo que pudiera incrementar la ganancia esperada. Un resultado fundamental de la teoría de martingalas formaliza esta idea.

Además de la Probabilidad y la teoría de Procesos Estocásticos, se pueden citar otras ramas donde se usan y aplican las martingalas: Integración Estocástica, Matemática Financiera, Geometría de los espacios de Banach.

La etimología de la palabra martingala es obscura[1], probablemente debido al hecho de que la palabra martingala no tiene un sentido único. La martingala tiene significado en el vocabulario ecuestre pero también en el vocabulario de las apuestas. Una martingala es una estrategia de juego (de dinero) que garantiza la ganancia de

[1] La información sobre la etimología de la palabra martingala proviene del articulo “Histoire de Martingales” de Roger Mansuy [3]

un jugador. Hubo una época donde era común que un jugador pretendiendo poseer una martingala consultara a un probabilista. Es así que la palabra martingala se trasmitió del vocabulario de los jugadores al de los probabilistas.

La palabra Martingala es de origen francés "Martingale" y Doob ha querido usar la misma palabra en inglés. Esa palabra proviene de una expresión francesa "Jouer a la Martingale" que significa jugar todo lo que se ha perdido, o ariesgarlo todo. Esta expresión descendería de una expresión provençal más antigua "Jouga a la Martegalo" que significa jugar de manera absurda e incomprensible. Para entender mejor el significado de esa expresión, ayuda saber que una de las primeras martingalas consiste en doblar la apuesta cada vez que se pierde hasta que se gane y así recuperar todo el dinero perdido. Se puede imaginar perfectamente que tal estrategia fuera considerada absurda. Teniendo en cuenta que en la práctica el presupuesto de un jugador es limitado, el riesgo de arruinarse es grande.

La combinación de la base intuitiva de la teoría de martingalas y su simplicidad matemática permite aplicaciones a situaciones concretas para obtener resultados nada triviales. En esta memoría sólo consideramos martingalas discretas (con índices en los naturales). Esta clase de procesos estocásticos permiten demostrar resultados centrales de la teoría de Probabilidad y del Análisis Matemático. Hemos elegido dos representativos que se presentan en el capítulo de las aplicaciones: Las demostraciones de la ley fuerte de los grandes números y del teorema de Radon-Nikodým.

El teorema de Radon-Nikodým es un teorema fundamental en teoría de la medida que tiene consecuencias importantes en Análisis. El teorema estipula que dado un espacio de medida σ-finito (Ω, Σ, μ), si ν es una medida σ-finita sobre Σ, absolutamente continua con respecto a μ, entonces existe una función Σ-medible $f : \Omega \longrightarrow [0, \infty)$ tal que $\nu(E) = \int_E f \; d\mu$ para cada $E \in \Sigma$. La función f lleva el nombre de derivada de Radon-Nikodým. Las demostraciones del teorema de Radon-Nikodým que se encuentran en libros de teoría de la medida o de probabilidad son a veces poco intuitivas. Se demuestra la existencia de la derivada de Radon-Nikodým de manera no constructiva. Dado un espacio de probabilidad, el teorema de Radon-Nikodým implica la existencia de la esperanza condicionada.

La esperanza condicionada puede verse como un operador lineal con propriedades muy similares a las de la integral. Demostraremos las versiones para la esperanza condicionada del lema de Fatou, del teorema de la convergencia monótona, del teorema de la convergencia dominada y de la desigualdad de Jensen. La esperanza condicionada es una herramienta indispensable de la teoría de las martingalas.

El objetivo de esta tesis de máster es demostrar el teorema de convergencia de martingalas. En vez de usar el teorema de Radon-Nikodým para demostrar la existencia de la esperanza condicionada, demostraremos en primer lugar su existencia para funciones de cuadrado integrable. Veremos que la esperanza condicionada de una función Σ-medible f de cuadrado integrable es la proyección ortogonal de f sobre el subespacio de las funciones de cuadrado integrable Σ_0-medible donde $\Sigma_0 \subset \Sigma$. Gracias al teorema de la convergencia monótona el resultado se extiende para todas las funciones integrables. Usaremos entonces el teorema de convergencia de martingalas para dar una demostración alternativa del teorema de Radon-Nikodým, construyendo de manera explicita una martingala convergiendo hacia la derivada de Radon-Nikodým.

En el capítulo préliminar, la sección teoría de la Medida e Integración sirve esencialmente para fijar la notación y la terminoligía. En esta sección están enunciados los resultados fundamentales necesarios para las secciones y capítulos siguientes. Las secciones modos de convergencia y integrabilidad uniforme son importantes y se completan. Una martingala que converge en casi todo punto, no tiene porque converger en media. Añadiendo la hypótesis de integrabilidad uniforme, gracias al teorema de convergencia de Vitali podemos decir que la martingala converge en media. La sección Espacio de Hilbert sirve para demostrar el teorema de la proyección ortogonal. Este teorema se usará para demostrar la existencia de la esperanza condicionada en $\mathcal{L}^2$.

El segundo capítulo empieza con un ejemplo intuitivo que ayuda a entender la definición abstracta de la esperanza condicionada y está dedicado a demostrar sus propriedades elementales. Este ejemplo es la base intuitiva de la demostración del teorema de Radon-Nikodým del último capítulo.

El tercer y cuarto capítulo se dedican a las martingalas y sus aplicaciones respectivamente. Hemos elegido seguir más o menos la estructura de los capítulos 10, 11, 12 y 14 en Williams [6].

Capítulo 1

Préliminares

1.1. Teoría de la Medida e Integración

Esta sección contiene los resultados fundamentales de teoría de la medida e integración que se van a usar con frecuencia. Los resultados serán enunciados con o sin demostraciones pero se pueden encontrar en Cohn [2], Rudin [4] y Ash [1]. El propósito de esta sección es fijar la notación y la terminología. Se suponen conocidas las nociones de σ-álgebra, medida, probabilidad y también el concepto de medibilidad e integrabilidad de funciones. Denotaremos por (Ω, Σ) a un espacio medible y por (Ω, Σ, μ) a un espacio de medida donde Ω es un conjunto abstracto, Σ una σ-álgebra sobre Ω y μ es una medida sobre Σ. Por comodidad supondremos todas las medidas consideradas, completas.

Otro concepto importante supuesto conocido, es el concepto de propriedad que se verifica en casi todo punto, es decir el conjunto de los puntos donde no se cumple dicha propriedad es nulo. Por ejemplo, considerando las funciones reales f y g definidas en (Ω, Σ, μ), la afirmación $f(\omega) = g(\omega)$ en casi todo punto (*en c.t.p*) significa que $\mu(\{\omega \in \Omega : f(\omega) \neq g(\omega)\}) = 0$.

1.1.1. Medida y Probabilidad

Sea Ω un conjunto abstracto. Diremos que $f : \Omega \longrightarrow F$ es $(\Sigma, \mathcal{F})$-medible si (Ω, Σ) y $(F, \mathcal{F})$ son dos espacios medibles y para cada $E \in \mathcal{F}$ el conjunto $f^{-1}(E) \in \Sigma$.

Definición 1.1.1. *Sean $\mathcal{H}$ una familia de subconjuntos de Ω y $\{\mathcal{S}_\alpha\}_{\alpha \in \mathcal{A}}$ la familia de todas las σ-álgebras que contienen a $\mathcal{H}$. La σ-álgebra generada por $\mathcal{H}$, denotada* *por $\sigma(\mathcal{H})$, es la mínima σ-álgebra que contiene a $\mathcal{H}$:*

$$\sigma(\mathcal{H}) = \bigcap_{\alpha\in\mathcal{A}} \mathcal{S}_\alpha.$$

Sea $\mathcal{B}(\mathbb{R})$ la σ-álgebra de Borel, es decir la σ-álgebra generada por la topología estandar sobre $\mathbb{R}$. Si f es $(\Sigma, \mathcal{B}(\mathbb{R}))$-medible, diremos simplemente que f es Σ-medible. Substituyendo el conjunto $\mathbb{R}$ por su ampliación $[-\infty, \infty]$, si f es $(\Sigma, \mathcal{B}([-\infty, \infty]))$-medible, también diremos que f es Σ-medible ($\mathcal{B}([-\infty, \infty])$ está generada por los abiertos de $[-\infty, \infty]$).

Definición 1.1.2. *Sean $\{(\Omega_\alpha, \Sigma_\alpha)\}_{\alpha\in\mathcal{A}}$ una familia de espacios medibles y $\{f_\alpha\}_{\alpha\in\mathcal{A}}$ una familia de funciones tal que $f_\alpha : \Omega \longrightarrow \Omega_\alpha$, para cada $\alpha \in \mathcal{A}$. La σ-álgebra generada por la familia $\{f_\alpha\}$, denotada por $\sigma(\{f_\alpha\}_{\alpha\in\mathcal{A}})$, es la mínima σ-álgebra Σ sobre Ω tal que f_α es (Σ, Σ_α)-medible para cada $\alpha \in \mathcal{A}$:*

$$\sigma(\{f_\alpha\}_{\alpha\in\mathcal{A}}) = \sigma\left(\bigcup_{\alpha\in\mathcal{A}} f_\alpha^{-1}(\Sigma_\alpha)\right).$$

Dada una sucesión de σ-álgebras $\{\Sigma_n\}$ sobre Ω, una σ-álgebra $\mathcal{G}$ sobre G y una función $f : \Omega \longrightarrow G$, definimos

$$\sigma(\Sigma_1, \Sigma_2, ..., \Sigma_k) = \sigma\left(\bigcup_{i=1}^{k} \Sigma_i\right),$$

$$\Sigma_\infty = \sigma\left(\bigcup_{i=1}^{\infty} \Sigma_i\right),$$

$$\sigma(f, \Sigma_i) = \sigma(f^{-1}(\mathcal{G}) \cup \Sigma_i).$$

Definición 1.1.3. *Una familia $\mathcal{D}$ de subconjuntos de Ω es un d-sistema sobre Ω si se verifican:*

(i) *$\Omega \in \mathcal{D}$,*

(ii) *$A, B \in \mathcal{D}$ y $B \subset A \Longrightarrow A \setminus B \in \mathcal{D}$,*

(iii) *$\{A_n\}$ es una sucesión creciente en $\mathcal{D} \Longrightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$.*

Definición 1.1.4. *Sean $\mathcal{H}$ una familia de subconjuntos de Ω y $\{\mathcal{D}_\alpha\}_{\alpha\in\mathcal{A}}$ la familia de todos los d-sistemas que contienen a $\mathcal{H}$. El d-sistema generado por $\mathcal{H}$, denotado por $d(\mathcal{H})$, es el mínimo d-sistema que contiene a $\mathcal{H}$:*

$$d(\mathcal{H}) = \bigcap_{\alpha\in\mathcal{A}} \mathcal{D}_\alpha.$$

Definición 1.1.5. *Una familia $\mathcal{H}$ de subconjuntos de Ω es un π-sistema sobre Ω si se verifica*

$$A_1, A_2, \ldots, A_n \in \mathcal{H} \Longrightarrow \bigcap_{i=1}^{n} A_i \in \mathcal{H}.$$

Lema 1.1.1. (Lema de Dynkin) *Sea $\mathcal{H}$ un π-sistema sobre un conjunto Ω. La σ-álegbra generada por $\mathcal{H}$ coincide con el d-sistema generado por $\mathcal{H}$. Es decir,*

$$\sigma(\mathcal{H}) = d(\mathcal{H}).$$

Demostración: (Véase Teorema 1.6.1 en Cohn [2]) ∎

Corolario 1.1.1. *Sean (Ω, Σ) un espacio medible y $\mathcal{H}$ un π-sistema sobre Ω tal que $\mathcal{H} \subset \Sigma$. Supongamos que μ_1 y μ_2 sean medidas sobre Σ tales que $\mu_1(\Omega) = \mu_2(\Omega) < \infty$ y $\mu_1(E) = \mu_2(E)$ para cada $E \in \mathcal{H}$. Entonces,*

$$\mu_1(E) = \mu_2(E) \qquad \forall\ E \in \sigma(\mathcal{H}).$$

Demostración: (Véase Corolario 1.6.2 en Cohn [2]) ∎

Definición 1.1.6. *Sea (Ω, Σ, μ) un espacio de probabilidad. Una familia $\{\mathcal{G}_\alpha\}_{\alpha\in\mathcal{A}}$ de sub-σ-álgebras de Σ se dice independiente si para cada conjunto finito $\{\alpha_1, \alpha_2, \ldots, \alpha_n\} \subset \mathcal{A}$ con $\alpha_i \neq \alpha_j$ para cada $(i,j) \in \{1,2,\ldots,n\} \times \{1,2,\ldots,n\}$ y cada familia finita $\{G_{\alpha_i}\}_{i=1}^n$ tal que $G_{\alpha_i} \in \mathcal{G}_{\alpha_i}$ para cada $i \in \{1, 2, \ldots n\}$, se verifica*

$$\mu\left(\bigcap_{i=1}^{n} G_{\alpha_i}\right) = \prod_{i=1}^{n} \mu(G_{\alpha_i}).$$

Definición 1.1.7. *Sea* (Ω, Σ, μ) *un espacio de probabilidad. Una familia* $\{f_\alpha\}_{\alpha\in\mathcal{A}}$*de funciones* Σ*-medibles tal que* $f_\alpha : \Omega \longrightarrow [-\infty, \infty]$ *para cada* $\alpha \in \mathcal{A}$*, es independiente si la familia de* σ*-álgebras* $\{\sigma(f_\alpha)\}_{\alpha\in\mathcal{A}}$ *es independiente.*

Definición 1.1.8. *Sean* (Ω, Σ, μ) *un espacio de probabilidad y* $\{f_\alpha\}_{\alpha\in\mathcal{A}}$ *una familia de funciones* Σ*-medibles tal que* $f_\alpha : \Omega \longrightarrow [-\infty, \infty]$ *para cada* $\alpha \in \mathcal{A}$*. Las funciones* f_α *son idénticamente distribuidas si para cada par* $(\alpha, \beta) \in \mathcal{A}\times\mathcal{A}$ *y cada* $B \in \mathcal{B}(\mathbb{R})$ *se cumple*

$$\mu(\{\omega \in \Omega : f_\alpha(\omega) \in B\}) = \mu(\{\omega \in \Omega : f_\beta(\omega) \in B\}).$$

1.1.2. Integración

Teorema 1.1.1. (Desigualdad de Chebyshev) *Sea* $f : \Omega \longrightarrow [0, \infty]$ *una función* Σ*-medible y* $A_t = \{\omega \in \Omega : |f(\omega)| \geq t\}$ *con* $t > 0$*, entonces*

$$\mu(A_t) \leq \frac{1}{t}\int_{A_t} f\ d\mu \leq \frac{1}{t}\int f\ d\mu.$$

Demostración: (Véase Proposición 2.3.9 en Cohn [2]) ■

Corolario 1.1.2. *Sea* $f : \Omega \longrightarrow [0, \infty]$ *una función* Σ*-medible.*

$$\int f\ d\mu = 0 \Longrightarrow f(\omega) = 0 \ \ \textit{en c.t.p.}$$

Demostración: (Véase Corolario 2.3.11 en Cohn [2]) ■

Corolario 1.1.3. *Sea* $f : \Omega \longrightarrow [-\infty, \infty]$ *una función* Σ*-medible.*

$$\int |f|\ d\mu < \infty \Longrightarrow |f(\omega)| < \infty \ \ \textit{en c.t.p.}$$

Demostración: (Véase Corolario 2.3.12 en Cohn [2]) ■

Denotaremos por $\mathcal{L}^p(\Omega, \Sigma, \mu, \mathbb{R})$, el espacio vectorial de las funciones Σ-medibles $f : \Omega \longrightarrow \mathbb{R}$ tales que $\int |f|^p\ d\mu < \infty$ con $1 \leq p < \infty$. El espacio $\mathcal{L}^p(\Omega, \Sigma, \mu, \mathbb{C})$ designará el espacio vectorial complejo de las funciones (complejas) Σ-medibles tales que $\int |f|^p\ d\mu < \infty$. Consideremos el subespacio vectorial $\mathcal{N}^p(\Omega, \Sigma, \mu, \mathbb{R}) = \{f \in \mathcal{L}^p(\Omega, \Sigma, \mu, \mathbb{R}) : f(\omega) = 0 \ \ \textit{en c.t.p}\}$, el espacio vectorial $L^p(\Omega, \Sigma, \mu, \mathbb{R})$ es el espacio cociente $\mathcal{L}^p(\Omega, \Sigma, \mu, \mathbb{R})/\mathcal{N}^p(\Omega, \Sigma, \mu, \mathbb{R})$, es decir el espacio vectorial cuyos elementos son clases de equivalencia $\boldsymbol{f}$ inducidas por la relación de equivalencia $\sim$ tal que $f \sim g \Longleftrightarrow f(\omega) - g(\omega) = 0$ *en c.t.p.* Para distinguir una función de su clase de equivalencia, escribiremos la clase de equivalencia en negrita.

Teorema 1.1.2. *Sean* $f, g \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$.

(i)

$$\forall\ E \in \Sigma \quad \int_E f\ d\mu \leq \int_E g\ d\mu \Longrightarrow f(\omega) \leq g(\omega) \quad \textit{en c.t.p.}$$

(ii)

$$\forall\ E \in \Sigma \quad \int_E f\ d\mu = \int_E g\ d\mu \Longrightarrow f(\omega) = g(\omega) \quad \textit{en c.t.p.}$$

Demostración:

(i) Sea $A = \{\omega \in \Omega : f(\omega) > g(\omega)\}$. Como $A \in \Sigma, \int_A f\ d\mu \leq \int_A g\ d\mu$. Luego $\int_A (f-g)\ d\mu = \int_A (f-g)^+\ d\mu - \int_A (f-g)^-\ d\mu \leq 0$. Es claro que $\int_A (f-g)^-\ d\mu = 0$, entonces

$$0 \leq \int_A (f-g)^+\ d\mu \leq 0 \Longrightarrow \int (f-g)^+ \chi_A\ d\mu = 0.$$

Por el corolario 1.1.2, $(f(\omega) - g(\omega))^+ \chi_A(\omega) = 0$ *en c.t.p*, luego $\mu(A) = 0$.

(ii) Si para cada $E \in \Sigma$ se cumple $\int_E f\ d\mu = \int_E g\ d\mu$, entonces se cumplen $\int_E f\ d\mu \leq \int_E g\ d\mu$ y $\int_E g\ d\mu \leq \int_E f\ d\mu$. Entonces por (i) se verifican

$$f(\omega) \leq g(\omega) \text{ y } g(\omega) \leq f(\omega) \quad \textit{en c.t.p,}$$

con lo cual $f(\omega) = g(\omega)$ *en c.t.p.*

■

En lo que sigue (Ω, Σ, μ) es un espacio de medida.

Teorema 1.1.3. (Convergencia Monótona) *Sea* $\{f_n\}$ *una sucesión creciente de funciones* Σ*-medibles tal que* $f_n : \Omega \longrightarrow [0, \infty]\ \forall\ n \in \mathbb{N}$. *Entonces se verifica*

$$\int \lim_{n\to\infty} f_n\ d\mu = \lim_{n\to\infty} \int f_n\ d\mu.$$

Demostración: (Véase Teorema 2.4.1 en Cohn [2]) ■

Notamos que el teorema 1.1.3 sigue siendo cierto si $\{f_n\}$ es creciente *en c.t.p*, la función $\lim_{n\to\infty} f_n$ sólo está definida *en c.t.p*, precisamente en el conjunto $\{\omega \in \Omega : m \leq n \Rightarrow f_m(\omega) \leq f_n(\omega)\}$.

Corolario 1.1.4. (Teorema de Beppo-Levi) *Sea $\sum_{k=1}^{\infty} f_k$ una serie de funciones Σ-medibles tal que $f_k : \Omega \longrightarrow [0,\infty]$ $\forall$ $k \in \mathbb{N}$. Entonces la suma $\sum_{k=1}^{\infty} f_k$ es Σ-medible y*

$$\int \sum_{k=1}^{\infty} f_k \; d\mu = \sum_{k=1}^{\infty} \int f_k \; d\mu.$$

Demostración: (Véase Corolario 2.4.2 en Cohn [2]) ■

Teorema 1.1.4. (Lema de Fatou) *Sea $\{f_n\}$ una sucesión de funciones Σ-medibles tal que $f_n : \Omega \longrightarrow [0,\infty]$ $\forall$ $n \in \mathbb{N}$. Entonces se cumple*

$$\int \liminf_{n} f_n \; d\mu \leq \liminf_{n} \int f_n \; d\mu.$$

Demostración: (Véase Teorema 2.4.3 en Cohn [2]) ■

Teorema 1.1.5. (Convergencia Dominada) *Sea $\{f_n\}$ una sucesión de funciones Σ-medibles tal que $f_n : \Omega \longrightarrow [-\infty,\infty]$ para cada $n \in \mathbb{N}$. Supongamos que exista $g \in \mathcal{L}^1(\Omega,\Sigma,\mu,\mathbb{R})$ tal que para cada $n \in \mathbb{N}$, $|f_n(\omega)| \leq g(\omega)$ en c.t.p . Si $\{f_n\}$ converge en c.t.p hacia la función Σ-medible $f : \Omega \longrightarrow [0,\infty]$, entonces $\{f_n\}\cup\{f\} \subset \mathcal{L}^1(\Omega,\Sigma,\mu,\mathbb{R})$ y*

$$\lim_{n\to\infty} \int f_n \; d\mu = \int f \; d\mu.$$

Demostración: (Véase Teorema 2.4.4 en Cohn [2]) ■

Teorema 1.1.6. (Desigualdad de Jensen) *Sea (Ω,Σ,μ) un espacio de probabilidad y $I \subset \mathbb{R}$ un intervalo abierto. Si $f \in \mathcal{L}^1(\Omega,\Sigma,\mu,\mathbb{R})$ con $f(\Omega) \subset I$ y si $\varphi : I \longrightarrow \mathbb{R}$ es convexa, entonces*

$$\varphi\left(\int f \; d\mu\right) \leq \int (\varphi \circ f) \; d\mu.$$

Demostración: (Véase Teorema 3.1.3 en Rudin [4]) ■

Teorema 1.1.7. *Sea* (Ω, Σ, μ) *un espacio de probabilidad y* $\{f_i\}_{i=1}^n \subset \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$, *una familia de funciones integrables independientes. Entonces* $\prod_{i=1}^n f_i \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ *y además,*

$$\int \prod_{i=1}^n f_i \; d\mu = \prod_{i=1}^n \int f_i \; d\mu.$$

Demostración: (Véase Teorema 5.10.8 en Ash [1]) ■

1.2. Modos de Convergencia

En esta sección definiremos y estudiaremos con detalle distintos modos de convergencia de sucesiones de funciones medibles. Los resultados enunciados se pueden encontrar en la sección 3.1 de Cohn [2].

Sea (Ω, Σ, μ) un espacio de medida, $\{f_n\}$ una sucesión de funciones Σ-medibles tal que $f_n : \Omega \longrightarrow \mathbb{R}$ para cada $n \in \mathbb{N}$ y $f : \Omega \longrightarrow \mathbb{R}$ también Σ-medible.

Definición 1.2.1. *La sucesión* $\{f_n\}$ *converge hacia* f *en casi todo punto (en c.t.p) si*

$$f(\omega) = \lim_{n\to\infty} f_n(\omega) \; \forall \; \omega \in \Omega \setminus N \quad \textit{con } \mu(N) = 0.$$

Definición 1.2.2. *La sucesión* $\{f_n\}$ *converge hacia* f *en medida si* $\forall \; k \in \mathbb{R}^+$

$$\lim_{n\to\infty} \mu\left(\{\omega \in \Omega : |f_n(\omega) - f(\omega)| > k\}\right) = 0.$$

Definición 1.2.3. *La sucesión* $\{f_n\}$ *converge casi uniformemente hacia* f *si* $\forall \; \epsilon > 0 \; \exists \; \Omega_\epsilon \in \Sigma$ *tal que* $\mu(\Omega \setminus \Omega_\epsilon) < \epsilon$ *y* $\{f_n\}$ *converge hacia* f *uniformemente sobre* Ω_ϵ.

Proposición 1.2.1. *Si* $\mu(\Omega) < \infty$ *y si* $\{f_n\}$ *converge hacia* f *en c.t.p. Entonces* $\{f_n\}$ *converge hacia* f *en medida.*

Demostración: Dado $k > 0$, definimos $A_n = \{\omega \in \Omega : |f_n(\omega) - f(\omega)| > k\}$ y $B_n = \bigcup_{i=n}^\infty A_i$. Es claro que B_n es una sucesión decreciente de conjuntos, es decir $B_{n+1} \subset B_n \; \forall \; n \in \mathbb{N}$. Sea $N = \{\omega \in \Omega : \lim_{n\to\infty} f_n(\omega) \neq f(\omega)\} \cup \{\omega \in \Omega : \liminf_n f_n(\omega) < \limsup_n f_n(\omega)\}$, notamos que $\mu(N) = 0$ por hypótesis.

$$\bigcap_{i=1}^\infty B_i = \bigcap_{i=1}^\infty \bigcup_{j=i}^\infty A_j.$$

Supongamos que $\omega \in \bigcap_{i=1}^{\infty} B_i$, eso implica que $\omega \in B_i$ para todo i. Para cada $i \in \mathbb{N}$, existe $j > i$ tal que $|f_j(\omega) - f(\omega)| > k$, luego $\omega \in N$. Entonces como $\bigcap_{i=1}^{\infty} B_i \subset N$, $\mu\left(\bigcap_{i=1}^{\infty} B_i\right) = 0$. Como μ es finita, $\mu(B_i) < \infty$ para cada $i \in \mathbb{N}$, lo cual implica que $0 = \mu\left(\bigcap_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} \mu(B_n)$. Como $A_n \subset B_n$, tenemos

$$0 \leq \lim_{n\to\infty} \mu(A_n) \leq \lim_{n\to\infty} \mu(B_n) = 0.$$

Entonces $\{f_n\}$ converge hacia f en medida. ■

Proposición 1.2.2. *Si $\{f_n\}$ converge en medida hacia f, entonces existe una subsucesión $\{f_{n_k}\}$ que converge hacia f en c.t.p.*

Demostración: Por hypótesis, $\lim_{n\to\infty} \mu\{\omega \in \Omega : |f_n(\omega) - f(\omega)| > t\} = 0 \ \forall\ t \in \mathbb{R}^+$. $\exists\ n_1 \in \mathbb{N}$ tal que $\mu\{\omega \in \Omega : |f_{n_1}(\omega) - f(\omega)| > 1\} \leq \frac{1}{2}$
$\exists\ n_2 > n_1 \in \mathbb{N}$ tal que $\mu\{\omega \in \Omega : |f_{n_2}(\omega) - f(\omega)| > \frac{1}{2}\} \leq \frac{1}{2^2}$. De forma inductiva, construimos una sucesión creciente $\{n_k\} \in \mathbb{N}$ tal que

$$\mu\{\omega \in \Omega : |f_{n_k}(\omega) - f(\omega)| > \frac{1}{k}\} \leq \frac{1}{2^k}.$$

Definimos $A_k = \{\omega \in \Omega : |f_{n_k}(\omega) - f(\omega)| > \frac{1}{k}\}$ y $B_j = \bigcup_{k=j}^{\infty} A_k$. Si $\omega \notin \bigcap_{j=1}^{\infty} \bigcup_{k=j}^{\infty} A_k = \bigcap_{j=1}^{\infty} B_j$, entonces existe $j \in \mathbb{N}$ tal que $\omega \notin A_k\ \forall\ k > j$, lo cual implica que

$$\lim_{n\to\infty} f_{n_k}(\omega) = f(\omega).$$

Entonces f_{n_k} converge puntualmente hacia f para todo $\omega \in \Omega \setminus \left(\bigcap_{j=1}^{\infty} B_j\right)$. Claramente $\bigcap_{j=1}^{\infty} B_j \subset \bigcup_{k=j}^{\infty} A_k$.

$$\mu\left(\bigcup_{k=j}^{\infty} A_k\right) \leq \sum_{k=j}^{\infty} \mu(A_k) \leq \sum_{k=j}^{\infty} \frac{1}{2^k} = \frac{1}{2^{j-1}}\ \forall\ j \in \mathbb{N}$$

$$\mu(B_j) \leq \frac{1}{2^{j-1}}\ \forall\ j \in \mathbb{N}.$$

$$\mu\left(\bigcap_{j=1}^{\infty} B_j\right) \leq \mu(B_j) \leq \frac{1}{2^{j-1}}\ \forall\ j \in \mathbb{N}.$$

Entonces $\mu\left(\bigcap_{j=1}^{\infty} B_j\right) = 0$ y $\{f_{n_k}\}$ converge hacia f *en c.t.p.* ■

Proposición 1.2.3. (Teorema de Egoroff) *Si $\mu(\Omega) < \infty$ y si $\{f_n\}$ converge hacia f en c.t.p, entonces $\{f_n\}$ converge hacia f casi uniformemente.*

Demostración: Por hypótesis $\lim_{n\to\infty} |f_n(\omega) - f(\omega)| = 0 \ \forall\ \omega \in \Omega \setminus N$ con $\mu(N) = 0$. Sea $\epsilon > 0$ y $g_n(\omega) = \sup_{j\geq n} |f_j(\omega) - f(\omega)|$. Claro que $g_n \searrow \limsup_n |f_n(\omega) - f(\omega)| = 0$ *en c.t.p.* Por proposición 1.2.1 $\{g_n\}$ converge hacia 0 en medida. Para cada $k \in \mathbb{N}$, existe $n_k \in \mathbb{N}$ tal que

$$\mu\left(\{\omega \in \Omega : g_{n_k}(\omega) > \frac{1}{k}\}\right) < \frac{\epsilon}{2^k}.$$

Definimos $B_k = \{\omega \in \Omega : g_{n_k}(\omega) \leq \frac{1}{k}\} \ \forall\ k \in \mathbb{N}$ y $B = \bigcap_{k=1}^{\infty} B_k$.

$$\mu(B^c) = \mu\left(\left(\bigcap_{k=1}^{\infty} B_k\right)^c\right) = \mu\left(\bigcup_{k=1}^{\infty} B_k^c\right) \leq \sum_{k=1}^{\infty} \mu(B_k^c) < \sum_{k=1}^{\infty} \frac{\epsilon}{2^k} = \epsilon.$$

Dado $\delta > 0$ eligimos $k \in \mathbb{N}$ tal que $\frac{1}{k} < \delta$, tenemos para todo $n > n_k$ y para cada $\omega \in B \subset B_k$

$$|f_n(\omega) - f(\omega)| \leq \sup_{j\geq n_k} |f_j(\omega) - f(\omega)| = g_{n_k}(\omega) \leq \frac{1}{k} < \delta.$$

Entonces $\{f_n\}$ converge hacia f uniformemente en B. Queda demostrado que $\{f_n\}$ converge hacia f casi uniformemente. ■

Sea $\{f_n\}$ una sucesión tal que $f_n \in \mathcal{L}^p(\Omega, \Sigma, \mu, \mathbb{R}) \ \forall\ n \in \mathbb{N}$ y $f \in \mathcal{L}^p(\Omega, \Sigma, \mu, \mathbb{R})$, con $p \geq 1$.

Definición 1.2.4. *La sucesión $\{f_n\}$ converge hacia f en media de orden p si*

$$\lim_{n\to\infty} \int |f_n(\omega) - f(\omega)|^p \, d\mu = 0.$$

A la convergencia en media de orden 1 se llama convergencia en media.

Proposición 1.2.4. *Si $\{f_n\}$ converge hacia f en media de orden p, entonces $\{f_n\}$ converge hacia f en medida.*

Demostración: Sea $A_t = \{\omega \in \Omega : |f_n(\omega) - f(\omega)|^p > t\}$. Por la desigualdad de Chebyshev

$$\mu(A_t) \leq \frac{1}{t} \int_{A_t} |f_n(\omega) - f(\omega)|^p \, d\mu \leq \frac{1}{t} \int |f_n(\omega) - f(\omega)|^p \, d\mu.$$

$$\lim_{n\to\infty} \mu\left(\{\omega \in \Omega : |f_n(\omega) - f(\omega)|^p > t\}\right) \leq \lim_{n\to\infty} \frac{1}{t} \int |f_n(\omega) - f(\omega)|^p \, d\mu = 0.$$

■

Proposición 1.2.5. *Si $\{f_n\}$ converge hacia f en c.t.p o en medida, y si existe una función no negativa $g : \Omega \longrightarrow [-\infty, \infty]$ tal que $g \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ que satisface para cada $n \in \mathbb{N}$ $|f_n(\omega)| \leq g(\omega)$ en c.t.p, entonces $\{f_n\}$ converge hacia f en media.*

Demostración: Supongamos en primer lugar que $\{f_n\}$ converja hacia f *en c.t.p.* Es claro que $|f_n(\omega) - f(\omega)| \longrightarrow 0$ *en c.t.p.* Por hypótesis y la desigualdad triangular $|f_n(\omega) - f(\omega)| \leq |f_n(\omega)| + |f(\omega)| \leq 2g(\omega)$ para cada $n \in \mathbb{N}$ *en c.t.p.* Por el teorema de la convergencia dominada $\int |f_n(\omega) - f(\omega)| \; d\mu \longrightarrow 0$. Ahora supongamos que $\{f_n\}$ converja hacia f en medida. Toda subsucesión $\{f_{n_k}\}$ también converge hacia f en medida simplemente porque toda subsucesión de una sucesión convergente de números reales converge al mismo límite. Entonces cada $\{f_{n_k}\}$ tiene por la proposición 1.2.2 una subsucesión $\{f_{n_{k_j}}\}$ que converge hacia f *en c.t.p*, lo cual implica por lo que demostramos en primer lugar que $f_{n_{k_j}} \longrightarrow f$ en media. Supongamos que $\{f_n\}$ no converja hacia f en media. Entonces existe $\epsilon > 0$ y un conjunto infinito $M \subset \mathbb{N}$ tal que $\int |f_m(\omega) - f(\omega)| \; d\mu \geq \epsilon$ para cada $m \in M$. La subsucesión $\{f_m\}_{m \in M}$ converge en medida hacia f y en virtud de la proposición 1.2.2 posee una subsucesión $\{f_{m_i}\}$ que converge hacia f *en c.t.p.* Entonces

$$\lim_{i \to \infty} \int |f_{m_i}(\omega) - f(\omega)| \; d\mu = 0,$$

pero esto es una contradicción. Queda demostrado que $\{f_n\}$ converge hacia f en media. ■

1.3. Integrabilidad Uniforme

El concepto de integrabilidad uniforme va a complementar la sección 1.2 y permitirnos demostrar el teorema de convergencia de Vitali que concluirá esta sección. Estos resultados se pueden encontrar en el capítulo 13 de Williams [6].

Proposición 1.3.1. *Sea $E_t(f) = \{\omega \in \Omega : |f(\omega)| \geq t\}$ con $t \in \mathbb{R}$. Cuando esté claro el contexto escribiremos simplemente E_t. Si $f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$, entonces*

$$\lim_{t \to \infty} \int_{E_t} |f| \; d\mu = 0.$$

Demostración: Sea $f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$. Consideremos la sucesión $\{E_n\}$. Es claro que $\{E_n\}$ es decreciente y $\bigcap_{n=1}^{\infty} E_n = \{\omega \in \Omega : |f(\omega)| \geq \infty\}$. Como f es integrable, $|f(\omega)| < \infty$ *en c.t.p*, lo cual implica que

$$\mu\left(\bigcap_{n=1}^{\infty} E_n\right) = \mu\left(\{\omega \in \Omega : |f(\omega)| \geq \infty\}\right) = 0.$$

Notamos que $|f(\omega)|\chi_{E_n} \leq |f(\omega)|$ para todo $\omega \in \Omega$ y que

$$\lim_{n\to\infty} |f|\chi_{E_n} = |f|\chi_{\cap_n E_n} = 0 \quad en \quad c.t.p,$$

entonces por el teorema de la convergencia dominada,

$$\lim_{t\to\infty} \int_{E_t} |f| \; d\mu = \lim_{n\to\infty} \int_{E_n} |f| \; d\mu = \lim_{n\to\infty} \int_{\cap_n E_n} |f| \; d\mu = 0.$$

■

Proposición 1.3.2. *Sea $f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$. Dado $\epsilon > 0$, existe $\delta > 0$ tal que*

$$[E \in \Sigma, \mu(E) < \delta] \Longrightarrow \int_E |f| \; d\mu < \epsilon.$$

Demostración: Supongamos que para algun $\epsilon > 0$ y cada $\delta > 0$ hay un conjunto $E \in \Sigma$ verificando $\mu(E) < \delta$ y $\int_E |f| \; d\mu \geq \epsilon$. En particular, tomando $\delta = \frac{1}{2^n}$, podemos encontrar una sucesión $\{E_n\} \subset \Sigma$ tal que $\mu(E_n) < \frac{1}{2^n}$ y $\int_{E_n} |f| \; d\mu \geq \epsilon$ para cada $n \in \mathbb{N}$. Definamos $F = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k$. Es claro que

$$\mu(F) \leq \mu\left(\bigcup_{k=n}^{\infty} E_k\right) \leq \sum_{k=n}^{\infty} \mu(E_k) \; \forall n \in \mathbb{N},$$

luego $\mu(F) = 0 = \int_F |f| \; d\mu$. Como

$$|f|\chi_{\cup_{k\geq n} E_k}(\omega) \leq |f(\omega)|$$

para cada $n \in \mathbb{N}$ y

$$\lim_{n\to\infty} |f|\chi_{\cup_{k\geq n} E_k} = |f|\chi_F,$$

por el teorema de la convergencia dominada,

$$0 = \int |f|\chi_F \; d\mu = \lim_{n\to\infty} \int |f|\chi_{\cup_{k\geq n} E_k} \; d\mu.$$

Notamos que para cada $n \in \mathbb{N}$, se verifica

$$\int_{\cup_{k\geq n} E_k} |f| \; d\mu \geq \int_{E_n} |f| \; d\mu \geq \epsilon,$$

luego

$$\lim_{n\to\infty} \int_{\cup_{k\geq n} E_k} |f| \; d\mu \geq \epsilon,$$

lo cual es una contradicción. ■

Definición 1.3.1. *Sea $\mathcal{F} \subset \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$. La familia $\mathcal{F}$ es uniformemente integrable si*

$$\lim_{t\to\infty} \sup_{f\in\mathcal{F}} \left\{ \int_{E_t} |f| \; d\mu \right\} = 0,$$

donde $E_t(f) = \{\omega \in \Omega : |f(\omega)| \geq t\}$ con $t \in \mathbb{R}$.

Teorema 1.3.1. *Sea (Ω, Σ, μ) un espacio de medida finito. La familia $\mathcal{F} \subset \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ es uniformemente integrable si y sólo si se cumplen*

(i) $\forall\ \epsilon > 0,\ \exists\ \delta > 0$ *tal que* $[E \in \Sigma, \mu(E) < \delta] \Longrightarrow \sup_{f\in\mathcal{F}} \left\{ \int_E |f| \; d\mu \right\} < \epsilon,$

(ii) $\sup_{f\in\mathcal{F}} \left\{ \int |f| \; d\mu \right\} < \infty.$

Demostración: Para cada $f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$, cada $E \in \Sigma$ y cada $t > 0$, se cumple

$$\int |f|\chi_E \; d\mu = \int |f|\chi_{E\cap E_t} \; d\mu + \int |f|\chi_{E\cap E_t^c} \; d\mu,$$

donde $E_t(f) = \{\omega \in \Omega : |f(\omega)| \geq t\}$ con $t \in \mathbb{R}$. Es claro que $\int |f|\chi_{E\cap E_t} \; d\mu \leq \int |f|\chi_{E_t} \; d\mu$ y que $\int |f|\chi_{E\cap E_t^c} \; d\mu \leq \int t\chi_E \; d\mu$, entonces

$$\int_E |f| \; d\mu \leq t\mu(E) + \int_{E_t} |f| \; d\mu \qquad (*)$$

Supongamos que $\mathcal{F}$ sea uniformemente integrable. Dado $\epsilon > 0$, $\exists\ t > 0$ tal que para cada $f \in \mathcal{F}$ se cumple $\int_{E_t} |f| \; d\mu \leq \frac{\epsilon}{2}$. Sea $\delta = \frac{\epsilon}{2t}$, y supongamos que $\mu(E) < \delta$, entonces si $f \in \mathcal{F}$, por $(*)$ se cumple,

$$\int_E |f| \; d\mu \leq t\mu(E) + \int_{E_t} |f| \; d\mu \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Eso demuestra el apartado (i). Para (ii) vemos que para cada $f \in \mathcal{F}$, se cumple

$$\int |f| \; d\mu \leq t\mu(\Omega) + \int_{E_t} |f| \; d\mu \leq t\mu(\Omega) + \frac{\epsilon}{2} < \infty.$$

Para demostrar el recíproco supongamos que se cumplan (i) y (ii). Dado $\epsilon > 0$, sea $\delta > 0$ el dado por (i). Sean $M = \sup_{f\in\mathcal{F}} \left\{ \int |f| \; d\mu \right\}$ y $t > \frac{M}{\delta}$. Para cada $f \in \mathcal{F}$ tenemos por la desigualdad de Chebyshev

$$\mu(E_t) \leq \frac{1}{t} \int_{E_t} |f| \; d\mu \leq \frac{1}{t} \int |f| \; d\mu \leq \frac{M}{t} < \delta.$$

Como $E_t \in \Sigma$ y $\mu(E_t) < \delta$, para cada $t > \frac{M}{\delta}$ se cumple por (i)

$$\sup_{f\in\mathcal{F}} \left\{ \int_{E_t} |f| \; d\mu \right\} < \epsilon.$$

■

Teorema 1.3.2. (Teorema de Vitali) *Sea* (Ω, Σ, μ) *un espacio de medida finito y* $\{f_n\} \subset \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ *una sucesión uniformemente integrable que converge en medida hacia la función* Σ*-medible* f*. Entonces* $f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ *y*

$$\lim_{n\to\infty} \int |f_n(\omega) - f(\omega)|\, d\mu = 0.$$

Demostración: Supongamos que $f_n \longrightarrow f$ *en c.t.p.* Como $\mu(\Omega) < \infty$, $f_n \longrightarrow f$ en medida. Dado $\epsilon > 0$, por el teorema 1.3.1, existe $\delta > 0$ tal que si $E \in \Sigma$ y $\mu(E) < \delta$ se cumple $\int_E |f_n|\, d\mu < \epsilon$ para todo $n \in \mathbb{N}$. Por el teorema de Egoroff, existe $A \in \Sigma$ tal que $\mu(\Omega \setminus A) < \delta$ y

$$\alpha_n = \sup_{\omega\in A}\{|f_n(\omega) - f(\omega)|\} \longrightarrow 0,$$

es decir $f_n \longrightarrow f$ uniformemente en A. Como $\Omega \setminus A \in \Sigma$ y $\mu(\Omega \setminus A) < \delta$, tenemos $\int_{\Omega\setminus A} |f_n|\, d\mu < \epsilon$ para todo $n \in \mathbb{N}$. Por la continuidad de $|\cdot| : \mathbb{R} \longrightarrow \mathbb{R}$, $|f_n| \longrightarrow |f|$ *en c.t.p.* Por el lema de Fatou,

$$\int_{\Omega\setminus A} |f|\, d\mu \leq \liminf_n \int_{\Omega\setminus A} |f_n|\, d\mu \leq \sup_{n\in\mathbb{N}} \int_{\Omega\setminus A} |f_n|\, d\mu \leq \epsilon.$$

Entonces

$$\int |f_n - f|\, d\mu = \int_A |f_n - f|\, d\mu + \int_{A^c} |f_n - f|\, d\mu$$

$$\leq \int \alpha_n \chi_A\, d\mu + \int_{A^c} |f_n|\, d\mu + \int_{A^c} |f|\, d\mu \leq \alpha_n \mu(A) + \epsilon + \epsilon.$$

Teniendo en cuenta que $\alpha_n \longrightarrow 0$, se cumple

$$\limsup_n \int |f_n - f|\, d\mu \leq 2\epsilon.$$

Como $\epsilon > 0$ es arbitrario se concluye

$$\lim_{n\to\infty} \int |f_n - f|\, d\mu = 0.$$

Ahora supongamos que $f_n \longrightarrow f$ en medida. Cada subsucesión $\{f_{n_k}\}$ converge hacia f en medida. Supongamos que $\{f_n\}$ no converja en media hacia f. Entonces existe $\epsilon > 0$ y un conjunto infinito $M \subset \mathbb{N}$ tal que $\int |f_m(\omega) - f(\omega)|\, d\mu \geq \epsilon$ para cada $m \in M$. La subsucesión $\{f_m\}_{m\in M}$ converge en medida hacia f. Por la proposición 1.2.2, $\{f_m\}_{m\in M}$ posee una una subsucesión $\{f_{m_i}\}$ que converge hacia f *en c.t.p* y por lo que hemos demostrado se verifica

$$\lim_{i\to\infty} \int |f_{m_i}(\omega) - f(\omega)|\, d\mu = 0.$$

Esto es una contradicción. Queda demostrado que $\{f_n\}$ converge hacia f en media. El hecho de que $f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ es consecuencia de que dado $\lambda > 0$, existe $N > 0$ tal que se cumple la desigualdad

$$\int |f| \, d\mu \leq \int |f - f_N| \, d\mu + \int |f_N| \, d\mu < \lambda + \int |f_N| \, d\mu < \infty.$$

■

1.4. Espacio de Hilbert

En esta sección vamos a introducir nociones elementales de la teoría de los espacios de Hilbert. El objetivo es demostrar el teorema de la proyección ortogonal. En el siguiente capítulo usaremos el teorema de la proyección ortogonal para demostrar la existencia de la esperanza condicionada en $\mathcal{L}^2$. Suponemos conocida la noción de espacio de Banach. Las fuentes principales son Young [7] y Rudin [4].

1.4.1. Teorema de Mejor Aproximación

Definición 1.4.1. *Sea X un espacio vectorial sobre el cuerpo $\mathbb{C}$. Un producto escalar sobre X es una aplicación $\langle \cdot, \cdot \rangle : X \times X \longrightarrow \mathbb{C}$ tal que para cualesquiera $x, y, z \in X$ y cualquier $k \in \mathbb{C}$ se verifican las siguientes propriedades:*

(i) $\langle x, y \rangle = \overline{\langle y, x \rangle}$,

(ii) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$,

(iii) $\langle kx, y \rangle = k\langle x, y \rangle$,

(iv) $\langle x, x \rangle \geq 0$,

(v) $\langle x, x \rangle = 0 \Longleftrightarrow x = 0$.

Definición 1.4.2. *Sean X un espacio vectorial sobre el cuerpo $\mathbb{C}$ y $\langle \cdot, \cdot \rangle$ un producto escalar sobre X. Se llama espacio prehilbertiano al par $(X, \langle \cdot, \cdot \rangle)$.*

Sea $(X, \langle \cdot, \cdot \rangle)$ un espacio prehilbertiano. La función $\| \cdot \| : X \longrightarrow \mathbb{R}^+$ con fórmula $\|x\| = \sqrt{\langle x, x \rangle}$ define una norma sobre X y se llama norma asociada al producto escalar $\langle \cdot, \cdot \rangle$. Es fácil ver que $d : X \times X \longrightarrow \mathbb{R}^+$ con fórmula $d(x, y) = \|x -$

$y\| = \sqrt{\langle x-y, x-y\rangle}$ define una métrica sobre X. Así que todo espacio hilbertiano $(X, \langle\cdot,\cdot\rangle)$ es también un espacio métrico (X, d) donde d es la métrica inducida por $\langle\cdot,\cdot\rangle$. En todo espacio prehilbertiano $(X, \langle\cdot,\cdot\rangle)$ se verifica la le del paralelogramo:

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \quad \forall\ x, y \in X$$

donde $\|\cdot\|$ es la norma asociada a $\langle\cdot,\cdot\rangle$.

Definición 1.4.3. *Sea $(X, \langle\cdot,\cdot\rangle)$ un espacio prehilbertiano. El espacio $(X, \langle\cdot,\cdot\rangle)$ se llama espacio de Hilbert si el espacio métrico (X, d) donde d es la métrica inducida por $\langle\cdot,\cdot\rangle$ es completo.*

Teorema 1.4.1. (Teorema de Mejor Aproximación) *Sea H un espacio de Hilbert y A un subconjunto convexo cerrado y no vacío de H. Para todo punto $x \in H$, existe un único punto $p \in A$ tal que*

$$\|x-p\| = \inf_{a\in A}\{\|x-a\|\}.$$

Demostración: Sea $x \in H$. Al estar acotado inferiormente, el conjunto $\{\|x-a\| : a \in A\}$ tiene un infimo. Definimos $M = \inf_{a\in A}\{\|x-a\|\}$. Sea $\{x-a_n\} \subset H$ una sucesión tal que para cada $n \in \mathbb{N}$ $a_n \in A$ y $\|x-a_n\| \longrightarrow M$. Por continuidad tenemos $\|x-a_n\|^2 \longrightarrow M^2$. Entonces dado $\epsilon > 0$, existe $N > 0$ tal que

$$n > N \Longrightarrow \|x-a_n\|^2 - M^2 < \frac{\epsilon}{4} \Longleftrightarrow \|x-a_n\|^2 < \frac{\epsilon}{4} + M^2.$$

Por la ley del paralelogramo, suponiendo que $n, m > N$ se cumple

$$\|x-a_m+x-a_n\|^2 + \|a_m-a_n\|^2 = 2\|x-a_n\|^2 + 2\|x-a_m\|^2$$
$$\|a_m-a_n\|^2 = 2\|x-a_n\|^2 + 2\|x-a_m\|^2 - 4\|x-(\frac{a_n}{2}+\frac{a_m}{2})\|^2.$$

Como A es convexo, $\frac{a_n}{2}+\frac{a_m}{2} \in A$ y por lo tanto $M \leq \|x-(\frac{a_n}{2}+\frac{a_m}{2})\|$. Tenemos

$$\|a_m-a_n\|^2 < 4\left(\frac{\epsilon}{4}+M^2\right) - 4M^2 = \epsilon,$$

lo cual implica que $\{a_n\}$ es una sucesión de Cauchy. Como H es un espacio completo, $a_n \longrightarrow p \in H$. Como A es cerrado $p \in A$, entonces $\|x-p\| = M$. Ahora supongamos que $q \in A$ y que se verifica $\|x-q\| = M$. Por la ley del paralelogramo,

$$\|x-q+x-p\|^2 + \|p-q\|^2 = 2\|x-q\|^2 + 2\|x-p\|^2$$
$$\|p-q\|^2 = 2\|x-q\|^2 + 2\|x-p\|^2 - 4\|x-\frac{p+q}{2}\|^2.$$

Por convexidad, $\frac{p+q}{2} \in A$ y $\|p-q\|^2 \leq 2M^2 + 2M^2 - 4M^2 = 0$, lo cual implica que $p = q$. ■

1.4.2. Ortogonalidad

La noción de ortogonalidad y en particular de proyección ortogonal sólo tiene sentido en un espacio prehilbertiano.

Definición 1.4.4. *Sea* $(X, \langle \cdot, \cdot \rangle)$ *un espacio prehilbertiano. Los vectores* $x, y \in X$ *se dicen ortogonales si* $\langle x, y \rangle = 0$.

Una familia $\{x_\alpha\}_{\alpha \in \mathcal{A}} \subset X \setminus \{\mathbf{0}\}$ se llama sistema ortogonal si $\langle x_\alpha, x_\beta \rangle = 0$ cuando $\alpha \neq \beta$, $\alpha, \beta \in \mathcal{A}$. Si $M \subset X$, entonces el conjunto $M^\perp = \{x \in X : \langle x, y \rangle = 0 \ \forall \ y \in M\}$ se llama complemento ortogonal de M.

Lema 1.4.1. (Teorema de Pytagoras) *Si* $\{x_i\}_{i=1}^n$ *es un sistema ortogonal en un espacio prehilbertiano* $(X, \langle \cdot, \cdot \rangle)$*, entonces*

$$\left\| \sum_{i=1}^n x_i \right\|^2 = \sum_{i=1}^n \|x_i\|^2.$$

Demostración: Usando las propriedades del producto escalar tenemos

$$\left\| \sum_{i=1}^n x_i \right\|^2 = \left\langle \sum_{i=1}^n x_i, \sum_{j=1}^n x_j \right\rangle = \sum_{i=1}^n \left\langle x_i, \sum_{j=1}^n x_j \right\rangle = \sum_{i=1}^n \sum_{j=1}^n \langle x_i, x_j \rangle = \sum_{i=1}^n \|x_i\|^2.$$

■

Lema 1.4.2. *Si* M *es un subespacio vectorial de un espacio prehilbertiano* $(X, \langle \cdot, \cdot \rangle)$ *y* $x \in X$*, entonces*

$$x \in M^\perp \Longleftrightarrow \|x - y\| \geq \|x\| \ \forall \ y \in M.$$

Demostración: (Véase Lema 4.23 en Young [7]) ■

Teorema 1.4.2. *Sea* M *un subespacio cerrado de un espacio de Hilbert* H *con* $x \in H$*. Existe* $y \in M$ *y* $z \in M^\perp$ *tal que* $x = y + z$.

Demostración: Sea $x \in H$. Es claro que M es convexo y cerrado, entonces existe un vector de mejor aproximación $y \in M$, es decir $\|x - y\| = \inf_{m \in M}\{\|x - m\|\}$. Sea $z = x - y$, tenemos $x = y + z$.

$$\|z\| = \|x - y\| \leq \|x - (y + m)\| = \|z - m\| \ \forall \ m \in M.$$

Por lema 1.4.2, $z \in M^\perp$. ■

Corolario 1.4.1. (Teorema de la Proyección Ortogonal) *Sea M un subespacio cerrado de un espacio de Hilbert H. Existe el operador proyección $P : H \longrightarrow M$ tal que $x - P(x) \in M^{\perp}$ para todo $x \in H$, donde $P(x)$ es vector de M que mejor aproxima a x.*

Demostración: Por el teorema anterior, cada $x \in H$ se puede descomponer de tal manera que $x = y + z$ donde y es el único vector de M que mejor aproxima el vector x, y $z \in M^{\perp}$. Sea $P : H \longrightarrow M$ tal que $P(x) = y$. Es claro que $x - P(x) = z \in M^{\perp}$. ■

Es fácil ver que $x \in M \Longrightarrow P(x) = x$ y $x \in M^{\perp} \Longrightarrow P(x) = 0$.

Capítulo 2

Esperanza Condicionada

En este capítulo vamos a dar la definición de esperanza condicionada y demostrar sus propriedades elementales. Los conceptos que se utilizan aquí se pueden encontrar en Williams [6], Ash [1] y Vera [5].

2.1. Probabilidad Condicionada

Para motivar la definición abstracta de esperanza condicionada que se dará más adelante, vamos a estudiar dos casos particulares. Empezamos primero con un ejemplo sencillo de probabilidad condicionada. Supongamos que hemos lanzado dos dados los ojos cerrados. La probabilidad de obtener un dos es igual a $\frac{1}{36}$. Abriendo un ojo descubrimos que el resultado del primer dado es uno, pero todavía no hemos visto el segundo dado. Teniendo en cuenta esa información nueva, la probabilidad de obtener un dos cambia y es igual a $\frac{1}{6}$. Si A es el suceso $\{\text{primer dado} = 1\}$ y B el suceso $\{\text{suma de dados} = 2\}$, acabamos de mostrar que la probabilidad que occure el suceso B dado que ha occurido el suceso A es $\frac{1}{6}$.

Sea (Ω, Σ, μ) un espacio de probabilidad. Dados $A, B \in \Sigma$ con $\mu(A) > 0$, la probabilidad del suceso B, condicionada por el suceso A, denotada $\mu(B|A)$ se define así

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

Se puede interpretar el valor $\mu(B|A)$ como una información parcial de un experimento representada por el suceso A. Es decir, el experimento no permite conocer el resultado $\omega \in \Omega$, pero permite saber si $\omega \in A$ ó $\omega \in \Omega \setminus A$. Sea $h_B = \mu(B|A)\chi_A + \mu(B|A^c)\chi_{A^c}$ una función que permite conocer la probabilidad del suceso B condicionada por la información parcial proporcionada por A. El valor $h_B(\omega)$ queda determinado gracias a la información parcial dada por la σ-álgebra $\{\emptyset, A, A^c, \Omega\} \subset \Sigma$. Es claro que h_B es $\{\emptyset, A, A^c, \Omega\}$-medible y se verifica para cada $E \in \{\emptyset, A, A^c, \Omega\}$

$$\int_E h_B \, d\mu = \mu(E \cap B) = \int_E \chi_B \, d\mu.$$

Observemos que aunque no se conozca el valor $\chi_B(\omega)$, conocemos $h_B(\omega)$ y sabemos que χ_B y h_B tienen la misma integral sobre los conjuntos de la σ-álgebra $\{\emptyset, A, A^c, \Omega\} \subset \Sigma$. Se puede considerar h_B como una aproximación de χ_B. Si $\mathcal{P} = \{A_n\}_{n \in M} \subset \Sigma$ es una partición numerable de Ω tal que $\mu(A_n) > 0$ para cada $n \in M$, se puede mejorar la aproximación de χ_B mediante la función $\sigma(\mathcal{P})$-medible

$$h_B = \sum_{n \in M} \mu(B|A_n) \chi_{A_n}.$$

Para cada conjunto $E \in \sigma(\mathcal{P})$ se verifica

$$\int_E h_B \, d\mu = \int_E \sum_{n \in M} \mu(B|A_n) \chi_{A_n} \, d\mu = \sum_{n \in M} \int_E \frac{\mu(B \cap A_n)}{\mu(A_n)} \chi_{A_n} \, d\mu,$$

teniendo en cuenta que $E = \cup_{n \in J} A_n$ con $J \subset M$, es una unión disjunta de elementos de $\mathcal{P}$, se cumple

$$\int_E h_B \, d\mu = \sum_{n \in J} \mu(A_n \cap B) = \mu(E \cap B) = \int_E \chi_B \, d\mu.$$

Consideremos ahora $f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ y $\mathcal{P} = \{A_1, \ldots, A_n\} \subset \Sigma$ una partición finita de Ω. Sea $h = \sum_{i=1}^n \gamma_i \chi_{A_i}$ una función simple tal que

$$\gamma_i = \begin{cases} \frac{1}{\mu(A_i)} \int_{A_i} f \, d\mu & si \quad \mu(A_i) > 0 \\ 0 & si \quad \mu(A_i) = 0. \end{cases}$$

Para cada $E \in \sigma(\mathcal{P})$ se verifica

$$\int_E h \, d\mu = \int_E \sum_{i=1}^n \gamma_i \chi_{A_i} = \sum_{i=1}^n \gamma_i \mu(E \cap A_i),$$

como E es una unión disjunta y finita de elementos de $\sigma(\mathcal{P})$ obtenemos

$$\int_E h \, d\mu = \sum_{i=1}^n \int_{A_i} f \, d\mu = \int_E f \, d\mu.$$

Podemos interpretar $\sigma(\mathcal{P}) \subset \Sigma$ como una información parcial que nos permite aproximar la función Σ-medible e integrable f mediante la función $\sigma(\mathcal{P})$-medible h. La aproximación se puede mejorar refinando la partición $\mathcal{P}$. Una vez que tengamos la noción de esperanza condicionada bien definida, diremos que la función h es una versión de la esperanza de f condicionada por la σ-álgebra $\sigma(\mathcal{P})$. Si consideramos una sucesión de particiones $\{\mathcal{P}_n\}$ tal que para cada $n > 1$, $\mathcal{P}_{n+1}$ es más fina que $\mathcal{P}_n$, entonces la sucesión $\{(h_n, \sigma(\mathcal{P}_n))\}$ donde para cada $n \in \mathbb{N}$, h_n es la esperanza de f condicionada por $\sigma(\mathcal{P}_n)$, es un ejemplo de martingala cuya definición se dará más adelante. Este ejemplo es importante porque da una idea intuitiva de lo que representa la esperanza condicionada, usando esa idea se demostrará el teorema de Radon-Nikodým en el último capítulo.

2.2. Existencia de la Esperanza Condicionada

En esta sección veremos que en el espacio $\mathcal{L}^2(\Omega, \Sigma, \mu, \mathbb{R})$, podemos considerar la esperanza condicionada como una proyección ortogonal. Pero antes necesitamos un resultado previo.

Sea (Ω, Σ, μ) un espacio de medida. Gracias a la desigualdad de Hölder se deduce que la función $\langle \cdot, \cdot \rangle : L^2(\Omega, \Sigma, \mu, \mathbb{C}) \longrightarrow \mathbb{C}$ con fórmula $\langle \boldsymbol{f}, \boldsymbol{g} \rangle = \int f\overline{g}\ d\mu$ define un producto escalar sobre $L^2(\Omega, \Sigma, \mu, \mathbb{C})$. Es conocido que para $p \geq 1$ el espacio $L^p(\Omega, \Sigma, \mu, \mathbb{C})$ con la norma $\|\cdot\|_p$ definida con fórmula

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

es un espacio de Banach. Es fácil ver que $\|\cdot\|_2$ es la norma asociada a $\langle \cdot, \cdot \rangle$, por lo tanto $L^2(\Omega, \Sigma, \mu, \mathbb{C})$ es un espacio de Hilbert.

Teorema 2.2.1. *Sea (Ω, Σ, μ) un espacio de medida. Si Σ_0 es una σ-álgebra tal que $\Sigma_0 \subset \Sigma$, entonces $L^p(\Omega, \Sigma_0, \mu, \mathbb{R})$ se identifica con un subespacio cerrado de $L^p(\Omega, \Sigma, \mu, \mathbb{R})$.*

Demostración: Sea $f \in \mathcal{L}^p(\Omega, \Sigma_0, \mu, \mathbb{R})$. Es claro que $f \in \mathcal{L}^p(\Omega, \Sigma, \mu, \mathbb{R})$ así que

$$\mathcal{L}^p(\Omega, \Sigma_0, \mu, \mathbb{R}) \subset \mathcal{L}^p(\Omega, \Sigma, \mu, \mathbb{R}).$$

Sean $\boldsymbol{f_0}$ y $\boldsymbol{f}$ las clases de equivalencia de f en $L^p(\Omega, \Sigma_0, \mu, \mathbb{R})$ y $L^p(\Omega, \Sigma, \mu, \mathbb{R})$ respectivamente. Claro que $\boldsymbol{f_0} \subset \boldsymbol{f}$. Consideremos el operador $T : L^p(\Omega, \Sigma_0, \mu, \mathbb{R}) \longrightarrow$

$L^p(\Omega,\Sigma,\mu,\mathbb{R})$ tal que $T(\boldsymbol{f_0}) = \boldsymbol{f}$. Definimos $L^p_0(\Omega,\Sigma,\mu,\mathbb{R}) = T(L^p(\Omega,\Sigma_0,\mu,\mathbb{R}))$. Cada elemento de este espacio es una clase de equivalencia de funciones Σ-medibles que tiene representantes Σ_0-medibles. Supongamos que $\boldsymbol{g_0}, \boldsymbol{h_0} \in L^p(\Omega,\Sigma_0,\mu,\mathbb{R})$ y que $\boldsymbol{g_0} \neq \boldsymbol{h_0}$. Entonces $T(\boldsymbol{g_0}) \neq T(\boldsymbol{h_0})$ porque si no fuera así, se cumpliría $\boldsymbol{g_0} = \boldsymbol{h_0}$, lo cual es una contradicción. Así que T es un operador inyectivo lo cual implica que $T : L^p(\Omega,\Sigma_0,\mu,\mathbb{R}) \longrightarrow L^p_0(\Omega,\Sigma,\mu,\mathbb{R})$ es una biyección. El operador T también es una isometría. Es claro que

$$\|\boldsymbol{h}\|_p = \|T(\boldsymbol{h})\|_p,$$

simplemente porque

$$[h_0 \in \boldsymbol{h}, h \in T(\boldsymbol{h})] \Longrightarrow h_0 = h \textit{ en c.t.p} \Longrightarrow \int |h_0|^p \, d\mu = \int |h|^p \, d\mu.$$

Hemos demostrado que $L^p(\Omega,\Sigma_0,\mu,\mathbb{R})$ se puede indentificar con $L^p_0(\Omega,\Sigma,\mu,\mathbb{R}) \subset L^p(\Omega,\Sigma,\mu,\mathbb{R})$, sólo falta demostrar que $L^p_0(\Omega,\Sigma,\mu,\mathbb{R})$ es cerrado. Sea $\{\boldsymbol{f}_n\}$ una sucesión en $L^p_0(\Omega,\Sigma,\mu,\mathbb{R})$ que converge en la norma p hacia $\boldsymbol{f} \in L^p(\Omega,\Sigma,\mu,\mathbb{R})$. Podemos elegir una sucesión $\{f_n\}$ de funciones Σ_0-medibles que converge en media de orden p hacia $f \in \boldsymbol{f}$ donde $f_n \in \boldsymbol{f}_n$ para cada $n \in \mathbb{N}$. Por la proposición 1.2.4, $\{f_n\}$ converge en medida hacia f y por la proposición 1.2.2 existe una subsucesión $\{f_{n_k}\}$ que converge *en c.t.p* hacia f. La función $g = \lim_{k\to\infty} f_{n_k}$, en principio sólo está definida en un conjunto $\Omega_0 \in \Sigma_0$ con $\mu(\Omega \setminus \Omega_0) = 0$. Definamos $g : \Omega \longrightarrow \mathbb{R}$ tal que

$$g(\omega) = \begin{cases} \lim_{k\to\infty} f_{n_k}(\omega), & si \quad \omega \in \Omega_0 \\ 0 & si \quad \omega \in \Omega \setminus \Omega_0. \end{cases}$$

El límite puntual de una sucesión de funciones Σ_0-medibles es Σ_0-medible. Entonces $\lim_{k\to\infty} f_{n_k}$ es Σ_0-medible. Teniendo en cuenta que $\Omega_0 \in \Sigma_0$ y

$$g = \left(\lim_{k\to\infty} f_{n_k}\right) \chi_{\Omega_0},$$

la función g también es Σ_0-medible. Como $g = f$ *en c.t.p*, hemos demostrado que la clase de equivalencia $\boldsymbol{f}$ tiene un representante Σ_0-medible. Por lo tanto $\boldsymbol{f} \in L^p_0(\Omega,\Sigma,\mu,\mathbb{R})$. Entonces $L^p_0(\Omega,\Sigma,\mu,\mathbb{R})$ es cerrado en $L^p(\Omega,\Sigma,\mu,\mathbb{R})$. ■

Hay otra forma de demostrar el teorema 2.2.1. Notamos que $L^p_0(\Omega,\Sigma,\mu,\mathbb{R})$ es completo por ser isométrico a $L^p(\Omega,\Sigma_0,\mu,\mathbb{R})$ que es completo. Por lo tanto $L^p_0(\Omega,\Sigma,\mu,\mathbb{R})$ es un subespacio cerrado de $L^p(\Omega,\Sigma,\mu,\mathbb{R})$.

A partir de aquí (Ω,Σ,μ) designará un espacio de probabilidad.

Lema 2.2.1. *Sea $f \in \mathcal{L}^2(\Omega, \Sigma, \mu, \mathbb{R})$ y Σ_0 una σ-álgebra tal que $\Sigma_0 \subset \Sigma$. Existe una función $f_0 \in \mathcal{L}^2(\Omega, \Sigma_0, \mu, \mathbb{R})$ que verifica*

$$\int_E f \ d\mu = \int_E f_0 \ d\mu \qquad \forall \ E \in \Sigma_0.$$

Demostración: Por el teorema 2.2.1, el espacio $L^2(\Omega, \Sigma_0, \mu, \mathbb{R})$ se puede identificar con un subespacio cerrado del espacio de Hilbert $L^2(\Omega, \Sigma, \mu, \mathbb{R})$. Entonces por el corolario 1.4.1, existe la proyección ortogonal $P : L^2(\Omega, \Sigma, \mu, \mathbb{R}) \longrightarrow L^2(\Omega, \Sigma_0, \mu, \mathbb{R})$ y se verifica $\langle \boldsymbol{f} - P(\boldsymbol{f}), \boldsymbol{\varphi} \rangle = 0$ para cada $\boldsymbol{\varphi} \in L^2(\Omega, \Sigma_0, \mu, \mathbb{R})$. Si $f_0 \in P(\boldsymbol{f})$, para todo $E \subset \Sigma_0$, podemos considerar la clase de equivalencia en $L^2(\Omega, \Sigma, \mu, \mathbb{R})$ de χ_E y se obtiene

$$\int (f - f_0) \ \chi_E \ d\mu = 0$$

$$\int_E f \ d\mu = \int_E f_0 \ d\mu.$$

■

Notamos que si existe otra función $h_0 \in \mathcal{L}^2(\Omega, \Sigma_0, \mu, \mathbb{R})$ que también satisface $\int_E f \ d\mu = \int_E h_0 \ d\mu \quad \forall \ E \in \Sigma_0$, entonces $h_0 = f_0$ *en c.t.p* en virtud del teorema 1.1.2.

Lema 2.2.2. *Sean $f, g \in \mathcal{L}^2(\Omega, \Sigma, \mu, \mathbb{R})$ tales que $f(\omega) \leq g(\omega)$ en c.t.p. Entonces existen $f_0, g_0 \in \mathcal{L}^2(\Omega, \Sigma_0, \mu, \mathbb{R})$ verificando*

$$\int_E f \ d\mu = \int_E f_0 \ d\mu, \ \int_E g \ d\mu = \int_E g_0 \ d\mu \qquad \forall \ E \in \Sigma_0.$$

y adémas $f_0(\omega) \leq g_0(\omega)$ en c.t.p.

Demostración: La existencia de funciones $f_0, g_0 \in \mathcal{L}^2(\Omega, \Sigma_0, \mu, \mathbb{R})$ verificando $\int_E f \ d\mu = \int_E f_0 \ d\mu$ y $\int_E g \ d\mu = \int_E g_0 \ d\mu \quad \forall \ E \in \Sigma_0$ está garantizada por el lema anterior. Por la monotonía del operador $\int_E \cdot \ d\mu : \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R}) \longrightarrow \mathbb{R}$ $(E \in \Sigma)$ tenemos

$$\int_E f_0 \ d\mu = \int_E f \ d\mu \leq \int_E g \ d\mu = \int_E g_0 \ d\mu,$$

lo cual implica por linealidad que $\int_E (f_0 - g_0) \ d\mu \leq 0$ para cada $E \in \Sigma_0$. Por el teorema 1.1.2, $f_0(\omega) - g_0(\omega) \leq 0$ *en c.t.p.* ■

Teorema 2.2.2. *Sea* $f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ *y* Σ_0 *una* σ*-álgebra tal que* $\Sigma_0 \subset \Sigma$*. Existe una función* $f_0 \in \mathcal{L}^1(\Omega, \Sigma_0, \mu, \mathbb{R})$ *que verifica*

$$\int_E f \; d\mu = \int_E f_0 \; d\mu \qquad \forall \; E \in \Sigma_0.$$

Demostración: Sin perdida de generalidad, podemos suponer que $f \geq 0$ *en c.t.p* porque utilizando la descomposición cánonica $f = f^+ - f^-$, el caso general reduce a este caso particular. Consideremos la sucesión $f_n = f \wedge n$. Es claro que $\{f_n\}$ está acotada y creciente en $\mathcal{L}^2(\Omega, \Sigma, \mu, \mathbb{R})$. Por el lema 2.2.1, existe para cada $n \in \mathbb{N}$, $g_n \in \mathcal{L}^2(\Omega, \Sigma, \mu, \mathbb{R})$ tal que $\int_E f_n \; d\mu = \int_E g_n \; d\mu \quad \forall \; E \in \Sigma_0$ y por el lema 2.2.2, $\{g_n\}$ es una sucesión creciente y no negativa para casi todo $\omega \in \Omega$.

$$0 \leq g_1(\omega) \leq g_2(\omega) \cdots \leq g_n(\omega) \leq g_{n+1}(\omega) \cdots \; \textit{en c.t.p.}$$

Existe un conjunto $N \in \Sigma_0$ con $\mu(N) = 0$ tal que si $\omega \in \Omega \setminus N$, $\{g_n(\omega)\}$ es creciente y converge hacia $f_0(\omega) = \text{lím}_{n\to\infty} g_n(\omega)$. En principio f_0 está definida en $\Omega \setminus N$. Podemos definir f_0 en Ω con fórmula

$$f_0(\omega) = \begin{cases} \text{lím}_{n\to\infty} g_n(\omega), & si \quad \omega \in \Omega \setminus N \\ 0 & si \quad \omega \in N. \end{cases}$$

Así f_0 es Σ_0-medible. Luego por el teorema de la convergencia monótona para cada $E \in \Sigma_0$ $\text{lím}_{n\to\infty} \int_E g_n \; d\mu = \int_E f_0 \; d\mu$. Es fácil ver que $\text{lím}_{n\to\infty} f_n(\omega) = f(\omega) \quad \forall \, \omega \in \Omega$ y usando otra vez el teorema de la convergencia monótona se cumple para cada $E \in \Sigma_0$

$$\int_E f \; d\mu = \lim_{n\to\infty} \int_E f_n \; d\mu = \lim_{n\to\infty} \int_E g_n \; d\mu = \int_E f_0 \; d\mu.$$

■

Definición 2.2.1. *Sea* (Ω, Σ, μ) *un espacio de probabilidad y* Σ_0 *una* σ*-álgebra tal que* $\Sigma_0 \subset \Sigma$*. La esperanza condicionada con respeto a* Σ_0 *es el operador* $\mathbb{E}(\, \cdot \, |\Sigma_0) : \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R}) \longrightarrow L^1(\Omega, \Sigma_0, \mu, \mathbb{R})$ *tal que* $\mathbb{E}(f|\Sigma_0) = \boldsymbol{f_0}$*, donde* $\boldsymbol{f_0}$ *es la clase de equivalencia de las funciones* $f_0 \in \mathcal{L}^1(\Omega, \Sigma_0, \mu, \mathbb{R})$ *que verifican*

$$\int_E f \; d\mu = \int_E f_0 \; d\mu \qquad \forall \; E \in \Sigma_0.$$

Gracias al teorema 2.2.2, el operador $\mathbb{E}(\, \cdot \, |\Sigma_0)$ está bien definido. La esperanza condicionada de la función integrable f, no es una función integrable sino una clase de equivalencia de funciones integrables iguales en casi todo puntos. Cualquier función f_0 que verifica la condición de la definición 2.2.1 se llama versión de la esperanza condicionada, es decir $f_0 \in \mathbb{E}(f|\Sigma_0)$.

2.3. Propriedades del Operador $\mathbb{E}(\ \cdot\ |\Sigma_0)$

A partir de aquí y adelante, cuando se refiera al operador $\mathbb{E}(\ \cdot\ |\Sigma_0)$, se tratará del operador definido en la definición de esperanza condicionada 2.2.1.

Proposición 2.3.1. *Sea $f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$. Si f_0 es una versión de $\mathbb{E}(f|\Sigma_0)$, entonces $\int f_0\ d\mu = \int f\ d\mu$.*

Demostración: Como $\Omega \in \Sigma_0$ y que por hypótesis f_0 es una versión de $\mathbb{E}(f|\Sigma_0)$, tenemos $\int f_0\ d\mu = \int f\ d\mu$. ■

Proposición 2.3.2. *Si $f \in \mathcal{L}^1(\Omega, \Sigma_0, \mu, \mathbb{R})$, entonces f es una versión de $\mathbb{E}(f|\Sigma_0)$.*

Demostración: El resultado es immediato por la definición de la esperanza condicionada. ■

Proposición 2.3.3. (Linealidad) *Sean $f, g \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$.*

$$\mathbb{E}(af + bg|\Sigma_0) = a\mathbb{E}(f|\Sigma_0) + b\mathbb{E}(g|\Sigma_0).$$

Demostración: Supongamos que f_0 y g_0 sean versiones de $\mathbb{E}(f|\Sigma_0)$ y $\mathbb{E}(g|\Sigma_0)$ respectivamente. Por la linealidad de $\int_E \cdot\ d\mu$ tenemos para cada $E \in \Sigma_0$,

$$\int_E (af+bg)\ d\mu = a\int_E f\ d\mu + b\int_E g\ d\mu = a\int_E f_0\ d\mu + b\int_E g_0\ d\mu = \int_E (af_0+bg_0)\ d\mu.$$

■

Proposición 2.3.4. (Monotonía) *Sean $f, g \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ tales que $f(\omega) \leq g(\omega)$ en c.t.p. Entonces $\mathbb{E}(f|\Sigma_0) \leq \mathbb{E}(g|\Sigma_0)$, es decir*

$$f_0 \in \mathbb{E}(f|\Sigma_0)\ \&\ g_0 \in \mathbb{E}(g|\Sigma_0) \Longrightarrow f_0(\omega) \leq g_0(\omega) \text{ en c.t.p.}$$

Demostración: Sean f_0 y g_0 versiones de $\mathbb{E}(f|\Sigma_0)$ y $\mathbb{E}(g|\Sigma_0)$ respectivamente. Por la monotonía de $\int_E \cdot\, d\mu$ tenemos para cada $E \in \Sigma_0$,

$$\int_E f_0\ d\mu = \int_E f\ d\mu \leq \int_E g\ d\mu = \int_E g_0\ d\mu,$$

y por el teorema 1.1.2 se sigue que $f_0(\omega) \leq g_0(\omega)$ *en c.t.p.* $\Longleftrightarrow \mathbb{E}(f|\Sigma_0) \leq \mathbb{E}(g|\Sigma_0)$.
■

Las tres proposiciones que siguen son para la esperanza condicional lo que son el lema de Fatou y los teoremas de la convergencia monótona y dominada para la integral. Cuando escribamos $\{\mathbb{E}(f_n|\Sigma_0)\}$ es una sucesión en $L^1(\Omega, \Sigma, \mu, \mathbb{R})$ tal que $\mathbb{E}(f_n|\Sigma_0) \nearrow \mathbb{E}(f|\Sigma_0)$, significará: existe una sucesión $\{g_n\}$ creciente *en c.t.p* tal que $g_n \in \mathbb{E}(f_n|\Sigma_0) \nearrow g \in \mathbb{E}(f|\Sigma_0)$ *en c.t.p.* Para la version condicional del lema de Fatou escribiremos, $\mathbb{E}(\liminf f_n|\Sigma_0) \leq \liminf \mathbb{E}(f_n|\Sigma_0)$, es decir, si $h \in \mathbb{E}(\liminf f_n|\Sigma_0)$ y $g_n \in \mathbb{E}(f_n|\Sigma_0)$ para cada $n \in \mathbb{N}$, entonces $h(\omega) \leq \liminf g_n(\omega)$ *en c.t.p.* Por último $\mathbb{E}(f_n|\Sigma_0) \longrightarrow \mathbb{E}(f|\Sigma_0)$ *en c.t.p*, significa existe una sucesión $\{g_n\}$ tal que $g_n \in \mathbb{E}(f_n|\Sigma_0) \longrightarrow g \in \mathbb{E}(f|\Sigma_0)$ *en c.t.p.*

Proposición 2.3.5. (Convergencia Monótona (*condicional*)) *Sea $\{f_n\}$ una sucesión creciente en $\mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ tal que $f_n : \Omega \longrightarrow [0, \infty]$ para cada $n \in \mathbb{N}$. Si $\{f_n\}$ converge hacia $f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ en casi todo $\omega \in \Omega$, entonces la sucesión $\{\mathbb{E}(f_n|\Sigma_0)\}$ es creciente en $L^1(\Omega, \Sigma_0, \mu, \mathbb{R})$ y converge hacia $\mathbb{E}(f|\Sigma_0)$.*

Demostración: Para cada $n \in \mathbb{N}$, existe $g_n \in \mathbb{E}(f_n|\Sigma_0)$ por el teorema 2.2.2, y $\{g_n\}$ es creciente y no negativa *en c.t.p* por la proposición 2.3.4. Existe un conjunto $N \in \Sigma_0$ con $\mu(N) = 0$ tal que si $\omega \in \Omega \setminus N$, $g_n(\omega) \nearrow \lim_{n\to\infty} g_n(\omega)$. Sea $f_0 : \Omega \longrightarrow [0, \infty]$ con fórmula

$$f_0(\omega) = \begin{cases} \lim_{n\to\infty} g_n(\omega), & si \quad \omega \in \Omega \setminus N \\ 0 & si \quad \omega \in N. \end{cases}$$

La función f_0 es Σ_0-medible y por el teorema de la convergencia monótona

$$\int_E f_0\ d\mu = \lim_{n\to\infty} \int_E g_n\ d\mu = \lim_{n\to\infty} \int_E f_n\ d\mu = \int_E f\ d\mu \qquad \forall\ E \in \Sigma_0.$$

Entonces $f_0 \in \mathbb{E}(f|\Sigma_0)$ y por definición de f_0, la sucesión $\{\mathbb{E}(f_n|\Sigma_0)\}$ es creciente en $L^1(\Omega, \Sigma_0, \mu, \mathbb{R})$ y converge hacia $\mathbb{E}(f|\Sigma_0)$. ■

Proposición 2.3.6. (Lema de Fatou (*condicional*)) *Sea* $\{f_n\}$ *una sucesión en* $\mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$*. Si* $f_n(\omega) \geq 0$ *en c.t.p para cada* $n \in \mathbb{N}$*, entonces* $\mathbb{E}(\liminf f_n|\Sigma_0) \leq \liminf \mathbb{E}(f_n|\Sigma_0)$*.*

Demostración: Sea $h_n = \inf_{k\geq n} f_k$. Claro que $h_n \nearrow \liminf f_n$. Por la proposición 2.3.5 $\mathbb{E}(h_n|\Sigma_0) \nearrow \mathbb{E}(\liminf f_n|\Sigma_0)$. Sean $h \in \mathbb{E}(\liminf f_n|\Sigma_0)$, $g_n \in \mathbb{E}(f_n|\Sigma_0)$ y $\varphi_n \in \mathbb{E}(h_n|\Sigma_0)$ para cada $n \in \mathbb{N}$.

$$h(\omega) = \lim_{n\to\infty} \varphi_n(\omega) = \liminf \varphi_n(\omega) \text{ en c.t.p.}$$

Como $h_n(\omega) \leq f_n(\omega)$, se cumple $\varphi_n(\omega) \leq g_n(\omega)$ *en c.t.p*, luego $\liminf \varphi_n(\omega) \leq \liminf g_n(\omega)$ *en c.t.p.* Entonces

$$h(\omega) = \liminf \varphi_n(\omega) \leq \liminf g_n(\omega) \text{ en c.t.p.}$$

■

Lema 2.3.1. *Sea* $f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$*,* f_0 *una versión de* $\mathbb{E}(f|\Sigma_0)$ *y* φ *una versión de* $\mathbb{E}(\ |f|\ |\Sigma_0)$*. Entonces* $|f_0(\omega)| \leq \varphi(\omega)$ *en c.t.p.*

Demostración: Es claro que $-|f(\omega)| \leq f(\omega) \leq |f(\omega)|$. Por proposición 2.3.4, $-\mathbb{E}(\ |f|\ |\Sigma_0) \leq \mathbb{E}(f|\Sigma_0) \leq\ \mathbb{E}(\ |f|\ |\Sigma_0)$, por lo tanto

$$-\varphi(\omega) \leq f_0(\omega) \leq \varphi(\omega) \text{ en c.t.p} \iff |f_0(\omega)| \leq \varphi(\omega) \text{ en c.t.p.}$$

■

Proposición 2.3.7. (Convergencia Dominada (*condicional*)) *Sea* $\{f_n\}$ *una sucesión en* $\mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ *que converge hacia* $f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ *en casi todo* $\omega \in \Omega$ *tal que para cada* $n \in \mathbb{N}$*,* $|f_n(\omega)| < g(\omega) \in \mathcal{L}^1(\Omega, \Sigma_0, \mu, \mathbb{R})$*. Entonces* $\{\mathbb{E}(f_n|\Sigma_0)\}$ *converge hacia* $\mathbb{E}(f|\Sigma_0)$ *en c.t.p.*

Demostración: Consideremos $h_n = \sup_{k\geq n} |f_k - f|$. Claro que $h_n(\omega) \leq 2g(\omega)$ y que $h_n \searrow 0$ *en c.t.p.* Por el teorema de la convergencia dominada $\lim_{n\to\infty} \int h_n \, d\mu = 0$. Sean $f_0 \in \mathbb{E}(f|\Sigma_0)$, $g_n \in \mathbb{E}(f_n|\Sigma_0)$, $\varphi_n \in \mathbb{E}(|f_n - f|\ |\Sigma_0)$ y $\xi_n \in \mathbb{E}(h_n|\Sigma_0)$. Por el lema 2.3.1 y la monótonia de $\mathbb{E}(\ \cdot\ |\Sigma_0)$,

$$|g_n(\omega) - f_0(\omega)| \leq \varphi_n(\omega) \leq \xi_n(\omega) \text{ en c.t.p.}$$

En casi todo $\omega \in \Omega$, $\{\xi_n\}$ es decreciente y acotada inferiormente por 0. Por tanto existe una función Σ_0-medible ξ tal que $\xi_n \searrow \xi \geq 0$ *en c.t.p.* Como $0 \leq h_n(\omega) \leq 2g(\omega)$ para cada $n \in \mathbb{N}$, tenemos

$$\int h_n \, d\mu = \int \xi_n \, d\mu \geq \int \xi \, d\mu \geq 0.$$

Es fácil ver que $\int \xi \, d\mu = 0$ porque $\int h_n \, d\mu \longrightarrow 0$, lo cual implica que $\xi = 0$ *en c.t.p.* Entonces tenemos, $g_n \in \mathbb{E}(f_n|\Sigma_0) \longrightarrow f_0 \in \mathbb{E}(f|\Sigma_0)$ *en c.t.p.* ■

El lema siguiente va a ser muy útil para demostrar la desigualdad de Jensen (version condicional). Sea f una función convexa real definida en un intervalo abierto I. Es conocido que f es continua en I. Luego las derivadas izquierda f_- y derecha f_+ de f existen y son crecientes en I. En particular, para cada $x \in I$ se verifica

$$f_-(x) = \sup_{y<x} \left\{ \frac{f(x) - f(y)}{x - y} \right\} \leq f_+(x) = \inf_{y>x} \left\{ \frac{f(y) - f(x)}{y - x} \right\}.$$

Lema 2.3.2. (Línea de Soporte) *Sea I un intervalo abierto de $\mathbb{R}$. Si la función $\varphi : I \longrightarrow \mathbb{R}$ es convexa, entonces existen sucesiones de números reales $\{a_n\}$ y $\{b_n\}$ tales que para cada $x \in I$ se cumple*

$$\varphi(x) = \sup_{n\in\mathbb{N}} (a_n x + b_n).$$

Demostración: Sean φ_- y φ_+ las derivadas izquierda y derecha de φ respectivamente. Para cada $x \in I$ se verifica

$$\varphi_-(x) = \sup_{y<x} \left\{ \frac{\varphi(x) - \varphi(y)}{x - y} \right\} \leq \varphi_+(x) = \inf_{y>x} \left\{ \frac{\varphi(y) - \varphi(x)}{y - x} \right\}.$$

Dado $x \in I$, elegimos a_x tal que $\varphi_-(x) \leq a_x \leq \varphi_+(x)$ y definimos en I la recta de pendiente a_x que pasa por el punto $(x, \varphi(x))$ con fórmula $L_x(y) = \varphi(x) + (y - x)a_x$. Entonces se cumple

$$\varphi(y) \geq L_x(y) \qquad \forall\, y \in I. \qquad\qquad (*)$$

La función L_x es la línea de soporte de φ en el punto x. Es inmediato que $\varphi(y) = \sup_{t\in I}\{L_t(y)\}$ para cada $y \in I$. Sea $\{r_n\}$ una sucesión en $\mathbb{Q} \cap I$ tal que $r_n \longrightarrow x$. Notamos que sobre cada subintervalo cerrado y finito de I, los conjuntos $\{\varphi_-(r_n)\}$ y $\{\varphi_+(r_n)\}$ son acotados. Luego el conjunto $\{a_{r_n}\}$ donde $\varphi_-(r_n) \leq a_{r_n} \leq \varphi_+(r_n)$ para cada $n \in \mathbb{N}$, también es acotado. Por la continuidad de φ se cumple

$$\lim_{n\longrightarrow\infty} L_{r_n}(x) = \lim_{n\longrightarrow\infty} \varphi(r_n) + (x - r_n)a_{r_n} = \varphi(x).$$

Por $(*)$ se cumple $\varphi(x) \geq L_{r_n}(x)$ para cada $n \in \mathbb{N}$. Luego $\varphi(x) = \sup_{n\in\mathbb{N}}\{L_{r_n}(x)\}$. Sea $A_x = \{r_n : n \in \mathbb{N}\} \subset \mathbb{Q} \cap I$. Notamos que

$$\varphi(x) = \sup_{r\in A_x}\{L_r(x)\} \leq \sup\{L_r(x) : r \in \mathbb{Q} \cap I\} \leq \sup_{r\in I}\{L_r(x)\} = \varphi(x),$$

con lo cual $\varphi(x) = \sup\{L_r(x) : r \in \mathbb{Q} \cap I\}$, donde $L_r(x) = a_r x + b_r$ y $b_r = \varphi(x) - a_r r$. Como $\mathbb{N}$ y $\mathbb{Q} \cap I$ son conjuntos de misma cardinalidad se obtiene que existen sucesiones $\{a_n\}$ y $\{b_n\}$ de números reales verificando $L_r(x) : r \in \mathbb{Q} \cap I\} = \{a_n x + b_n : n \in \mathbb{N}\}$. ■

Proposición 2.3.8. (Desigualdad de Jensen (*condicional*)) *Sea $f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ con $f(\Omega) \subset I$, donde $I \subset \mathbb{R}$ es un intervalo abierto. Sea $\varphi : I \longrightarrow \mathbb{R}$ una función convexa tal que $\varphi \circ f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$. Si $\xi \in \mathbb{E}(\varphi \circ f|\Sigma_0)$ y $g \in \mathbb{E}(f|\Sigma_0)$, entonces*

(i) *$g(\omega) \in I$ en c.t.p*

(ii) *$\varphi(g(\omega)) \leq \xi(\omega)$ en c.t.p.*

Demostración: Supongamos que $I = (a, b)$ y $g \in \mathbb{E}(f|\Sigma_0)$. Consideremos el conjunto $E = \{\omega \in \Omega : g(\omega) \leq a\}$. Es claro que $E \in \Sigma_0$. Por linealidad $g - a \in \mathbb{E}(f - a|\Sigma_0)$, entonces se cumple

$$0 \leq \int_E (f - a)\ d\mu = \int_E (g - a)\ d\mu \leq 0,$$

luego $\int (f - a)\chi_E\ d\mu = 0$, como $f(\omega) - a \geq 0$, esto implica que $(f - a)\chi_E = 0$ *en c.t.p.* Sea N el conjunto nulo tal que $\omega \in N \Longrightarrow (f(\omega) - a)\chi_E(\omega) \neq 0$. Si $\omega \in E$, $(f(\omega) - a)\chi_E(\omega) > 0$ con lo cual $\omega \in N$. Entonces $\mu(E) = 0$ y $g(\omega) > a$ *en c.t.p.* Analogamente se prueba que $g(\omega) < b$ *en c.t.p* y así se queda demostrado el apartado (i). Para demostrar el apartado (ii) usamos el lema 2.3.2. Consideremos la sucesión de funciones afines $\{a_n x + b_n\}$ dada por el lema 2.3.2 tal que para cada $x \in I$ se verifica

$$\varphi(x) = \sup_{n\in\mathbb{N}}\{a_n x + b_n\}.$$

Para cada $n \in \mathbb{N}$ cada $\omega \in \Omega$ se cumple

$$a_n f(\omega) + b_n \leq \varphi(f(\omega)).$$

Por la monotonía y la linealidad de $\mathbb{E}(\ \cdot\ |\Sigma_0)$, tenemos

$$a_n g(\omega) + b_n \leq \xi(\omega)\ \textit{en c.t.p.}$$

Sea N_n el conjunto nulo tal que si $\omega \in \Omega \setminus N_n$ se cumple la desigualdad anterior. Considerando el supremo cuando n recorre $\mathbb{N}$ se obtiene que para cada $\omega \in \Omega \setminus (\bigcup_{n=1}^{\infty} N_n)$ se cumple

$$\varphi(g(\omega)) = \sup_{n\in\mathbb{N}}\{a_n g(\omega) + b_n\} \leq \xi(\omega).$$

■

Corolario 2.3.1. *Sea* $f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ *con* $p \geq 1$. *Si* $g \in \mathbb{E}(f|\Sigma_0)$ *y* $\xi \in \mathbb{E}(|f|^p|\Sigma_0)$, *entonces*

$$|g(\omega)|^p \leq \xi(\omega) \text{ en c.t.p}$$

Demostración: Basta ver que $|\cdot|^p : \mathbb{R} \longrightarrow \mathbb{R}$ es convexa y utilizar la desigualdad de Jensen. ■

Proposición 2.3.9. *Sea* $f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$. *Si* Σ_0 *y* Σ_1 *son dos sub-*σ*-álgebras de* Σ *tal que* $\Sigma_0 \subset \Sigma_1$ *y* f_0, f_1 *son versiones de* $\mathbb{E}(f|\Sigma_0)$ *y* $\mathbb{E}(f|\Sigma_1)$ *respectivamente, entonces*

$$f_0 \in \mathbb{E}(f_1|\Sigma_0).$$

Demostración: Sean $f_0 \in \mathbb{E}(f|\Sigma_0)$ y $f_1 \in \mathbb{E}(f|\Sigma_1)$. Entonces,

$$\int_E f_0 \, d\mu = \int_E f \, d\mu \qquad \forall\ E \in \Sigma_0$$

$$\int_E f_1 \, d\mu = \int_E f \, d\mu \qquad \forall\ E \in \Sigma_1.$$

Sea $g \in \mathbb{E}(f_1|\Sigma_0)$. Entonces,

$$\int_E g \, d\mu = \int_E f_1 \, d\mu = \int_E f \, d\mu = \int_E f_0 \, d\mu \qquad \forall\ E \in \Sigma_0,$$

con lo cual $g = f_0$ *en c.t.p*, lo que significa que $f_0 \in \mathbb{E}(f_1|\Sigma_0)$. ■

Proposición 2.3.10. *Sea* $f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ *y* $g : \Omega \longrightarrow [-\infty, \infty]$ *una función* Σ_0*-medible tal que* $gf \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$. *Si* $f_0 \in \mathbb{E}(f|\Sigma_0)$ *y* $\varphi \in \mathbb{E}(gf|\Sigma_0)$, *entonces*

$$\varphi(\omega) = g(\omega) f_0(\omega) \text{ en c.t.p.}$$

Demostración: Supongamos que $g = \chi_A$ con $A \in \Sigma_0$ y $f_0 \in \mathbb{E}(f|\Sigma_0)$. Entonces

$$\int_E f \; d\mu = \int_E f_0 \; d\mu \qquad \forall \; E \in \Sigma_0.$$

$$\int_E \chi_A f \; d\mu = \int_{E\cap A} f \; d\mu = \int_{E\cap A} f_0 \; d\mu = \int_E \chi_A f_0 \; d\mu \qquad \forall \; E \in \Sigma_0.$$

Por lo tanto $\chi_A f_0 \in \mathbb{E}(\chi_A f|\Sigma_0)$. Si g es una función simple, $g = \sum_{k=1}^n \alpha_k \chi_{A_k}$ con $\alpha_k \in \mathbb{R}$ y $A_k \in \Sigma_0 \; \forall \; k \in \{1, \ldots, n\}$.

$$\int_E \left(\sum_{k=1}^n \alpha_k \chi_{A_k} \right) f \; d\mu = \sum_{k=1}^n \alpha_k \int_E \chi_{A_k} f \; d\mu$$

$$= \sum_{k=1}^n \alpha_k \int_E \chi_{A_k} f_0 \; d\mu = \int_E \left(\sum_{k=1}^n \alpha_k \chi_{A_k} \right) f_0 \; d\mu \qquad \forall \; E \in \Sigma_0.$$

Supongamos ahora que $g : \Omega \longrightarrow [0, \infty]$ es Σ_0-medible. Existe una sucesión de funciones simples $\{g_n\}$ tal que $g_n \nearrow g$. Para cada $n \in \mathbb{N}$ se cumple $g_n f_0 \in \mathbb{E}(g_n f|\Sigma_0)$, por la proposición 2.3.5 si $\varphi \in \mathbb{E}(gf|\Sigma_0)$ entonces $g_n f_0 \nearrow \varphi = g f_0$ *en c.t.p.* El caso general, es decir, $g : \Omega \longrightarrow [-\infty, \infty]$ es Σ_0-medible, se reduce al caso anterior utilisando la descomposición $g = g^+ - g^-$. ■

Proposición 2.3.11. *Sea $f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ y sean Σ_0 y Σ_1 dos sub-σ-álgebras de Σ. Si $\sigma(f, \Sigma_0)$ es independiente de Σ_1, entonces*

$$\mathbb{E}\left(f|\sigma(\Sigma_0, \Sigma_1)\right) = \mathbb{E}(f|\Sigma_0).$$

Demostración: Sea $f_0 \in \mathbb{E}(f|\Sigma_0)$. Entonces por definición se cumple,

$$\int_E f_0 \; d\mu = \int_E f \; d\mu \qquad \forall \; E \in \Sigma_0.$$

Queremos demostrar que se cumple lo mismo para cada $E \in \sigma(\Sigma_0, \Sigma_1)$. Consideremos la familia $\mathcal{H} = \{E_0 \cap E_1 : E_0 \in \Sigma_0, E_1 \in \Sigma_1\}$. Es fácil ver que $\mathcal{H}$ es un π-sistema. Si $A, B \in \mathcal{H}$, entonces $A = A_0 \cap A_1$ y $B = B_0 \cap B_1$ donde $A_0, B_0 \in \Sigma_0$ y $A_1, B_1 \in \Sigma_1$. Por lo tanto $A \cap B = (A_0 \cap B_0) \cap (A_1 \cap B_1) \in \mathcal{H}$. Es evidente que $\mathcal{H} \subset \sigma(\mathcal{H}) \subset \sigma(\Sigma_0, \Sigma_1)$. Todo conjunto de $\Sigma_0 \cup \Sigma_1$ es intersección de si mismo con $\Omega \in \Sigma_0 \cap \Sigma_1$ así que $\Sigma_0 \cup \Sigma_1 \subset \mathcal{H}$ luego $\sigma(\mathcal{H}) = \sigma(\Sigma_0, \Sigma_1)$. Consideremos ahora la familia $\mathcal{D} = \{E \in \sigma(\Sigma_0, \Sigma_1) : \int_E f_0 \; d\mu = \int_E f \; d\mu\}$. Es claro que $\mathcal{D} \subset \sigma(\Sigma_0, \Sigma_1)$. Sean $E_0 \in \Sigma_0$ y $E_1 \in \Sigma_1$.

$$\int_{E_1 \cap E_0} f_0 \; d\mu = \int \chi_{E_1} \chi_{E_0} f_0 \; d\mu.$$

Por el teorema 1.1.7, como χ_{E_1} es Σ_1-medible, $\chi_{E_0} f_0$ es Σ_0-medible y las σ-álgebras Σ_0, Σ_1 son independientes

$$\int_{E_1 \cap E_0} f_0 \, d\mu = \int \chi_{E_1} \chi_{E_0} f_0 \, d\mu = \left(\int \chi_{E_1} \, d\mu \right) \left(\int \chi_{E_0} f_0 \, d\mu \right) = \left(\int \chi_{E_1} \, d\mu \right) \left(\int_{E_0} f_0 \, d\mu \right)$$

$$= \left(\int \chi_{E_1} \, d\mu \right) \left(\int_{E_0} f \, d\mu \right) = \left(\int \chi_{E_1} \, d\mu \right) \left(\int \chi_{E_0} f \, d\mu \right).$$

Como $\sigma(f, \Sigma_0)$ es indpendiente de Σ_1 y $\chi_{E_0} f$ es $\sigma(f, \Sigma_0)$-medible,

$$\left(\int \chi_{E_1} \, d\mu \right) \left(\int \chi_{E_0} f \, d\mu \right) = \int \chi_{E_1} \chi_{E_0} f \, d\mu = \int_{E_1 \cap E_0} f \, d\mu.$$

Entonces $\mathcal{H} \subset \mathcal{D}$. Obviamente $\Omega \in \mathcal{D}$. Notamos además que si A y $B \in \mathcal{D}$, entonces $A \setminus B \in \mathcal{D}$ porque

$$\begin{aligned} \int_{A \setminus B} f_0 \, d\mu &= \int_A f_0 \, d\mu - \int_B f_0 \, d\mu \\ &= \int_A f \, d\mu - \int_B f \, d\mu = \int_{A \setminus B} f \, d\mu. \end{aligned}$$

Si $\{A_n\}$ es una sucesión creciente en $\mathcal{D}$, como $f_0 \chi_{A_n} \nearrow f_0 \chi_{\cup_n A_n}$ y $f \chi_{A_n} \nearrow f \chi_{\cup_n A_n}$, por el teorema de la convergencia monótona $\cup_{n=1}^{\infty} A_n \in \mathcal{D}$. Acabamos de probar que $\mathcal{D}$ es un d-sistema. Por el lema de Dynkin, $d(\mathcal{H}) = \sigma(\mathcal{H})$. Entonces $\mathcal{H} \subset d(\mathcal{H}) \subset \mathcal{D} \subset \sigma(\Sigma_0, \Sigma_1) = \sigma(\mathcal{H}) = d(\mathcal{H})$. Como $\mathcal{D} = \sigma(\Sigma_0, \Sigma_1)$, queda demostrado que $f_0 \in \mathbb{E}(f | \sigma(\Sigma_0, \Sigma_1))$. ■

Corolario 2.3.2. *Si $f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ es independiente de la σ-álgebra $\Sigma_0 \subset \Sigma$, entonces la función constante igual a $\int f \, d\mu$ es una versión de $\mathbb{E}(f|\Sigma_0)$.*

Demostración: Basta aplicar el teorema 2.3.11 con $\Sigma_0 = \{\emptyset, \Omega\}$ y $\Sigma_1 = \Sigma_2$. ■

Teorema 2.3.1. *Sea $f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ y $\{\mathcal{G}_\alpha\}_{\alpha \in \mathcal{A}}$ una familia arbitraria de sub-σ-álgebras de Σ. Si $g_\alpha \in \mathbb{E}(f|\mathcal{G}_\alpha)$ para cada $\alpha \in \mathcal{A}$, entonces $\{g_\alpha\}$ es uniformemente integrable.*

Demostración: Supongamos que $g_\alpha \in \mathbb{E}(f|\mathcal{G}_\alpha)$ para cada $\alpha \in \mathcal{A}$ y $\varphi_\alpha \in \mathbb{E}(|f||\mathcal{G}_\alpha)$. Dado $\epsilon > 0$, por la proposición 1.3.2, existe $\delta > 0$ tal que

$$[E \in \Sigma, \mu(E) < \delta] \Longrightarrow \int_E |f| \, d\mu < \epsilon.$$

Eligimos $t > 0$ tal que $\frac{1}{t}\int |f|\ d\mu < \delta$. Fijemos $\alpha \in \mathcal{A}$, por la desigualdad de Jensen, $|g_\alpha(\omega)| \leq \varphi_\alpha(\omega)$ *en c.t.p.* Entonces,

$$\int |g_\alpha|\ d\mu \leq \int \varphi_\alpha\ d\mu = \int |f|\ d\mu.$$

Sea $E_t = \{\omega \in \Omega : |g_\alpha(\omega)| \geq t\}$. Por la desigualdad de Chebyshev,

$$\mu(E_t) \leq \frac{1}{t}\int |g_\alpha|\ d\mu \leq \frac{1}{t}\int |f|\ d\mu < \delta.$$

Como $E_t \in \mathcal{G}_\alpha \subset \Sigma$, es claro que

$$\int_{E_t} |g_\alpha|\ d\mu \leq \int_{E_t} \varphi_\alpha\ d\mu = \int_{E_t} |f|\ d\mu < \epsilon.$$

Esto se verifica para cualquier $\alpha \in \mathcal{A}$, entonces la familia $\{g_\alpha\}$ es uniformemente integrable. ∎

Capítulo 3

Martingalas

La noción de esperanza condicionada es fundamental en la teoría de martingalas. Por esa misma razón en primer lugar vamos a recordar sus propriedades principales. A partir de aquí se supone que el lector sabe que la esperanza condicionada es una clase de equivalencia en L^1, entonces para simplificar la notación en vez de escribir la frase f_0 es una versión de la esperanza condicionada $\mathbb{E}(f|\Sigma_0)$ o $f_0 \in \mathbb{E}(f|\Sigma_0)$, escribiremos, $f_0 = \mathbb{E}(f|\Sigma_0)$. Cada vez que se vea $\mathbb{E}(f|\Sigma_0)$, (segun el contexto) se supondrá que se trata de una versión de la esperanza condicionada y no de la clase de equivalencia. La desigualdad $\mathbb{E}(f|\Sigma_0) \leq \mathbb{E}(g|\Sigma_0)$ siempre se supondrá en casi todo punto. Cuando hablemos de la proyección ortogonal $P(\boldsymbol{f})$ donde $\boldsymbol{f} \in L^2$ escribiremos simplemente $P(f)$ donde f es un representante de la clase de equivalencia $\boldsymbol{f}$. Las referencias principales son Williams [6] y Ash [1].

3.1. Lista de las propiedades de $\mathbb{E}(\,\cdot\,|\Sigma_0)$

Sean $f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ y Σ_0, Σ_1 dos sub-σ-álgebras de Σ. Según hemos visto en el capítulo anterior se verifican:

(1) $g = \mathbb{E}(f|\Sigma_0) \Longrightarrow \int g\ d\mu = \int f\ d\mu$

(2) $f \in \mathcal{L}^1(\Omega, \Sigma_0, \mu, \mathbb{R}) \Longrightarrow f = \mathbb{E}(f|\Sigma_0)$

(3) $\mathbb{E}(af + bg|\Sigma_0) = a\mathbb{E}(f|\Sigma_0) + b\mathbb{E}(g|\Sigma_0)$

(4) $f \leq g\ en\ c.t.p \Longrightarrow \mathbb{E}(f|\Sigma_0) \leq \mathbb{E}(g|\Sigma_0)$

(5) $0 \leq f_n \nearrow f\ en\ c.t.p \Longrightarrow \mathbb{E}(f_n|\Sigma_0) \nearrow \mathbb{E}(f|\Sigma_0)$

(6) $f_n \geq 0 \Longrightarrow \mathbb{E}(\liminf f_n|\Sigma_0) \leq \liminf \mathbb{E}(f_n|\Sigma_0)$

(7) $|f_n| \leq g \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})\ \forall\ n\ \&\ f_n \to f$ *en c.t.p* $\Longrightarrow \mathbb{E}(f_n|\Sigma_0) \to \mathbb{E}(f|\Sigma_0)$

(8) $\varphi : \mathbb{R} \to \mathbb{R}$ convexa & $\varphi \circ f \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R}) \Longrightarrow \mathbb{E}(\varphi \circ f|\Sigma_0) \leq \varphi(\mathbb{E}(f|\Sigma_0))$

(9) $\Sigma_0 \subset \Sigma_1 \Longrightarrow \mathbb{E}(\mathbb{E}(f|\Sigma_1)|\Sigma_0) = \mathbb{E}(f|\Sigma_0)$

(10) $h \in \mathcal{L}^1(\Omega, \Sigma_0, \mu, \mathbb{R})\ \&\ |h| \leq M > 0 \Longrightarrow \mathbb{E}(hf|\Sigma_0) = h\mathbb{E}(f|\Sigma_0)$

(11) Σ_1 es independiente de $\sigma(\sigma(f), \Sigma_0) \Longrightarrow \mathbb{E}(f|\sigma(\Sigma_0, \Sigma_1)) = \mathbb{E}(f|\Sigma_0)$

(12) f es independiente de $\Sigma_0 \Longrightarrow \mathbb{E}(f|\Sigma_0) = \int f\ d\mu$

3.2. Proceso Adaptado a una Filtración

3.2.1. Definición de Martingala

Definición 3.2.1. *Sea (Ω, Σ, μ) un espacio de probabilidad y $(E, \mathcal{E})$ un espacio medible y $T \subset [-\infty, \infty]$. Un proceso estocástico X es una aplicación*

$$X : \Omega \times T \longrightarrow E.$$

Definición 3.2.2. *Sea (Ω, Σ, μ) un espacio de probabilidad. Una filtración en (Ω, Σ, μ) es una sucesión creciente $\{\Sigma_n\}_{n=0}^{\infty}$ de sub-σ-álgebras de Σ:*

$$\Sigma_0 \subset \Sigma_1 \subset \cdots \subset \Sigma.$$

En este caso definimos

$$\Sigma_\infty = \sigma\left(\bigcup_{n=0}^{\infty} \Sigma_n\right) \subset \Sigma.$$

Si $X = \{f_n\}_{n=0}^{\infty}$ un proceso estocástico, entonces la sucesión $\Sigma_n = \sigma(f_0, f_1, \ldots, f_n)$ para cada $n \in \mathbb{Z}^+$ se llama filtración natural del proceso.

Definición 3.2.3. *Un espacio filtrado $(\Omega, \Sigma, \{\Sigma_n\}, \mu)$ es un espacio de probabilidad (Ω, Σ, μ) dotado de una filtración $\{\Sigma_n\}$.*

Definición 3.2.4. *Un proceso estocástico $X = \{f_n\} : \Omega \times \mathbb{Z}^+ \longrightarrow [-\infty, \infty]$ es adaptado a la filtración $\{\Sigma_n\}$ si f_n es Σ_n-medible para cada $n \in \mathbb{Z}^+$.*

Si consideramos que el parametro $n \in \mathbb{Z}^+$ es el tiempo, entonces la filtración $\{\Sigma_n\}$ representa la cantidad de información sobre $\omega \in \Omega$ que tenemos al instante n y esa información crece con el tiempo. Si $\{f_n\}$ es un proceso adaptado, entonces cada funcion f_n sólo depende de la historia del proceso hasta el tiempo n incluido.

Definición 3.2.5. *Sea* $(\Omega, \Sigma, \{\Sigma_n\}, \mu)$ *un espacio filtrado. Un proceso estocástico* $X = \{f_n\} : \Omega \times \mathbb{Z}^+ \longrightarrow [-\infty, \infty]$ *se llama martingala con respeto a* $(\{\Sigma_n\}, \mu)$ *si*

(i) X *es adaptado a* $\{\Sigma_n\}$,

(ii) $f_n \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})\ \forall\ n \in \mathbb{Z}^+$,

(iii) $\mathbb{E}(f_{n+1}|\Sigma_n) = f_n\ \forall\ n \in \mathbb{Z}^+$.

Si reemplazamos la condicion (iii) por $\mathbb{E}(f_{n+1}|\Sigma_n) \geq f_n\ \forall\ n \in \mathbb{Z}^+$, entonces X se llama submartingala y si $\mathbb{E}(f_{n+1}|\Sigma_n) \leq f_n\ \forall\ n \in \mathbb{Z}^+$, entonces X es una supermartingala.
Según el contexto puede ser que sea más conveniente considerar procesos o martingalas con índices en $\mathbb{N} = \{1, 2, 3, \ldots\}$ en vez de $\mathbb{Z}^+ = \{0, 1, 2, 3, \ldots\}$.

3.2.2. Propriedades Elementales de las Martingalas

Sea $M = \{f_n\}_{n=0}^{\infty}$ un proceso estocástico.

(1) M es una supermartingala $\Longleftrightarrow -M$ es una submartingala

(2) M es una martingala $\Longleftrightarrow M$ es una submartingala y una supermartinala

(3) M es una martingala $\Longleftrightarrow M - f_0$ es una martingala
M es una submartingala $\Longleftrightarrow M - f_0$ es una submartingala
M es una supermartingala $\Longleftrightarrow M - f_0$ es una supermartingala

(4) M es una martingala; $m < n \Longrightarrow \mathbb{E}(f_n|\Sigma_m) = (\geq / \leq) f_m$
M es una submartingala; $m < n \Longrightarrow \mathbb{E}(f_n|\Sigma_m) \geq f_m$
M es una supermartingala; $m < n \Longrightarrow \mathbb{E}(f_n|\Sigma_m) \leq f_m$

(5) M es una martingala $\Longleftrightarrow \mathbb{E}(f_{n+1} - f_n|\Sigma_n) = 0$
M es una submartingala $\Longleftrightarrow \mathbb{E}(f_{n+1} - f_n|\Sigma_n) \geq 0$
M es una supermartingala $\Longleftrightarrow \mathbb{E}(f_{n+1} - f_n|\Sigma_n) \leq 0$

Demostración:

(1) La demostración es inmediata por la linealidad de $\mathbb{E}(\,\cdot\,|\Sigma_0)$ y por definición de submartingala y supermartingala.

(2) Es cierto por definición.

(3) Supongamos que $M = \{f_n\}$ sea una martingala. El proceso $M - f_0 = \{f_n - f_0\}$ verifica para cada $n \in \mathbb{Z}^+$, $\mathbb{E}(f_{n+1} - f_0|\Sigma_n) = \mathbb{E}(f_{n+1}|\Sigma_n) - \mathbb{E}(f_0|\Sigma_n) = f_n - f_0$ por hypótesis y la propriedad (2) de $\mathbb{E}(\,\cdot\,|\Sigma_0)$. En el caso de que M sea una submartingala o una supermartingala la demostración es similar.

(4) Si M es una martingala se verifica $\mathbb{E}(f_{m+1}|\Sigma_m) = f_m$. Supongamos que para $k \in \mathbb{N}$ se cumpla también $\mathbb{E}(f_{m+k}|\Sigma_m) = f_m$. Como se cumple $\mathbb{E}(f_{m+(k+1)}|\Sigma_{m+k}) = f_{m+k}$, por la propriedad (9) de $\mathbb{E}(\,\cdot\,|\Sigma_0)$ y por inducción se verifica

$$\mathbb{E}(f_{m+(k+1)}|\Sigma_m) = \mathbb{E}(\mathbb{E}(f_{m+(k+1)}|\Sigma_{m+k})|\Sigma_m) = \mathbb{E}(f_{m+k}|\Sigma_m) = f_m.$$

En el caso de que M sea una submartingala o una supermartingala la demostración es similar.

(5) Sea M una martingala. Se cumple por linealidad y la propriedad (2) de $\mathbb{E}(\,\cdot\,|\Sigma_0)$ $\mathbb{E}(f_{n+1} - f_n|\Sigma_n) = \mathbb{E}(f_{n+1}|\Sigma_n) - \mathbb{E}(f_n|\Sigma_n) = f_n - f_n = 0$. El caso de una submartingala o supermartingala es similar.

■

Ejemplo 3.2.1. Sea $\{f_n\} \subset \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ una sucesión de funciones medibles independientes tal que para cada $n \in \mathbb{N}$ se verifique

$$\int f_n \, d\mu = 0.$$

Definamos $S_0 = 0$, $\Sigma_0 = \{\emptyset, \Omega\}$ y

$$S_n = \sum_{k=1}^{n} f_k, \;\; \Sigma_n = \sigma(f_1, f_2, \ldots, f_n).$$

Entonces para cada $n \in \mathbb{N}$ se cumple por linealidad

$$\mathbb{E}(S_{n+1}|\Sigma_n) = \mathbb{E}(S_n|\Sigma_n) + \mathbb{E}(f_{n+1}|\Sigma_n),$$

y por la propiedades (2) y (12) de $\mathbb{E}(\,\cdot\,|\Sigma_0)$ se cumple

$$\mathbb{E}(S_{n+1}|\Sigma_n) = S_n + \int f_{n+1}\, d\mu = S_n.$$

Luego, el proceso estocástico $S = \{S_n\}$ es una martingala.

Ejemplo 3.2.2. Sea $\{f_n\}$ una sucesión de funciones medibles independientes tal que $f_n : \Omega \longrightarrow [0, \infty]$ con

$$\int f_n\, d\mu = 1,$$

para cada $n \in \mathbb{N}$. Definamos $p_0 = 0$, $\Sigma_0 = \{\emptyset, \Omega\}$ y

$$p_n = \prod_{k=1}^{n} f_k,\ \ \Sigma_n = \sigma(f_1, f_2, \ldots, f_n).$$

Por el teorema 1.1.7, $p_n \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ para cada $n \in \mathbb{N}$ y $p_i p_j \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ para cada $(i, j) \in \{1, 2, \ldots, n\} \times \{1, 2, \ldots, n\}$. Entonces se cumple por la proposición 2.3.10 y la propriedad (11) de $\mathbb{E}(\,\cdot\,|\Sigma_0)$

$$\mathbb{E}(p_{n+1}|\Sigma_n) = p_n \mathbb{E}(f_{n+1}|\Sigma_n) = p_n \int f_n\, d\mu = p_n.$$

El proceso estocástico $P = \{p_n\}$ es una martingala.

Ejemplo 3.2.3. Sean $(\Omega, \Sigma, \{\Sigma_n\}, \mu)$ un espacio filtrado y $\xi \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$. Definamos $f_n = \mathbb{E}(\xi|\Sigma_n)$. Por la propriedad (9) de $\mathbb{E}(\,\cdot\,|\Sigma_0)$, se cumple

$$\mathbb{E}(f_{n+1}|\Sigma_n) = \mathbb{E}(\mathbb{E}(\xi|\Sigma_{n+1})|\Sigma_n) = \mathbb{E}(\xi|\Sigma_n) = f_n.$$

El proceso $M = \{f_n\}$ es una martingala.

Ayuda mucho interpretar las martingalas en términos de juegos de dinero o apuestas (igual que en los casinos). Supongamos que en cada momento $n \in \mathbb{N}$, occure un suceso y que el jugador tenga que apostar una cantidad $x > 0$ de dinero sobre la realización del suceso. El jugador tiene también la opción de "pasar", es decir, en cada momento n apuesta una cantidad $\chi_{P^c} x$, (P es el suceso "pasar"). Dependiendo de el suceso que occure en el momento n el jugador gana o pierde lo que ha apostado. Sea $\{f_n(\omega)\}$ la cantidad de dinero que tiene el jugador en su bolsillo en el tiempo n ($f_n(\omega) < 0 =$ deuda). Se dice que el juego es equitativo si $\{f_n\}$ es una martingala, favorable si $\{f_n\}$ es una submartingala y no favorable si $\{f_n\}$ es una supermartingala. En el caso de una submartingala, la cantidad de dinero en el bolsillo del jugador tiende a crecer en función del tiempo. En el caso de una supermartingala, el bolsillo se vacía y la deuda del jugador tiende a crecer con el tiempo y en el caso de una martingala la cantidad de dinero en el bolsillo tiende hacia una constante.

3.3. Transformación de Martingala

Volvamos al casino y supongamos que un jugador adopte una estrategia de juego que le ayuda a tomar la decisión de pasar o apostar dinero. Esta estrategia de juego sólo puede depender de la historia del proceso del juego, luego el jugador decide si va apostar o pasar. Por ejemplo imaginemos que durante un partido de baloncesto, alguien tenga que apostar dinero cada vez que el jugador A vaya a tirar un tiro libre. La estrategia puede ser la siguiente: esperar que A meta 2 tiros libres seguidos, si eso occure, entonces apostar una cantidad fija de dinero hasta que A falle, así succesivamente. Esta estrategia es un ejemplo de proceso previsible.

Definición 3.3.1. *Un proceso estocástico $H = \{h_n\}_{n=1}^{\infty}$ es previsible si h_n es Σ_{n-1}-medible para cada $n \in \mathbb{N}$.*

Definición 3.3.2. *Sea M una martingala y H un proceso previsible. La transformación de martingala $H \bullet M$ es el proceso estocástico $H \bullet M : \Omega \times \mathbb{N} \longrightarrow \mathbb{R}$ tal que $(H \bullet M)_0(\omega) = 0$ y*

$$(H \bullet M)_n(\omega) = \sum_{k=1}^{n} h_k(\omega)(f_k(\omega) - f_{k-1}(\omega)) \quad \forall n \geq 1.$$

Definición 3.3.3. *Sea $X = \{f_n\}_{n=0}^{\infty}$ un proceso estocástico. El proceso X se dice uniformemente acotado si existe $k \in \mathbb{R}^+$ tal que $|f_n(\omega)| \leq k$ para cada $n \in \mathbb{Z}$ y para cada $\omega \in \Omega$.*

Teorema 3.3.1. *Sea H un proceso previsible.*

(i) *Supongamos que H sea uniformemente acotado y no negativo.*
M es una supermartingala $\Longrightarrow H \bullet M$ es una supermartingala.

(ii) *Supongamos que H sea uniformemente acotado.*
M es una martingala $\Longrightarrow H \bullet M$ es una martingala.

(iii) *Supongamos que $H, M \subset \mathcal{L}^2(\Omega, \Sigma, \mu, \mathbb{R})$.*
M es una martingala $\Longrightarrow H \bullet M$ es una martingala.

Demostración:

(i)

$$\mathbb{E}\left((H \bullet M)_n - (H \bullet M)_{n-1} | \Sigma_{n-1}\right) = \mathbb{E}\left(\sum_{k=1}^{n} h_k(f_k - f_{k-1}) - \sum_{k=1}^{n-1} h_k(f_k - f_{k-1}) | \Sigma_{n-1}\right)$$
$$= \mathbb{E}\left(h_n(f_n - f_{n-1})|\Sigma_{n-1})\right)$$

Por hypótesis h_n es acotada, no negativa y Σ_{n-1}-medible para cada $n \in \mathbb{N}$. Entonces por la propriedad (10) de $\mathbb{E}(\,\cdot\,|\Sigma_0)$ y la propriedad (5) de las supermartingalas,

$$\mathbb{E}(h_n(f_n - f_{n-1})|\Sigma_{n-1}) = h_n\mathbb{E}((f_n - f_{n-1})|\Sigma_{n-1}) \leq 0.$$

(ii)
$$\mathbb{E}\left((H \bullet M)_n - (H \bullet M)_{n-1}|\Sigma_{n-1}\right) = h_n\mathbb{E}((f_n - f_{n-1})|\Sigma_{n-1}) = 0.$$

(iii)
$$\mathbb{E}\left((H \bullet M)_n - (H \bullet M)_{n-1}|\Sigma_{n-1}\right) = \mathbb{E}(h_n(f_n - f_{n-1})|\Sigma_{n-1}).$$

Como para cada $n \in \mathbb{N}$

$$h_n(\omega)(f_n(\omega) - f_{n-1}(\omega)) \leq (h_n(\omega) \vee f_n(\omega))^2,$$

$h_n f_n \in \mathcal{L}^2(\Omega, \Sigma, \mu, \mathbb{R})$, entonces por la proposición 2.3.10

$$\mathbb{E}(h_n(f_n - f_{n-1})|\Sigma_{n-1}) = h_n\mathbb{E}((f_n - f_{n-1})|\Sigma_{n-1}) = 0.$$

■

Si M es una (sub)martingala y H un proceso previsible acotado no negativo. Entonces $H \bullet M$ es una (sub)martingala.

Se puede interpretar el teorema 3.3.1 de la siguiente manera: Sean $H = \{h_n\}$ una estrategia de juego que depende de la historia del proceso y $X = \{f_n\}$ el dinero del jugador en su bolsillo. La cantidad h_n es el valor de la apuesta en el momento

n. La ganancia en el momento n es $h_n(f_n - f_{n-1})$, y la ganancia total en el tiempo n es

$$(H \bullet X)_n = \sum_{k=1}^{n} h_k(f_k - f_{k-1}).$$

Si el juego no es favorable y apostamos (sólo en función de la historia del proceso de juego) una cantidad de dinero superior a cero pero inferior a un valor maximo, entonces el juego sigue siendo no favorable. Es decir, el sistema siempre gana.

3.4. Martingalas Paradas

Definición 3.4.1. *Un tiempo de parada es una aplicación $T : \Omega \longrightarrow \mathbb{Z}^+ \cup \{\infty\}$ que verifica* (i) *o* (ii)*:*

(i) $\{\omega \in \Omega : T(\omega) \leq n\} \in \Sigma_n \ \forall \ n \in \mathbb{Z}^+ \cup \{\infty\}$,

(ii) $\{\omega \in \Omega : T(\omega) = n\} \in \Sigma_n \ \forall \ n \in \mathbb{Z}^+ \cup \{\infty\}$.

Es fácil ver que (i) $\Longleftrightarrow$ (ii) simplemente porque $\{\omega \in \Omega : T(\omega) \leq n\} = \cup_{k=0}^{n}\{\omega \in \Omega : T(\omega) \leq k\}$ y $\{\omega \in \Omega : T(\omega) = n\} = \{\omega \in \Omega : T(\omega) \leq n\} \setminus \{\omega \in \Omega : T(\omega) \leq n-1\}$.

El tiempo de parada es el momento en que un jugador decide parar el juego pero la decisión sigue dependiendo de la historia del juego. Por ejemplo el jugador de baloncesto A ha metido 11 tiros libres seguidos y el jugador (él que apuesta) satisfecho de su ganancia no quiere ariesguarse y decide parar las apuestas.

Proposición 3.4.1. *Sea X un proceso estocástico adaptado y $B \subset \mathbb{R}$ un conjunto de Borel. La función $T : \Omega \longrightarrow \mathbb{Z}^+ \cup \{\infty\}$ tal que*

$$T(\omega) = \begin{cases} \inf\{n \geq 0 : f_n(\omega) \in B\}, & \textit{si} \quad \exists\ n \textit{ tal que } f_n(\omega) \in B \\ \infty, & \textit{si} \quad f_n(\omega) \notin B \ \forall \ n \end{cases}$$

es un tiempo de parada.

Demostración:

$$\{\omega \in \Omega : T(\omega) \leq n\} = \bigcup_{k=0}^{n}\{\omega \in \Omega : f_k(\omega) \in B\} \in \Sigma_n.$$

■

Proposición 3.4.2. *Sean T un tiempo de parada y $H^{(T)}$ un proceso estocástico tal que*

$$\left(H^{(T)}\right)_n = \chi_{\{n\leq T\}} = \begin{cases} 1, & si \quad n \leq T(\omega) \\ 0, & si \quad n > T(\omega). \end{cases}$$

Entonces $H^{(T)}$ es un proceso previsible.

Demostración: Tenemos que demostrar que cada $\chi_{n\leq T}$ es Σ_{n-1}-medible.

$$\{n \leq T\} = \{w \in \Omega : T(\omega) \geq n\} = \{w \in \Omega : T(\omega) > n-1\}$$

$$= \{w \in \Omega : T(\omega) \leq n-1\}^c \in \Sigma_{n-1}.$$

■

Como $H^{(T)}$ es un proceso previsible, se puede considerar la transformación de martingala $H^{(T)} \bullet M$ donde

$$\left(H^{(T)} \bullet M\right)_n = \sum_{k=1}^{n} \chi_{\{n\leq T\}}(f_k - f_{k-1})$$

$$= \sum_{k=1}^{n\wedge T}(f_k - f_{k-1}) = f_{n\wedge T} - f_0.$$

Definición 3.4.2. *Sea M una martingala y T un tiempo de parada. Una martingala parada al tiempo T es el proceso M^T tal que $(M^T)_n = f_{n\wedge T}$.*

Teorema 3.4.1. *Sea M una supermartingala y T un tiempo de parada. El proceso M^T es una supermartingala y*

$$\int f_{n\wedge T}\; d\mu \leq \int f_0\; d\mu.$$

Demostración: Es claro que el proceso H^T tal como está definido en la proposición 3.4.2 es acotado y no negativo. Entonces por el teorema 3.3.1 $H^T \bullet M = M^T - f_0$ es una supermartingala y por la propriedad (3) de las martingalas M^T también es una supermartingala. En particular, por la propriedad (4)

$$\mathbb{E}(f_{n\wedge T}|\Sigma_0) \leq f_0$$

$$\int \mathbb{E}(f_{n\wedge T}|\Sigma_0)\; d\mu = \int f_{n\wedge T}\; d\mu \leq \int f_0\; d\mu.$$

■

Teorema 3.4.2. *Sea T un tiempo de parada y M una supermartingala. Si se cumplen una de las tres siguientes condiciones:*

(i) *T es acotado*

(ii) *M es acotado y $T(\omega) \neq \infty$ en c.t.p*

(iii) *$T \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ y para algun $K \in \mathbb{R}^+$ $|f_n(\omega) - f_{n-1}(\omega)| \leq K \; \forall \; n \in \mathbb{Z}^+$.*

Entonces, $f_T \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ y $\int f_T \, d\mu \leq \int f_0 \, d\mu$.

Demostración: Por el teorema 3.4.1, M^T es una supermartingala, $f_{n\wedge T} \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ y $\int (f_{n\wedge T} - f_0) \, d\mu \leq 0$.

(i) Si $T(\omega) \leq N \in \mathbb{Z}^+ \; \forall \; \omega \in \Omega$
$f_{N\wedge T} = f_T$, con lo cual

$$\int (f_T - f_0) \, d\mu \leq 0 \Longleftrightarrow \int f_T \, d\mu \leq \int f_0 \, d\mu.$$

(ii) Si $|f_n(\omega)| \leq K \in \mathbb{R}^+$ y $T(\omega) \neq \infty$ *en c.t.p*
$\int f_{n\wedge T} \, d\mu \leq \int f_0 \, d\mu$ para cada $n \in \mathbb{Z}^+$. Es fácil ver que $f_{n\wedge T} \nearrow f_T$ *en c.t.p*, lo cual implica por el teorema de la convergencia monótona que $\int f_{n\wedge T} \, d\mu \nearrow \int f_T \, d\mu \leq \int f_0 \, d\mu$.

(iii) Sea $H^{(T)}$ el proceso previsble definido en la proposición 3.4.2. Si $T \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ y $|f_n(\omega) - f_{n-1}(\omega)| \leq K \in \mathbb{R}^+ \; \forall \; n \in \mathbb{Z}^+$

$$(H^{(T)} \bullet M)_n(\omega) = |f_{n\wedge T(\omega)}(\omega) - f_0(\omega)| = \left| \sum_{k=1}^{n\wedge T} (f_k(\omega) - f_{k-1}(\omega)) \right| \leq T(\omega) K \; \forall \; \omega \in \Omega.$$

Como $TK \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$, por el teorema de la converencia dominada

$$\int (f_T - f_0) \, d\mu = \lim_{n\to\infty} \int (f_{n\wedge T} - f_0) \, d\mu \leq 0.$$

■

3.5. Lema de los Cruces por Arriba

El lema de los Cruces por Arriba (desigualdad de Doob) es una idea de Doob que se usa para demostrar el teorema de convergencia de martingalas. La figura 1 nos da una idea intuitiva de la desigualdad que se deduce.

Proposición 3.5.1. *Sea* $[a, b] \subset \mathbb{R}$. *Sean* $X = \{f_n\}_{n=0}^{\infty}$ *un proceso adaptado a una filtración* $\{\Sigma_n\}$ *y* $H = \{h_n\}$ *un proceso estocástico tal que*

$$h_1 = \chi_{\{f_0 < a\}},$$

$$h_n = \chi_{\{h_{n-1}=1\}}\chi_{\{f_{n-1}\leq b\}} + \chi_{\{h_{n-1}=0\}}\chi_{\{f_{n-1}<a\}} \ \forall\ n \geq 2.$$

Entonces H *es un proceso previsible.*

Demostración: Como X es un proceso adaptado, f_0 es Σ_0-medible por definición. Así que $\{f_0 < a\} \in \Sigma_0$, con lo cual, $h_1 = \chi_{\{f_0<a\}}$ es Σ_0-medible. El conjunto $\{h_1 = 1\} = \{h_1 < 2\} \in \Sigma_0 \subset \Sigma_1$. Del mismo modo se deduce que $\{h_1 = 0\} \in \Sigma_1$. Obviamente $\{f_1 < a\}$ y $\{f_1 \leq b\} \in \Sigma_1$. La función h_2 es la suma de productos de funciones Σ_1-medibles, por lo tanto es Σ_1-medible. Supongamos que h_n sea Σ_{n-1}-medible. Tenemos

$$h_{n+1} = \chi_{\{h_n=1\}}\chi_{\{f_n\leq b\}} + \chi_{\{h_n=0\}}\chi_{\{f_n<a\}},$$

es fácil ver que h_{n+1} es la suma de productos de funciones Σ_n-medibles. Como h_n es Σ_{n-1}-medible para cada $n \in \mathbb{N}$, H es previsible. ■

Consideremos la sucesión de tiempos de parada $\{T_n\}$ tal que

$$\begin{gathered}
T_1 = \alpha_1 = \text{ínf}\{0 \leq k \leq N : f_k < a\} \\
T_2 = \beta_1 = \text{ínf}\{\alpha_1 \leq k \leq N : f_k > b\} \\
\vdots \\
T_{2n-1} = \alpha_n = \text{ínf}\{\beta_{n-1} \leq k \leq N : f_k < a\} \\
T_{2n} = \beta_n = \text{ínf}\{\alpha_n \leq k \leq N : f_k > b\}
\end{gathered}$$

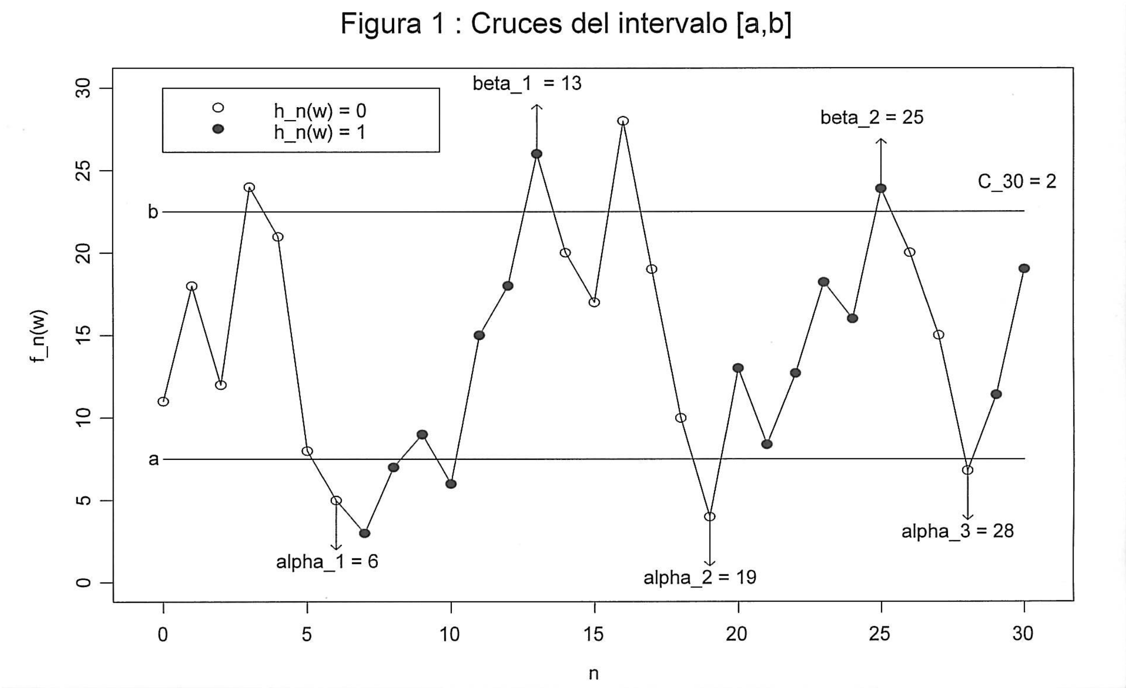
donde $N \in \mathbb{Z}^+ \cup \{\infty\}$. La función $C_N[a, b] : \Omega \longrightarrow \mathbb{Z}^+ \cup \{\infty\}$ tal que

$$C_N[a, b](\omega) = \begin{cases} \sup\{k : \beta_k(\omega) \leq N\}, & si \quad \beta_k(\omega) \leq N \\ 0, & si \quad \beta_k(\omega) > N \end{cases}$$

reprensenta el número de cruces por arriba del intervalo [a,b] con respecto a $\{f_1, f_2, \ldots, f_N\}$. Notamos que $C_N[a, b]$ se puede escribir como combinación lineal de funciones caracteristicas. Para cada $m \in \mathbb{N}$, sean $B_m^- = \{\omega \in \Omega : \beta_m(\omega) \leq N\}$ y $B_m^+ = \{\omega \in \Omega : \beta_m(\omega) > N\}$.

$$C_N[a, b](\omega) = \sum_{m=1}^{N} m\chi_{B_m^- \cap B_{m+1}^+}.$$

La función β_m es un tiempo de parada para cada $m \in \mathbb{N}$, luego $B_m^-, B_{m+1}^+ \subset \Sigma_{m+1}$, por lo tanto $C_N[a, b]$ es Σ-medible.



Lema 3.5.1. *Sea* $[a,b] \subset \mathbb{R}$. *Si* H *es el proceso previsible definido en la proposición* 3.5.1 *y* $X = \{f_n\}$ *es un proceso adaptado, entonces*

$$(H \bullet X)_N = \sum_{k=1}^{N} h_k(f_k - f_{k-1}) \geq (b-a)C_N[a,b] - (f_N - a)^-.$$

Demostración:

Caso 1: $C_N[a,b] = 0$. No se ha completado un cruce.

a) Si $\alpha_1 = 0$, entonces $f_0 < a$ y $h_k = 1 \ \forall \ k \in \{1, \ldots, N\}$. Como $f_N - a = (f_N - a)^+ - (f_N - a)^-$, se cumple

$$(H \bullet X)_N = f_N - f_0 > f_N - a \geq -(f_N - a)^-.$$

b) Si $0 < \alpha_1 \leq N$, entonces $f_{\alpha_1} < a$, $h_{\alpha_1+1}, h_{\alpha_1+2}, \ldots, h_N = 1$ y se cumple

$$(H \bullet X)_N = f_N - f_{\alpha_1} > f_N - a \geq -(f_N - a)^-.$$

c) Si $\alpha_1 > N$, entonces $h_k = 0 \ \forall \ k \leq N$ y se cumple

$$(H \bullet X)_N = 0 \geq -(f_N - a)^-.$$

Caso 2: $m = C_N[a,b] > 0$. Se ha completado un cruce.

a) Si $\alpha_{m+1} \geq N$ (no se comienza un último cruce)

$$(H \bullet X)_N = \sum_{k=1}^{C_N[a,b]} (f_{\beta_k} - f_{\alpha_k}) \geq C_N[a,b](b-a) \geq C_N[a,b](b-a) - (f_N - a)^-$$

b) Si $\alpha_{m+1} < N$ (se comienza un último cruce incompleto)

$$(H \bullet X)_N = \sum_{k=1}^{m}(f_{\beta_k} - f_{\alpha_k}) + \sum_{k=\alpha_{m+1}}^{N} h_k(f_k - f_{k-1})$$

$$= \sum_{k=1}^{m}(f_{\beta_k} - f_{\alpha_k}) + f_N - f_{\alpha_m}$$

y como $f_N - f_{\alpha_m} > f_N - a \geq -(f_N - a)^-$ se cumple

$$(H \bullet X)_N \geq C_N[a,b](b-a) - (f_N - a)^-.$$

■

Teorema 3.5.1. (Lema de los Cruces por Arriba (Desigualdad de Doob)) *Sea $[a,b] \subset \mathbb{R}$. Si $M = \{f_n\}$ es una supermartingala y H el proceso previsible definido en la proposición* 3.5.1, *entonces se verifica*

$$(b-a)\int C_N[a,b]\ d\mu \leq \int (f_N - a)^-\ d\mu.$$

Demostración: Por el lema 3.5.1,

$$(H \bullet M)_N \geq (b-a)C_N[a,b] - (f_N - a)^-.$$

Por el teorema 3.3.1, $H \bullet M$ es una supermartingala. Entonces

$$\mathbb{E}((H \bullet M)_N|\Sigma_0) \leq (H \bullet M)_0 = 0,$$

lo cual implica que

$$\int \mathbb{E}((H \bullet M)_N|\Sigma_0)\ d\mu = \int (H \bullet M)_N\ d\mu \leq 0.$$

Entonces

$$0 \geq \int (H \bullet M)_N \geq (b-a) \int C_N[a,b]\ d\mu - \int (f_N - a)^-\ d\mu$$

$$(b-a) \int C_N[a,b]\ d\mu \leq \int (f_N - a)^-\ d\mu.$$

■

Corolario 3.5.1. *Sea M una supermartingala acotada en $\mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$, es decir* $\sup_{n\in\mathbb{N}} \left\{\int |f_n|\ d\mu\right\} < \infty$. *Entonces,*

$$(b-a) \int C_\infty[a,b]\ d\mu \leq |a| + \sup_{n\in\mathbb{N}} \left\{ \int |f_n|\ d\mu \right\} < \infty.$$

Demostración:

$$(b-a) \int C_N[a,b]\ d\mu \leq \int (f_N - a)^-\ d\mu \leq \int (f_N - a)^+ - (f_N - a)^-\ d\mu$$

$$\int (f_N - a)^+ - (f_N - a)^-\ d\mu = \int |f_N - a|\ d\mu \leq \int |a| + |f_N|\ d\mu$$

Entonces para cada N,

$$(b-a) \int C_N[a,b]\ d\mu \leq |a| \int +|f_N|\ d\mu \leq |a| + \sup_{n\in\mathbb{N}} \left\{ \int |f_n|\ d\mu \right\}.$$

Como $C_N[a,b] \nearrow C_\infty[a,b]$, por la convergencia monótona

$$\int C_N[a,b]\ d\mu \nearrow \int C_\infty[a,b]\ d\mu \leq |a| + \sup_{n\in\mathbb{N}} \left\{ \int |f_n|\ d\mu \right\} < \infty.$$

■

3.6. Convergencia de Martingalas

Ahora disponemos de las herramientas necesarias para demostrar el teorema central de esta tesis.

Teorema 3.6.1. (Convergencia de Martingalas) *Sea $M = \{f_n\}$ una supermartingala acotada en $\mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$, es decir,* $\sup_{n\in\mathbb{N}} \left\{\int |f_n|\ d\mu\right\} < \infty$. *Entonces existe una función $f_\infty \in \mathcal{L}^1(\Omega, \Sigma_\infty, \mu, \mathbb{R})$ tal que*

$$\lim_{n\to\infty} f_n(\omega) = f_\infty(\omega) \text{ en c.t.p.}$$

Demostración: Sea $\Lambda = \{\omega \in \Omega : \liminf f_n(\omega) < \limsup f_n(\omega)\}$, es decir, el conjunto de los puntos ω donde la sucesión $\{f_n(\omega)\}$ no converge hacia un límite $L \in [-\infty, \infty]$. Si $\liminf f_n(\omega) < \limsup f_n(\omega)$, existen $a, b \in \mathbb{Q}$ tal que $\liminf f_n(\omega) < a < b < \limsup f_n(\omega)$. Dado $a, b \in \mathbb{Q}$ tal que $\liminf f_n(\omega) < a < b < \limsup f_n(\omega)$, sea $\Lambda_{a,b} = \{\omega \in \Omega : \liminf f_n(\omega) < a < b < \limsup f_n(\omega)\}$. Es claro que $\Lambda_{a,b} \subset \Lambda$ y si $\omega \in \Lambda$, siempre se puede encontrar $r, s \in \mathbb{Q}$ tal que $\omega \in \Lambda_{r,s}$, entonces

$$\Lambda = \bigcup_{a,b \in \mathbb{Q}} \Lambda_{a,b}.$$

Supongamos que $\omega \in \Lambda_{a,b}$. entonces, existe una subsucesión $\{f_{n_k}\}$ que converge hacia $\limsup f_n(\omega) > b$ y también existe una subsucesión $\{f_{m_k}\}$ que converge hacia $\liminf f_n(\omega) < a$. Esto implica que hay un número infinito de valores de $n \in \mathbb{N}$ tales que $f_n(\omega) > b$ y también un número infinito de valores de $n \in \mathbb{N}$ tales que $f_n(\omega) < a$. En otras palabras, hay un número infinito de cruces por arriba del intervalo $[a, b]$. Entonces tenemos

$$\Lambda_{a,b} \subset \{\omega \in \Omega : C_\infty[a, b](\omega) = \infty\} \qquad (*).$$

Por el corolario 3.5.1, $\int C_\infty[a, b]\ d\mu < \infty$, con lo cual

$$C_\infty[a, b] < \infty \ \textit{en c.t.p} \Longleftrightarrow \mu\left(\{\omega \in \Omega : C_\infty[a, b] = \infty\}\right) = 0.$$

Como se verifica $(*)$, $\mu(\Lambda_{a,b}) = 0$ y

$$\mu(\Lambda) = \mu\left(\bigcup_{a,b \in \mathbb{Q}} \Lambda_{a,b}\right) = 0.$$

Así que $\{f_n\}$ converge hacia $f_\infty : \Omega \longrightarrow [-\infty, \infty]$ *en c.t.p.* Por el lema de Fatou,

$$\int |f_\infty|\ d\mu = \int \liminf |f_n|\ d\mu \leq \liminf \int |f_n|\ d\mu \leq \sup_{n \in \mathbb{N}} \left\{\int |f_n|\ d\mu\right\} < \infty.$$

■

Corolario 3.6.1. *Si $M = \{f_n\}$ es una supermartingala no negativa, entonces existe una función $f_\infty \in \mathcal{L}^1(\Omega, \Sigma_\infty, \mu, \mathbb{R})$ tal que*

$$\lim_{n \to \infty} f_n(\omega) = f_\infty(\omega) \ \textit{en c.t.p.}$$

Demostración: Como $\int |f_n|\, d\mu = \int f_n\ d\mu \leq \int f_0\ d\mu$ para todo $n \in \mathbb{N}$, por el teorema 3.6.1,

$$\lim_{n \to \infty} f_n(\omega) = f_\infty(\omega) \ \textit{en c.t.p.}$$

■

3.7. Martingalas en $\mathcal{L}^2$

Proposición 3.7.1. *Sea $M = \{f_n\}$ una martingala en $\mathcal{L}^2(\Omega, \Sigma, \mu, \mathbb{R})$. Si $m, n, p, q \in \mathbb{Z}^+$ tal que $m \leq n \leq p \leq q$, entonces*

$$\int (f_n - f_m)(f_q - f_p)\ d\mu = 0.$$

Demostración: Es claro que $\mathbb{E}(f_n|\Sigma_m) = P_m(f_n)$, $\mathbb{E}(f_q|\Sigma_p) = P_p(f_q)$, donde $P_m : L^2(\Omega, \Sigma, \mu, \mathbb{R}) \longrightarrow L^2(\Omega, \Sigma_m, \mu, \mathbb{R})$ y $P_p : L^2(\Omega, \Sigma, \mu, \mathbb{R}) \longrightarrow L^2(\Omega, \Sigma_p, \mu, \mathbb{R})$ son proyecciones ortogonales. Por lo tanto, $\langle f_n - P_m(f_n), \varphi\rangle = \langle f_q - P_p(f_q), \xi\rangle = 0$ para cada $\varphi \in L^2(\Omega, \Sigma_m, \mu, \mathbb{R})$ y cada $\xi \in L^2(\Omega, \Sigma_p, \mu, \mathbb{R})$. Como $f_n - P_m(f_n) \in L^2(\Omega, \Sigma_p, \mu, \mathbb{R})$,

$$\langle f_n - P_m(f_n), f_q - P_p(f_q)\rangle = \int (f_n - f_m)(f_q - f_p)\ d\mu = 0.$$

■

Lema 3.7.1. *Sea $M = \{f_n\}$ una martingala en $\mathcal{L}^2(\Omega, \Sigma, \mu, \mathbb{R})$. Entonces para cada $n \in \mathbb{N}$ se verifica*

$$\int f_n^2\ d\mu = \int f_0^2\ d\mu + \sum_{k=1}^{n} \int (f_k - f_{k-1})^2\ d\mu.$$

Demostración: Es claro que $f_n = f_0 + \sum_{k=1}^{n}(f_k - f_{k-1})$. Usando la proposición 3.7.1, es fácil ver que $f_0 \cup \{f_k - f_{k-1}\}_{k=1}^{\infty}$ es un sistema ortogonal. Entonces por el teorema de Pytagoras

$$\begin{aligned}\|f_n\|^2 = \left\| f_0 + \sum_{k=1}^{n}(f_k - f_{k-1}) \right\|^2 &= \|f_0\|^2 + \sum_{k=1}^{n} \|(f_k - f_{k-1})\|^2 \\ = \int f_n^2\ d\mu &= \int f_0^2\ d\mu + \sum_{k=1}^{n} \int (f_k - f_{k-1})^2\ d\mu.\end{aligned}$$

■

Teorema 3.7.1. *Sea $M = \{f_n\}$ una martingala en $\mathcal{L}^2(\Omega, \Sigma, \mu, \mathbb{R})$. La martingala M es acotada en $\mathcal{L}^2(\Omega, \Sigma, \mu, \mathbb{R})$ si y sólo si*

$$\sum_{k=1}^{\infty} \int (f_k - f_{k-1})^2\ d\mu < \infty. \qquad (*)$$

Si se verifica $(*)$, entonces $\{f_n\}$ converge hacia $f_\infty \in \mathcal{L}^2(\Omega, \Sigma_\infty, \mu, \mathbb{R})$ en c.t.p y en media de orden 2.

Demostración: Supongamos que M es acotada en $\mathcal{L}^2(\Omega, \Sigma, \mu, \mathbb{R})$. Usando el lema 3.7.1, se cumple para cada $n \in \mathbb{N}$,

$$\sum_{k=1}^{n} \int (f_k - f_{k-1})^2 \, d\mu \leq \int f_n^2 \, d\mu - \int f_0^2 \, d\mu < \infty,$$

entonces

$$\sum_{k=1}^{\infty} \int (f_k - f_{k-1})^2 \, d\mu < \infty.$$

Ahora supongamos que se cumple $\sum_{k=1}^{\infty} \int (f_k - f_{k-1})^2 \, d\mu < \infty$. Entonces para cada $n \in \mathbb{N}$ se cumple

$$\begin{aligned} \int f_n^2 \, d\mu &= \int f_0^2 \, d\mu + \sum_{k=1}^{n} \int (f_k - f_{k-1})^2 \, d\mu \\ &\leq \int f_0^2 \, d\mu + \sum_{k=1}^{\infty} \int (f_k - f_{k-1})^2 \, d\mu < \infty, \end{aligned}$$

es decir, M es acotada en $\mathcal{L}^2(\Omega, \Sigma, \mu, \mathbb{R})$. Por la desigualdad de Jensen, se verifica $\left(\int |f| \, d\mu\right)^2 \leq \int f^2 \, d\mu$ para toda función integrable, así que

$$\sup_{n \in \mathbb{N}} \left\{ \int |f_n| \, d\mu \right\} \leq \sup_{n \in \mathbb{N}} \left\{ \sqrt{\int f_n^2 \, d\mu} \right\} < \infty.$$

Por lo tanto $f_n \longrightarrow f_\infty \in \mathcal{L}^1(\Omega, \Sigma_\infty, \mu, \mathbb{R})$ *en c.t.p.* Como

$$\begin{aligned} f_{n+r} - f_n &= f_0 + \sum_{k=1}^{n+r} (f_k - f_{k-1}) - f_0 - \sum_{k=1}^{n} (f_k - f_{k-1}) \\ &= \sum_{k=n+1}^{n+r} (f_k - f_{k-1}), \end{aligned}$$

entonces por el lema 3.7.1

$$\int (f_{n+r} - f_n)^2 \, d\mu = \sum_{k=n+1}^{n+r} \int (f_k - f_{k-1})^2 \, d\mu.$$

Como $\lim_{r \to \infty} (f_{n+r} - f_n)^2 = (f_\infty - f_n)^2$, por el lema de Fatou

$$\int (f_\infty - f_n)^2 \, d\mu \leq \liminf \sum_{k=n+1}^{n+r} \int (f_k - f_{k-1})^2 \, d\mu = \sum_{k=n+1}^{\infty} \int (f_k - f_{k-1})^2 \, d\mu.$$

Como $\lim_{n \to \infty} \sum_{k=n+1}^{\infty} \int (f_k - f_{k-1})^2 \, d\mu = 0$ por hypótesis, también se cumple

$$\lim_{n \to \infty} \int (f_\infty - f_n)^2 \, d\mu = 0.$$

■

3.8. Martingalas Uniformemente Integrables

Todavía sólo hemos econtrado una condición necesaria para garantizar la convergencia de martingalas en casi todo punto. Esto no implica la convergencia en media, la cual es necesaria para demostrar el teorema de Radon-Nikodým. Nos falta la hypótesis de integrabilidad uniforme.

Teorema 3.8.1. *Sea $M = \{f_n\}$ una martingala uniformemente integrable. Entonces se cumplen:*

(i) $\exists\ f_\infty \in \mathcal{L}^1(\Omega, \Sigma_\infty, \mu, \mathbb{R})$ *tal que* $\lim_{n\to\infty} f_n(\omega) = f_\infty(\omega)$ *en c.t.p,*

(ii) $\lim_{n\to\infty} \int |f_n(\omega) - f_\infty(\omega)|\ d\mu = 0,$

(iii) $\mathbb{E}(f_\infty|\Sigma_n) = f_n\ \forall\ n \in \mathbb{N}.$

Demostración: Por el teorema 1.3.1, M es acotada en $\mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$, entonces por el teorema 3.6.1, $\lim_{n\to\infty} f_n(\omega) = f_\infty(\omega)$ *en c.t.p.* Como (Ω, Σ, μ) es finito, por la proposición 1.2.1, $f_n \longrightarrow f_\infty$ en medida. Luego por el teorema de Vitali,

$$\lim_{n\to\infty} \int |f_n(\omega) - f_\infty(\omega)|\ d\mu = 0.$$

Si $m \geq n$, entonces $\mathbb{E}(f_m|\Sigma_n) = f_n$ por la propriedad (4) de las martingalas y se verifica

$$\int_E f_m\ d\mu = \int_E f_n\ d\mu \qquad \forall\ E \in \Sigma_n.$$

Como

$$\left|\int_E f_m\ d\mu - \int_E f_\infty\ d\mu\right| = \left|\int_E (f_m - f_\infty)\ d\mu\right| \leq \int_E |f_m - f_\infty|\ d\mu$$

$$\leq \int |f_m - f_\infty|\ d\mu \longrightarrow 0,$$

se cumple $\int_E f_m\ d\mu \longrightarrow \int_E f_\infty\ d\mu = \int_E f_n\ d\mu$ para todo $E \in \Sigma_n$. Entonces $\mathbb{E}(f_\infty|\Sigma_n) = f_n$. ■

Teorema 3.8.2. *Sean $\xi \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ y $M = \{f_n\}$ un proceso estocástico tal que para cada $n \in \mathbb{Z}^+$ $f_n = \mathbb{E}(\xi|\Sigma_n)$. Entonces se verifican:*

(i) $\lim_{n\to\infty} f_n = f_\infty \in \mathcal{L}^1(\Omega, \Sigma_\infty, \mu, \mathbb{R})$ *en c.t.p,*

(ii) $\lim_{n\to\infty} \int |f_n - f_\infty|\ d\mu = 0,$

(iii) $f_\infty = \mathbb{E}(\xi|\Sigma_\infty)$.

Demostración: El ejemplo 3.2.3 demuestra que M es una martingala, y por el teorema 2.3.1 M es uniformemente integrable. Entonces por el teorema 3.8.1 $\lim_{n\to\infty} f_n = f_\infty \in \mathcal{L}^1(\Omega, \Sigma_\infty, \mu, \mathbb{R})$ *en c.t.p* y en media. Sea φ una version de $\mathbb{E}(\xi|\Sigma_\infty)$, podemos suponer que $\xi(\omega) \geq 0\ \forall\omega \in \Omega$ (en el caso general, usamos la descomposición $\xi = \xi^+ - \xi^-$). Entonces φ es positiva *en c.t.p.* Gracias al teorema de Beppo-Levi, podemos definir las medidas $\nu_1 : \Sigma_\infty \longrightarrow [0,\infty]$ y $\nu_2 : \Sigma_\infty \longrightarrow [0,\infty]$ tales que $\nu_1(E) = \int_E \varphi\, d\mu$ y $\nu_2(E) = \int_E f_\infty\, d\mu$. Por la propriedad (9) del operador $\mathbb{E}(\,\cdot\,|\Sigma_0)$y el teorema 3.8.1, como $\mathbb{E}(\varphi|\Sigma_\infty) = \mathbb{E}(\mathbb{E}(\xi|\Sigma_\infty)|\Sigma_n) = \mathbb{E}(\xi|\Sigma_n)$, si $E \in \Sigma_n$,

$$\nu_1(E) = \int_E \varphi\, d\mu = \int_E f_n\, d\mu = \int_E f_\infty\, d\mu = \nu_2(E) \qquad \forall\ E \in \bigcup_{n=1}^{\infty} \Sigma_n.$$

Como $\cup_{n=1}^{\infty}\Sigma_n$ es un π-sistema, por el corolario 1.1.1, $\nu_1 = \nu_2 \Longrightarrow \varphi = f_\infty$ *en c.t.p.* Entonces $f_\infty = \mathbb{E}(\xi|\Sigma_\infty)$. ■

Hasta ahora siempre hemos estudiado el comportamiento de los procesos estocásticos cuando el n crece, es decir viajando en el futuro. De una manera similar se puede estudiar el comportamiento de un proceso estocástico en el pasado, considerando $-n$ en vez de n, y una sucesión decreciente de σ-álgebras con índices negativos (perdiendo información).

Definición 3.8.1. *Sean (Ω, Σ, μ) un espacio de probabilidad y $\{\mathcal{G}_{-n}\}_{n=1}^{\infty}$ una sucesión decreciente de sub-σ-álgebras de Σ tal que*

$$\mathcal{G}_{-\infty} = \bigcap_{n=1}^{\infty} \mathcal{G}_{-n} \subset \cdots\ \mathcal{G}_{-(n+1)} \subset \mathcal{G}_{-n} \subset \cdots\ \mathcal{G}_1 \subset \Sigma.$$

Un proceso estocástico $X = \{f_n\} : \Omega \times \mathbb{N} \longrightarrow [-\infty, \infty]$ se llama martingala inversa con respeto a $(\{\mathcal{G}_{-n}\}, \mu)$ si

(i) *X es adaptado a $\{\mathcal{G}_{-n}\}$,*

(ii) *$f_{-n} \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})\ \forall\ n \in \mathbb{N}$,*

(iii) *$\mathbb{E}(f_{-n}|\mathcal{G}_{-(n+1)}) = f_{-(n+1)}\ \forall\ n \in \mathbb{N}$.*

Lema 3.8.1. *Sean* (Ω, Σ, μ) *un espacio de probabilidad y* $\{\mathcal{G}_{-n}\}_{n=1}^{\infty}$ *una sucesión decreciente de sub-*σ*-álgebras de* Σ *tal que*

$$\mathcal{G}_{-\infty} = \bigcap_{n=1}^{\infty} \mathcal{G}_{-n} \subset \cdots\ \mathcal{G}_{-(n+1)} \subset \mathcal{G}_{-n} \subset \cdots\ \mathcal{G}_1 \subset \Sigma.$$

Supongamos que $\gamma \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ *y sea* $M = \{f_{-n}\}$ *un proceso estocástico tal que para cada* $n \in \mathbb{N}$ *se verifica* $f_{-n} = \mathbb{E}(\gamma|\mathcal{G}_{-n})$*. Entonces* M *es una martingala inversa.*

Demostración: Por hypótesis se cumplen los apartados (i) y (ii) de la definición de martingala inversa. Como

$$\mathbb{E}(f_{-n}|\mathcal{G}_{-(n+1)}) = \mathbb{E}(\mathbb{E}(\xi|\mathcal{G}_{-n})|\mathcal{G}_{-(n+1)}) = \mathbb{E}(\xi|\mathcal{G}_{-(n+1)}) = f_{-(n+1)},$$

M es una martingala inversa. ■

Teorema 3.8.3. *Sea* $M = \{f_{-n}\}$ *la martingala inversa definida en el lema 3.8.1. Entonces se verifican:*

(i) $\text{lím}_{n\to\infty} f_{-n} = f_{-\infty} \in \mathcal{L}^1(\Omega, \mathcal{G}_{-\infty}, \mu, \mathbb{R})$ *en c.t.p,*

(ii) $\text{lím}_{n\to\infty} \int |f_{-n} - f_{-\infty}|\ d\mu = 0,$

(iii) $f_{-\infty} = \mathbb{E}(\gamma|\mathcal{G}_{-\infty})$.

Demostración: Aplicando el lema de los cruces por arriba al proceso $\{f_k, \mathcal{G}_k : -N \leq k \leq -1\}$, de manera análoga a la demostración teorema de la convergencia de martingala 3.6.1, se demuestra que $f_{-n} \longrightarrow f_{-\infty} \in \mathcal{L}(\Omega, \mathcal{G}_{-\infty}, \mu, \mathbb{R})$ *en c.t.p.* Por el teorema 2.3.1 M es uniformemente integrable, luego $f_{-n} \longrightarrow f_{-\infty}$ en media. Para demostrar el apartado (iii), basta ver que para cada $G \in \mathcal{G}_{-\infty} \subset \mathcal{G}_{-n}$, se cumple

$$\int_G \gamma\ d\mu = \int_G f_{-n}\ d\mu.$$

Cuando $n \to \infty$ se cumple (iii). ■

Capítulo 4

Aplicaciones

4.1. Ley Fuerte de los Grandes Números

Vamos a aplicar la teoría de martingalas para demostrar la ley fuerte de los grandes números. Primero se demostrará la ley 0-1 de Kolmogorov (también utilizando la teoría de martingalas) que se necesitará luego. Supondremos conocido la noción de medida producto y el teorema de Fubini. Este teorema se utilizará en la demostración de la ley fuerte de los grandes números. Las referencias son Williams [6] y Ash [1].

Teorema 4.1.1. (Ley 0-1 de Kolmogorov) *Sea* (Ω, Σ, μ) *un espacio de probabilidad. Sea* $\{f_n\}$ *una sucesión de funciones* Σ*-medibles independientes tal que* $f_n : \Omega \longrightarrow [-\infty, \infty]$ *para cada* $n \in \mathbb{Z}^+$*. Si*

$$\mathcal{T}_k = \sigma(\{f_n\}_{n=k+1}^{\infty}), \ \ y\ \mathcal{T} = \bigcap_{k=0}^{\infty} \mathcal{T}_k,$$

entonces se cumple

$$E \in \mathcal{T} \Longrightarrow \mu(E) \in \{0, 1\}.$$

Demostración: Definamos $\Sigma_n = \sigma(f_0, f_1, \ldots, f_n)$. Supongamos que $E \in \mathcal{T}$. Claro que $\chi_E \in \mathcal{L}^1(\Omega, \Sigma_\infty, \mu, \mathbb{R})$, entonces por el teorema 3.8.2, $\mathbb{E}(\chi_E|\Sigma_n) \longrightarrow \mathbb{E}(\chi_E|\Sigma_\infty)$. Como χ_E es Σ_∞-medible, χ_E es una version de $\mathbb{E}(\chi_E|\Sigma_\infty)$. Para cada $n \in \mathbb{Z}^+$, $E \in \mathcal{T}_n$, entonces χ_E es $\mathcal{T}_n$-medible, pero por hypótesis χ_E es independiente de Σ_n. Por la propriedad (12) de $\mathbb{E}(\ \cdot\ |\Sigma_0)$, $\mathbb{E}(\chi_E|\Sigma_n) = \int \chi_E \, d\mu = \mu(E)$ para cada $n \in \mathbb{Z}^+$. Como $\mathbb{E}(\chi_E|\Sigma_n) \longrightarrow \chi_E$, se cumple $\chi_E = \mu(E)$ *en c.t.p*, lo cual implica que $\mu(E) \in \{0, 1\}$. ■

Lema 4.1.1. *Sea (Ω, Σ, μ) un espacio de probabilidad. Sea $\{f_n\}$ una sucesión de funciones Σ-medibles independientes tal que $f_n \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ para cada $n \in \mathbb{N}$. Supongamos que las funciones f_n sean idénticamente distribuidas. Entonces se cumple*

$$\mathbb{E}(f_k|\sigma(S_n)) = \frac{S_n}{n} \qquad \forall\ k \in \{1, 2, \ldots, n\},$$

donde $S_n = \sum_{k=1}^{n} f_k$.

Demostración: Si $B \in \mathcal{B}(\mathbb{R})$ y $B_n = \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : x_1 + x_2 + \cdots + x_n \in B\}$ entonces $S_n^{-1}(B) = \boldsymbol{f}^{-1}(B_n)$ donde $\boldsymbol{f} = (f_1, f_2, \ldots, f_n)$. Sea ν la distribución común de las f_i. En virtud de la hypótesis de independencia la medida producto $\mu \circ f_1^{-1} \times \cdots \times \mu \circ f_n^{-1}$ es igual a $\mu \circ \boldsymbol{f}^{-1}$, donde $\nu = \mu \circ f_i^{-1}$ para cada $i \in \{1, \ldots, n\}$. Luego se cumple para cada $i \in \{1, \ldots, n\}$

$$\int_{S_n^{-1}(B)} f_i \ d\mu = \int_{B_n} x_j \ d\nu(x_1) \cdots d\nu(x_n).$$

Por el teorema de Fubini, se obtiene

$$\int_{S_n^{-1}(B)} f_1 \ d\mu = \int_{S_n^{-1}(B)} f_2 \ d\mu = \cdots = \int_{S_n^{-1}(B)} f_n \ d\mu,$$

luego

$$\begin{aligned} \mathbb{E}(f_1|\sigma(S_n)) &= \mathbb{E}(f_2|\sigma(S_n)) = \cdots = \mathbb{E}(f_n|\sigma(S_n)) \\ &= \frac{1}{n}(n \ \mathbb{E}(f_1|\sigma(S_n)) \ ) = \frac{S_n}{n}. \end{aligned}$$

■

Lema 4.1.2. *Sea (Ω, Σ, μ) un espacio de probabilidad. Sean $\{f_n\}$ una sucesión de funciones Σ-medibles y $S_n = \sum_{k=1}^{n} f_k$. Entonces se cumple*

$$\sigma(S_n, S_{n+1}, S_{n+2}, \ldots) = \sigma(S_n, f_{n+1}, f_{n+2}, \ldots).$$

Demostración: Como $f_{n+k} = S_{n+k} - S_{n+k-1}$ es claro que $S_n, f_{n+1}, f_{n+2}, \ldots$ son $\sigma(S_n, S_{n+1}, S_{n+2}, \ldots)$-medibles, entonces $\sigma(S_n, f_{n+1}, f_{n+2}, \ldots) \subset \sigma(S_n, S_{n+1}, S_{n+2}, \ldots)$. De manera análoga, como $S_{n+k} = S_n + f_{n+1} + \cdots + f_{n+k}$, las funciones $S_n, S_{n+1}, S_{n+2}, \ldots$ son $\sigma(S_n, f_{n+1}, f_{n+2}, \ldots)$-medibles, luego $\sigma(S_n, f_{n+1}, f_{n+2}, \ldots) \subset \sigma(S_n, f_{n+1}, f_{n+2}, \ldots)$.
■

Teorema 4.1.2. (Ley Fuerte de los Grandes Números) *Sea* (Ω, Σ, μ) *un espacio de probabilidad. Sea* $\{f_n\}$ *una sucesión de funciones* Σ*-medibles independientes tal que* $f_n \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ *con* $\int f_n \, d\mu = m$ *para cada* $n \in \mathbb{N}$. *Supongamos que las funciones* f_n *sean idénticamente distribuidas. Entonces se cumplen:*

(i)
$$\lim_{n\to\infty} \frac{S_n}{n} = m \textit{ en c.t.p},$$

(ii)
$$\lim_{n\to\infty} \int \left| \frac{S_n}{n} - m \right| \, d\mu = 0,$$

donde $S_n = \sum_{k=1}^{n} f_k$.

Demostración: Definamos

$$\mathcal{G}_{-n} = \sigma(S_n, S_{n+1}, \ldots), \; \mathcal{G}_{-\infty} = \bigcap_{n=1}^{\infty} \mathcal{G}_{-n}.$$

Por el lema 4.1.2, $\mathcal{G}_{-n} = \sigma(S_n, f_{n+1}, f_{n+2}, \ldots)$. Notamos que $\sigma(f_{n+1}, f_{n+2}, \ldots)$ es independiente de $\sigma(f_1, S_n)$, entonces por la propriedad (11) de $\mathbb{E}(\,\cdot\,|\Sigma_0)$, $\mathbb{E}(f_1|\mathcal{G}_{-n}) = \mathbb{E}(f_1|\sigma(S_n))$. Luego $\mathbb{E}(f_1|\sigma(S_n)) = \frac{S_n}{n}$ por el lema 4.1.1. Por el teorema 3.8.3, $\frac{S_n}{n} = \mathbb{E}(f_1|\mathcal{G}_{-n})$ converge hacia $f_{-\infty} \in \mathcal{L}^1(\Omega, \mathcal{G}_{-\infty}, \mu, \mathbb{R})$ *en c.t.p* y en media. Sea N el conjunto nulo donde no se cumple $\lim_{n\to\infty} \frac{S_n}{n}(\omega) = f(\omega)$ y definamos f_∞ en todo Ω tal que

$$f_{-\infty}(\omega) = \begin{cases} \lim_{n\to\infty} \frac{S_n}{n}(\omega), & si \quad \omega \in \mathcal{G}_\infty \setminus N \\ 0 & si \quad \omega \in N. \end{cases}$$

Como
$$\lim_{n\to\infty} \frac{S_n}{n}(\omega) = \lim_{n\to\infty} \frac{f_{k+1} + \cdots + f_{k+n}}{n}(\omega),$$
La función $f_{-\infty}$ es $\mathcal{T}$-medible donde $\mathcal{T} = \bigcap_{n=1}^{\infty} \sigma(f_{n+1}, f_{n+2}, \ldots)$. Por la ley 0-1 de Kolmogorov, $\mu(\{\omega \in \Omega : f_{-\infty}(\omega) \le x\}) \in \{0, 1\}$ para cada $x \in \mathbb{R}$. Sea $c = \sup\{x : \mu(\{\omega \in \Omega : f_{-\infty}(\omega) \le x\}) = 0\}$. Supongamos en primer lugar que $c \in \{-\infty, \infty\}$. Si $c = -\infty$, entonces $\mu(\{\omega \in \Omega : f_{-\infty}(\omega) = \infty\}) = 1$ lo cual contradice el hecho de que $f_{-\infty}$ es integrable. Si $c = \infty$, entonces $\mu(\{\omega \in \Omega : f_{-\infty}(\omega) \le n\}) = 0$ para cada $n \in \mathbb{N}$ lo cual implica que $f_{-\infty} = \infty$ *en c.t.p* y eso contradice el hecho de que $f_{-\infty}$ es integrable. Concluimos que $c \in \mathbb{R}$, luego $\mu(\{\omega \in \Omega : f_{-\infty}(\omega) \le c - \frac{1}{n}\}) = 0$ para cada $n \in \mathbb{N}$. Entonces

$$\mu\left(\bigcup_{n=1}^{\infty} \{\omega \in \Omega : f_{-\infty}(\omega) \le c - \frac{1}{n}\}\right) = \mu(\{\omega \in \Omega : f_{-\infty}(\omega) < c\}) = 0.$$

Para cada $n \in \mathbb{N}$ también se cumple $\mu(\{\omega \in \Omega : f_{-\infty}(\omega) \leq c + \frac{1}{n}\}) = 1$, entonces

$$\mu\left(\bigcap_{n=1}^{\infty}\{\omega \in \Omega : f_{-\infty}(\omega) \leq c + \frac{1}{n}\}\right) = \mu(\{\omega \in \Omega : f_{-\infty}(\omega) \leq c\}) = 1,$$

por lo tanto $\mu(\{\omega \in \Omega : f_{-\infty}(\omega) = c\}) = 1$, para algun $c \in \mathbb{R}$. Como $\frac{S_n}{n} \longrightarrow f_{-\infty}$ en media y $|\int S_n/n \; d\mu - \int f_{-\infty} \; d\mu| \leq \int |S_n/n - f_{-\infty}| \; d\mu$,

$$m = \int f_1 \; d\mu = \int \frac{S_n}{n} \; d\mu = \int f_{-\infty} \; d\mu = c,$$

entonces $c = m$, lo que significa que $\frac{S_n}{n}$ converge hacia m *en c.t.p* y en media. ∎

4.2. El Teorema de Radon-Nikodým

En esta última sección demostraremos el teorema de Radon-Nikodým usando el teorema de convergencia de martingalas. Para simplicar la demostración del teorema de Radon-Nikodým, supondremos en primer lugar, que la σ-álgebra del espacio de probabilidad considerado es numerable. Luego extendremos el resultado al caso general usando como herramienta la noción de red de Cauchy.

Lema 4.2.1. *Sea (Ω, Σ, μ) un espacio de probabilidad. Supongamos que $\nu : \Sigma \longrightarrow [0, \infty)$ sea una medida finita absolutamente continua con respecto a μ, es decir,*

$$\mu(E) = 0 \Longrightarrow \nu(E) = 0 \qquad \forall \; E \in \Sigma.$$

Entonces dado $\epsilon > 0$, existe $\delta > 0$ tal que

$$\mu(E) < \delta \Longrightarrow \nu(E) < \epsilon \qquad \forall \; E \in \Sigma.$$

Demostración: Supongamos que exista $\epsilon > 0$ tal que para cada $\delta > 0$, $\mu(E) < \delta \Longrightarrow \nu(E) \geq \epsilon$. Entonces existe una sucesión $\{E_n\} \subset \Sigma$ tal que $\mu(E_n) < \frac{1}{2^n} \Longrightarrow \nu(E_n) \geq \epsilon$. El conjunto $\cap_{k=1}^{\infty} \cup_{n=k}^{\infty} E_n$ es Σ-medible. Es fácil ver que $\mu\left(\cap_{k=1}^{\infty} \cup_{n=k}^{\infty} E_n\right) = 0$, pero $\nu\left(\cap_{k=1}^{\infty} \cup_{n=k}^{\infty} E_n\right) \geq \epsilon$, lo cual contradice la hypótesis de continuidad absoluta de ν. ∎

Teorema 4.2.1. *Sea (Ω, Σ, μ) un espacio de probabilidad con $\Sigma = \sigma(\{E_n\})$, donde $E_n \in \Omega$ para cada $n \in \mathbb{Z}^+$. Supongamos que $\nu : \Sigma \longrightarrow [0, \infty)$ sea una medida finita absolutamente continua con respecto a μ. Entonces existe $g \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ tal que*

$$\nu(E) = \int_E g \; d\mu \qquad \forall \; E \in \Sigma.$$

Demostración: Sea $\Sigma_n = \sigma(E_0, E_1, \ldots, E_n)$. En este caso $\Sigma_n = \sigma(\mathcal{P}_n)$ donde $\mathcal{P}_n = \{A_{n_1}, A_{n_2}, \ldots, A_{n_k}\}$ es una partición finita de Ω. Además cada A_{n_i} es un átomo, es decir, el único subconjunto propio Σ_n-medible de cada A_{n_i} es el vacío. Definamos, $f_n : \Omega \longrightarrow [0, \infty]$ tal que

$$f_n(\omega) = \sum_{i=1}^{k} \alpha_{n_i} \chi_{A_{n_i}}(\omega),$$

donde

$$\alpha_{n_i} = \begin{cases} \frac{\nu(A_{n_i})}{\mu(A_{n_i})} & si \quad \mu(A_{n_i}) > 0 \\ 0 & si \quad \mu(A_{n_i}) = 0. \end{cases}$$

Es fácil ver que f_n es Σ_n-medible y integrable porque es constante sobre los átomos de Σ_n.

$$\int_{A_{n_j}} f_n \, d\mu = \int_{A_{n_j}} \sum_{i=1}^{k} \alpha_{n_i} \chi_{A_{n_i}} \, d\mu = \sum_{i=1}^{k} \alpha_{n_i} \mu(A_{n_j} \cap A_{n_i}) = \nu(A_{n_i}).$$

Esto es claro cuando $\mu(A_{n_j}) > 0$,

$$\sum_{i=1}^{k} \alpha_{n_i} \mu(A_{n_j} \cap A_{n_i}) = \frac{\nu(A_{n_j})}{\mu(A_{n_j})} \mu(A_{n_j}) = \nu(A_{n_j}),$$

y cuando $\mu(A_{n_j}) = 0$,

$$\sum_{i=1}^{k} \alpha_{n_i} \mu(A_{n_j} \cap A_{n_i}) = 0 = \nu(A_{n_j}),$$

por la hypótesis de continuidad absoluta de ν. Como cada $E \in \Sigma_n$, es una unión finita de átomos, es decir $E = \cup_{i=1}^{s} A_{n_i}$ (con $A_{n_i} \cap A_{n_j} = \emptyset$ para cada $i \neq j$), se cumple

$$\int_E f_n \, d\mu = \sum_{i=1}^{s} \int_{A_{n_i}} f_n \, d\mu = \sum_{i=1}^{s} \nu(A_{n_i}) = \nu\left(\bigcup_{i=1}^{s} A_{n_i}\right) = \nu(E).$$

Supongamos que $m < n$, entonces $\Sigma_m \subset \Sigma_n$ y se verifica

$$\int_E f_n \, d\mu = \nu(E) = \int_E f_m \, d\mu \qquad \forall \; E \in \Sigma_m,$$

con lo cual, $\mathbb{E}(f_n | \Sigma_m) = f_m$ y por la propriedad (4) de las martingalas, $M = \{f_n\}$ es una martingala. Claro que M es no negativa, entonces por el corolario 3.6.1, $f_n \longrightarrow f_\infty \in \mathcal{L}^1(\Omega, \Sigma_\infty, \mu, \mathbb{R})$ *en c.t.p.* Dado $\epsilon > 0$, por el lema 4.2.1, existe $\delta > 0$

tal que $\mu(E) < \delta \Longrightarrow \nu(E) < \epsilon$. Como $\nu(\Omega) < \infty$, existe $t > 0$ tal que $\frac{\nu(\Omega)}{t} < \delta$. Sea $E_t = \{\omega \in \Omega : f_n(\omega) > t\} \in \Sigma_n$. Por la desigualdad de Chebyshev,

$$\mu(E_t) \leq \frac{1}{t}\int f_n \, d\mu = \frac{\nu(\Omega)}{t} < \delta,$$

lo cual implica que para cada $n \in \mathbb{Z}^+$

$$\int_{E_t} f_n \, d\mu = \nu(E_t) < \epsilon.$$

Entonces M es uniformemente integrable. Por el teorema 3.8.1, $f_n \longrightarrow f_\infty$ en media y $\mathbb{E}(f_\infty|\Sigma_n) = f_n$ para cada $n \in \mathbb{Z}^+$. Sea $\int_{(\cdot)} f_\infty \, d\mu$ una medida sobre Σ. Es claro que

$$\int_E f_\infty \, d\mu = \nu(E) \qquad \forall\ E \in \bigcup_{n=0}^{\infty} \Sigma_n,$$

entonces por el corolario 1.1.1,

$$\int_E f_\infty \, d\mu = \nu(E) \qquad \forall\ E \in \Sigma_\infty = \Sigma.$$

■

Antes de demostrar la versión general del Teorema de Radon-Nikodým conviene definir algunas nociones de análisis funcional. Un conjunto D se dice dirigido si existe una relación de orden parcial $\prec$ sobre D tal que se verifica

$$\alpha, \beta \in D \Longrightarrow \exists\ \gamma \in D \text{ tal que } \alpha \prec \gamma \text{ y } \ \beta \prec \gamma.$$

Una red en un conjunto X es una función $\phi : D \longrightarrow X$ donde D es un conjunto dirigido. Denotaremos la red ϕ por $\{x_\alpha\}_{\alpha \in D}$, donde $x_\alpha = \phi(\alpha)$. Si (X, d) es un espacio métrico, diremos que ϕ es una red de Cauchy en (X, d) si para cada $\epsilon > 0$, existe $\alpha \in D$ tal que se verifica

$$\beta, \gamma \in D; \alpha \prec \beta, \alpha \prec \gamma \Longrightarrow d(\phi(\beta), \phi(\gamma)) < \epsilon.$$

Sean (Ω, Σ, μ) un espacio de probabilidad y Π el conjunto de las particiones finitas de Ω perteneciendo a Σ. Dirijamos el conjunto Π por refinamiento, es decir, dado los elementos $\Gamma, \Delta \in \Pi$, $\Gamma \prec \Delta$ significa que cada $E \in \Gamma$ es unión de algunos conjuntos de Δ. Sea $\pi \in \Pi$ y consideremos la función simple $f_\pi = \sum_{E \in \pi} \alpha_E \chi_E$ donde

$$\alpha_E = \begin{cases} \frac{\nu(E)}{\mu(E)} & si \quad \mu(E) > 0 \\ 0 & si \quad \mu(E) = 0, \end{cases}$$

con ν siendo una medida finita sobre Σ, absolutamente continua con respecto a μ.

Lema 4.2.2. *La función $\phi : \Pi \longrightarrow L^1(\Omega, \Sigma, \mu, \mathbb{R})$ tal que $\phi(\pi) = \boldsymbol{f}_\pi$ es una red de Cauchy.*

Demostración: Supongamos que exista $\epsilon > 0$ tal que para cada $\Gamma \in \Pi$ podemos encontrar $\Delta, \Upsilon \in \Pi$ tal que $\Gamma \prec \Delta$, $\Gamma \prec \Upsilon$ y $\int |f_\Delta - f_\Upsilon|\ d\mu \geq \epsilon$. Por la desigualdad triangular se verifica

$$\epsilon \leq \int |(f_\Delta - f_\Gamma) + (f_\Gamma - f_\Upsilon)|\ d\mu \leq \int |f_\Delta - f_\Gamma|\ d\mu + \int |f_\Gamma - f_\Upsilon|\ d\mu$$

$$\leq 2\ max\left(\int |f_\Delta - f_\Gamma|\ d\mu, \int |f_\Gamma - f_\Upsilon|\ d\mu\right).$$

Luego o bien $\int |f_\Delta - f_\Gamma|\ d\mu \geq \frac{\epsilon}{2}$ o bien $\int |f_\Gamma - f_\Upsilon|\ d\mu \geq \frac{\epsilon}{2}$. Entonces para cada $\Gamma \in \Pi$ podemos encontrar Φ tal que $\Gamma \prec \Phi$ y verificando

$$\int |f_\Gamma - f_\Phi|\ d\mu \geq \frac{\epsilon}{2} \qquad\qquad (*).$$

Fijemos $\Phi_1 \in \Pi$ en virtud de $(*)$. Entonces existe Φ_2 tal que $\Phi_1 \prec \Phi_2$ y verificando $\int |f_{\Phi_1} - f_{\Phi_2}|\ d\mu \geq \frac{\epsilon}{2}$. De forma inductiva obtenemos una sucesión $\{\Phi_n\}$ tal que $\Phi_1 \prec \Phi_2 \prec \Phi_3 \cdots$ y verificando $\int |f_{\Phi_n} - f_{\Phi_{n+1}}|\ d\mu \geq \frac{\epsilon}{2}$ para cada $n \in \mathbb{N}$. Es claro que $\{\sigma(\Phi_n)\}$ es una filtración de (Ω, Σ, μ) con $\Phi_n \in \Pi$ para cada $n \in \mathbb{N}$. Razonando como en la demostración del teorema 4.2.1 se verifica

$$\int_E f_{\Phi_n}\ d\mu = \int_E f_{\Phi_{n+1}}\ d\mu \qquad \forall\ E \in \sigma(\Phi_n),$$

luego $\{f_{\Phi_n}\}$ es una martingala uniformemente integrable entonces por el teorema 3.8.1

$$\lim_{n\to\infty} \int |f_{\Phi_n} - f_{\Phi_{n+1}}|\ d\mu = 0,$$

lo cual es una contradicción. ■

Teorema 4.2.2. (Teorema de Radon-Nikodým) *Sea (Ω, Σ, μ) un espacio de probabilidad. Supongamos que $\nu : \Sigma \longrightarrow [0, \infty)$ sea una medida finita absolutamente continua con respecto a μ. Entonces existe $g \in \mathcal{L}^1(\Omega, \Sigma, \mu, \mathbb{R})$ tal que*

$$\nu(E) = \int_E g\ d\mu \qquad \forall\ E \in \Sigma.$$

Demostración: Por el lema 4.2.2 $\{\boldsymbol{f}_\pi\}_{\pi\in\Pi}$ es una red de Cauchy en $L^1(\Omega,\Sigma,\mu,\mathbb{R})$. Como $L^1(\Omega,\Sigma,\mu,\mathbb{R})$ es un espacio de Banach, dado $\epsilon>0$, existe $f\in\mathcal{L}^1(\Omega,\Sigma,\mu,\mathbb{R})$ tal que podemos encontrar $\Gamma\in\Pi$ verificando $\Gamma\prec\pi\Longrightarrow\int|f_\pi-f|\ d\mu<\epsilon$. Como

$$\left|\int f_\pi\ d\mu-\int f\ d\mu\right|\leq\int|f_\pi-f|\ d\mu<\epsilon,$$

$\{\int f_\pi\ d\mu\}_{\pi\in\Pi}$ es una red de Cauchy en $\mathbb{R}$ que converge hacia $\int f\ d\mu$. Sea $\Gamma\in\Pi$, razonando como en la demostración del teorema 4.2.1, vemos que si $\Phi\in\Pi$ tal que $\Gamma\prec\Phi$

$$\int_E f_\Gamma\ d\mu=\nu(E)=\int_E f_\Phi\ d\mu\qquad\forall\ E\in\sigma(\Gamma).$$

Dejando fijo Γ y pasando al límite de la red se obtiene

$$\int_E f\ d\mu=\lim_\Phi\int_E f_\Phi\ d\mu=\int_E f_\Gamma\ d\mu\qquad\forall\ E\in\sigma(\Gamma),$$

luego $\mathbb{E}(f|\sigma(\Gamma))=f_\Gamma$.
Sean A un elemento arbitrario de Σ y $\Gamma=\{A,A^c\}\in\Pi$. Entonces se verifica

$$\int_A f\ d\mu=\int_A f_\Gamma\ d\mu=\nu(A).$$

■

Para demostrar el teorema de Radon-Nikodým en el caso donde μ y ν son medidas σ-finitas, basta formar una partición de Ω donde μ y ν son finitas sobre cada átomo y luego aplicar el teorema 4.2.2.

Bibliografía

[1] R.B. Ash, *Real Analysis and Probability.* Academic Press 1972.

[2] D.L. Cohn, *Measure Theory.* Birkhäuser 1980.

[3] R. Mansuy, *Histoire de Martingales.* Math & Sci. hum. / Mathematical Social Sciences (43e année, n° 169, 2005(1), p. 105-113).

[4] W. Rudin, *Analisis Real y Complejo.* Alhambra 1979.

[5] G. Vera, *Medidas, Integrales y Martingalas.* Apuntes de Clase, Universidad de Murcia 2007.

[6] D. Williams, *Probability with Martingales.* Cambridge University Press 1991.

[7] N. Young, *An Introduction to Hilbert Space.* Cambridge University Press 1988.