Universidad Nacional Autónoma de México

Facultad de Ciencias

Leyes de los grandes números en espacios de Banach

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

Matemático

PRESENTA:

Ezequiel Severiano Guadalupe

TUTORA

Dra. Ana Meda Guardiola

Ciudad Universitaria, Cd. Mx., 2018

Hoja de datos del jurado

1. **Datos del alumno**
   Severiano
   Guadalupe
   Ezequiel
   (55) 48 56 29 30
   Universidad Nacional Autónoma de México
   Facultad de Ciencias
   Matemáticas
   414072343

2. **Datos del tutor**
   Dra.
   Ana
   Meda
   Guardiola

3. **Datos del sinodal 1**
   Dr.
   Sergio Iván
   López
   Ortega

4. **Datos del sinodal 2**
   Dra.
   María de los Ángeles
   Sandoval
   Romero

5. **Datos del sinodal 3**
   Mat.
   Rodrigo
   Quijón
   Hipólito

6. **Datos del sinodal 4**
   Mat.
   Rafael
   Miranda
   Cordero

7. **Datos del trabajo escrito**
   Leyes de los grandes números en espacios de Banach
   160 p.
   2018

A mi madre y a mi padre

Agradecimientos

Primeramente, agradezco a mis padres por darme la oportunidad de estudiar en la Ciudad de México, por el enorme sacrificio que eso representa. Sobre todo a mi mamá, por su infinito amor y por estar a mi lado todos los días a pesar de la distancia.

A mi papá, por darme su apoyo y confianza y por siempre estar al pendiente de mi bienestar.

A Ananda, Yadira y Orlando, por su invaluable amistad y porque, de un modo u otro, me apoyaron en la realización de este trabajo.

Finalmente, mas no al final, quiero expresar mi más sincero agradecimiento a la Dra. Ana Meda Guardiola, por haber aceptado dirigir esta tesis, por su tiempo y paciencia y por sus sugerencias que sin duda hicieron de este un mejor trabajo.

Índice general

Introducción

Uno de los resultados centrales de la Teoría de Probabilidad es la Ley de los Grandes Números. La primera versión de este teorema apareció en 1713 en el *Ars Conjectandi* de Jacob Bernoulli, a quien le tomó más de 20 años desarrollar una prueba rigurosa de dicho teorema [33]. Desde entonces han aparecido múltiples versiones.

En términos modernos, una ley de grandes números es un enunciado que tiene la siguiente forma:

Enunciado. *Sea $(V_n)_n$ una sucesión de variables aleatorias definidas en un espacio de probabilidad sujetas a un cierto conjunto de condiciones y sea $S_n = \sum_{k=1}^{n} V_k$. Entonces $\frac{1}{n}S_n$ converge en algún sentido a una constante (que suele ser una esperanza).*

Si la convergencia se da en el sentido casi seguro al enunciado anterior se le llama Ley Fuerte, mientras que si la convergencia se da en probabilidad, el resultado será llamado Ley Débil. Las condiciones que generalmente se suelen imponer sobre las variables aleatorias son las de independencia y distribución idéntica, pero por supuesto no se limitan a estas; muchos resultados interesantes han sido obtenidos con restricciones del «tipo ortogonalidad» sobre las variables aleatorias o del «tipo martingalas» sobre la sucesión de sumas parciales.[1]

En 1953, en su tesis doctoral, Edith Mourier probó la primera Ley Fuerte de los Grandes Números para variables aleatorias que toman valores en un espacio de Banach, iniciando [5] con ello el estudio de la Probabilidad en Espacios de Banach. Los primeros trabajos en esta dirección parecían en gran parte paralelos a los teoremas para variables aleatorias reales o complejas, y por esta razón no atrajeron mucha atención (de acuerdo a A. Beck [5]). Sin embargo, a principio de los años sesenta comenzó a desarrollarse toda una teoría que vinculaba la veracidad o falsedad de tales teoremas probabilísticos en los espacios de Banach con la geometría del mismo espacio.

Cada vez más, la teoría parecía indicar que los teoremas con los que estabamos familiarizados eran en realidad teoremas de tipo probabilísticos-geométricos, donde la geometría accidental de los numeros reales y complejos desempeñaba un papel importante. Reducir las hipótesis geométricas requería ajustar

[1]Los conceptos antes mencionados pueden encontrarse en [9].

las probabilísticas.

La búsqueda de teoremas límites llevó a considerar nuevas clases de espacios de Banach, y el análisis minucioso de estos teoremas ha llevado a considerar nuevas condiciones probabilísticas. En términos genéricos, una Ley de Grandes Números en un espacio de Banach tiene la siguiente forma:

Enunciado. *Sea X un espacio de Banach que cumple un cierto conjunto de condiciones. Sea $(V_n)_n$ una sucesión de variables aleatorias con valores en el espacio de Banach X sujetas a un cierto conjunto de condiciones. Entonces $n^{-1}\sum_{k=1}^{n} V_k$ converge (según la topología de la norma) en algún sentido a una función constante.*

Nuestro objetivo en este texto es presentar el material necesario para estudiar algunos de estos teoremas límites en espacios de Banach. En concreto, y siguiendo el artículo de G. Pisier y J. Hoffmann-Jørgensen [17], estudiaremos una versión del teorema anterior cuando la condición geométrica que cumple el espacio es ser de «tipo p» y las condiciones a las que están sujetos las variables aleatorias son la de independencia y una condición sobre sus p-momentos del tipo Kolmogorov llamada condición de Chung.

Remarcamos que nuestro interés en este trabajo es probabilístico por lo que no elaboramos mucho en propiedades que luego no necesitamos.

A continuación presentamos a modo de resumen el contenido de los capítulos en los que se divide este trabajo.

El capítulo 1 está dedicado a presentar definiciones y resultados preliminares que usaremos a lo largo del texto y ha sido dividido en cuatro secciones. En la primera sección introducimos la definición de elemento aleatorio (que es como llamaremos a las variables aleatorias que toman sus valores en un espacio de Banach arbitrario), estudiamos algunas de sus propiedades básicas y formas de caracterizarlos, también introducimos las nociones de convergencia de estos objetos. En la segunda sección estudiamos los conceptos de independencia e indéntica distribución para elementos aleatorios y les caracterizamos a partir del espacio dual.

El estudio de la teoría de probabilidad en espacios más abstractos se hizo posible con la introducción de teorías de integración en dichos espacios, por ello, las últimas dos secciones de este capítulo están dedicadas a estudiar dos formas de definir la esperanza de un elemento aleatorio; primero presentamos el enfoque según la integral de Pettis y después el enfoque según la integral de Bochner. Tales secciones ilustran la fuerte relación existente entre el análisis funcional y la teoría de probabilidad en espacios de Banach. Adicionalmente también dedicamos dos apartados para hablar de la esperanza condicional (existencia y propiedades) según una u otra integral. La bibliografía de este capítulo principalmente se basa en los libros [25] y [36], adicionalmente nos auxiliamos de [2], [7] y [9]. En la sección 1.3 nos apoyamos de [38] mientras que para la sección 1.4 la bibliografía principal fue [11], también se consultaron [12] y [27].

En el capítulo 2 aplicamos parte de la teoría desarrollada en el capítulo anterior para estudiar algunas leyes de los grandes números en espacios de Hilbert. Introducimos también los conceptos de ortogonalidad y no correlación de elementos aleatorios. De entre los teoremas límites que probamos destacan la ley fuerte de los grandes números de Rademacher-Menshov y la ley fuerte de los grandes números de Kolmogorov (caso independiente), así también, incluimos las desigualdades que implican dichos teoremas. Para este capítulo hemos consultado principalmente [29] y [36], también nos auxiliamos en [9].

Finalmente, en el capítulo 3 de este trabajo, hacemos la discusión de teoremas límite a espacios en Banach más generales. El primer resultado que estudiamos es la ley fuerte de los grandes números de Mourier para elementos aleatorios independientes e idénticamente distribuidos con primer momento finito. Posteriormente, en las secciones 3.2 y 3.3, presentamos teoremas límites en los que ahora también imponemos condiciones sobre el espacio. Ejemplo de ello es la ley débil de los grandes números de Taylor para elementos aleatorios idénticamente distribuidos con valores en un espacio de Banach con base de Schauder.

Se ha dedicado la sección 3.3 para presentar la ley fuerte de los grandes números de Pisier y Hoffmann-Jørgensen [17], pero previo a esto y con el fin de seguir un desarrollo cronólogico, presentamos la ley fuerte de los grandes números de Beck [4] (aunque no lo probamos) que establece una condición de convexidad sobre el espacio que resulta equivalente a la extensión de la clásica ley fuerte de los grandes números de Kolmogorov. Este resultado marcó un precedente en el desarrollo de la teoría de probabilidad en espacios de Banach al que se sumaron los trabajos de B. Maurey, G. Pisier [28] y J. Hoffmann-Jørgensen [15], quienes introdujeron [23] la noción del *tipo* y *cotipo* de un espacio de Banach. Es precisamente para presentar el tipo de un espacio de Banach que en la sección 3.3 también nos dedicamos de hablar sobre sumas de Rademacher, así como de algunos resultados concernientes a sumas de elementos aleatorios independientes y/o simétricos. Resultados relevantes al respecto son las desigualdades de Lévy [21], el teorema de Itô-Nisio [19] y las desigualdades de Hoffmann-Jørgensen [15].

Culminamos pues, con la ley fuerte de los grandes números de Pisier y Hoffmann-Jørgensen [17] que establece la una equivalencia entre el *tipo* de un espacio de Banach y la ley fuerte de los grandes números para elementos aleatorios independientes que satisfacen la condición de Chung.

Además de la bibliografía antes indicada, para este capítulo también nos apoyamos en [1], [3], [10], [16], [18] y [20].

Símbolos

$\bigcup_{k=1}^{n} A_k$	Unión de los conjuntos $A_1, \ldots, A_n$
$\biguplus_{k=1}^{n} A_k$	Unión disjunta de los conjuntos $A_1, \ldots, A_n$
$\bigcap_{k=1}^{n} A_k$	Intersección de los conjuntos $A_1, \ldots, A_n$
$A \times B$	Producto cartesiano de los conjuntos A y B
$A \subset B$	El conjunto A está contenido en el conjunto B
$A \subsetneq B$	El conjunto A está contenido propiamente en el conjunto B
$\mathcal{P}(A)$	El conjunto potencia del conjunto A
$\lvert A\rvert, \operatorname{card}(A)$	Cardinalidad del conjunto A
$X \oplus Y$	Suma directa de X e Y, con X, Y espacios vectoriales
$\emptyset$	El conjunto vacío
$\mathbb{N}$	El conjunto de números naturales
$\mathbb{R}$	El conjunto de números reales
$\mathbb{R}_+$	El conjunto de números reales positivos
$\mathfrak{c}$	Cardinalidad de $\mathbb{R}$
$(x_\alpha)_{\alpha\in\Lambda}$	Sucesión de elementos en X indexados por el conjunto $\Lambda \neq \emptyset$, con X un espacio vectorial
$(x_n)_n$	Sucesión de elementos en X indexados por los números naturales, con X un espacio vectorial.
$\ell_p(X)$	Espacio de sucesiones $(x_n)_n$ en X tales que $\sum_{n=1}^{\infty} \lVert x\rVert^p < \infty$
ℓ_p	$\ell_p(X)$ para $X = \mathbb{R}$
$\lVert x\rVert_{\ell_p}$	Norma de un elemento $x = (x_n)_n \in \ell_p(X)$ dada por la expresión $\left(\sum_{n=1}^{\infty} \lVert x_n\rVert^p\right)^{1/p}$
c_0	Espacio de sucesiones (reales) convergentes a 0
$(\Omega, \mathcal{F}, \mathbb{P})$	Espacio de probabilidad
$\mathbb{E}\left[V\right]$	Integral de Lebesgue de la variable aleatoria (real) V en $(\Omega, \mathcal{F}, \mathbb{P})$ respecto a la medida de probabilidad $\mathbb{P}$
$L_p, L_p(\Omega, \mathcal{F}, \mathbb{P})$	Espacio de variables aleatorias reales V en $(\Omega, \mathcal{F}, \mathbb{P})$ tales que $\lVert V\rVert_p := \left(\mathbb{E}\lvert V\rvert^p\right)^{1/p} < \infty$
$d(x, A)$	Distancia del elemento x al conjunto A dada por $\inf\{d(x,a) \,:\, a \in A\}$
$\log_b x$	Logaritmo base b del número real x
$\ln x$	Logaritmo natural del número real x
$\lfloor x \rfloor$	Parte entera del número real x

Capítulo 1

Preliminares

Este capítulo tiene como propósito introducir los conceptos y resultados básicos que usaremos a lo largo del texto. Para ello, vale la pena hacer algunas aclaraciones respecto a nuestros objetos de trabajo así como sobre la notación que utilizaremos.

Convendremos que todos los espacios métricos, normados y de Hilbert con los cuales trabajaremos serán reales, y respecto a los normados estos siempre serán de Banach, a menos que se especifique lo contrario.

Siguiendo la notación del análisis funcional, las letras mayúsculas X, Y representarán espacios normados; usaremos M, N para métricos y H para los de Hilbert. El espacio dual topológico de un espacio normado X estará denotado por X^* y sus elementos serán comúnmente designados por f, g ó x^*.

Si M es un espacio métrico y r es un real positivo, usaremos el símbolo $B(x,r)$ para denotar a la bola abierta con centro en $x \in M$ y radio r, esto es

$$B(x,r) = \{z \in M \mid d(x,z) < r\}.$$

Similarmente, $B[x,r] = \{z \in M \mid d(x,z) \leq r\}$ denota a la bola cerrada con centro en x y radio r, mientras que el símbolo $\mathbb{S}(x,r)$ representará a la esfera con centro en x y radio r, es decir, el conjunto $\{z \in M \mid d(x,z) = r\}$.

La terna $(\Omega, \mathcal{F}, \mathbb{P})$ denotará a un espacio de probabilidad[1] y a los elementos de la σ-álgebra $\mathcal{F}$ les llamaremos eventos (aleatorios). Si $A \in \mathcal{F}$, el número real $\mathbb{P}(A)$ se llama *probabilidad* del evento A. Dado X un espacio vectorial topológico, denotaremos por $\mathcal{B}(X)$ a la σ-álgebra de Borel de X, esto es, la σ-álgebra generada por los subconjuntos abiertos en X.

Una función $T\colon \Omega \to X$ se dice es Borel medible si la imagen inversa $T^{-1}(B) \in \mathcal{F}$ para cada $B \in \mathcal{B}(X)$, y en general, si $(X, \mathcal{S})$ es un espacio de medida, diremos que $T\colon \Omega \to X$ es $(\mathcal{S}, \mathcal{F})$-medible si $T^{-1}(S) \in \mathcal{F}$ para todo $S \in \mathcal{S}$. Si es obvio por el contexto cuáles σ-álgebras estamos considerando simplemente diremos que T es una función medible.

[1]Donde Ω es un conjunto no vacío, $\mathcal{F}$ es una σ-álgebra de subconjuntos de Ω y $\mathbb{P}$ una medida de probabilidad en $\mathcal{F}$.

Hemos preparado dos apéndices (A y B) donde precisamos definiciones y propiedades de los objetos antes mencionados. El resto de la notación será introducido a lo largo del texto.

1.1. Definiciones básicas

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Recordemos que una variable aleatoria es una función real, $V : \Omega \to \mathbb{R}$, tal que para cualquier elemento B de $\mathcal{B}(\mathbb{R})$ se tiene que $V^{-1}(B)$ pertenece a $\mathcal{F}$.

Definiremos a nuestros objetos de estudio, los elementos aleatorios, como toda función Borel medible del espacio Ω en M, donde M es un espacio métrico. Aunque podemos generalizar este concepto a cualquier espacio vectorial topológico vamos a concentrarnos en estudiar el caso para espacios métricos; como veremos más adelante esto nos dará el primer acercamiento a la condición de separabilidad.

Definición 1.1.1. Una función $V : \Omega \to M$ es un **elemento aleatorio** en M, (y lo abreviaremos como e.a.), si $V^{-1}(B) = \{\omega \in \Omega \,|\, V(\omega) \in B\} \in \mathcal{F}$ para cada elemento $B \in \mathcal{B}(M)$.

Bajo esta definición, V es un elemento aleatorio en $\mathbb{R}$ si y solo si V es variable aleatoria. Lo análogo ocurre en $\mathbb{R}^n$, todo elemento aleatorio en $\mathbb{R}^n$ es un vector aleatorio de tamaño n.

Ejemplo 1.1.1. Consideremos $\mathbb{R}^\infty = \{x = (x_1, x_2, \dots) : x_i \in \mathbb{R}\}$ el espacio de las sucesiones de números reales, con la métrica ρ definida para cada $x, y \in \mathbb{R}$ como

$$\rho(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{|x_n - y_n|}{1 + |x_n - y_n|}.$$

Sea $V(\omega) = (V_1(\omega), V_2(\omega), \dots)$ una función Ω en $\mathbb{R}^\infty$, donde $V_n : \Omega \to \mathbb{R}$ para cada n. Entonces V es un elemento aleatorio en $\mathbb{R}^\infty$ si y solo si V_n es variable aleatoria para cada n.

En efecto, como la topología generada por la métrica ρ coincide con la topología producto tenemos que las proyecciones $\pi_n : \mathbb{R}^\infty \to \mathbb{R}$ son continuas para cada n y por lo tanto medibles. Como consecuencia $\pi_n \circ V = V_n$ es una variable aleatoria para toda n siempre que V sea un elemento aleatorio.

Por otro lado, sea $B = \{(\omega_1, \omega_2, \dots) : \omega_j \in A_j, j = 1, \dots, n\}$ donde $A_j \in \mathcal{B}_j(\mathbb{R})$. Entonces B es un rectángulo medible en la σ-álgebra producto $\bigotimes_{j=1}^{\infty} \mathcal{B}_j(\mathbb{R})$. Y como suponemos que V_j es medible para cada j, entonces

$$\begin{aligned} V^{-1}(B) &= \{\omega \in \Omega : V(\omega) \in B\} \\ &= \bigcap_{j=1}^{n} \{x \in \Omega : V_j(\omega) \in A_j\} = \bigcap_{j=1}^{n} V_j^{-1}(A_j) \in \mathcal{F}. \end{aligned}$$

Probando así que V es elemento aleatorio en $\mathbb{R}^\infty$ siempre que sus funciones coordenadas sean variables aleatorias.

Proposición 1.1.1. *Sean M, N espacios métricos y $f\colon M \to N$ una función Borel medible. Entonces si V es un elemento aleatorio en M, $f \circ V = f(V)$ es un elemento aleatorio en N.*

Demostración. Sea B un elemento en $\mathcal{B}(N)$, como f es una función Borel medible, $f^{-1}(B) \in \mathcal{B}(M)$, además por ser V un e.a. en M, $V^{-1}(f^{-1}(B)) \in \mathcal{F}$. Por lo que $(f \circ V)^{-1}(B) = V^{-1}(f^{-1}(B)) \in \mathcal{F}$. Así, $f \circ V = f(V)$ es un e.a. en N. ■

Proposición 1.1.2. *Sea $(E_n)_n$ una partición numerable de Ω, es decir, una sucesión de eventos en $\mathcal{F}$ tales que $E_n \cap E_m = \emptyset$ si $n \neq m$, y $\bigcup_{n=1}^\infty E_n = \Omega$. Si $(x_n)_n$ una sucesión en M, entonces la función $V\colon \Omega \to M$ definida como $V(\omega) = x_n$ si $\omega \in E_n$, es un elemento aleatorio en M.*

Demostración. Sea B un elemento en $\mathcal{B}(M)$. Como $(x_n)_n$ es una sucesión en M y $B \subset M$, denotemos por $(x_{n_k})_k$ a los términos de la sucesión $(x_n)_n$ que se quedan contenidos en B. Afirmamos entonces que

$$V^{-1}(B) = \bigcup_{k=1}^{\infty} E_{n_k}.$$

En efecto, si $\omega \in V^{-1}(B)$, entonces $V(\omega) \in B$. Además, como $(E_n)_n$ forma una partición para Ω tenemos que $\omega \in E_{n_0}$ para alguna n_0, así que $V(\omega) = x_{n_0}$. Pero si $x_{n_0} \in B$, entonces $n_0 = n_i$, para alguna $i \geq 1$ y por tanto $\omega \in E_{n_i} \subset \bigcup_{k=1}^\infty E_{n_k}$.

Por otro lado, si $\omega \in \bigcup_{k=1}^\infty E_{n_k}$, entonces $\omega \in E_{n_j}$ para algún $j \geq 1$. Luego $V(\omega) = x_{n_j} \in B$, lo cual implica que $\omega \in V^{-1}(B)$.

Finalmente, como para cada k el conjunto E_{n_k} pertenece a la σ-álgebra $\mathcal{F}$ concluimos que $V^{-1}(B) \in \mathcal{F}$. ■

Proposición 1.1.3. *Sea $(V_n)_n$ una sucesión de elementos aleatorios en (M, d) tal que para cada $\omega \in \Omega$, $\lim_{n\to\infty} V_n(\omega) = V(\omega)$. Entonces V es un elemento aleatorio en M.*

Demostración. Como $\mathcal{B}(M)$ es también la σ-álgebra generada por los subconjuntos cerrados de M (ver corolario B.1.4), es suficiente probar que $V^{-1}(C) \in \mathcal{F}$ para todo $C \subset M$ cerrado.

Así pues, sea C un conjunto cerrado de M. Veamos que $V^{-1}(C) \in \mathcal{F}$.
Como por hipótesis V_n converge a V puntualmente entonces, para todo entero positivo k, existe $N = N(\omega, k) \in \mathbb{N}$ tal que

$$d(V_n(\omega), V(\omega)) < \frac{1}{k}$$

si $n \geq N$, y esto para cada $\omega \in \Omega$. Por esta razón, para cada $k \in \mathbb{N}$ definimos el conjunto

$$C_k = \bigcup_{x \in C} B\left(x, k^{-1}\right).$$

Entonces C_k es abierto y por lo tanto $C_k \in \mathcal{B}(M)$. Afirmamos que

$$V^{-1}(C) = \bigcap_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} V_n^{-1}(C_k). \tag{1.1}$$

En efecto, si $\omega \in V^{-1}(C)$, entonces $V(\omega) \in C$ y como $V_n(\omega)$ converge a $V(\omega)$ para todo $\omega \in \Omega$ obtenemos que para todo k existe $N > 0$ tal que $d(V_n(\omega), V(\omega)) < \frac{1}{k}$ siempre que $n \geq N$, es decir,

$$V_n(\omega) \in B\left(V(\omega), k^{-1}\right) \subset C_k$$

siempre que $n \geq N$. Así, para cada $k \in \mathbb{N}$ tenemos que

$$\omega \in \bigcap_{n=N}^{\infty} V_n^{-1}(C_k) \subset \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} V_n^{-1}(C_k)$$

y por lo tanto

$$\omega \in \bigcap_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} V_n^{-1}(C_k).$$

Por otro lado, si $\omega \in \bigcap_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} V_n^{-1}(C_k)$, entonces para todo $k \geq 1$, existe $m \geq 1$ tal que $d(V_n(\omega), x) < \dfrac{1}{k}$ para todo $n \geq m$, y para algún $x \in C$. Luego,

$$d(V(\omega), x) \leq d(V(\omega), V_n(\omega)) + d(V_n(\omega), x) < \frac{1}{k} + \frac{1}{k} = \frac{2}{k}$$

si $n \geq \text{máx}\{m, N\}$.

Hemos probado entonces que $0 \leq d(V(w), C) \leq d(V(\omega), x) < \frac{2}{k}$ para toda k, por lo tanto $V(\omega) \in \overline{C} = C$, pues C es cerrado. Concluimos así que ω es elemento de $V^{-1}(C)$.

Finalmente, como V_n es elemento aleatorio en M para cada n y C_k es un conjunto abierto, $V_n^{-1}(C_k) \in \mathcal{F}$. Por la igualdad (1.1) tenemos que $V(C) \in \mathcal{F}$ para todo C cerrado en M. ■

Las leyes de grandes números para espacios normados fueron primero probadas para elementos aleatorios que toman a lo más una cantidad numerable de valores mediante el siguiente lema (el cual dice que la función identidad es aproximable por funciones escalonadas).

Lema 1.1.4. *Sea (M, d) un espacio métrico separable. Dado $\varepsilon > 0$, existe una función $f_\varepsilon \colon M \to M$ (Borel) medible que toma a lo más un conjunto numerable de valores y $d(f_\varepsilon(x), x) < \varepsilon$ para cada $x \in M$.*

Demostración. Sea $\varepsilon > 0$. Como M es separable, existe $D = \{d_n \mid n \geq 1\}$ denso numerable en M. Definimos los conjuntos

$$E_1 = B(d_1, \varepsilon)$$

$$E_n = B(d_n, \varepsilon) \setminus \bigcup_{i=1}^{n-1} B(d_i, \varepsilon) \quad (\text{para } n \geq 2).$$

Afirmamos que $(E_n)_n$ forma una partición para M.

1. Primero veamos que (E_n) forma una sucesión de eventos ajenos por pares. Es claro que para cada n, E_n es un elemento de la σ-álgebra $\mathcal{F}$. Ahora consideremos $m, n \in \mathbb{N}$ índices con $m \neq n$; sin pérdida de generalidad supongamos que $n < m$. Entonces

$$\begin{aligned}
E_n \cap E_m &= \left(B(d_n, \varepsilon) \setminus \bigcup_{i=1}^{n-1} B(d_i, \varepsilon) \right) \cap \left(B(d_m, \varepsilon) \setminus \bigcup_{i=1}^{m-1} B(d_i, \varepsilon) \right) \\
&= \left(B(d_n, \varepsilon) \cap \bigcap_{i=1}^{n-1} (M \setminus B(d_i, \varepsilon)) \right) \cap \left(B(d_m, \varepsilon) \cap \bigcap_{i=1}^{m-1} (M \setminus B(d_i, \varepsilon)) \right) \\
&\subset B(d_n, \varepsilon) \cap \left(\bigcap_{i=1}^{n-1} (M \setminus B(d_i, \varepsilon) \cap (M \setminus B(d_n, \varepsilon) \cap (\bigcap_{i=n+1}^{m-1} (M \setminus B(d_i, \varepsilon)) \right) \\
&\subset B(d_n, \varepsilon) \cap (M \setminus B(d_n, \varepsilon)) \\
&= \emptyset.
\end{aligned}$$

2. Ahora veamos que $M = \bigcup_{n=1}^{\infty} E_n$. Para ello probaremos que

$$\bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} B(d_n, \varepsilon).$$

Observemos que para cada $n \geq 1$, E_n está contenido en $B(d_n, \varepsilon)$ y por tanto $\bigcup_{n=1}^{\infty} E_n \subset \bigcup_{n=1}^{\infty} B(d_n, \varepsilon)$. Por otro lado, si $x \in \bigcup_{n=1}^{\infty} B(d_n, \varepsilon)$ entonces existe $m \in \mathbb{N}$ tal que $x \in B(d_m, \varepsilon)$. Sea m_0 el primer índice tal que $x \in E_{m_0}$ y $x \notin E_j$ para cada $1 \leq j < m_0$, entonces $x \in B(d_{m_0}, \varepsilon) \setminus \bigcup_{j=1}^{m_0-1} E_j = E_{m_0} \subset \bigcup_{n=1}^{\infty} E_n$.

Por la densidad de D en M, tenemos que $M = \bigcup_{n=1}^{\infty} B(d_n, \varepsilon)$. Así, por el paso 2, podemos concluir que $M = \bigcup_{n=1}^{\infty} E_n$.

Entonces por la proposición 1.1.2, para $\Omega = M$ y $\mathcal{F} = \mathcal{B}(M)$, la función $f_\varepsilon(x) = d_n$ si $x \in E_n$ es medible. Más aún, dado cualquier $x \in M$, $x \in E_n$ para algún $n \in \mathbb{N}$ y entonces $f_\varepsilon(x) = d_n$ con $d(f_\varepsilon(x), x) = d(d_n, x) < \varepsilon$. ■

Un elemento aleatorio que toma a lo más una cantidad numerable de valores se llama **elemento aleatorio discreto**. Cuando el espacio es separable el siguiente resultado nos caracteriza a todo elemento aleatorio como límite uniforme de elementos aleatorios discretos.

Proposición 1.1.5. *Sea (M, d) un espacio métrico separable. Una función $V : \Omega \to M$ es un elemento aleatorio en M, si y solo si existe una sucesión $(V_n)_n$ de elementos aleatorios discretos en M tal que V_n converge uniformemente*[2] *a V.*

[2]La sucesión de funciones $(V_n)_n$ converge uniformemente a la función V, si para cada $\varepsilon > 0$ existe $N = N(\varepsilon) \in \mathbb{N}$ tal que $d\,(V_n(\omega) - V(\omega)) < \varepsilon$ para toda $n \geq N$ y toda $\omega \in \Omega$.

Demostración. La implicación directa es consecuencia de la proposición 1.1.3, pues el límite de elementos aleatorios es otra vez un elemento aleatorio.

Por otro lado, supongamos que $V\colon \Omega \to M$ es e.a. en M. Como M es separable, por el lema 1.1.4, para cada $n > 0$, existe una función medible $f_n\colon M \to M$, que toma a lo más una cantidad numerable de valores y satisface que $d(f_n(x), x) < \frac{1}{n}$ para todo $x \in M$.

Para cada $n \geq 1$ definimos la función $V_n = f_n(V)$. De la proposición 1.1.1 tenemos que V_n es un e.a. en M. Sea $\omega \in \Omega$, entonces

$$d(V_n(\omega)), V(\omega)) = d(f_n(V(\omega)), V(\omega)) < \frac{1}{n}$$

para cada $n \geq 1$, por lo tanto $(V_n) = (f_n(V))$ es una sucesión de elementos aleatorios discretos en M que converge uniformemente a V. ■

Comentario. Si X es un espacio de Banach la definición de elemento aleatorio en X que manejaremos en este texto es la dada en la definición 1.1.1, sin embargo esta no es la única a considerar. No son pocos los autores que definen el concepto de elemento aleatorio en un espacio de Banach como una función fuertemente medible del espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ en X. Sin entrar mucho en detalles, una función $V\colon \Omega \to X$ es *fuertemente medible* si existe una sucesión de *funciones discretas (simples)*, $(V_n)_n$, tal que $\lim_n \|V_n(\omega) - V(\omega)\| = 0$ para casi toda $\omega \in \Omega$.

Para espacios de Banach separables, la proposición 1.1.5 muestra que estas dos definiciones de elemento aleatorio son (casi seguramente) la misma. Cuando el espacio X no es separable el rango de V debe ser un subconjunto separable.

Esto último será consecuencia directa del Teorema de Medibilidad de Pettis. Aclararemos esto más adelante, en la sección 1.4, cuando nos interese estudiar la esperanza de un elemento aleatorio mediante la integral de Bochner.

Dado (M, d) espacio métrico, se sabe que la función distancia $d\colon M \times M \to \mathbb{R}$ es continua. Sin embargo si $U, V\colon \Omega \to M$ son elementos aleatorios en M, la composición $d(U, V)$ no es necesariamente una variable aleatoria, es decir, no es una función medible respecto a la σ-álgebra producto en el espacio (producto) $M \times M$. Una condición necesaria para que d sea variable aleatoria es que el espacio M sea separable.

Proposición 1.1.6. *Sean U, V elementos aleatorios en (M,d) espacio métrico. Si M es separable, entonces d(U,V) es una variable aleatoria (en $\mathbb{R}$).*

Demostración. Recordemos que la σ-álgebra producto en el espacio $M \times M$ es $\mathcal{B}(M) \otimes \mathcal{B}(M) = \sigma(\{A \times B \ : \ A, B \in \mathcal{B}(M)\})$. Como por hipótesis M es separable, entonces $\mathcal{B}(M \times M) = \mathcal{B}(M) \otimes \mathcal{B}(M)$ (ver B.2.1). Y debido a que $U, V\colon \Omega \to M$ son elementos aleatorios en M, la función $(U, V)\colon \Omega \to M \times M$ definida como $(U, V)(\omega) = (U(\omega), V(\omega))$ es un elemento aleatorio en $M \times M$. Al ser $d\colon M \times M \to [0, \infty)$ una función continua, es Borel medible y por lo tanto la composición $d \circ (U, V) = d(U, V)\colon \Omega \to M$ es una variable aleatoria. ■

Si (M, d) no es separable, $d(U, V)$ puede no ser variable aleatoria como muestra el siguiente ejemplo.

Ejemplo 1.1.2. Consideremos M el espacio de funciones acotadas de $\mathbb{R}$ en sí mismo, es decir, $M = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ es acotada}\}$ dotado de la norma $\|f\|_M := \sup\{|f(t)| : t \in \mathbb{R}\}$ [3]. Sea N el espacio de funciones de $\mathbb{R}$ en $\{0, 1\}$, $N = \{f : \mathbb{R} \to \{0, 1\}\}$. Entonces, $|M| \geq |N| = 2^{\mathfrak{c}} > \mathfrak{c}$, donde $\mathfrak{c}$ representa la cardinalidad del continuo.

Y consideremos (M, d) el espacio métrico cuya métrica está inducida por la norma, es decir, para cada $f, g \in M$, $d(f, g) = ||f - g||_M$.

Para cada $A \in \mathcal{B}(M) \otimes \mathcal{B}(M)$ definamos

$$\mathbb{P}(A) = \begin{cases} 1 & \text{si} \quad (0,0) \in A \\ 0 & \text{si} \quad (0,0) \notin A. \end{cases}$$

Entonces $(M \times M, \mathcal{B}(M) \otimes \mathcal{B}(M), \mathbb{P})$ es un espacio de probabilidad. Definimos las variables aleatorias $U, V : M \times M \to M$ como las proyecciones en la primera y segunda entrada respectivamente, es decir, para cada (f, g) en $M \times M$,

$$U(f, g) = f \quad \text{y} \quad V(f, g) = g.$$

Note que en efecto U y V son elementos aleatorios en M, pues las proyecciones son funciones continuas.

Afirmamos que $d(U, V) = ||U - V||_M$ no es variable aleatoria.
En efecto, como $\{0\}$ es cerrado en $\mathbb{R}$ entonces $\{0\} \in \mathcal{B}(\mathbb{R})$. Además,

$$\begin{aligned} d(U, V)^{-1}(\{0\}) &= \{(f, g) \in M \times M : d(U(f, g), V(f, g)) = 0\} \\ &= \{(f, g) \in M \times M : |f(t) - g(t)| = 0 \text{ para toda } t \in \mathbb{R}\} \\ &= \{(f, g) \in M \times M : f(t) = g(t) \text{ para toda } t \in \mathbb{R}\} \\ &= \{(f, f) \in M \times M : f \in M\} = \Delta \end{aligned}$$

pero como la cardinalidad de M excede a la cardinalidad del continuo, la proposición B.2.4 nos garantiza que $\Delta \notin \mathcal{B}(M) \otimes \mathcal{B}(M)$ (ver Apéndice B para la prueba de esta afirmación). Por lo tanto $d(U, V)$ no es variable aleatoria.

1.1.1. Algunos modos de convergencia

Debido a lo anterior en las siguientes definiciones de modos de convergencias supondremos que M es un espacio métrico separable. Esto es de utilidad por ejemplo, cuando se está interesado en estudiar convergencia de sucesiones de elementos aleatorios en espacios métricos como $D[0, 1]$: el espacio de funciones reales en $[0, 1]$ que son continuas por la derecha cuyos límites por la izquierda existen (cadlag).[4]

[3] Que no es (norma) separable.

[4] En este caso varias consideraciones deben ser tomadas en cuenta ya que $D[0, 1]$ es en general un espacio «mal portado». Dependiendo de la métrica definida en $D[0, 1]$ este espacio puede ser separable o no, así también con la completez. Al respecto puede consultarse [8].

Definición 1.1.2. Sean $V_1, V_2, \ldots$ y V elementos aleatorios en un espacio métrico separable (M, d). Decimos que V_n **converge a** V

1. **casi seguramente**, $V_n \overset{c.s.}{\to} V$, si
$$\mathbb{P}(\lim_{n\to\infty} d(V_n, V) = 0) = \mathbb{P}(\{\omega \in \Omega : \lim_{n\to\infty} d(V_n(\omega), V(\omega)) = 0\}) = 1$$

2. **en probabilidad**, $V_n \overset{\mathbb{P}}{\to} V$, si para todo $\epsilon > 0$
$$\lim_{n\to\infty} \mathbb{P}\left(d(V_n, V) \geq \epsilon\right) = \lim_{n\to\infty} \mathbb{P}\left(\{\omega \in \Omega : d(V_n(\omega), V(\omega)) \geq \epsilon\}\right) = 0$$

3. **en media r-ésima** o en L^r, $V_n \overset{r}{\to} V$, con $r > 0$, si
$$\lim_{n\to\infty} \mathbb{E}\left[d(V_n, V)^r\right] = 0.$$

Cabe resaltar que hasta ahora no hemos definido la esperanza de un elemento aleatorio, de esta cuestión nos ocupamos en las secciones 1.3 y 1.4. Aquí el símbolo «$\mathbb{E}$» debe intepretarse como la integral de Lebesgue.

A la ley de los grandes números para elementos aleatorios con convergencia en el sentido *casi seguro* la llamaremos **ley fuerte de los grandes números**, mientras que si la convergencia se da en *probabilidad* nos referiremos a ella como **ley débil de los grandes números**.

Como consecuencia de las definiciones encontramos que las relaciones entre modos de convergencia de elementos aleatorios son las mismas que las existentes para variables aleatorias, y no solo eso, los métodos de prueba son los mismos. Ilustraremos esto verificando algunas propiedades entre la convergencia casi segura y en probabilidad.

Primero hagamos notar que $d(U, V)$ es variable aleatoria, más aún, $d(U, V)$ es no negativa, de modo que podemos extender de manera inmediata resultados importantes como el siguiente:

Desigualdad de Markov. Sea $\varepsilon > 0$ y sea $r > 0$ tal que $\mathbb{E}\left[d(U, V)^r\right] < \infty$ (es decir existe), donde U, V son elementos aleatorios en M. Entonces
$$\mathbb{P}[d(U, V) \geq \varepsilon] \leq \frac{1}{\varepsilon^r}\mathbb{E}\left[d(U, V)^r\right] \tag{1.2}$$

Para $r = 2$ la desigualdad (1.2) recibe el nombre de **desigualdad de Chebyshev**. Mediante la desigualdad de Markov obtenemos que convergencia en r-media implica convergencia en probabilidad.

Lema 1.1.7. *Si $(V_n)_n$ es una sucesión de elementos aleatorios en un espacio métrico separable (M, d) tal que $V_n \overset{r}{\to} V$ para alguna $r > 0$, entonces $V_n \overset{\mathbb{P}}{\to} V$.*

Demostración. Sea $\varepsilon > 0$. Por hipótesis tenemos que $\lim_n \mathbb{E}[d(V_n, V)^r] = 0 < \infty$ para alguna $r > 0$, utilizando la desigualdad de Markov obtenemos que
$$\lim_{n\to\infty} \mathbb{P}[d(V_n, V) \geq \varepsilon] \leq \frac{1}{\varepsilon^r} \lim_{n\to\infty} \mathbb{E}[d(V_n, V)^r] = 0.$$

Esto es, $V_n \overset{\mathbb{P}}{\to} V$. ∎

Enseguida extendemos el resultado que dice que convergencia casi segura implica convergencia en probabilidad. Algunas equivalencias para la convergencia casi segura y su relación con el lema de Borel-Cantelli son también establecidas.

Lema 1.1.8. *Sea $(V_n)_n$ una sucesión de elementos aleatorios en un espacio métrico separable (M,d) , tal que $V_n \overset{c.s.}{\to} V$, entonces $V_n \overset{\mathbb{P}}{\to} V$.*

Demostración. Como $\mathbb{P}\left(\lim_{n\to\infty} d(V_n,V)=0\right)=1$, existe un conjunto $\mathcal{N}$ nulo tal que si $\omega \in \Omega\setminus\mathcal{N}$, entonces $\lim_{n\to\infty} d(V_n(\omega),V(\omega))=0$. Por tanto, dados $\omega\in\Omega\setminus\mathcal{N}$ y $\varepsilon>0$, existe $N=N(\omega,\varepsilon)$ tal que $d(V_n(\omega),V(\omega))<\varepsilon$ si $n\geq N$, y por tanto

$$\omega \in \bigcap_{n=N}^{\infty}\{\omega\in\Omega\,:\,d(V_n(\omega),V(\omega))<\varepsilon\}.$$

Luego, si $\omega\in\Omega\setminus\mathcal{N}$, entonces para cada $\varepsilon>0$

$$\omega \in \bigcup_{m=1}^{\infty}\bigcap_{n=m}^{\infty}\{\omega\in\Omega\,:\,d(V_n(\omega),V(\omega))<\varepsilon\},$$

y por consiguiente

$$1=\mathbb{P}\left(\Omega\setminus\mathcal{N}\right)\leq\mathbb{P}\left(\bigcup_{m=1}^{\infty}\bigcap_{n=m}^{\infty}\{\omega\in\Omega\,:\,d(V_n(\omega),V(\omega))<\varepsilon\}\right)\leq 1.$$

Así, para cada m definimos $B_m=\bigcap_{n=m}^{\infty}\{\omega\in\Omega\,:\,d(V_n(\omega),V(\omega))<\varepsilon\}$, entonces $(B_m)_m$ es una sucesión en $\mathcal{F}$ creciente y por tanto

$$\lim_{m\to\infty}\mathbb{P}\left(B_m\right)=\mathbb{P}\left(\bigcup_{m=1}^{\infty}B_m\right)=1,$$

lo cuál implica que $\lim_{m\to\infty}\mathbb{P}\left(\Omega\setminus B_m\right)=0$, pero

$$\Omega\setminus B_m=\bigcup_{n=m}^{\infty}\{\omega\in\Omega\,:\,d(V_n(\omega),V(\omega))\geq\varepsilon\}$$

así que $\{d(V_m(\omega),V(\omega))\geq\varepsilon\}\subset\Omega\setminus B_m$, y por lo tanto

$$\lim_{m\to\infty}\mathbb{P}\left(d(V_m(\omega),V(\omega))\geq\varepsilon\right)\leq\lim_{m}\mathbb{P}\left(\Omega\setminus B_m\right)=0.$$

Esto prueba que V_n converge a V de forma casi segura. ∎

Lema 1.1.9. *Sea $(V_n)_n$ una sucesión de elementos aleatorios en un espacio métrico separable (M,d), entonces $V_n \overset{c.s.}{\to} V$ si y solo si para cada $\varepsilon>0$*

$$\mathbb{P}\left(\{\omega\in\Omega : d(V_n(\omega),V(\omega))>\varepsilon \text{ para una infinidad de } n\}\right)=0.$$

Demostración. Para cada $\varepsilon > 0$ definamos

$$A_\varepsilon = \{\omega \in \Omega : d(V_n(\omega), V(\omega)) > \varepsilon \text{ para una infinidad de } n\}.$$

Hagamos primero la implicación de «ida»: como $V_n \overset{c.s.}{\to} V$, existe un conjunto nulo $\mathcal{N} \in \mathcal{F}$ tal que para cualquier $\omega \notin \mathcal{N}$ se tiene que $V_n(\omega) \to V(\omega)$. Es decir, dado $\varepsilon > 0$ existe $N > 0$ tal que $d(V_n(\omega), V(\omega)) < \varepsilon$ para cualquier $n \geq N$, y para cada ω afuera de $\mathcal{N}$. Ahora obsérvese que esto significa que $\omega \in \Omega \setminus A_\varepsilon$ y en consecuencia $\Omega \setminus \mathcal{N} \subset \Omega \setminus A_\varepsilon$ o equivalentemente, $A_\varepsilon \subset \mathcal{N}$, lo cual implica que $\mathbb{P}(A_\varepsilon) \leq \mathbb{P}(\mathcal{N}) = 0$.
Para la implicación de «regreso»: como para cada $\varepsilon > 0$,

$$\mathbb{P}\left(\{\omega \in \Omega : d(V_n(\omega), V(\omega)) > \varepsilon \text{ para una infinidad de } n\}\right) = 0,$$

entonces para cada $k \in \mathbb{N}$ definimos

$$B_k = \left\{\omega \in \Omega \,:\, d(V_n(\omega), V(\omega)) > \frac{1}{k} \text{ para una infinidad de } n\right\}.$$

Así que $(B_k)_k$ es una sucesión creciente de eventos en $\mathcal{F}$ $(B_k \subset B_{k+1})$ tales que para cada k, $\mathbb{P}(B_k) = 0$. Entonces,

$$\mathbb{P}\left(\bigcup_{k=1}^{\infty} B_k\right) = \lim_{k\to\infty} \mathbb{P}(B_k) = 0.$$

Por lo tanto $\bigcap_{k=1}^{\infty} \Omega \setminus B_k$ es un conjunto de probabilidad 1, donde

$$\Omega \setminus B_k = \left\{\omega \,:\, \text{existe } N(\omega) \text{ tal que } d(V_n(\omega), V(\omega)) \leq \frac{1}{k} \text{ para toda } n \geq N(\omega)\right\}.$$

Así, si $\omega \in \bigcap_{k=1}^{\infty} \Omega \setminus B_k$ entonces para todo $k \geq 1$, existe $N = N(\omega)$ tal que $d(V_n(\omega), V(\omega)) \leq \frac{1}{k}$ para todo $n \geq N$. En particular, dado $\varepsilon > 0$ sea $k \geq 1$ tal que $1 < \varepsilon k$, luego $d(V_n(\omega), V(\omega)) \leq \frac{1}{k}$ para toda $n \geq N(\omega)$ y para alguna $N(\omega) > 0$, y por tanto $d(V_n(\omega), V(\omega)) < \varepsilon$ siempre que $n \geq N(\omega)$. Es decir, $V_n(\omega)$ converge a $V\omega)$ para cada ω en la intersección de los complementos de los B_k, así que

$$\bigcap_{k=1}^{\infty} \Omega \setminus B_k \subset \{\omega \in \Omega \,:\, \lim_{n\to\infty} d(V_n(\omega), V(\omega)) = 0\},$$

y por tanto

$$1 = \mathbb{P}\left(\bigcap_{k=1}^{\infty} \Omega \setminus B_k\right) \leq \mathbb{P}\Big(\{\omega \in \Omega \,:\, \lim_{n\to\infty} d(V_n(\omega), V(\omega)) = 0\}\Big) \leq 1.$$

■

A continuación recordamos una de las versiones del lema de Borel-Cantelli.

Lema 1.1.10 (Borel-Cantelli). *Sea* $(A_n)_n$ *una sucesión en* $\mathcal{F}$ *para la cual* $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$*, entonces*

$$\mathbb{P}\left(\{\omega \in \Omega \,:\, \omega \in A_n \textit{ para una infinidad de } n\}\right) = 0.$$

Demostración. Denotamos por $A = \{\omega \in \Omega \,:\, \omega \in A_n$ para una infinidad de $n\}$. Ahora, para cada $m \geq 1$ definimos $B_m = \bigcup_{n=m}^{\infty} A_n$. Luego $(B_m)_m$ es una sucesión decreciente tal que $A = \bigcap_{m=1}^{\infty} B_m$ y por tanto

$$\begin{aligned}\mathbb{P}(A) = \mathbb{P}\left(\bigcap_{m=1}^{\infty} B_m\right) &= \lim_{m\to\infty} \mathbb{P}(B_m) \\ &= \lim_{m\to\infty} \mathbb{P}\left(\bigcup_{n=m}^{\infty} A_n\right) \leq \lim_{m\to\infty} \sum_{n=m}^{\infty} \mathbb{P}\left(A_n\right) = 0.\end{aligned}$$

■

Corolario 1.1.11. *Sean* $V_1, V_2, \ldots, V$ *elementos aleatorios en un espacio métrico separable* (M, d) *tales que para cada* $\varepsilon > 0$,

$$\sum_{n=1}^{\infty} \mathbb{P}\left(d(V_n(\omega), V(\omega)) > \varepsilon\right) < \infty.$$

Entonces $V_n \overset{c.s.}{\to} V$.

Demostración. Para cada $\varepsilon > 0$ sea

$$A_\varepsilon = \{\omega \in \Omega : d(V_n(\omega), V(\omega)) > \varepsilon \text{ para una infinidad de } n\}.$$

Aplicando lema de Borel-Cantelli (donde $A_n = \{\omega \,:\, d(V_n(\omega), V(\omega)) > \varepsilon\}$) obtenemos que $\mathbb{P}\left(A_\varepsilon\right) = 0$. El lema 1.1.9 garantiza el enunciado. ■

A partir de estos últimos resultados podemos establecer otros bien conocidos, como los siguientes.

Lema 1.1.12. *Sea* $(V_n)_n$ *una sucesión de elementos aleatorios en un espacio métrico separable* (M, d)*. Si existen* $r > 0$ *y* V *un elemento aleatorio en* M *tales que* $\sum_{n=1}^{\infty} \mathbb{E}\left[d(V_n, V)^r\right] < \infty$*, entonces* $V_n \overset{c.s.}{\to} V$.

Demostración. Sea $\varepsilon > 0$, entonces

$$\begin{aligned}\mathbb{P}\left(\bigcup_{m=n}^{\infty} \{\omega \in \Omega \,:\, d\left(V_m(\omega), V(\omega)\right) \geq \varepsilon\}\right) &\leq \sum_{m=n}^{\infty} \mathbb{P}\left(d\left(V_m(\omega), V(\omega)\right) \geq \varepsilon\right) \\ &\leq \sum_{m=n}^{\infty} \frac{1}{r} \mathbb{E}\left[d\left(V_m(\omega), V(\omega)\right)^r \geq \varepsilon\right].\end{aligned}$$

Tomando el límite cuando $n \to \infty$ obtenemos que

$$\mathbb{P}\left(\bigcup_{m=n}^{\infty} \{\omega \in \Omega \,:\, d\left(V_m(\omega), V(\omega)\right) \geq \varepsilon\}\right) = 0,$$

esto es, el conjunto de $\omega \in \Omega$ tal que $d(V_m(\omega), V(\omega)) \geq \varepsilon$ para una infinidad de m tiene probabilidad 0. Por el lema 1.1.9 concluimos que $V_n \overset{c.s.}{\to} V$. ■

Lema 1.1.13. *Sea* $(V_n)_n$ *es una sucesión de elementos aleatorios en un espacio métrico separable* (M, d) *tal que* $V_n \overset{\mathbb{P}}{\to} V$*, entonces existe una subsucesión* $(V_{n_k})_k$ *tal que* $\lim_{k\to\infty} V_{n_k} = V$ *c.s.*

Demostración. Por hipótesis, dados $\varepsilon > 0$ y $\delta > 0$ existe $N = N(\delta) \in \mathbb{N}$ tal que

$$\mathbb{P}(d(V_n, V) > \varepsilon) < \delta.$$

Escojamos entonces una sucesión creciente de indices $n_1 < n_2 < n_3 < \dots$ tal que

$$\mathbb{P}\left(d(V_{n_k}, V) > \frac{1}{k}\right) < \frac{1}{2^k}$$

para cada $k \geq 1$. Entonces por el lema 1.1.10 de Borel-Cantelli

$$\mathbb{P}\left(d(V_{n_k}, V) > \frac{1}{k} \text{ para una infinidad de } k\right) = 0,$$

y por lo tanto afuera de este conjunto nulo tenemos que $\lim_{k\to\infty} V_{n_k} = V$ puntualmente. ■

Recordemos que todo espacio normado $(X, \|\cdot\|)$ es un espacio métrico, en particular la definición 1.1.2 debería ser válida para $d(V_n, V) = \|V_n - V\|$, sin embargo aún no hemos justificado que la suma de dos elementos aleatorios esté bien definida (recuerde el ejemplo 1.1.2). A continuación abordamos esta cuestión por lo que dejamos de lado la discusión para espacios métricos.

1.2. Elementos aleatorios en espacios normados

Sea $(X, \|\cdot\|)$ un espacio vectorial normado y consideremos X^* su espacio dual topológico formado por las funcionales lineales continuas $f\colon X \to \mathbb{R}$. Dada $f \in X^*$, f es continua y por lo tanto Borel medible, de modo que si $V\colon \Omega \to X$ es un elemento aleatorio en X, $f \circ V\colon \Omega \to \mathbb{R}$ es una variable aleatoria. Más aún, si X es separable el siguiente resultado nos dice que el recíproco de esta afirmación es cierto.

Proposición 1.2.1. *Sea* $(X, \|\cdot\|)$ *un espacio normado y separable. Entonces,* V *es un elemento aleatorio en* X *si y solo si* $f(V)$ *es variable aleatoria para cada* $f \in X^*$*.*

Demostración. Es claro que si V es un e.a. en X, $f(V)$ es variable aleatoria para todo $f \in X^*$. Probemos entonces la otra implicación.

Sea $B \in \mathcal{B}(X)$, queremos mostrar que $V\colon \Omega \to X$ es elemento aleatorio, esto es, $V^{-1}(B) \in \mathcal{F}$. Por hipótesis sabemos que dado $f \in X^*$, $f(V)\colon \Omega \to \mathbb{R}$ es una variable aleatoria, por lo que $(f(V))^{-1}(A) = V^{-1}(f^{-1}(A)) \in \mathcal{F}$ para todo $A \in \mathcal{B}(\mathbb{R})$.

De modo que si $B = f^{-1}(B')$ para algún $B' \in \mathcal{B}(\mathbb{R})$ ya habremos terminado. Probaremos entonces que bajo la hipótesis de separabilidad de X, todo elemento de la σ-álgebra de Borel de X es de la forma antes mencionada.

Consideremos la siguiente familia de conjuntos

$$\mathcal{C} = \{f^{-1}(B) \,:\, f \in X^*,\, B \in \mathcal{B}(\mathbb{R})\}$$

y sea $\mathcal{B}(\mathcal{C})$ la σ-álgebra generada por $\mathcal{C}$, $\mathcal{B}(\mathcal{C}) := \sigma(\mathcal{C})$.
Afirmamos que $\mathcal{B}(\mathcal{C}) = \mathcal{B}(X)$.

1. Es claro que $\mathcal{B}(\mathcal{C}) \subset \mathcal{B}(X)$ pues si $A \in \mathcal{C}$ entonces $A = f^{-1}(A')$ donde $f \colon X \to \mathbb{R}$ es continua y $A' \in \mathcal{B}(\mathbb{R})$, por tanto $A \in \mathcal{B}(X)$.

2. Para la otra contención, como X es separable existe $B = \{b_n \,:\, n \geq 1\}$ denso numerable en X. Consideremos además la sucesión $(f_n)_n$ en X^* tal que

$$\|f_n\| = 1, \qquad f_n(b_n) = \|b_n\|, \qquad n = 1, 2, \ldots$$

Tales funciones f_n existen para cada $n \geq 1$ por el teorema de extensión de Hanh-Banach (ver corolario A.2.5).
Probaremos primero que

$$\{x \in X \,:\, \|x\| \leq r\} = \bigcap_{n=1}^{\infty} \{x \in X \,:\, f_n(x) \leq r\}. \tag{1.3}$$

Llamemos B_1 y B_2 respectivamente a los conjuntos en ambos lados de la igualdad anterior. Es claro que $B_1 \subset B_2$, pues si $x \in B_1$ entonces $\|x\| \leq r$, luego $f_n(x) \leq |f_n(x)| \leq \|f_n\| \, \|x\| \leq r$ para todo $n \geq 1$ y por lo tanto $x \in B_2$. Para la otra contención veamos que $X \setminus B_1 \subset X \setminus B_2$.
Sea $x \in X \setminus B_1$, entonces $\|x\| > r$. Como por hipótesis $B = \{b_n\}$ es denso en X, existe $n \in \mathbb{N}$ tal que

$$\|x - b_n\| < \frac{1}{2}(\|x\| - r).$$

Sabemos que $|\|x\| - \|b_n\|| \leq \|x - b_n\|$, entonces

$$\begin{aligned} \|b_n\| &\geq \|x\| - \|x - b_n\| \geq \|x\| - \frac{1}{2}(\|x\| - r) \\ &= \frac{1}{2}(\|x\| + r). \end{aligned}$$

Además

$$\begin{aligned} |f_n(x) - \|b_n\|| &= |f_n(x) - f_n(b_n)| = |f_n(x - b_n)| \\ &\leq \|f_n\| \cdot \|x - b_n\| < \frac{1}{2}(\|x\| - r). \end{aligned}$$

De modo que

$$\begin{aligned} f_n(x) &> \|b_n\| - \frac{1}{2}\left(\|x\| - r\right) \\ &\geq \frac{1}{2}\left(\|x\| + r\right) - \frac{1}{2}\left(\|x\| - r\right) \\ &= r \end{aligned}$$

Por lo tanto $f_n(x) > r$ para alguna $n \in \mathbb{N}$, así $x \in X \setminus B_2$, lo cuál muestra que $X \setminus B_1 \subset X \setminus B_2$, si y solo si $B_2 \subset B_1$. Esto prueba la identidad 1.3.

Notemos además que $B_2 = \bigcap_{n=1}^{\infty}\{x \in X \,:\, f_n(x) \leq r\} = \bigcap_{n=1}^{\infty} f_n^{-1}((-\infty, r])$ y por tanto $B_2 \in \mathcal{B}(\mathcal{C})$, es decir $B_1 \in \mathcal{B}(\mathcal{C})$.
Y como $\mathcal{B}(\mathcal{C})$ es **invariante bajo traslaciones**, entonces tenemos que

$$\{x \in X \,:\, \|x - a\| \leq r\} = B[a, r] \in \mathcal{B}(\mathcal{C}).$$

Al ser X separable, $\mathcal{B}(X) = \sigma(\{\{x \in X \,:\, \|x - a\| \leq r\} \,:\, r > 0,\, a \in X\})$ (ver lema B.1.5) y por la minimalidad de la σ-álgebra de Borel de X concluimos que $\mathcal{B}(X) \subset \mathcal{B}(\mathcal{C})$. Por lo tanto $\mathcal{B}(X) = \mathcal{B}(\mathcal{C})$, probando así el enunciado.

Observación 1.2.1. $\mathcal{B}(\mathcal{C})$ es invariante bajo traslaciones en el siguiente sentido: sea $T_y\colon X \to X$ la traslación por $y \in X$ definida como $T_y(x) = x + y$. Consideremos la traslación por $-a$, T_{-a}, entonces

$$\begin{aligned} B[a, r] = T_{-a}^{-1}(B[0, r]) = T_{-a}^{-1}(B_1) &= T_{-a}^{-1}\Big(\bigcap_{n=1}^{\infty}\{x \in X \,:\, f_n(x) \leq r\}\Big) \\ &= \bigcap_{n=1}^{\infty} T_{-a}^{-1}(\{x \in X \,:\, f_n(x) \leq r\}) = \bigcap_{n=1}^{\infty} T_{-a}^{-1}\left(f_n^{-1}(-\infty, r]\right) \\ &= \bigcap_{n=1}^{\infty} (f_n \circ T_{-a})^{-1}\left((-\infty, r]\right) = \bigcap_{n=1}^{\infty}\{x \in X \,:\, f_n(T_{-a})(x) \leq r\} \\ &= \bigcap_{n=1}^{\infty}\{x \in X \,:\, f_n(x - a) \leq r\} = \bigcap_{n=1}^{\infty}\{x \in X \,:\, f_n(x) - f_n(a) \leq r\} \\ &= \bigcap_{n=1}^{\infty}\{x \in X \,:\, f_n(x) \leq r + f_n(a)\} = \bigcap_{n=1}^{\infty} f_n^{-1}((-\infty, r_n]) \in \mathcal{B}(\mathcal{C}) \end{aligned}$$

donde $r_n = r + f_n(a)$.

■

Corolario 1.2.2. *Sea $(X, \|\cdot\|)$ un espacio normado separable. Si $\alpha \in \mathbb{R}$ y U, V son elementos aleatorios en X, entonces $\alpha U + V$ es un elemento aleatorio en X.*

Demostración. Sea $f \in X^*$, entonces $f(\alpha U + V) = f(\alpha U) + f(V) = \alpha f(U) + f(V)$. Como $f(U), f(V)$ son variables aleatorias, entonces $\alpha f(U) + f(V)$ es variable aleatoria, así, en virtud de la proposición 1.2.1 concluimos que $\alpha U + V$ es elemento aleatorio en X. ■

El ejemplo 1.1.2 también ilustra que la suma de elementos aleatorios en un espacio no separable no necesariamente vuelve a ser elemento aleatorio.

No podemos establecer un resultado análogo para el producto de elementos aleatorios pues en X no tenemos definida una operación de «multiplicación» entre sus elementos, sin embargo sí podemos multiplicar una variable aleatoria con un elemento aleatorio para construir un nuevo elemento aleatorio.

Proposición 1.2.3. *Si V es elemento aleatorio en $(X, \|\cdot\|)$ y A es una variable aleatoria, entonces AV es elemento aleatorio en X.*

Demostración. Sea una sucesión $(A_n)_n$ de variables aleatorias discretos tales que A_n converge puntualmente a A (sabemos que esta existe por la proposición 1.1.5). Como todo espacio normado es un espacio vectorial topológico (ver proposición A.2.1), la operación de multiplicación por escalares es continua, y entonces A_nV es un elemento aleatorio en X, más aún, $\lim_{n\to\infty} A_nV = AV$ (puntualmente). Finalmente, la proposición 1.1.3 nos asegura que AV es un elemento aleatorio en X. ■

Ejemplo 1.2.1. Usando las proposiciones 1.2.1 y 1.2.3 podemos describir a los elementos aleatorios en ciertos espacios. Por ejemplo, consideremos el espacio de las sucesiones reales absolutamente sumables $\ell_1 = \{(x_n)_n : \sum_{n=1}^{\infty} |x_n| < \infty\}$ equipado con la norma $\|x\|_{\ell_1} = \sum_{n=1}^{\infty} |x_n|$ donde $x := (x_n)_n$. Sabemos que $(\ell_1, \|\cdot\|_{\ell_1})$ es un espacio de Banach separable. Entonces, $V = (V_1, V_2, \dots)$ es un elemento aleatorio en ℓ_1 si y solamente si $(V_n)_n$ es una sucesión de variables aleatorios absolutamente sumables.

En efecto, si V es elemento aleatorio en ℓ_1 entonces componiendo con las proyecciones canónicas $\pi_n\colon \ell_1 \to \mathbb{R}$ (que son continuas y por tanto medibles) obtenemos que $V_n = \pi_n \circ V$ es variable aleatoria en $\mathbb{R}$.

Por la proposición 1.2.1, para ver que V es un elemento aleatorio en ℓ_1 basta ver que $f(V)$ es variable aleatoria para cada $f \in (\ell_1)^*$. Pero sabemos que $(\ell_1)^* = \ell_\infty$ (son isométricamente isomorfos), por tanto, si $f \in (\ell_1)^*$ entonces $f((x_n)_n) = \sum_{n=1}^{\infty} a_n x_n$ donde $(a_n)_n =: a$ es una sucesión en ℓ_∞.

Entonces,

$$f(V) = \sum_{n=1}^{\infty} a_n V_n \leq \|a\|_{\ell_\infty} \sum_{n=1}^{\infty} |V_n| < \infty$$

por ser $(V_n)_n$ una sucesión de variables aleatorias absolutamente sumables y $a \in \ell_\infty$. Concluimos que $f(V)$ es variable aleatoria por la proposición 1.1.3 (es el límite de las variables aleatorias $(\sum_{n=1}^{m} a_n V_n)_m$).

1.2.1. Independencia e idéntica distribución

En la proposición 1.2.1 hemos dado una caracterización de todos elementos aleatorios que toman valores en un espacio de Banach separable. Lo siguiente que haremos será extender los conceptos de *idéntica distribución* e *independencia* a elementos aleatorios, y como antes, caracterizarlos mediante el espacio dual.

Dado $V\colon \Omega \to X$ un elemento aleatorio en X, V induce una medida de probabilidad en el espacio de Borel $(X, \mathcal{B}(X))$ de la siguiente manera: para cada $B \in \mathcal{B}(X)$

$$\mathbb{P}_V(B) = \mathbb{P}(V^{-1}(B)) = \mathbb{P}(V \in B).$$

Llamaremos **distribución** o **ley** de V a la medida de probabilidad $\mathbb{P}_V$.

Definición 1.2.1. Sean U y V elementos aleatorios en X. Decimos que U y V son **idénticamente distribuidos** si para cada $B \in \mathcal{B}(X)$

$$\mathbb{P}(U \in B) = \mathbb{P}(V \in B).$$

Diremos que una familia arbitraria de elementos aleatorios en X es idénticamente distribuida si cualesquiera dos elementos de la familia son idénticamente distribuidos.

Definición 1.2.2. Una familia finita de elementos aleatorios $\{V_1, V_2, \ldots, V_n\}$ en X, se dice **independiente** si

$$\mathbb{P}(V_1 \in B_1, \ldots, V_n \in B_n) = \mathbb{P}(V_1 \in B_1)\mathbb{P}(V_2 \in B_2) \cdot \ \ldots \ \cdot \mathbb{P}(V_n \in B_n)$$

para cada $B_1, B_2, \ldots, B_n \in \mathcal{B}(X)$. Una familia arbitraria de elementos aleatorios en M es independiente si todo subconjunto finito de ésta es independiente.

Utilizaremos la abreviación **e.a.i.i.d.** para denotar a una familia de **e**lementos **a**leatorios **i**ndependientes e **i**dénticamente **d**istribuidos.

Lema 1.2.4. *Sean X, Y espacios normados y $\Lambda \neq \emptyset$ un conjunto de índices.*

a) Si $(V_\alpha)_{\alpha\in\Lambda}$ una familia de elementos aleatorios idénticamente distribuidos y $T\colon X \to Y$ es una función medible, entonces $(T(V_\alpha))_{\alpha\in\Lambda}$ es una familia de elementos aleatorios idénticamente distribuidos en Y.

b) Si $(V_\alpha)_{\alpha\in\Lambda}$ una familia de elementos aleatorios independientes y $(T_\alpha)_{\alpha\in\Lambda}$ es una familia de funciones medibles de X en Y, entonces $(T_\alpha(V_\alpha))_{\alpha\in\Lambda}$ es una familia de elementos aleatorios independientes en Y.

Demostración.

a) Sean $\alpha_1, \alpha_2 \in \Lambda$ y $B \in \mathcal{B}(Y)$, como $T\colon X \to Y$ es medible, la proposición 1.1.1 nos garantiza que $T(V_{\alpha_i})$ es elemento aleatorio en Y para $i = 1, 2$. Luego,

$$\begin{aligned}\mathbb{P}[T(V_{\alpha_1}) \in B] &= \mathbb{P}\left[V_{\alpha_1} \in T^{-1}(B)\right] = \mathbb{P}\left[V_{\alpha_2} \in T^{-1}(B)\right] \\ &= \mathbb{P}[T(V_{\alpha_2}) \in B],\end{aligned}$$

por lo tanto $T(V_{\alpha_1})$ y $T(V_{\alpha_2})$ tienen la misma distribución.

b) Sean $\alpha_1, \alpha_2, \ldots, \alpha_n \in \Lambda$ y $B_1, B_2, \ldots, B_n \in \mathcal{B}(Y)$, entonces

$$\begin{aligned}
&\mathbb{P}\left[T_{\alpha_1}(V_{\alpha_1}) \in B_1, T_{\alpha_2}(V_{\alpha_2}) \in B_2, \ldots, T_{\alpha_n}(V_{\alpha_n}) \in B_n\right] \\
&\qquad = \mathbb{P}\left[V_{\alpha_1} \in T_{\alpha_1}^{-1}(B_1), V_{\alpha_2} \in T_{\alpha_2}^{-1}(B_2), \ldots, V_{\alpha_n} \in T_{\alpha_n}^{-1}(B_n)\right] \\
&\qquad = \mathbb{P}\left[V_{\alpha_1} \in T_{\alpha_1}^{-1}(B_1)\right] \cdot \mathbb{P}\left[V_{\alpha_2} \in T_{\alpha_2}^{-1}(B_2)\right] \cdot \ldots \cdot \mathbb{P}\left[V_{\alpha_n} \in T_{\alpha_n}^{-1}(B_n)\right] \\
&\qquad = \mathbb{P}\left[T_{\alpha_1}(V_{\alpha_1}) \in B_1\right] \cdot \mathbb{P}\left[T_{\alpha_2}(V_{\alpha_2}) \in B_2\right] \cdot \ldots \cdot \mathbb{P}\left[T_{\alpha_2}(V_{\alpha_n}) \in B_n\right]
\end{aligned}$$

donde la segunda igualdad se sigue de la hipótesis ya que $T_{\alpha_i}^{-1}(B_i) \in \mathcal{B}(X)$. Por lo tanto $(T_\alpha(V_\alpha))_{\alpha\in\Lambda}$ es una familia de e.a. independientes en Y.

■

Corolario 1.2.5. *Si U y V son e.a.i.i.d. en X, y $f\colon X \to Y$ una función medible, entonces $f(U)$ y $f(V)$ son e.a.i.i.d. en Y.*

Bajo la condición de separabilidad los siguientes dos resultados, el 1.2.6 y el 1.2.7, nos dan el recíproco del lema anterior. Con este propósito introducimos la siguiente definición:

Definición 1.2.3. Sea $\mathcal{U} \subset \mathcal{B}(X)$ y sean $\mathbb{P}_1$ y $\mathbb{P}_2$ dos medidas de probabilidad en $\mathcal{B}(X)$. Diremos que $\mathcal{U}$ es una **familia de unicidad** o **clase determinante** para $\mathcal{B}(X)$ si $\mathbb{P}_1(D) = \mathbb{P}_2(D)$ para cada $D \in \mathcal{U}$ implica que $\mathbb{P}_1 = \mathbb{P}_2$ en $\mathcal{B}(M)$.

Respecto a esta última definición se sabe que bajo la hipótesis de separabilidad del espacio X, la familia

$$\mathcal{C} = \{\{x \in X \,:\, f(x) < t\} \,:\, f \in X^* \text{ y } t \in \mathbb{R}\} \tag{1.4}$$

es una familia de unicidad para X (véase en [14] pág. 128). Nos apoyaremos en este hecho para probar los siguientes resultados.

Proposición 1.2.6. *Sea $(X, \|\cdot\|)$ un espacio normado y separable. Los elementos aleatorios U y V en X son idénticamente distribuidos, si y solo si f(U) y f(V) son variables aleatorias idénticamente distribuidas para cada $f \in X^*$.*

Demostración. Por el lema 1.2.4 si U y V son elementos aleatorios idénticamente distribuidos, entonces $f(U)$ y $f(V)$ son variables aleatorias idénticamente distribuidas para cada $f \in X^*$.

Ahora supongamos que para cada $f \in X^*$, $f(U)$ y $f(V)$ son variables aleatorias idénticamente distribuidas, entonces se verifica que

$$\mathbb{P}\left(f(U) \in B\right) = \mathbb{P}\left(f(V) \in B\right)$$

para cada $B \in \mathcal{B}(\mathbb{R})$ y $f \in X^*$. Si $\mathbb{P}_U = \mathbb{P}_V$ en la familia de unicidad dada en 1.4, donde $\mathbb{P}_U$ y $\mathbb{P}_V$ son las medidas de probabilidad inducidas en $\mathcal{B}(X)$ por U

y V respectivamente, entonces U y V son idénticamente distribuidas. Así,

$$\begin{aligned}
\mathbb{P}_V(\{x : f(x) < t\}) &= \mathbb{P}(V \in \{x : f(x) < t\}) \\
&= \mathbb{P}\big(V \in f^{-1}(-\infty, t)\big) \\
&= \mathbb{P}(f(V) \in (-\infty, t)) \\
&= \mathbb{P}(f(U) \in (-\infty, t)) \\
&= \mathbb{P}\big(U \in f^{-1}(-\infty, t)\big) \\
&= \mathbb{P}(U \in \{x : f(x) < t\}) = \mathbb{P}_U(\{x : f(x) < t\}).
\end{aligned}$$

Por lo tanto, V y U son elementos aleatorios idénticamente distribuidos en X. ■

Proposición 1.2.7. *Sea $(X, \|\cdot\|)$ normado y separable. Los elementos aleatorios V y Z son independientes, si y solo si $f(V)$ y $g(Z)$ son variables aleatorias independientes para cada $f, g \in X^*$.*

Demostración. La parte «solo si» es consecuencia de lema 1.2.4 b). Ahora la parte «si»: Suponemos que para cada $f, g \in X^*$, $f(V)$ y $g(Z)$ son variables aleatorias independientes.
Dados $f \in X^*$ y $t \in \mathbb{R}$ fijos, definimos para cada $B \in \mathcal{B}(X)$,

$$\mathbb{P}_{f,t}(B) = \frac{\mathbb{P}(V \in \{x : f(x) < t\}\ ,\ Z \in B)}{\mathbb{P}(V \in \{x : f(x) < t\})} \tag{1.5}$$

Al respecto, notemos que $\mathbb{P}_{f,t} = \mathbb{P}_Z$ en la familia

$$\mathcal{C} = \{\{x : h(x) < s\} : h \in X^*, s \in \mathbb{R}\}.$$

En efecto, si $C \in \mathcal{C}$

$$\begin{aligned}
\mathbb{P}_{t,f}(C) &= \mathbb{P}_{t,f}(\{x : h(x) < s\}) \\
&= \frac{\mathbb{P}(V \in \{x : f(x) < t\}\ ,\ Z \in \{x : h(x) < s\})}{\mathbb{P}(V \in \{x : f(x) < t\})} \\
&= \frac{\mathbb{P}\big(V \in f^{-1}(-\infty, t)\ ,\ Z \in h^{-1}(-\infty, s)\big)}{\mathbb{P}(V \in f^{-1}(-\infty, t))} \\
&= \frac{\mathbb{P}(f(V) \in (-\infty, t)\ ,\ h(Z) \in (-\infty, s))}{\mathbb{P}(f(V) \in (-\infty, t))} \\
&= \frac{\mathbb{P}(f(V) \in (-\infty, t)) \cdot \mathbb{P}(h(Z) \in (-\infty, s))}{\mathbb{P}(f(V) \in (-\infty, t))} \qquad \text{(por hipótesis)} \\
&= \mathbb{P}(h(Z) \in (-\infty, s)) = \mathbb{P}\big(Z \in h^{-1}(-\infty, s)\big) \\
&= \mathbb{P}_Z(\{x : h(x) < s\}).
\end{aligned}$$

Pero como X es separable y $\mathcal{C}$ es una familia de unicidad para $\mathcal{B}(X)$, entonces $\mathbb{P}_{f,t} = \mathbb{P}_Z$ en $\mathcal{B}(X)$. Utilizando la identidad (1.5) obtenemos que

$$\mathbb{P}_Z(B)\mathbb{P}[V \in \{x : f(x) < t\}] = \mathbb{P}[V \in \{x : f(x) < t\}\ ,\ Z \in B]$$

o equivalentemente,

$$\mathbb{P}_Z(B)\mathbb{P}_V\left[\{x\,:\,f(x)<t\}\right]=\mathbb{P}\left[V\in\{x\,:\,f(x)<t\}\ ,\ Z\in B\right] \tag{1.6}$$

y esto para cada $B\in\mathcal{B}(X), f\in X^*$ y $t\in X^*$.
Ahora fijemos $B_1\in\mathcal{B}(X)$ y definamos para cada $B\in B(X)$

$$\mathbb{P}_1(B)=\frac{\mathbb{P}\left(V\in B\ ,\ Z\in B_1\right)}{\mathbb{P}\left(Z\in B_1\right)}. \tag{1.7}$$

Por la igualdad obtenida en 1.6, obtenemos que

$$\begin{aligned}\mathbb{P}_1(\{x\,:\,h(x)<s\}) &= \frac{\mathbb{P}\left(V\in\{x\,:\,h(x)<s\}\ ,\ Z\in B_1\right)}{\mathbb{P}\left(Z\in B_1\right)}\\ &= \frac{\mathbb{P}\left(V\in\{x\,:\,h(x)<s\}\right)\mathbb{P}_Z\left(B_1\right)}{\mathbb{P}_Z\left(B_1\right)}\\ &= \mathbb{P}_V\left(\{x\,:\,h(x)<s\}\right)\end{aligned} \tag{1.8}$$

para cada $h\in X^*$ y $s\in\mathbb{R}$, es decir, $\mathbb{P}_1=\mathbb{P}_V$ en $\mathcal{B}(X)$ (ya que $\mathcal{C}$ es una familia de unicidad para $\mathcal{B}(X)$). Por lo tanto, para cada $B\in\mathcal{B}(X)$ y para cada $B_1\in\mathcal{B}(X)$ fijo (pero arbitrario)

$$\mathbb{P}_1(B)=P_V(B)=\frac{\mathbb{P}\left(V\in B\ ,\ Z\in B_1\right)}{\mathbb{P}\left(Z\in B_1\right)}$$

es decir,

$$\mathbb{P}\left(V\in B\ ,\ Z\in B_1\right)=\mathbb{P}_V(B)\cdot\mathbb{P}\left(Z\in B_1\right)=\mathbb{P}_V(B)\cdot\mathbb{P}_Z\left(B_1\right)$$

para todo $B,B_1\in\mathcal{B}(X)$, lo cuál prueba que V y Z son idependientes. ■

Ejemplo 1.2.2. A diferencia de la proposición 1.2.6, en el resultado anterior (1.2.7) no es suficiente pedir que para cada $f\in X^*$, $f(V)$ y $f(Z)$ sean variables aleatorias independientes para concluir que V y Z son elementos aleatorios independientes. Consideremos por ejemplo A y B variables aleatorias independientes con distribución normal estándar, i.e. $A,B\sim N(0,1)$, y definamos en $X=\mathbb{R}^2$ los elementos (vectores) aleatorios

$$V=(A,B)\quad \text{y}\quad Z=(B,-A).$$

Recuerde que $(\mathbb{R}^2)^*=\mathbb{R}^2$. Luego, si $f\in(\mathbb{R}^2)^*$ entonces $f((x_1,x_2))=ax_1+bx_2$ donde $a,b\in\mathbb{R}$. Entonces $f(V)=aA+bB$ y $f(Z)=aB-bA$ son variables aleatorias independientes (para ver esta afirmación puede utilizarse el teorema de cambio de variable, ya que $f(V),f(Z)\sim N(0,a^2+b^2)$). Pero $\pi_1((x_1,x_2))=x_1$ y $\pi_2((x_1,x_2))=x_2$ son elementos de $(\mathbb{R}^2)^*$ y sin embargo $\pi_2(V)=B=\pi_1(Z)$ no son independientes, lo cuál implica que V y Z no son elementos aleatorios independientes de acuerdo a la proposición anterior.

Cuando los elementos aleatorios no necesariamente toman valores en el mismo espacio de Banach la noción de independencia dada en la definición 1.2.2 se puede extender de la siguiente forma:

Definición 1.2.4. Una familia de elementos aleatorios $(V_\alpha)_{\alpha\in\Lambda}$ donde $V_\alpha\colon \Omega \to X_\alpha$, X_α es un espacio de Banach (separable), y Λ un conjunto de índices, es independiente si para cada elección (finita) de índices $\alpha_1, \alpha_2, \ldots, \alpha_n \in \Lambda$ y cada elección de borelianos $B_1 \in \mathcal{B}(X_{\alpha_1}), B_2 \in \mathcal{B}(X_{\alpha_2}), \ldots, B_n \in \mathcal{B}(X_{\alpha_n})$ se tiene

$$\mathbb{P}(V_{\alpha_1} \in B_1, \ldots, V_{\alpha_2} \in B_2) = \mathbb{P}(V_{\alpha_1} \in B_1) \cdot \ \ldots \ \cdot \mathbb{P}(V_{\alpha_n} \in B_n).$$

Así pues, para verificar la independencia de una familia arbitraria de elementos aleatorias es suficiente verificar la condición en sus subfamilias finitas, entonces consideremos el *vector aleatorio* formado por elementos elatorios $(V_1, V_2, \ldots, V_n)$ en el espacio $X_1 \times X_2 \times \cdots \times X_n$ y detonemos por $\mathbb{P}_{(V_1,V_2,\ldots,V_n)}$ a su distribución, que llamaremos **distribución conjunta** de los elementos aleatorios $V_1, V_2, \ldots, V_n$.

Proposición 1.2.8. *Los elementos aleatorios $V_1, V_2, \ldots, V_n$ son independientes si y solo si*

$$\mathbb{P}_{(V_1,V_2,\ldots,V_n)} = \mathbb{P}_{V_1} \times \mathbb{P}_{V_2} \times \cdots \times \mathbb{P}_{V_n}.$$

Demostración. Recuerde que $\mathbb{P}_{V_1} \times \mathbb{P}_{V_2} \times \cdots \times \mathbb{P}_{V_n}$ es la medida producto del espacio producto $(X_1 \times \cdots \times X_n, \mathcal{B}(X_1) \otimes \cdots \otimes \mathcal{B}(X_n))$ (por supuesto, la distribución conjunta también está definida en este espacio).

Como $X_1 \times \cdots \times X_n$ es también un espacio de Banach separable, entonces

$$\mathcal{B}(X_1 \times \cdots \times X_n) = \mathcal{B}(X_1) \otimes \cdots \otimes \mathcal{B}(X_n). \tag{1.9}$$

Además, vimos en la prueba de la proposición 1.2.1 que para verificar que dos medidas de probabilidad coinciden en un σ-álgebra generada basta verificar que las medidas que coinciden en una subfamilia cerrada bajo intersecciones finitas y que genere a la σ-álgebra (véase [7] teorema 3.3, pág. 42). Así que por (1.9) basta verificar que $\mathbb{P}_{(V_1,V_2,\ldots,V_n)}(B) = \mathbb{P}_{V_1} \times \mathbb{P}_{V_2} \times \cdots \times \mathbb{P}_{V_n}(B)$ para cada B de la forma $B_1 \times B_2 \times \cdots \times B_n$ donde $B_i \in \mathcal{B}(X_i)$. Pero entonces

$$\begin{aligned}
\mathbb{P}_{(V_1,V_2,\ldots,V_n)}(B) &= \mathbb{P}((V_1, V_2, \ldots, V_n) \in B_1 \times B_2 \times \cdots \times B_n) \\
&= \mathbb{P}(V_1 \in B_1, V_2 \in B_2, \ldots, V_n \in B_n) \\
&= \mathbb{P}(V_1 \in B_1) \cdot \mathbb{P}(V_2 \in B_2) \cdot \ \ldots \ \cdot \mathbb{P}(V_n \in B_n) \\
&= \mathbb{P}_{V_1}(B_1) \cdot \mathbb{P}_{V_2}(B_2) \cdot \ \ldots \ \cdot \mathbb{P}_{V_n}(B_n) \\
&= \mathbb{P}_{V_1} \times \mathbb{P}_{V_2} \times \cdots \times \mathbb{P}_{V_n}(B)
\end{aligned}$$

si y solo si $V_1, V_2, \ldots, V_n$ son independientes. ∎

Vale la pena tener presente las siguientes definiciones de independencia y idéntica distribución de sucesiones de elementos aleatorios (vistos como procesos estocásticos) que nos serán de utilidad más adelante.

Definición 1.2.5 ([18]). Diremos que **dos sucesiones de elementos aleatorios** $(V_i)_{i\in I}$ y $(U_i)_{i\in I}$ (donde $V_i, U_i\colon \Omega \to X_i$ son elementos aleatorios en X_i, X_i es un espacio de Banach, I un conjunto de índices) **son idénticamente distribuidas** si para cada elección (finita) de índices $i_1, i_2, \dots, i_n \in I$ los elememtos aleatorios (*vectores*) $(V_{i_1}, V_{i_2}, \dots, V_{i_n})$ y $(U_{i_1}, U_{i_2}, \dots, U_{i_n})$ son idénticamente distribuidos.

Obsérvese que por la proposición 1.2.8 si $(V_i)_{i\in I}$ y $(U_j)_{j\in J}$ son sucesiones cada una de elementos aleatorios independientes tales que para cada $i \in I$, V_i y U_i son idénticamente distribuidas, entonces $(V_i)_{i\in I}$ y $(U_j)_{j\in J}$ son idénticamente distribuidas.

En efecto, sea $n \in \mathbb{N}$, por independencia de $(V_i)_{i\in\mathbb{N}}$ y la proposición 1.2.8: $\mathbb{P}_{(V_1,\dots,V_n)} = \mathbb{P}_{V_1} \times \cdots \times \mathbb{P}_{V_n}$, análogamente $\mathbb{P}_{(U_1,\dots,U_n)} = \mathbb{P}_{U_1} \times \cdots \times \mathbb{P}_{U_n}$. El resultado entonces se sigue por la idéntica distribución entre los elementos aleatorios V_i y U_i para cada $i \in \mathbb{N}$.

Definición 1.2.6 ([18]). **Dos sucesiones de elementos aleatorios** $(V_i)_{i\in I}$ y $(U_j)_{j\in J}$ (donde $V_i\colon \Omega \to X_i$, $U_j\colon \Omega \to Y_j$ con X_i y Y_j espacios de Banach, I y J conjuntos de índices) **son independientes** si para cada elección (finita) de índices $i_1, i_2, \dots, i_n \in I$ y $j_1, j_2, \dots, j_m \in J$ los elememtos aleatorios (*vectores*) $(V_{i_1}, V_{i_2}, \dots, V_{i_n})$ y $(U_{j_1}, U_{j_2}, \dots, U_{j_m})$ son independientes.

También será de utilidad considerar el concepto de variable aletoria Rademacher.

Definición 1.2.7. Una sucesión de variables aleatorias $(\varepsilon_n)_n$ es una **sucesión de Rademacher** si es independiente y para cada n,

$$\mathbb{P}(\varepsilon_n = +1) = \frac{1}{2} = \mathbb{P}(\varepsilon_n = -1).$$

Y a ε_n se le llama **variable aleatoria Rademacher.**

1.3. Esperanza como integral de Pettis

En la literatura existen (principalmente) dos maneras de definir el valor esperado o esperanza de un elemento aleatorio con valores en un espacio de Banach: mediante la integral de Pettis y mediante la integral de Bochner. En esta sección estudiaremos a la esperanza de un elemento aleatorio mediante la integral de Pettis. Como antes, consideremos $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y X un espacio de Banach separable (aunque para la siguiente definición esta hipótesis no es requerida).

Definición 1.3.1. Una función $\xi\colon \Omega \to X$ es **Pettis integrable sobre** un conjunto $E \in \mathcal{F}$ si y solo si existe x_E en X tal que

$$f(x_E) = \int_E f(\xi(\omega))\, d\mathbb{P}$$

para toda $f \in X^*$. Al elemento x_E se le llama integral de Pettis sobre E de ξ y se denota por $\int_E \xi(\omega)\, d\mathbb{P} = x_E$.[5] Diremos que f es **Pettis integrable** si y solo si es Pettis integrable sobre cada $E \in \mathcal{F}$.

Obsérvese que la integral del lado derecho de la igualdad anterior es la de Lebesgue (la composición $f(\xi)$ es una función de Ω en $\mathbb{R}$). Antes de proceder a probar propiedades de esta integral introduciremos el concepto de esperanza de un elemento aleatorio, esto nos facilitará la notación.

Definición 1.3.2. Sea V elemento aleatorio en X. Supongamos que existe un elemento $m \in X$ tal que

$$\mathbb{E}[f(V)] = \int_\Omega f(V)(\omega)\, d\mathbb{P} = f(m)$$

para todo $f \in X^*$. Llamaremos a m el **valor esperado** o **esperanza** del elemento aleatorio V, y lo denotaremos por $\mathbb{E}[V]$.[6]

Notemos que si $\mathbb{E}[V] = m$ existe, entonces es único. En efecto, supongamos que $m \neq x_0$, como todo espacio normado tiene suficientes funcionales continuos (ver corolario A.2.6 del teorema de Hahn-Banach) entonces existe $\varphi \in X^*$ tal que $\mathbb{E}[\varphi(V)] = \varphi(m) \neq \varphi(x_0)$. Por tanto, $f(x_0)$ no es igual a $\mathbb{E}[f(V)]$ para todo $f \in X^*$.

El problema entonces radica en garantizar la existencia de la esperanza (integral de Pettis) de V. Aunque no profundizaremos en la teoría de integración de Pettis sí daremos dos criterios de existencia. El más útil de ellos, el lema 1.3.3, relaciona a la integral de Bochner (que estudiamos en la siguente sección) con la integral de Pettis.

Definición 1.3.3. La **varianza** de un elemento aleatorio $V\colon \Omega \to X$ se define como

$$\operatorname{Var}(V) = \sigma_V^2 = \mathbb{E}\left[\|V - \mathbb{E}\,[V]\,\|^2\right] = \int_\Omega \|V - \mathbb{E}\,[V]\,\|^2\, d\mathbb{P}$$

y la **desviación estándar** de V, denotada por σ_V, es la raíz cuadrada no negativa de la varianza.

Observe que $\operatorname{Var}(V)$ es un número real mientras que $\mathbb{E}\,[V]$ es un vector (elemento de X).

1.3.1. Propiedades de la esperanza

Como consecuencia de la definición 1.3.2 obtenemos las siguientes propiedades.

[5]Es común encontrar la notación $(P)\int_E \xi(\omega)\, d\mathbb{P}$ para denotar a la integral de Pettis de ξ sobre E. Resulta ser útil para evitar confusiones entre ésta y la integral de Lebesgue (u otras). Creemos que tal notación no es necesaria si especificamos el codominio de la función a integrar.

[6]Obsérvese que estamos usando el símbolo «$\mathbb{E}$» tanto para la integral de Lebesgue como para la integral de Pettis. El abuso de notación queda justificado por el hecho de que la integral de Pettis generaliza a la de Lebesgue.

Teorema 1.3.1. *Sea X un espacio normado y separable. Sean V, W y Z elementos aleatorios en X y sea $x \in X$.*

a) Si $\mathbb{E}[W]$ y $\mathbb{E}[Z]$ existen y $W + Z$ es un elemento aleatorio[7] *en X, entonces $\mathbb{E}[W] + \mathbb{E}[Z]$ existe y además $\mathbb{E}[W + Z] = \mathbb{E}[W] + \mathbb{E}[Z]$.*

b) Si $\mathbb{E}[V]$ existe, entonces $\mathbb{E}[bV] = b\mathbb{E}[V]$ para cada $b \in \mathbb{R}$.

c) Si $V = x$ c.s., entonces $\mathbb{E}[V] = x$.

d) Si $V = x$ c.s. y A es una variable aleatoria tal que $\mathbb{E}[A] < \infty$, entonces $\mathbb{E}[AV] = \mathbb{E}[A]x$.

e) Sea $h \colon X \to Y$ una función lineal y continua donde Y es cualquier espacio de Banach. Si $\mathbb{E}[V]$ existe, entonces $\mathbb{E}[h(V)] = h(\mathbb{E}[V])$.

f) Si $\mathbb{E}[V]$ existe, entonces $\|\mathbb{E}[V]\| \leq \mathbb{E}\left[\|V\|\right]$.

Demostración.

a) Ya que $\mathbb{E}[W]$ y $\mathbb{E}[Z]$ existen, consideremos el elemento $x_0 = \mathbb{E}[W] + \mathbb{E}[Z]$ en X. Así, sea $f \in X^*$, entonces

$$\begin{aligned} f(x_0) &= f(\mathbb{E}[W] + \mathbb{E}[Z]) = f(\mathbb{E}[W]) + f(\mathbb{E}[Z]) \\ &= \mathbb{E}[f(W)] + \mathbb{E}[f(Z)] \quad \text{(pues por hipótesis } \mathbb{E}[W] \text{ y } \mathbb{E}[Z] \text{ existen)} \\ &= \mathbb{E}[f(W) + f(Z)] \qquad \text{(por la linealidad de la integral de Lebesgue)} \\ &= \mathbb{E}[f(W + Z)] \end{aligned}$$

de modo que existe un elemento $x_0 \in X$, tal que para todo $f \in X^*$, $f(x_0) = \mathbb{E}[f(W + Z)]$ y por lo tanto $\mathbb{E}[W + Z] = x_0 = \mathbb{E}[W] + \mathbb{E}[Z]$.

b) Tomemos $x_0 = b\mathbb{E}[V]$ elemento en X y sea $f \in X^*$, entonces

$$\begin{aligned} f(x_0) &= f(b\mathbb{E}[V]) = bf(\mathbb{E}[V]) = b\mathbb{E}[f(V)] = \mathbb{E}[bf(V)] \\ &= \mathbb{E}[f(bV)] \end{aligned}$$

Por tanto $x_0 = b\mathbb{E}[V]$ es la esperanza del elemento aleatorio bV.

c) Observemos que dado $f \in X^*$ se satisface

$$\{\omega \in \Omega \,:\, V(\omega) = x\} \subseteq \{\omega \in \Omega \,:\, f(V(\omega)) = f(x)\}$$

Además, como por hipótesis $V = x$ c.s., entonces

$$1 = \mathbb{P}(\{\omega \in \Omega \,:\, V(\omega) = x\}) \leq \mathbb{P}(\{\omega \in \Omega \,:\, f(V(\omega)) = f(x)\}) \leq 1$$

y por lo tanto $\mathbb{P}(\{\omega \in \Omega \,:\, f(V(\omega)) = f(x)\}) = \mathbb{P}(f(V) = f(x)) = 1$, así que $f(V) = f(x)$ c.s. y en consecuencia

$$\mathbb{E}[f(V)] = \int_\Omega f(V)(\omega)\, d\mathbb{P}(\omega) = \int_\Omega f(x)\, d\mathbb{P}(\omega) = f(x),$$

por lo tanto $x = \mathbb{E}[V]$.

[7] Esta última hipótesis solo es necesaria cuando no suponemos la separabilidad del espacio X.

d) Por hipótesis $\mathbb{E}[A] \in \mathbb{R}$, entonces elegimos $x_0 := \mathbb{E}[A]x \in X$ y sea $f \in X^*$, entonces

$$\begin{aligned} f(x_0) = f(\mathbb{E}[A]x) = \mathbb{E}[A]f(x) = \mathbb{E}[f(x)A] &= \int_\Omega f(x)A(\omega)\, d\mathbb{P}(\omega) \\ = \int_\Omega f(V)(\omega)A(\omega)\, d\mathbb{P}(\omega) &= \int_\Omega f(VA)(\omega)\, d\mathbb{P}(\omega) = \mathbb{E}[f(AV)] \end{aligned}$$

por lo tanto $\mathbb{E}[AV] = x_0 = \mathbb{E}[A]x$.

e) Como $h\colon X \to Y$ es lineal y continua, por la proposición 1.1.1 $h(V)$ es un elemento aleatorio en Y. Ahora, sea $g \in Y^*$, $g\colon Y \to \mathbb{R}$, entonces la composición $g \circ h\colon X \to \mathbb{R}$ es lineal y continua, de modo que $g \circ h \in X^*$ para cada $g \in Y^*$. Entonces,

$$\begin{aligned} g(h(\mathbb{E}[V])) &= (g \circ h)(\mathbb{E}[V]) \\ &= \mathbb{E}[(g \circ h)(V)] \quad \text{(pues por hipótesis } \mathbb{E}[V] \text{ existe y } g \circ h \in X^*) \\ &= \mathbb{E}[g(h(V))]. \end{aligned}$$

Por lo tanto $h(\mathbb{E}[V])$ es la esperanza del elemento aleatorio $h(V)$ en Y, es decir, $h(\mathbb{E}[V]) = \mathbb{E}[h(V)]$.

f) Claramente si $\mathbb{E}[V] = \mathbf{0}$ entonces $\|\mathbb{E}[V]\| = 0 \leq \mathbb{E}\|V\|$. Así, supongamos que $\mathbb{E}[V] \neq 0$, entonces por el corolario A.2.5 del Teorema de Extensión de Hahn-Banach, existe $f \in X^*$ tal que $\|f\| = 1$ y $f(\mathbb{E}[V]) = \|\mathbb{E}[V]\|$. Entonces

$$\begin{aligned} \|\mathbb{E}V\| = f(\mathbb{E}[V]) &\leq |f(\mathbb{E}[V])| = |\mathbb{E}[f(V)]| \\ &\leq \mathbb{E}|f(V)| \leq \mathbb{E}\left[\, \|f\|\|V\| \,\right] = \mathbb{E}\|V\| \end{aligned}$$

por lo tanto $\|\mathbb{E}V\| \leq \mathbb{E}\|V\|$.

■

Definición 1.3.4. Diremos que un elemento aleatorio V en X es de **orden p débil** o **débilmente p-integrable** si

$$\mathbb{E}\left[|f(V)|^p\right] = \int_\Omega |f(V)|^p \; d\mathbb{P} < \infty$$

para todo $f \in X^*$. Y será de **orden p fuerte** o **fuertemente p-integrable** si

$$\mathbb{E}\left[\|V\|^p\right] = \int_\Omega \|V\|^p \, d\mathbb{P} < \infty.$$

Como podemos intuir, ser de orden fuerte implica ser de orden débil. En efecto, como cada $f \in X^*$ es una función continua, existe $M > 0$ tal que $|f(x)| \leq M\|x\|$ para cada $x \in X$, y por monotonía de la integral de Lebesgue,

$$\mathbb{E}[|f(V)|^p] \leq \mathbb{E}[M^p\|V\|^p] = M^p\mathbb{E}[\|V\|^p].$$

Comentario. También se dice que la distribución de V es de orden p débil. Respectivamente para el orden p fuerte. Y se puede probar que el orden p débil implica el orden p fuerte cuando el espacio X es de dimensión finita (ver por ejemplo [38] pág. 104 corolario de Teo. 2.1).

Introducimos ahora los espacios «L_p» pero para elementos aleatorios (funciones de Ω en X).

Definición 1.3.5. Denotamos por $L_p(\Omega, \mathcal{F}, \mathbb{P}; X)$, o en breve $L_p(\Omega; X)$, al conjunto de clases de equivalencias de elementos aleatorios en X (con X separable) que tienen orden p fuerte para $p \in (0, \infty)$.

$L_p(\Omega; X)$ resulta tener estructura de espacio vectorial y es normado con la L_p-norma (para $p \geq 1$) dada por

$$\|V\|_p := \left(\mathbb{E}\left[\|V\|^p\right]\right)^{\frac{1}{p}} = \left(\int_\Omega \|V\|^p \, d\mathbb{P}\right)^{\frac{1}{p}}$$

para cada $V \in L_p(\Omega; X)$. Más aún, $L_p(\Omega; X)$ es de Banach si X lo es.

Por $L_\infty(\Omega, \mathcal{F}, \mathbb{P}; X) = L_\infty(\Omega; X)$ entenderemos al espacio de clases de equivalencias de elementos aleatorios esencialmente acotados con la norma

$$\|V\|_\infty = \operatorname{ess\,sup} \|V(\omega)\|.$$

A los espacios espacios de Lebesgue clásicos (cuando $X = \mathbb{R}$) los denotaremos por $L_p(\Omega; \mathbb{R}), L_p(\mathbb{R})$ o simplemente L_p. Para más información sobre el orden débil, fuerte, y los espacios $L_p(\Omega; X)$ remitimos al lector al capítulo 2 sección 2 de [38].

Un caso particularmente interesante cuando $\mathbb{E}[V]$ existe es cuando V es débilmente integrable y el espacio X es reflexivo.

Proposición 1.3.2. *Sea X un espacio de Banach reflexivo. Si V es un elemento aleatorio de orden 1 débil, entonces $\mathbb{E}[V]$ existe. Más aún, V es Pettis integrable.*

Demostración. Para cada $A \in \mathcal{F}$, definimos el operador lineal $T\colon X^* \to \mathbb{R}$ dado por

$$T(f) = \int_A f(V) \, d\mathbb{P}$$

para cada $f \in X^*$. Afirmamos que T es continuo. Para ello veamos que tiene gráfica cerrada. Sea $(f_n)_n$ una sucesión en X^* tal que $\text{lím}_n f_n = f$ y $\text{lím}_n T(f_n) = y$. Como V es de orden 1 débil y $f_n \in X^*$ para cada n, entonces la composición $f_n(V)$ es integrable, así, por el teorema de convergencia dominada

$$\begin{aligned} y &= \lim_{n\to\infty} T(f_n) = \lim_{n\to\infty} \int_A f_n(V) \, d\mathbb{P} \\ &= \int_A \lim_{n\to\infty} f_n(V) \, d\mathbb{P} = \int_A f(V) \, d\mathbb{P} \\ &= T(f). \end{aligned}$$

Por lo tanto T tiene gráfica cerrada. El teorema de la gráfica cerrada (ver A.2.3) garantiza que T es continuo y por tanto $T \in X^{**}$. Pero como X es reflexivo, existe $m \in X$ tal que $T(f) = f(m)$. Es decir, $\mathbb{E}[V\mathbb{I}_A]$ existe para cada $A \in \mathcal{F}$ y por lo tanto V es Pettis integrable. En particular, si $A = \Omega$, $\mathbb{E}[V]$ existe. ■

El siguiente es un criterio que utilizaremos a lo largo del texto para garantizar la existencia de la esperanza de un elemento aleatorio, e ilustra la relación entre la integral de Pettis y la integral de Bochner (que presentamos en la siguiente sección).

Lema 1.3.3. *Sea X un espacio de Banach separable, y sea $V : \Omega \to X$ un elemento aleatorio. Si $\mathbb{E}[\|V\|] < \infty$, entonces $\mathbb{E}[V]$ existe.*

Demostración. Hagamos dos casos.

Caso 1. Si $V : \Omega \to X$ es un elemento aleatorio discreto, digamos, $V(\omega) = a_n$ si $\omega \in A_n$, donde $a_k \neq a_m$ si $k \neq m$ y $\Omega = \bigcup_{n=1}^{\infty} A_n$. Por hipótesis $\mathbb{E}\|V\| < \infty$, es decir

$$\mathbb{E}\|V\| = \int_{\Omega} \|V\| \, d\mathbb{P} = \sum_{n=1}^{\infty} \|a_n\| \mathbb{P}[V = a_n] < \infty.$$

Como X es de Banach, toda serie absolutamente convergente es convergente y por lo tanto

$$\sum_{n=1}^{\infty} a_n \mathbb{P}[V = a_n] \text{ converge en } X$$

denotemos por x a este elemento, es decir, $x = \sum_{n=1}^{\infty} a_n \mathbb{P}[V = a_n]$. Ahora, sea $f \in X^*$, entonces

$$\begin{aligned} f(x) &= f\left(\sum_{n=1}^{\infty} a_n \mathbb{P}[V = a_n]\right) = f\left(\lim_{m\to\infty} \sum_{n=1}^{m} a_n \mathbb{P}[V = a_n]\right) \\ &= \lim_{m\to\infty} f\left(\sum_{n=1}^{m} a_n \mathbb{P}[V = a_n]\right) = \lim_{m\to\infty} \sum_{n=1}^{m} f(a_n) \mathbb{P}[V = a_n] \\ &= \sum_{n=1}^{\infty} f(a_n) \mathbb{P}[V = a_n] = \mathbb{E}[f(V)] \end{aligned}$$

Por lo tanto $x \in X$ es la esperanza de V.

Caso 2. Sea $V : \Omega \to X$ un e.a. arbitrario, por la proposición 1.1.5 existe $(V_n)_n$ una sucesión de e.a. discretos tales que V_n converge uniformemente V. Entonces dado $\varepsilon > 0$ existe $N > 0$ tal que $\|V - V_n\| < \varepsilon$ siempre que $n \geq N$. Utilizando la desigualdad del triángulo inversa; $|\,\|V\| - \|V_n\|\,| \leq \|V - V_n\|$, tenemos lo siguiente

$$\begin{aligned}
-\varepsilon < \|V\| - \|V_n\| < \varepsilon \\
-\varepsilon < \mathbb{E}[\|V\| - \|V_n\|] < \varepsilon \\
-\varepsilon < \mathbb{E}\|V\| - \mathbb{E}\|V_n\| < \varepsilon
\end{aligned}$$

De modo que $|\mathbb{E}\|V\| - \mathbb{E}\|V_n\|| < \varepsilon$ para cada $n \geq N$, y como por hipótesis $\mathbb{E}\|V\| < \infty$, entonces $\mathbb{E}\|V_n\|$ es finito (existe) para cada $n \geq N$. Así, consideramos la sucesión $(V_m)_{m \geq N}$, por el Caso 1 tenemos que $\mathbb{E}[V_m]$ existe para cada m.

Afirmamos que $(\mathbb{E}[V_m])_m$ es una sucesión de Cauchy. En efecto, como $(V_n)_n$ es una sucesión convergente, $(V_m)_{m \geq N}$ también lo es, y por lo tanto es de Cauchy. Luego, dado $\varepsilon > 0$, $\|V_p - V_q\| < \varepsilon$ si $p, q \geq N_0$ para alguna $N_0 \geq 1$, entonces por el teorema 1.3.1 incisos (a) y (f),

$$\|\mathbb{E}[V_p] - \mathbb{E}[V_q]\| = \|\mathbb{E}[V_p - V_q]\| \leq \mathbb{E}\left[\|V_p - V_q\|\right] < \varepsilon.$$

Pero como X es de Banach existe $x \in X$ tal que $\lim\limits_{m \to \infty} \mathbb{E}[V_m] = x$. Veamos ahora que $x = \mathbb{E}[V]$. Sea $f \in X^*$, entonces

$$\begin{aligned}
f(x) &= f\left(\lim_{m \to \infty} \mathbb{E}[V_m]\right) \\
&= \lim_{m \to \infty} f(\mathbb{E}[V_m]) \\
&= \lim_{m \to \infty} \mathbb{E}[f(V_m)] \qquad \text{(pues la esperanza de } V_m \text{ existe para cada } m) \\
&= \mathbb{E}\left[\lim_{m \to \infty} f(V_n)\right] \qquad \text{(por teo. de Convergencia Dominada)} \\
&= \mathbb{E}[f(V)]
\end{aligned}$$

Por tanto $x = \mathbb{E}[V]$. Concluimos entonces que $\mathbb{E}[V]$ existe. ∎

Observación 1.3.1. Podemos aplicar el **Teorema de Convergencia Dominada de Lebesgue** a la sucesión de funciones $(f(V_n))_m$ pues:

1. $f(V_m)\colon \Omega \to \mathbb{R}$ es medible para cada m, al ser composición de funciones medibles.

2. Como $V_m \to V$ y f es continua, entonces $f(V_m) \to f(V)$, donde V es medible por ser límite de funciones medibles. Luego $f(V)$ es también medible.

3. Como $f \in X^*$, f satisface que $|f(V_m(\omega))| \leq \|f\|\|V_m(\omega)\|$.
Observemos que por ser f continuo, $\|f\| \leq M$. Por otra parte, $(V_m)_m$ converge a V, así que para $\varepsilon = 1$ tenemos que $\|V_m - V\| < 1$ si $m \geq K$ para alguna $K > 0$. Utilizando la desigualdad del triángulo inversa, si $m \geq K$ entonces $\|V_m\| - \|V\| < 1$, es decir, $\|V_m\| < 1 + \|V\|$.

De modo que eligiendo $L = \max\{1 + \|V\|, \|V_N\|, \|V_{N+1}\|, \ldots, \|V_{K-1}\|\}$ obtenemos que $\|V_m\| < L$ para toda m, es decir, es una sucesión acotada.

Entonces, $|f(V_m(\omega))| \leq \|f\|\|V_m(\omega)\| < ML = L_0$ para cada $\omega \in \Omega$. Al ser L_0 una constante, es Lebesgue integrable.

Por lo tanto $\lim_{m\to\infty} \mathbb{E}[f(V_m)] = \mathbb{E}\left[\lim_{m\to\infty} f(V_m)\right]$.

1.3.2. Esperanza condicional I

Adicionalmente podemos definir el concepto de esperanza condicional para elementos aleatorios. En contraste con el caso clásico, esta nueva esperanza condicional no debe su existencia al teorema de Radon-Nikodým (pues en primer lugar deberíamos justificar que tal teorema es posible para funciones definidas de Ω en X), sino a un teorema de extensión de operadores lineales.

En concreto, la idea es definir la esperanza condicional primero en un subespacio denso de elementos aleatorios, mostrar que actúa de manera lineal y continua en él, para después mediante el teorema A.2.7 extenderlo a todo el espacio. El subespacio del que hablamos es el de las funciones simples. Recordamos a continuación esta definición.

Definición 1.3.6. Un elemento aleatorio $V\colon \Omega \to X$ se llama simple si y solo si existe $n \in \mathbb{N}$, existen $x_1, x_2, \ldots, x_n \in X$ distintos, y existen $A_1, A_2, \ldots, A_n \in \mathcal{F}$ ajenos dos a dos con $\mathbb{P}(A_i) < \infty$ tal que

$$V(\omega) = \sum_{i=1}^{n} x_i \mathbb{I}_{A_i}.$$

A la expresión del lado derecho de la igualdad anterior se le conoce como *representación de V*.

Como consecuencia de la definición todo elemento aleatorio simple es Pettis integrable, más aún,

$$\mathbb{E}[V\mathbb{I}_A] = \sum_{i=1}^{n} x_i \mathbb{P}(A \cap A_i)$$

para cada $A \in \mathcal{F}$. En efecto, si $f \in X^*$ entonces

$$\begin{aligned}
\mathbb{E}[V\mathbb{I}_A] &= \int_A f(V)\, d\mathbb{P} = \int_A f\left(\sum_{i=1}^{n} x_i \mathbb{I}_{A_i}\right) d\mathbb{P} = \int_A \sum_{i=1}^{n} f\left(x_i \mathbb{I}_{A_i}\right) d\mathbb{P} \\
&= \sum_{i=1}^{n} \int_A f\left(x_i\right) \mathbb{I}_{A_i}\, d\mathbb{P} = \sum_{i=1}^{n} \int_{A\cap A_i} f\left(x_i\right)\, d\mathbb{P} = \sum_{i=1}^{n} f\left(x_i\right) \mathbb{P}(A \cap A_i) \\
&= f\left(\sum_{i=1}^{n} x_i \mathbb{P}(A \cap A_i)\right).
\end{aligned}$$

Ahora, para que esta sea una buena definición, no debe depender de la representación de V, es decir, si $V = \sum_{j=1}^{m} y_j \mathbb{I}_{B_j}$ (donde $y_j \in X$, $B_j \in \mathcal{F}$ y $\mathbb{P}(B_j) < \infty$ y $B_k \cap B_l = \emptyset$ si $l \neq k$) entonces para cada $A \in \mathcal{F}$,

$$\mathbb{E}[V\mathbb{I}_A] = \sum_{i=1}^{n} x_i \mathbb{P}(A \cap A_i) = \sum_{j=1}^{m} y_j \mathbb{P}(B_j \cap A).$$

Para ver esto podemos suponer sin pérdida de generalidad que $(A_i)_{i=1}^n$ y $(B_j)_{j=1}^m$ son particiones de A (cada sucesión es de elementos disjuntos cuya unión da A), luego $A_i = A \cap A_i = \uplus_{j=1}^m B_j \cap A_i$ y por tanto

$$\begin{aligned}
\sum_{i=1}^{n} x_i \mathbb{P}(A \cap A_i) &= \sum_{i=1}^{n} x_i \mathbb{P}(\uplus_{j=1}^m B_j \cap A_i) = \sum_{i=1}^{n} x_i \sum_{j=1}^{m} \mathbb{P}(A_i \cap B_j) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} x_i \mathbb{P}(A_i \cap B_j) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} y_i \mathbb{P}(A_i \cap B_j) \\
&= \sum_{j=1}^{m} y_j \sum_{i=1}^{n} \mathbb{P}(A_i \cap B_j) = \sum_{j=1}^{m} y_j \mathbb{P}(A \cap B_j)
\end{aligned}$$

Note que la cuarta igualdad se da pues si $A_i \cap B_j \neq \emptyset$, entonces $x_i = y_j$.

Proposición 1.3.4. *Sea X un espacio de Banach separable. El conjunto de elementos aleatorios simples, $S = \{s \colon \Omega \to X \ : \ s \text{ es simple}\}$, es denso en $L_1(\Omega; X)$.*

Demostración. Sea $\varepsilon > 0$, como X es separable sea $D = \{d_n \, : \, n \geq 1\}$ un denso numerable. Para cada $n \in \mathbb{N}$ definimos los conjuntos $A_n = \{x \, : \, \|x - d_n\| < \frac{\varepsilon}{4}\}$, observe que son tales que $\bigcup_{n=1}^{\infty} A_n = X$. Y como en el lema 1.1.4 los ajenizamos de la siguiente manera:

$$\begin{aligned}
B_1 &= A_1 \\
B_n &= A_n \setminus \bigcup_{i=1}^{n-1} A_i \quad \text{(para } n \geq 2).
\end{aligned}$$

De este modo $\{B_n\}$ forma una partición numerable de X que además satisface que $\operatorname{diam}(B_n) < \frac{\varepsilon}{2}$ para cada n (pues $B_n \subset A_n$ donde A_n es una bola abierta de radio $\frac{\varepsilon}{4}$).

Sea $V \in L_1(\Omega; X)$, y definamos $S \colon \Omega \to X$ como $S(\omega) = \sum_{n=1}^{\infty} x_n \mathbb{I}_{\{V \in B_n\}}(\omega)$ donde $x_n \in B_n$. Sea $\omega \in \Omega$, entonces

$$\|V(\omega) - S(\omega)\| = \left\| V(\omega) - \sum_{n=1}^{\infty} x_n \mathbb{I}_{\{V \in B_n\}}(\omega) \right\| < \frac{\varepsilon}{2},$$

pues necesariamente $V(\omega) \in B_{n_0}$ para alguna $n_0 \in \mathbb{N}$, como además $x_{n_0} \in B_{n_0}$ y $\operatorname{diam}(B_{n_0}) < \frac{\varepsilon}{2}$,

$$\|V(\omega) - S(\omega)\| = \|V(\omega) - x_{n_0}\| < \frac{\varepsilon}{2},$$

por lo tanto $S \in L_1(\Omega; X)$, o sea, $\mathbb{E}[\|S\|] = \sum_{n=1}^{\infty} \|x_n\| \mathbb{P}(V \in B_n) < \infty$. Así que existe $N > 0$ tal que

$$\sum_{n=N+1}^{\infty} \|x_n\| \mathbb{P}(V \in B_n) < \frac{\varepsilon}{2}.$$

Entonces elegimos a la función simple $S' = \sum_{n=1}^{N} x_n \cdot \mathbb{I}_{\{V \in B_n\}}$, esta satisface que

$$\|V - S'\|_1 = \mathbb{E}[\|V - S'\|] \leq \mathbb{E}[\|V - S\|] + \mathbb{E}[\|S - S'\|] < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Por tanto, el conjunto de elementos aleatorios simples es denso en $L_1(\Omega; X)$. ∎

Corolario 1.3.5. *El conjunto de funciones simples es denso en $L_p(\Omega; X)$ para cada $p \in [1, \infty)$.*

Demostración. Para ello basta observar que $L_p(\Omega; X) \subset L_1(\Omega; X)$ y además $S \subset L_p(\Omega; X)$ para cada $p \geq 1$. ∎

Implícitamente, en la prueba del lema 1.3.3, usamos que el espacio de las funciones simples es denso en L_1.

Definición 1.3.7. Sea $V\colon (\Omega, \mathcal{F}, \mathbb{P}) \to X$ un elemento aleatorio, y sea $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$. La esperanza condicional de V dada la σ-álgebra $\mathcal{G}$ es el elemento aleatorio $Z\colon \Omega \to X$ único c.s., medible respecto a $\mathcal{G}$ y tal que satisface

$$\mathbb{E}[Z\mathbb{I}_B] := \int_B Z \, d\mathbb{P} = \int_B V \, d\mathbb{P} =: \mathbb{E}[V\mathbb{I}_B] \tag{1.10}$$

para cada $B \in \mathcal{G}$. Designaremos por $\mathbb{E}[V | \mathcal{G}]$ o también $\mathbb{E}^{\mathcal{G}}[V]$ a la esperanza condicional de V dado $\mathcal{G}$.

Hacemos hincapié en que las integrales de la identidad (1.10) son de Pettis.

Teorema 1.3.6. *Sea X un espacio de Banach separable y V un elemento aleatorio en X. Sea $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$. Si $V \in L_1(\Omega; X)$, entonces $\mathbb{E}[V | \mathcal{G}]$ existe.*

Demostración. Supongamos que V es simple, es decir, $V = \sum_{n=1}^{m} x_n \cdot \mathbb{I}_{A_n}$ donde $x_n \in A_n$, $A_n \in \mathcal{F}$ y $A_i \cap A_j = \emptyset$ si $i \neq j$. Afirmamos que

$$\mathbb{E}[V | \mathcal{G}] = \sum_{n=1}^{m} x_n \mathbb{E}[\mathbb{I}_{A_n} | \mathcal{G}],$$

donde $\mathbb{E}[\mathbb{I}_A|\,\mathcal{G}]$ es la esperanza condicional de $\mathbb{I}_A\colon \Omega \to \mathbb{R} \in L_p(\mathbb{R})$, la cual existe debido al teorema de Radon-Nikodým (y que además define a la probabilidad condicional). En efecto,

$$\begin{aligned}\int_B \sum_{n=1}^m x_n \mathbb{E}[\mathbb{I}_{A_n}|\,\mathcal{G}]\,d\mathbb{P} &= \sum_{n=1}^m \int_B x_n \mathbb{E}[\mathbb{I}_{A_n}|\,\mathcal{G}]\,d\mathbb{P} = \sum_{n=1}^m x_n \int_B \mathbb{E}[\mathbb{I}_{A_n}|\,\mathcal{G}]\,d\mathbb{P}\\ &\overset{(*)}{=} \sum_{n=1}^m x_n \int_B \mathbb{P}(A_n|\,\mathcal{G})\,d\mathbb{P} \overset{(*)}{=} \sum_{n=1}^m x_n \mathbb{P}(A_n\cap B)\,d\mathbb{P}\\ &= \int_B V\,d\mathbb{P}.\end{aligned}$$

(Comprobar que $\mathbb{E}[V|\mathcal{G}]$ no depende de la representación V es análogo a lo que hicimos cuando probamos que $\mathbb{E}[V]$ no dependía de tal representación). Más aún, $\mathbb{E}[\,\cdot\,|\mathcal{G}]$ es lineal en $\mathcal{S} = \{S\colon \Omega \to X \;:\; S \text{ es simple}\}$. Tomemos U y W funciones simples, $a, b \in \mathbb{R}$ y sea $C \in \mathcal{G}$, entonces

$$\begin{aligned}\int_C (aU + bW)\,d\mathbb{P} &= \int_C aU\,d\mathbb{P} + \int_C bW\,d\mathbb{P}\\ &= a\int_C U\,d\mathbb{P} + b\int_C W\,d\mathbb{P}\\ &= \int_C a\mathbb{E}[\,U|\mathcal{G}]\,d\mathbb{P} + \int_C b\mathbb{E}[\,W|\mathcal{G}]\,d\mathbb{P}\\ &= \int_C \left(a\mathbb{E}[\,U|\mathcal{G}] + b\mathbb{E}[\,W|\mathcal{G}]\right)\,d\mathbb{P}\end{aligned}$$

Es decir, $\mathbb{E}[aU + bW|\mathcal{G}] = a\mathbb{E}[\,U|\mathcal{G}] + b\mathbb{E}[\,W|\mathcal{G}]$ casi seguramente[8]. Además, si $V \in S$, entonces

$$\begin{aligned}\|\mathbb{E}[\,V\,|\mathcal{G}]\|_1 &= \int_\Omega \left\|\sum_{n=1}^m x_n \mathbb{E}[\mathbb{I}_{A_n}|\,\mathcal{G}]\right\|\,d\mathbb{P} \le \int_\Omega \sum_{n=1}^m \|x_n\|\,\mathbb{E}[\mathbb{I}_{A_n}|\,\mathcal{G}]\,d\mathbb{P}\\ &= \sum_{n=1}^m \|x_n\| \int_\Omega \mathbb{E}[\mathbb{I}_{A_n}|\,\mathcal{G}]\,d\mathbb{P} = \sum_{n=1}^m \|x_n\|\,\mathbb{P}(A_n\cap\Omega) = \|V\|_1.\end{aligned}$$

Es decir, $\mathbb{E}[\,\cdot\,|\mathcal{G}]\colon S \to L_1(\Omega, \mathcal{G}, \mathbb{P}; X)$ es un operador lineal y acotado (por tanto continuo) definido un subespacio denso de $L_1(\Omega; X)$, por lo tanto $\mathbb{E}[\,\cdot\,|\mathcal{G}]$ tiene una extensión lineal y continua en $L_1(\Omega; X)$ que igualmente denotaremos por $\mathbb{E}[\,\cdot\,|\mathcal{G}]\colon L_1(\Omega, \mathcal{F}, \mathbb{P}; X) \to L_1(\Omega, \mathcal{G}, \mathbb{P}; X)$. Tal operador es llamado **operador de esperanza condicional con respecto a la σ-álgebra $\mathcal{G}$.** ■

***Comentario.** Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, $\mathcal{G}$ una sub-σ-álgebra de de $\mathcal{F}$; y sea $A \in \mathcal{F}$ fijo pero arbitrario. Entonces existe una función

[8] Note que solo usamos la linealidad de la integral de Pettis establecida en el teorema 1.3.1 y no que U y W sean simples, así que esta prueba funciona para el caso general una vez establecida la existencia de la esperanza condicional.

$\mathbb{P}(A|\mathcal{G})\colon (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, llamada probabilidad condicional de A dado $\mathcal{G}$, tal que

$$\int_B \mathbb{P}(A|\mathcal{G})\, d\mathbb{P} = \mathbb{P}(B \cap A)$$

para cada $B \in \mathcal{G}$. Cualesquiera dos funciones que cumplan lo anterior deben coincidir casi seguramente, y de hecho $\mathbb{P}(A|\mathcal{G}) = \mathbb{E}[\mathbb{I}_A|\mathcal{G}]$ c.s.

Como hemos comentado antes, la existencia de la probabilidad condicional está garantizada por el teorema de Radon-Nikodým (tomando $\lambda(B) = \mathbb{P}(A \cap B)$ para cada $B \in \mathcal{G}$, λ resulta ser una función numerablemente aditiva en $\mathcal{G}$ y absolutamente continua respecto a $\mathbb{P}$).

En la siguiente sección estudiaremos más en detalle el teorema de Radon-Nikodým para funciones de Ω en X y justificaremos que el método de prueba del teorema anterior es necesario pues daremos un ejemplo (en la proposición 1.4.14) que muestra que existe una función $\mathbf{F}$ σ-aditiva con valores en X y defina en $\mathcal{F}$ tal que $\mathbf{F}(A) = \mathbf{0}$ siempre que $\mathbb{P}(A) = 0$ pero que no es la integral indefinida de ninguna función con respecto a $\mathbb{P}$.

Teorema 1.3.7. *Sea X un espacio de Banach separable. Sea $V \in L_1(\Omega, \mathcal{F}, \mathbb{P}; X)$ y consideremos $\mathcal{H} \subset \mathcal{G}$ dos sub-σ-álgebras de $\mathcal{F}$. Entonces,*

1. $\mathbb{E}\left[\mathbb{E}\left[V|\mathcal{G}\right]\right] = \mathbb{E}\left[V\right]$ *c.s.*
2. $\mathbb{E}[aU + bV|\mathcal{G}] = a\mathbb{E}[U|\mathcal{G}] + b\mathbb{E}[V|\mathcal{G}]$ *c.s., donde $a, b \in \mathbb{R}$ y $U \in L_1(\Omega, \mathcal{F}, \mathbb{P}; X)$.*
3. *Si $\mathcal{A} = \{\emptyset, X\}$, entonces $\mathbb{E}\left[V|\mathcal{A}\right] = \mathbb{E}[V]$ c.s.*
4. *Si $V \in L_1(\Omega, \mathcal{G}, \mathbb{P}; X)$, entonces $\mathbb{E}\left[V|\mathcal{G}\right] = V$ c.s.*
5. $\mathbb{E}\left[\mathbb{E}\left[V|\mathcal{G}\right]|\mathcal{H}\right] = \mathbb{E}\left[V|\mathcal{H}\right]$ *c.s.*
6. *Si $T\colon X \to Y$ es un operador lineal continuo, donde Y es otro espacio de Banach separable, entonces $\mathbb{E}\left[T(V)|\mathcal{G}\right] = T\left(\mathbb{E}\left[V|\mathcal{G}\right]\right)$ c.s.*

Demostración.

1. Es claro haciendo $B = \Omega$ en (1.10), es decir,

$$\int_\Omega \mathbb{E}\left[V|\mathcal{G}\right]\, d\mathbb{P} = \int_\Omega X\, d\mathbb{P} = \mathbb{E}\left[V\right].$$

2. Ya está hecho (ver prueba del teorema 1.3.6).
3. Basta observar que

$$\int_A V\, d\mathbb{P} = \int_A \mathbb{E}\left[V\right]\, d\mathbb{P}$$

   para $A \in \mathcal{A} = \{\emptyset, \Omega\}$.

4. Como V es medible respecto a $\mathcal{G}$, V ya satisface (1.10), i.e.,

$$\int_B V\, d\mathbb{P} = \int_B V\, d\mathbb{P}$$

para cada $B \in \mathcal{G}$, así que $\mathbb{E}[V|\mathcal{G}] = V$. Por cumplir esta propiedad se dice que el operador $\mathbb{E}[\cdot|\mathcal{G}]$ define una proyección de $L_1(\Omega, \mathcal{F}, \mathbb{P}; X)$ en $L_1(\Omega, \mathcal{G}, \mathbb{P}; X)$, y no solo eso, $\|\mathbb{E}[\cdot|\mathcal{G}]\| = 1$.

5. Dado $B \in \mathcal{H} \subset \mathcal{G}$, entonces

$$\int_B \mathbb{E}[\mathbb{E}[V|\mathcal{G}]|\mathcal{H}]\, d\mathbb{P} = \int_B \mathbb{E}[V|\mathcal{G}]\, d\mathbb{P} = \int_B V\, d\mathbb{P}.$$

Por tanto $\mathbb{E}[\mathbb{E}[V|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[V|\mathcal{H}]$. Adicionalmente, por el inciso anterior: $\mathbb{E}[\mathbb{E}[V|\mathcal{H}]|\mathcal{G}] = \mathbb{E}[V|\mathcal{H}]$.

6. Es consecuencia de la definición 1.3.7 y del teorema 1.3.1 inciso (e) (la integral de Pettis «conmuta» con operadores lineales continuos). Así:

$$\begin{aligned}
\int_B T(\mathbb{E}[V|\mathcal{G}])\, d\mathbb{P} &= T\left(\int_B \mathbb{E}[V|\mathcal{G}]\, d\mathbb{P}\right) \\
&= T\left(\int_B V\, d\mathbb{P}\right) \\
&= \int_B T(V)\, d\mathbb{P} = \int_B \mathbb{E}[T(V)|\mathcal{G}]\, d\mathbb{P}.
\end{aligned}$$

■

Vale la pena mencionar que el operador de esperanza condicional puede no estar definido para elementos aleatorios que solo son Pettis integrables (observe que en el teorema 1.3.6 hemos pedido una condición más fuerte que ser Pettis integrable (según lo visto en el lema 1.3.3), esta es: $V \in L_1(\Omega; X)$, más adelante veremos que esta condición es equivalente a ser Bochner integrable). Este problema fue estudiado primero (según N. Vakhania [38]) por V. I. Rybakov en [31] y posteriormente por G. E. F. Thomas en [37]. La importancia de la esperanza condicional para elementos aleatorios con valores en un espacio de Banach radica en que ahora se puede definir (extender) el concepto de martingalas a espacios de Banach, y no solo eso, también se pueden estudiar teoremas límites para estos objetos. No tocaremos este tema en esta tesis pero referimos al lector interesado a los libros [11], [12] y [38] para un primer acercamiento.

1.4. Esperanza como integral de Bochner

Como hemos comentado antes, hay otra manera de definir la esperanza de un elemento aleatorio con valores en un espacio de Banach y es mediante mediante la integral de Bochner. Como veremos a continuación, la integral de Bochner resulta ser una generalización de la integral de Lebesgue por lo que algunos autores afirman que se trata de la integral de Lebesgue con signos de norma en lugar de valor absoluto. Este en general es el caso, sin embargo un resultado prueba que no es así: la integral de Bochner no cumple el Teorema de Radon-Nikodým.

Para llegar a estudiar el teorema de Radon-Nikodým damos una pequeña introducción al tema de medidas vectoriales, posteriormente introducimos dos nuevos conceptos de medibilidad de funciones con valores en un espacio de Banach para dar un teorema que les caracteriza: el teorema de medibilidad de Pettis, el cual nos permitirá dar otra la definición (equivalente) de elemento aleatorio.

Vamos a denotar por $\mathcal{E}(\Omega, \mathcal{F}, \mathbb{P}; X) := \mathcal{E}$ al espacio de funciones $f\colon \Omega \to X$ donde $(X, \|\cdot\|_X)$ es un espacio de Banach (no necesariamente separable) y $(\Omega, \mathcal{F}, \mathbb{P})$ es un espacio de probabilidad (o más general, un espacio de medida (σ-)finito).

1.4.1. Medidas vectoriales

Definición 1.4.1. Una **medida vectorial** en X es una función $\mathbf{F}\colon \mathcal{F} \to X$ tal que

$$\mathbf{F}\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mathbf{F}(E_n)$$

para cada sucesión $(E_n)_n$ en $\mathcal{F}$ de elementos disjuntos dos a dos, i.e., $E_i \cap E_j = \emptyset$ si $i \neq j$, y donde la convergencia de la serie se da respecto a la norma de X.

En tal caso se dice que $\mathbf{F}$ es numerablemente aditiva. Cuando la igualdad anterior se da solamente para una cantidad finita de elementos disjuntos de $\mathcal{F}$ se dice que $\mathbf{F}$ es una *medida vectorial finitamente aditiva.*

Ejemplo 1.4.1. Sea $T\colon L_1([0,1]) \to X$ una función lineal y continua, y para cada subconjunto E Lebesgue medible del intervalo $[0,1]$ definamos $\mathbf{F}(E) = T(\chi_E)$. Afirmamos que $\mathbf{F}$ es una medida vectorial.

Primero notemos que $\mathbf{F}$ es finitamente aditiva, pues si E_1, E_2 son dos subconjuntos disjuntos y Lebesgue medibles del intervalo unitario, entonces $\chi_{E_1 \cup E_2} = \chi_{E_1} + \chi_{E_2}$ y por tanto

$$\begin{aligned}\mathbf{F}(E_1 \cup E_2) &= T(\chi_{E_1 \cup E_2}) = T(\chi_{E_1} + \chi_{E_2}) = T(\chi_{E_1}) + T(\chi_{E_2}) \\ &= \mathbf{F}(E_1) + \mathbf{F}(E_2).\end{aligned}$$

Ahora observemos que para cada $E \subset [0,1]$ Lebesgue medible se satisface que $\|\mathbf{F}(E)\| \leq \lambda(E)\|T\|$, donde λ denota a la medida de Lebesgue. En efecto,

$$\|\mathbf{F}(E)\| = \|T(\chi_E)\| \leq \|T\|\, \|\chi_E\|_1 = \|T\|\lambda(E).$$

Luego, si $(E_n)_n$ es una sucesión de conjuntos Lebesgue medibles disjuntos dos a dos en el $[0,1]$, entonces

$$\begin{aligned}
\lim_{m\to\infty}\left\|\mathbf{F}\left(\bigcup_{n=1}^{\infty}E_n\right)-\sum_{n=1}^{m}\mathbf{F}(E_n)\right\| &= \lim_{m\to\infty}\left\|\mathbf{F}\left(\bigcup_{n=1}^{\infty}E_n\right)-\mathbf{F}\left(\bigcup_{n=1}^{m}E_n\right)\right\| \\
&= \lim_{m\to\infty}\left\|\mathbf{F}\left(\bigcup_{n=m+1}^{\infty}E_n\right)\right\| \\
&\leq \lim_{m\to\infty}\lambda\left(\bigcup_{n=m+1}^{\infty}E_n\right)\|T\| = 0
\end{aligned}$$

Y por lo tanto

$$\mathbf{F}\left(\bigcup_{n=1}^{\infty}E_n\right)=\sum_{n=1}^{\infty}\mathbf{F}(E_n).$$

Definición 1.4.2. Decimos que la medida vectorial $\mathbf{F}$ es **absolutamente continua** con respecto a $\mathbb{P}$, y lo denotamos por $\mathbf{F} \ll \mathbb{P}$, si y solo si $\mathbf{F}(A) = \mathbf{0}$ para todo $A \in \mathcal{F}$ con $\mathbb{P}(A) = 0$.

También se dice que $\mathbf{F}$ es una medida vectorial $\mathbb{P}$-continua.

Definición 1.4.3. La **variación** de $\mathbf{F}$ en un conjunto $A \in \mathcal{F}$ es

$$|\mathbf{F}|(A) = \sup \sum_{i=1}^{n} \|\mathbf{F}(A_i)\|,$$

donde el supremo es tomado sobre todas las sucesiones finitas $A_1, A_2, \ldots, A_n \subset A$ de elementos disjuntos dos a dos en $\mathcal{F}$ con $A = \bigcup_{k=1}^{n} A_n$. A toda sucesión finita de subconjuntos de A que cumpla lo anterior se le llama **partición** (finita) de A. En términos de particiones de A, la variación de $\mathbf{F}$ se escribe como

$$|\mathbf{F}|(A) = \sup_{\pi} \sum_{B\in\pi} \|\mathbf{F}(B)\|$$

donde el supremo se toma sobre todas las particiones finitas π de A.

Cuando $|\mathbf{F}|(\Omega) < \infty$ diremos que $\mathbf{F}$ es una medida de **variación finita.**

En la siguiente subsección definiremos a la integral de Bochner y daremos el ejemplo más importante de medida vectorial de variación finita que nos ocupa: A saber, si V es un elemento aleatorio Bochner integrable, $\mathbf{F}(E) = \mathbb{E}\left[V\,\mathbb{I}_E\right]$ para cada $E \in \mathcal{F}$. No solo eso, mostraremos que su variación en Ω es exactamente $\mathbb{E}\left[\|V\|\right]$. (Aquí el símbolo $\mathbb{E}[\,\cdot\,]$ representa a la integral de Bochner de V.)

Definición 1.4.4. La **semivariación** de una medida vectorial $\mathbf{F}$ en un conjunto $A \in \mathcal{F}$ es

$$\|\mathbf{F}\|(A) = \sup\left\{|x^*(\mathbf{F})|(A) \,:\, x^* \in X^*, \|x\| \leq 1\right\},$$

donde $|x^*(\mathbf{F})|$ es la variación de la medida *real-valuada* $x^*(\mathbf{F})$.

Cuando $\|\mathbf{F}\|(\Omega) < \infty$ diremos que $\mathbf{F}$ es una medida de **semivariación finita.**

Varios resultados pueden ser establecidos respecto a la variación y la semivariación de una medida vectorial $\mathbf{F}$ pero no es nuestro propósito presentarlos aquí. Pueden consultarse en [11] (ver capítulo 1).

1.4.2. La integral de Bochner

La integral de Bochner es la integral de una función con valores en un espacio de Banach respecto a una medida con valores escalares. Pertenece a la familia de las así llamadas **integrales fuertes** (la integral de Pettis es también llamada integral débil).

En contraste con la integral de Pettis, las funciones simples tienen un papel más importante para definir a la integral de Bochner, recordemos su definición.

Definición 1.4.5. Una función $s \in \mathcal{E}$ es **simple**, si y solo si para alguna $n \in \mathbb{N}$ existen $x_1, x_2, \ldots, x_n \in X$ y existen $E_1, E_2, \ldots, E_n \in \mathcal{F}$, con $\mathbb{P}(E_i) < \infty$ para cada $1 \leq i \leq n$, tales que

$$s(\omega) = \sum_{i=1}^{n} \chi_{E_i}(\omega) \cdot x_i \quad \text{donde} \quad \chi_{E_i}(\omega) = \begin{cases} 1 & \text{si} \quad \omega \in E_i \\ 0 & \text{si} \quad \omega \notin E_i \end{cases} .$$

Note que de entre estas formas de representar a s existe una única representación, llamada **representación estándar** de s, caracterizada por el hecho de que cada x_i es distinto y los conjuntos E_i son disjuntos. En efecto, si $x_1, x_2, \ldots, x_n$ son los distintos valores que toma s, para cada i tomamos $E_i = s^{-1}(\{x_i\}) = \{\omega \in \Omega \,:\, s(\omega) = x_i\}$.
Aunque no es necesario, vamos a pedir que la sucesión $(E_i)_{i=1}^n$ en la definición 1.4.5 sea de elementos disjuntos dos a dos.

Definición 1.4.6. Una función $f \in \mathcal{E}$ es llamada **$\mathbb{P}$-medible** o **fuertemente medible** si existe una sucesión de funciones simples $(f_n)_n$ tales que

$$\lim_{n\to\infty} \|f_n - f\| = 0 \quad c.s.$$

Definición 1.4.7. Una función $f \in \mathcal{E}$ es llamada **débilmente $\mathbb{P}$-medible** o **escalarmente medible** si para todo $x^* \in X^*$ la composición $x^*(f) \colon \Omega \to \mathbb{R}$ es $\mathbb{P}$-medible.

Como podemos intuir medibilidad fuerte implica medibilidad débil. Esto se justifica de la siguiente manera: Tomemos f una función fuertemente medible, entonces existe $(f_n)_n$ una sucesión de funciones simples tal que $\lim_{n\to\infty} \|f - f_n\| = 0$ c.s., donde cada f_n es de la forma $\sum_{i=1}^{m_n} \chi_{E_{i,n}} \cdot x_{i,n}$. Tomemos ahora $x^* \in X^*$, entonces, como ésta es lineal, se cumple que para cada entero positivo n,

$$x^*(f_n) = x^* \left(\sum_{i=1}^{m_n} \chi_{E_{i,n}} \cdot x_{i,n} \right) = \sum_{i=1}^{m_n} \chi_{E_{i,n}} \cdot x^* \left(x_{i,n}\right),$$

por tanto $x^*(f_n)$ es simple para todo n. Más aún, como cada x^* es también continuo,

$$\lim_{n\to\infty} \|x^*(f) - x^*(f_n)\| = \lim_{n\to\infty} \|x^*(f - f_n)\| \leq \lim_{n\to\infty} \|x^*\|\|f - f_n\| = 0 \quad \text{c.s.}$$

y por lo tanto f es débilmente medible. La implicación inversa en general no se da, el resultado más conocido en esta dirección es el Teorema de Medibilidad de Pettis, el cuál no solo se limita a dar las condiciones para la ya antes comentada implicación de regreso sino que proporciona una forma de caracterizar a todas las funciones fuertemente medibles de Ω en X. Antes de presentarlo necesitaremos de algunos resultado, entre ellos está la generalización del clásico teorema de Egorov.

Lema 1.4.1. *Si $f\colon \Omega \to X$ es $\mathbb{P}$-medible, entonces $\|f\|\colon \Omega \to \mathbb{R}$ es $\mathbb{P}$-medible.*

Demostración. Como f es $\mathbb{P}$-medible existe $(f_n)_n$ sucesión de funciones simples tal que $\lim_{n\to\infty} f_n(\omega) = f(\omega)$ para casi todo $\omega \in \Omega$. Note entonces que la función real valuada $\|f_n(\omega)\|$ es simple para cada n. Más aún,

$$|\|f(\omega)\| - \|f_n(\omega)\|| \leq \|f(\omega) - f_n(\omega)\|$$

por lo tanto $\lim_{n\to\infty} \|f_n(\omega)\| = \|f(\omega)\|$, es decir, $\|f\|$ es $\mathbb{P}$-medible. ■

Teorema 1.4.2 (Egorov)**.** *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de medida finito. Supongamos $f \in \mathcal{E}$ es una función $\mathbb{P}$-medible y que $(f_n)_n$ es una sucesión de funciones $\mathbb{P}$-medibles de Ω en X que convergen c.s. a f en algún subconjunto medible E (y de medida finita), entonces para cada $\varepsilon > 0$, existe un subconjunto medible $F \subset E$ con $\mathbb{P}(F) < \varepsilon$ y tal que $(f_n)_n$ converge uniformemente a f en $E \setminus F$.*

Demostración. Dado $\varepsilon > 0$ y $n \in \mathbb{N}$ tomemos $(\varepsilon_n)_n$ una sucesión decreciente a 0, $\lim_{n\to\infty} \varepsilon_n = 0$, y sea $\delta_n = \dfrac{\varepsilon}{2^n}$. Sea $\mathcal{N}$ el subconjunto nulo de E donde la sucesión $(f_n)_n$ no convergeH a f. Definimos

$$G_{n,k} = \{\omega \in E \setminus \mathcal{N} \,:\, \|f_k(\omega) - f(\omega)\| \geq \varepsilon_n\},$$

observemos que $G_{n,k}$ es efecto medible: aplicando el lema 1.4.1 a la función $f_k - f$, que es $\mathbb{P}$-medible, obtenemos que $\|f_k - f\|\colon \Omega \to \mathbb{R}$ es también $\mathbb{P}$-medible en $\mathbb{R}$ y por tanto medible. Luego el conjunto de puntos donde una función (real) medible pierde o le gana a una constante es medible. También definimos

$$E_{n,K} = \bigcup_{k=K}^{\infty} G_{n,k} = \{\omega \in E \setminus \mathcal{N} \,:\, \|f_k(\omega) - f(\omega)\| \geq \varepsilon_n \text{ para alguna } k \geq K\},$$

para cada $K \geq 1$. Entonces $E_{n,K}$ también es medible y claramente cumple que $E_{n,K+1} \subset E_{n,K}$. Por otra parte notemos que si $\omega \in E \setminus \mathcal{N}$, entonces $f_n(\omega) \to f(\omega)$, por lo tanto existe $N_0 \geq 1$ tal que $\omega \notin E_{n,N_0}$, es decir, $\bigcap_{K=1}^{\infty} E_{n,K} = \emptyset$. En consecuencia

$$0 = \mathbb{P}\left(\bigcap_{K=1}^{\infty} E_{n,K}\right) = \lim_{K\to\infty} \mathbb{P}(E_{n,K}),$$

entonces existe un $K_n \geq 1$ (depende de ε_n y por tanto de n) tal que

$$\mathbb{P}(E_{n,K_n}) = \mathbb{P}(\{\omega \in E \setminus \mathcal{N} : \|f_k(\omega) - f(\omega)\| \geq \varepsilon_n \text{ p.a. } k \geq K_n\}) < \delta_n.$$

Sea $F_n = E_{n,K_n}$, entonces $\mathbb{P}(F_n) < \delta_n$ y además

$$(E \setminus \mathcal{N}) \setminus F_n = \{\omega \in E \setminus \mathcal{N} : \|f_k(\omega) - f(\omega)\| < \varepsilon_n \text{ para todo } k \geq K_n\}.$$

Tomando $F = \mathcal{N} \cup (\bigcup_{n=1}^{\infty} F_n)$, tenemos que

$$\mathbb{P}(F) \leq \sum_{n=1}^{\infty} \mathbb{P}(F_n) < \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon$$

y además $E \setminus F = \bigcap_{n=1}^{\infty} (E \setminus \mathcal{N}) \setminus F_n$. Como $\varepsilon_n \to 0$, dado $\delta > 0$ existe $n \geq 1$ tal que $\varepsilon_n < \delta$ y en consecuencia

$$\|f_k(\omega) - f(\omega)\| < \delta$$

si $k \geq K_n$, y esto ocurre para todo $\omega \in E \setminus F$. Por tanto la sucesión $(f_n)_n$ converge uniformemente a f en $E \setminus F$. ■

Teorema 1.4.3 (de Medibilidad de Pettis). *Una función $f \in \mathcal{E}$ es $\mathbb{P}$-medible, si y solo si*

i. f es de valores $\mathbb{P}$-escencialmente separables, es decir, existe $E \in \mathcal{F}$ con $\mathbb{P}(E) = 0$ tal que $f(\Omega \setminus E)$ es un subconjunto separable (respecto a la topología de la norma) de X.

ii. f es débilmente $\mathbb{P}$-medible.

Demostración. ($\Rightarrow$) Como f es $\mathbb{P}$-medible existe una sucesión de funciones simples $(f_n)_n$ tal que

$$\lim_{n\to\infty} \|f_n(\omega) - f(\omega)\| = 0 \quad \text{c.s.}$$

Por el teorema de Egorov, para cada $n \geq 1$ existe un conjunto $E_n \in \mathcal{F}$ tal que $\mathbb{P}(E_n) < \frac{1}{n}$ y $\lim_{n\to\infty} f_n = f$ uniformemente en $\Omega \setminus E_n$.

Notemos f_n tiene rango a lo más numerable pues es simple para cada n, y si $x \in f(\Omega \setminus E_m)$ entonces $x = f(\omega)$ para algún $\omega \in \Omega \setminus E_m$, recordemos que $f_n \to f$ uniformemente en $\Omega \setminus E_m$ (y esto para cada m), en particular $f_n(\omega) \to f(\omega) = x$. Así, si $x_n := f_n(\omega)$ para cada n, entonces $(x_n)_n$ es una sucesión en $\bigcup_{n\in\mathbb{N}} f_n(\Omega)$ tal que $x_n \to x$, por lo tanto

$$f(\Omega \setminus E_m) \subset \overline{\bigcup_{n\in\mathbb{N}} f_n(\Omega)}.$$

Es decir, el rango $f(\Omega \setminus E_m)$ está contenido en un métrico separable y por tanto también separable[9], y esto para cada m. Y como

$$f\left(\bigcup_{m=1}^{\infty} \Omega \setminus E_m\right) = \bigcup_{m=1}^{\infty} f(\Omega \setminus E_m)$$

entonces $f\left(\bigcup_{m=1}^{\infty} \Omega \setminus E_m\right)$ es separable. Además, $\Omega \setminus (\bigcup_{m=1}^{\infty} \Omega \setminus E_m) = \bigcap_{m=1}^{\infty} E_m$ es un conjunto de medida cero pues $\mathbb{P}(E_m) < \frac{1}{m}$ para cada m. Esto prueba la parte (i) del enunciado. La parte (ii) ya la hemos comentado previamente (ver comentario previo al lema 1.4.1).

($\Leftarrow$) Sea $E \in \mathcal{F}$ tal que $\mathbb{P}(E) = 0$ y $f(\Omega \setminus E)$ es separable. Sea $D = \{d_n : n \geq 1\}$ un subconjunto denso numerable en $f(\Omega \setminus E)$. Por el corolario A.2.5 de Hanh-Banach existe una sucesión $(x_n^*)_n$ en X^* tal que

$$\|x_n^*\| = 1, \qquad x_n^*(d_n) = \|d_n\|, \qquad n = 1, 2, \ldots$$

Recordemos que (como consecuencia del corolario A.2.5) para cada $x \in X$,

$$\|x\| = \sup\{|x^*(x)| : x^* \in X^*, \|x^*\| \ \leq 1\}.$$

Afirmamos que para cada $\omega \in \Omega \setminus E$,

$$\|f(\omega)\| := \sup\{|x^*(f(\omega))| : x^* \in X^*, \|x^*\| \ \leq 1\} = \sup_{n \geq 1} |x_n^*(f(\omega))|.$$

En efecto, sea $\omega \in \Omega \setminus E$. Como D es denso en $f(\Omega \setminus E)$ existe una sucesión $(d_{n_k})_k$ en D tal que $d_{n_k} \to f(\omega)$. Así, para cada $\varepsilon > 0$ existe $K > 0$ tal que

$$|\|f(\omega)\| - \|d_{n_k}\|| \leq \|f(\omega) - d_{n_k}\| < \frac{\varepsilon}{2}$$

si $k \geq K$. Entonces, para k suficientemente grande

$$\begin{aligned}
|x_{n_k}^*(f(\omega))| \leq \|f(\omega)\| \leq \|d_{n_k}\| + \frac{\varepsilon}{2} &= x_{n_k}^*(d_{n_k}) + \frac{\varepsilon}{2} \\
&= x_{n_k}^*(d_{n_k} - f(\omega) + f(\omega)) + \frac{\varepsilon}{2} \\
&= x_{n_k}^*(d_{n_k} - f(\omega)) + x_{n_k}^*(f(\omega)) + \frac{\varepsilon}{2} \\
&\leq \|x_{n_k}^*\|\|d_{n_k} - f(\omega)\| + |x_{n_k}^*(f(\omega))| + \frac{\varepsilon}{2} \\
&\leq 1 \cdot \frac{\varepsilon}{2} + |x_{n_k}^*(f(\omega))| + \frac{\varepsilon}{2} \\
&= |x_{n_k}^*(f(\omega))| + \varepsilon.
\end{aligned}$$

[9]Todo subconjunto de un métrico separable es también separable: Si es métrico separable, entonces es segundo numerable (las bolas con radio racional centradas en puntos del denso forman una base numerable del espacio), y por tanto todos sus subconjuntos son también segundo numerables en la topología relativa. En esta nueva base del subconjunto basta tomarse un punto en cada abierto básico para construir un denso numerable en dicho subconjunto.

Haciendo tender ε a cero obtenemos que $\|f(\omega)\| = \sup_n |x_n^*(f(\omega))|$. Ahora, como por hipótesis f es débilmente medible, entonces $x_n^*(f(\omega))$ es una función real $\mathbb{P}$-medible y por tanto $\sup_{n\in\mathbb{N}} |x_n^*(f(\omega))| = \|f(\omega)\|$ es también $\mathbb{P}$-medible. Por el mismo argumento de antes, las funciones $g_n(\cdot) = \|f(\cdot) - d_n\|$ son también $\mathbb{P}$-medibles para cada $n \in \mathbb{N}$.
Sea $\varepsilon > 0$, definimos los conjuntos medibles

$$E_n = \{\omega \in \Omega \setminus E \,:\, g_n(\omega) < \varepsilon\} = \{\omega \in \Omega \setminus E \,:\, \|f(\omega) - d_n\| < \varepsilon\},$$

por la densidad de D en $f(\Omega \setminus E)$ podemos observar que $\bigcup_{n=1}^{\infty} E_n = \Omega \setminus E$. Como en la proposición 1.1.4, definamos los conjuntos

$$F_1 = E_1$$
$$F_n = E_n \setminus \bigcup_{k=1}^{n-1} E_k \quad \text{(para } n \geq 2).$$

Entonces $\{F_n\}$ forma una colección numerable de conjuntos medibles, ajenos dos a dos y tales que

$$\bigcup_{n=1}^{\infty} F_n = \bigcup_{n=1}^{\infty} E_n = \Omega \setminus E.$$

Definimos la función $f_\varepsilon \colon \Omega \to X$ como

$$f_\varepsilon(\omega) = \begin{cases} d_n & \text{si} \quad \omega \in F_n \\ 0 & \text{si} \quad \omega \in E. \end{cases} = \sum_{n=1}^{\infty} d_n \cdot \chi_{F_n}$$

Entonces f_ε está bien definida, toma una cantidad a lo más numerable de valores y es tal que $\|f(\omega) - f_\varepsilon(\omega)\| = \|f(\omega) - d_n\| < \varepsilon$ para cada $\omega \in \Omega \setminus E$, es decir, $\|f(\omega) - f_\varepsilon(\omega)\| < \varepsilon$ c.s.

Dicho de otra forma, mostramos que para cada $\varepsilon > 0$ hemos podido aproximar (de forma casi segura) a f mediante una función f_ε que toma a lo más un número numerable de valores. Tomando $\varepsilon = \frac{1}{n}$ con $n \in \mathbb{N}$ obtenemos una sucesión de funciones $(f_n)_n$ que toman a lo más una cantidad numerable de valores tales que $\|f_n - f\| < \frac{1}{n}$ c.s. para cada n.

Como cada f_n es de la forma $f_n = \displaystyle\sum_{m=1}^{\infty} d_{n,m} \cdot \chi_{E_{n,m}}$ donde $E_{n,i} \cap E_{n,j} = \emptyset$ si $i \neq j$ y $E_{n,m} \in \mathcal{F}$, basta truncar la serie anterior hasta cierto $k_n \in \mathbb{N}$ para obtener una función simple f_n' de la forma $\displaystyle\sum_{m=1}^{k_n} d_{n,m} \cdot \chi_{E_{n,m}}$. Esto prueba que f es el límite casi seguro de funciones simples y por tanto es $\mathbb{P}$-medible. ■

Como consecuencia de la última parte de la prueba del teorema de medibilidad de Pettis 1.4.3 obtenemos el siguiente corolario:

Corolario 1.4.4. *Una función $f \in \mathcal{E}$ es $\mathbb{P}$-medible, si y solo si f es el límite uniforme casi seguro de una sucesión de funciones $\mathbb{P}$-medibles que toman una cantidad numerable de valores.*

Comentario. El teorema de medibilidad de Pettis así como su corolario son todo lo que necesitamos para definir a un elemento aleatorio $f\colon \Omega \to X$ en un espacio de Banach como una función $\mathbb{P}$-medible y por tanto, de valores $\mathbb{P}$-esencialmente separables. Y esta definición es así porque en toda esta sección no hemos asumido la hipótesis de separabilidad del espacio X. El lema 1.1.5 nos da la equivalencia entre esta nueva definición y la dada en 1.1.1 cuando X es separable. En este sentido esta definición es una extensión de la dada en 1.1.1, así que en lo sucesivo cuando escribamos «$f \in \mathcal{E}$ una función $\mathbb{P}$-medible» podemos entender «f un elemento aleatorio en X» . Siendo este el caso vamos a definir la esperanza de f mediante la integral de Bochner.

Y para definir la integral de Bochner vamos a proceder como en el caso de la integral de Lebesgue. Primero daremos la definición para funciones simples y después para cualquier función $\mathbb{P}$-medible aproximando a ésta mediante simples.

Definición 1.4.8. Sea $s \in \mathcal{E}$ una función simple, con representación estándar $s(\omega) = \sum_{i=1}^{n} \chi_{E_i}(\omega)x_i$. La integral de Bochner de s sobre un conjunto $E \in \mathcal{F}$ se define como

$$\int_E s\, d\mathbb{P} := \sum_{i=1}^{n} \mathbb{P}(E \cap E_i)x_i.$$

Veamos que la integral de Bochner para una función simple está bien definida. Es decir, si $\sum_{j=1}^{m} \chi_{F_j}(\omega)y_j$ es otra representación para s (con $y_j \in X$, $F_j \in \mathcal{F}$ y $\mathbb{P}(F_j) < \infty$ para toda $j = 1, \dots, m$ y $F_k \cap F_j = \emptyset$ para todo $k \neq j$), entonces $\int_E s\, d\mathbb{P} = \sum_{j=1}^{m} \mathbb{P}(E \cap F_j)x_j$.

En efecto, supongamos que $(E_n)_{i=1}^{n}$ y $(F_j)_{j=1}^{m}$ son sucesiones de elementos ajenos dos a dos, con $\cup_{i=1}^{n} E_i = \cup_{j=1}^{m} F_j = E$. Consideremos el refinamiento $(G_{i,j})$ donde $G_{i,j} = E_i \cap F_j$ para cada i, j. Entonces,

$$s(\omega) = \sum_{i=1}^{n}\sum_{j=1}^{m} \chi_{G_{i,j}} x_i = \sum_{i=1}^{n}\sum_{j=1}^{m} \chi_{G_{i,j}} y_j.$$

Cuando calculamos la integral s sobre E obtenemos que

$$\sum_{i=1}^{n} \mathbb{P}(E \cap E_i)x_i = \sum_{i=1}^{n}\sum_{j=1}^{m} \mathbb{P}(E \cap G_{i,j})x_i,$$

y

$$\sum_{j=1}^{m} \mathbb{P}(E \cap F_i)y_j = \sum_{i=1}^{n}\sum_{j=1}^{m} \mathbb{P}(E \cap G_{i,j})y_j.$$

Resta ver que ambas sumas son iguales, pero observe que si $G_{i,j} \neq \emptyset$ entonces debe ocurrir que $x_i = y_j$, y esto para cada i, j. Por lo tanto ambas sumas coinciden y las integrales de ambas representaciones son iguales.

Proposición 1.4.5. *Sean $f, g \in \mathcal{E}$ funciones simples y $\alpha \in \mathbb{R}$. Entonces para cada $E \in \mathcal{F}$,*

I. $\displaystyle\int_E \alpha f \, d\mathbb{P} = \alpha \int_E f \, d\mathbb{P}$

II. $\displaystyle\int_E f + g \, d\mathbb{P} = \int_E f \, d\mathbb{P} + \int_E g \, d\mathbb{P}.$

Demostración. Si $f(\omega) = \sum_{i=1}^n \chi_{E_i}(\omega) \cdot x_i$ y $g(\omega) = \sum_{j=1}^m \chi_{F_j}(\omega) \cdot y_j$,

I. $\displaystyle\int_E \alpha f \, d\mathbb{P} = \sum_{i=1}^n \mathbb{P}(E \cap E_i)(\alpha \cdot x_i) = \alpha \sum_{i=1}^n \mathbb{P}(E \cap E_i) x_i = \alpha \int_E f \, d\mathbb{P}$

II. Parecido a lo que hicimos en la sección anterior; para cada i, j tomamos el refinamiento $G_{i,j} = E_i \cap F_j$. Luego $f + g = \displaystyle\sum_{i=1}^n \sum_{j=1}^m \chi_{G_{i,j}}(x_i + y_j)$, y entonces

$$\begin{aligned}
\int_E f \, d\mathbb{P} + \int_E g \, d\mathbb{P} &= \sum_{i=1}^n \mathbb{P}(E \cap E_i) x_i + \sum_{j=1}^m \mathbb{P}(E \cap F_j) y_j \\
&= \sum_{i=1}^n \sum_{j=1}^m \mathbb{P}(E \cap E_i \cap F_j)(x_i + y_j) \\
&= \sum_{i=1}^n \sum_{j=1}^m \mathbb{P}(E \cap G_{i,j})(x_i + y_j) = \int_E (f + g) \, d\mathbb{P}.
\end{aligned}$$

La tercera igualdad se da debido a que si $E_i \cap F_j \neq \emptyset$, entonces $x_i = y_j$.

■

Como consecuencia de este último resultado tenemos que la integral de Bochner de funciones simples es aditiva.

Corolario 1.4.6. *Sea E un elemento en $\mathcal{F}$ y sea $(E_i)_{i=1}^n$ una partición E, entonces*

$$\sum_{i=1}^n \int_{E_i} f \, d\mathbb{P} = \int_E f \, d\mathbb{P}.$$

Demostración. Sea $f_i = f \cdot \chi_{E_i}$, entonces

$$\begin{aligned}
\sum_{i=1}^n \int_{E_i} f \, d\mathbb{P} = \sum_{i=1}^n \int_\Omega f_i \, d\mathbb{P} &= \int_\Omega \sum_{i=1}^n f_i \, d\mathbb{P} = \int_\Omega f \cdot \sum_{i=1}^n \chi_{E_i} \, d\mathbb{P} \\
&= \int_\Omega f \cdot \chi_{\bigcup_{i=1}^n E_i} \, d\mathbb{P} = \int_\Omega f \chi_E \, d\mathbb{P}.
\end{aligned}$$

■

Si $s \in \mathcal{E}$ es función simple con representación estándar $s(\omega) = \sum_{i=1}^{n} \chi_{E_i}(\omega) \cdot x_i$, entonces la función $\|s\| : \Omega \to \mathbb{R}$ definida como $\|s\|(\omega) = \|s(\omega)\|$ es también simple y con representación estándar

$$\|s\| = \sum_{i=1}^{n} \chi_{E_i}(\omega) \cdot \|x_i\|.$$

Más aún, para cada $E \in \mathcal{F}$,

$$\int_E \|s\| \, d\mathbb{P} = \sum_{i=1}^{n} \mathbb{P}(E \cap E_i) \cdot \|x_i\| \quad \text{y} \quad \left\| \int_E s \, d\mathbb{P} \right\| \leq \int_E \|s\| \, d\mathbb{P}.$$

Probamos esto último en un caso más general en el teorema 1.4.10. Dicho lo anterior ya podemos definir la integral de Bochner para cualquier función $f \in \mathcal{E}$ fuertemente medible.

Definición 1.4.9. Una función $\mathbb{P}$-medible $f \in \mathcal{E}$ es **Bochner integrable** si existe una sucesión de funciones simples $(f_n)_n$ tal que

$$\lim_{n\to\infty} \int_\Omega \|f_n(\omega) - f(\omega)\| \, d\mathbb{P}(\omega) = 0.$$

En tal caso $\displaystyle\int_\Omega f \, d\mathbb{P}$ está definida para cada $E \in \mathcal{F}$ como

$$\int_E f \, d\mathbb{P} = \lim_{n\to\infty} \int_\Omega \chi_E \cdot f_n \, d\mathbb{P} = \lim_{n\to\infty} \int_E f_n \, d\mathbb{P}.$$

Observación 1.4.1. Aunque no lo hicimos explícito, en la definición anterior estamos asumiendo que la norma de la diferencia $\|f_n(\omega) - f(\omega)\|$ es Lebesgue integrable para cada $n \geq 1$.

Veamos que la definición 1.4.9 tiene sentido, i.e., tal límite existe y no depende de la elección de la sucesión $(f_n)_n$ de funciones simples. Es decir, si $(g_n)_n$ es otra sucesión funciones de simples tal que $\|f - g_n\|$ es Lebesgue integrable para cada n y $\lim_{n\to\infty} \int_\Omega \|f - g_n\| \, d\mathbb{P} = 0$, entonces $\lim_{n\to\infty} \int_E f_n \, d\mathbb{P} = \lim_{n\to\infty} \int_E g_n \, d\mathbb{P}$ para cada $E \in \mathcal{F}$.

Primero, el límite existe pues la sucesión $\left(\int_E f_n \, d\mathbb{P}\right)_n$ es de Cauchy en X que es de Banach (y esto para cada $E \in \mathcal{F}$), esto es consecuencia de la proposición 1.4.5 y de la desigualdad del triángulo para la integral de Bochner de funciones simples. En efecto,

$$\begin{aligned}\left\| \int_E f_n \, d\mathbb{P} - \int_E f_m \, d\mathbb{P} \right\| &= \left\| \int_E (f_n - f_m) \, d\mathbb{P} \right\| \leq \int_E \|f_n - f_m\| \, d\mathbb{P} \\ &\leq \int_E \|f_n - f\| \, d\mathbb{P} + \int_E \|f - f_m\| \, d\mathbb{P}\end{aligned}$$

y por tanto, la diferencia tiende a 0 cuando n, m son suficientemente grandes.

La unicidad del límite ocurre de manera análoga, a saber,

$$\begin{aligned}\left\|\int_E f_n\, d\mathbb{P} - \int_E g_n\, d\mathbb{P}\right\| &= \left\|\int_E (f_n - g_n)\, d\mathbb{P}\right\| \leq \int_E \|f_n - g_n\|\, d\mathbb{P}\\ &= \int_E \|f_n - f\|\, d\mathbb{P} + \int_E \|f - g_n\|\, d\mathbb{P}\end{aligned}$$

para n suficientemente grande la diferencia tiende a 0, por lo que podemos concluir que $\lim_{n\to\infty}\int_E f_n\, d\mathbb{P} = \lim_{n\to\infty}\int_E g_n\, d\mathbb{P}$.

El siguiente resultado caracteriza a todas las funciones $\mathbb{P}$-medibles que son Bochner integrables utilizando la (Lebesgue) integrabilidad de su norma. Al mismo tiempo, y en conjunción con el lema 1.3.3, nos dice que Bochner integrabilidad implica Pettis integrabilidad (y en dado caso ambas integrales coinciden).

Teorema 1.4.7. *Sea $f \in \mathcal{E}$ una función $\mathbb{P}$-medible.*

$$f \text{ es Bochner integrable, si y solo si } \int_\Omega \|f(\omega)\|\, d\mathbb{P}(\omega) < \infty$$

Demostración. 1. Si f es Bochner integrable, entonces existe $(f_n)_n$ una sucesión de simples tal que $\lim_n \int_\Omega \|f - f_n\|\, d\mathbb{P} = 0$. Entonces,

$$\int_\Omega \|f\|\, d\mathbb{P} \leq \int_\Omega \|f - f_n\|\, d\mathbb{P} + \int_\Omega \|f_n\|\, d\mathbb{P} < \infty$$

para n suficientemente grade.

2. Por otra parte, como f es $\mathbb{P}$-medible, por el corolario 1.4.4 existe una sucesión $(f_n)_n$ de funciones $\mathbb{P}$-medibles que toman a lo más una cantidad numerable de valores (que nosotros hemos definido como **discretas** en el caso separable) tales que $\|f(\omega) - f_n(\omega)\| < \frac{1}{n}$ para casi todo $\omega \in \Omega$ y para toda n. Ello implica que $\|f_n\| \leq \|f\| + \frac{1}{n}$ para casi todo $\omega \in \Omega$ y toda n. Como además $\mathbb{P}$ es una medida finita, entonces $\int_\Omega \|f_n\|\, d\mathbb{P} \leq \int_\Omega \|f\| + \frac{\mathbb{P}(\Omega)}{n} < \infty$ para toda n.

Ahora bien, para cada entero positivo n, la función f_n es de la forma

$$f_n = \sum_{m=1}^{\infty} x_{n,m}\chi_{E_{n,m}}$$

donde $E_{n,i} \cap E_{n,j} = \emptyset$ para toda $i \neq j$, con $E_{n,m} \in \mathcal{F}$ y $x_{n,m} \in X$. Así, para cada n, escojamos[10] p_n suficientemente grande tal que

$$\int_{\bigcup_{m=p_n+1}^{\infty} E_{n,m}} \|f_n\|\, d\mathbb{P} < \frac{\mathbb{P}(\Omega)}{n}.$$

Definamos $g_n = \sum_{m=1}^{p_n} x_{n,m}\chi_{E_{n,m}}$. Entonces para cada n, la función g_n es simple y es tal que

$$\begin{aligned}
\int_\Omega \|f - g_n\|\, d\mathbb{P} &\leq \int_\Omega \|f - f_n\|\, d\mathbb{P} + \int_\Omega \|f_n - g_n\|\, d\mathbb{P} \\
&\leq \frac{\mathbb{P}(\Omega)}{n} + \int_\Omega \left\| \sum_{m=1}^{\infty} x_{n,m}\chi_{E_{n,m}} - \sum_{m=1}^{p_n} x_{n,m}\chi_{E_{n,m}} \right\| d\mathbb{P} \\
&\leq \frac{\mathbb{P}(\Omega)}{n} + \int_\Omega \left\| \sum_{m=p_n+1}^{\infty} x_{n,m}\chi_{E_{n,m}} \right\| d\mathbb{P} \\
&\leq \frac{\mathbb{P}(\Omega)}{n} + \int_\Omega \sum_{m=p_n+1}^{\infty} \|x_{n,m}\|\chi_{E_{n,m}}\, d\mathbb{P} < \frac{2\mathbb{P}(\Omega)}{n}.
\end{aligned}$$

Por lo tanto f es Bochner integrable. ■

Con ayuda del resultado anterior podemos probar que la integral de Bochner es lineal.

Proposición 1.4.8. *Si f, g son Bochner integrables y $\alpha \in \mathbb{R}$, entonces*

I. $\int \alpha f\, d\mathbb{P} = \alpha \int f\, d\mathbb{P}$,

II. $\int f + g\, d\mathbb{P} = \int f\, d\mathbb{P} + \int g\, d\mathbb{P}$.

Demostración.

I. Por el teorema 1.4.7, αf es Bochner integrable (pues $\alpha\|f\|$ es Lebesgue integrable). Sea $(f_n)_n$ la sucesión de funciones simples que converge c.s. a f. Entonces $(\alpha f_n)_n$ es una sucesión funciones simples que converge a αf c.s., además debido a que f es Bochner integrable (y a la linealidad de la integral de Lebesgue),

$$\lim_{n\to\infty} \int \|\alpha f - \alpha f_n\|\, d\mathbb{P} = \alpha \left(\lim_{n\to\infty} \int \|f - f_n\|\, d\mathbb{P} \right) = 0.$$

[10]Este natural existe (y depende de n) pues,

$$\int_\Omega \|f_n\|\, d\mathbb{P} = \int_{\bigcup_{m=1}^{\infty} E_{n,m}} \|f_n\|\, d\mathbb{P},$$

y $\lim_{i\to\infty} \int_{\bigcup_{i=m+1}^{\infty} E_{n,i}} \|f_n\|\, d\mathbb{P} = 0.$

Por tanto,

$$\int \alpha f \, d\mathbb{P} = \lim_{n\to\infty} \int \alpha f_n \, d\mathbb{P} = \alpha \lim_{n\to\infty} \int f_n \, d\mathbb{P} = \alpha \int f \, d\mathbb{P}.$$

Observe que la segunda igualdad se da por la parte (i) de la proposición 1.4.5.

II. Como f y g son Bochner integrables,

$$\int \|f+g\| \, d\mathbb{P} \leq \int \|f\| \, d\mathbb{P} + \int \|g\| \, d\mathbb{P} < \infty,$$

así, por el teorema 1.4.7, $f+g$ es Bochner integrable. Ahora, si $(f_n)_n$ y $(g_n)_n$ son sucesiones de simples que convergen a f y g c.s. respectivamente, entonces $(f_n+g_n)_n$ converge a $f+g$ c.s. y por tanto

$$\lim_{n\to\infty} \int \|(f+g)-(f_n+g_n)\| \, d\mathbb{P} = 0 \quad \text{y} \quad \int_\Omega f+g \, d\mathbb{P} = \lim_{n\to\infty} \int f_n + g_n \, d\mathbb{P}.$$

Veamos entonces que $\int f \, d\mathbb{P} + \int g \, d\mathbb{P} = \lim_{n\to\infty} \int f_n + g_n \, d\mathbb{P}$. En efecto,

$$\begin{aligned} &\left\| \int f \, d\mathbb{P} + \int g \, d\mathbb{P} - \left(\int f_n + g_n \, d\mathbb{P} \right) \right\| \\ &\qquad \leq \left\| \left(\int f \, d\mathbb{P} - \int f_n \, d\mathbb{P} \right) + \left(\int g \, d\mathbb{P} - \int g_n \, d\mathbb{P} \right) \right\| \\ &\qquad \leq \left\| \int f \, d\mathbb{P} - \int f_n \, d\mathbb{P} \right\| + \left\| \int g \, d\mathbb{P} - \int g_n \, d\mathbb{P} \right\|, \end{aligned}$$

pero para n suficientemente grande, los sumandos en el lado derecho de la desigualdad anterior tienden a cero, por lo que concluimos el resultado.

■

Teorema 1.4.9 (Convergencia Dominada de Bochner)**.** *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad (de medida finita) y $(f_n)_n$ una sucesión de funciones Bochner integrables con $f_n \in \mathcal{E}$ para cada $n \geq 1$. Si $\lim_n f_n = f$ en en probabilidad (medida), es decir, $\lim_{n\to\infty} \mathbb{P}(\{\omega \in \Omega : \|f_n - f\| \geq \varepsilon\}) = 0$ para todo $\varepsilon > 0$, o bien puntualmente c.s., es decir, $\lim_n f_n(\omega) = f(\omega)$ para casi todo $\omega \in \Omega$, y existe una función $g\colon \Omega \to \mathbb{R}$ Lebesgue integrable en Ω tal que $\|f_n\| \leq g$ c.s., entonces*

I. *f es Bochner integrable*

II. $\lim_{n\to\infty} \int_E f_n \, d\mathbb{P} = \int_E f \, d\mathbb{P}$ *para cada $E \in \mathcal{F}$*

Más aún, $\lim_{n\to\infty} \int \|f - f_n\| \, d\mathbb{P} = 0.$

Demostración. Probaremos solo el caso cuando la convergencia es puntual, el otro caso es análogo.
Observése que $\|f - f_n\|: \Omega \to \mathbb{R}$ es dominada por $2g$, y es tal que $\|f - f_n\| \to 0$ c.s., así que usando el teorema clásico de Convergencia Dominada obtenemos que

$$\lim_{n\to\infty} \int_\Omega \|f - f_n\| \, d\mathbb{P} = \int_\Omega \lim_{n\to\infty} \|f - f_n\| \, d\mathbb{P} = 0.$$

Por otra parte, como para cada n se tiene que f_n es Bochner integrable podemos defininir $(s_m^{(n)})_m$ una sucesión de simples tales que $\int \|f_n - s_m^{(n)}\| \to 0$ cuando $m \to \infty$. Entonces, para cada n eligimos a la función simple ϕ_n tal que $\int \|f_n - \phi_n\| \, d\mathbb{P} < \frac{1}{n}$. Y por lo tanto

$$\int \|f - \phi_n\| \, d\mathbb{P} < \int \|f - f_n\| \, d\mathbb{P} + \int \|f_n - \phi_n\| \, d\mathbb{P} < \frac{2}{n}.$$

Esto prueba que el límite f es también Bochner integrable, y en consecuencia

$$\lim_{n\to\infty} \int_E f_n \, d\mathbb{P} = \lim_{n\to\infty} \int_E \phi_n \, d\mathbb{P} = \int_E f \, d\mathbb{P}$$

para cada $E \in \mathcal{F}$. ■

Teorema 1.4.10. *Si f es una función Bochner integrable, entonces*

i. $\displaystyle\lim_{\mathbb{P}(E)\to 0} \int_E f \, d\mathbb{P} = 0$

ii. $\displaystyle\left\| \int_E f \, d\mathbb{P} \right\| \leq \int_E \|f\| \, d\mathbb{P}$ *para todo* $E \in \mathcal{F}$

iii. Si (E_n) es una sucesión de elementos de $\mathcal{F}$ disjuntos dos a dos tales que $E = \bigcup_{n=1}^\infty E_n$, entonces

$$\int_E f \, d\mathbb{P} = \sum_{n=1}^\infty \int_{E_n} f \, d\mathbb{P}$$

iv. Para todo $E \in \mathcal{F}$ definimos $\mathbf{F}(E) = \int_E f \, d\mathbb{P}$, entonces $\mathbf{F}$ es una medida vectorial de variación acotada y además

$$|\mathbf{F}|(E) = \int_E \|f\| d\mathbb{P}.$$

Demostración. (i) Como f es Bochner integrable, entonces $\|f\|$ es integrable, luego

$$0 = \lim_{\mathbb{P}(E)\to 0} \int_E \|f\| \, d\mathbb{P} \geq \lim_{\mathbb{P}(E)\to 0} \left\| \int_E f \, d\mathbb{P} \right\| = \left\| \lim_{\mathbb{P}(E)\to 0} \int_E f \, d\mathbb{P} \right\| \geq 0,$$

por lo que el resultado se sigue de (ii).

(ii) Observemos primero que esta afirmación se cumple para las funciones simples. En efecto, si $f = \displaystyle\sum_{i=1}^{n} \chi_{E_i} x_i$, entonces dado $E \in \mathcal{F}$

$$\begin{aligned}\left\| \int_E f \, d\mathbb{P} \right\| &= \left\| \int_E \left(\sum_{i=1}^{n} \chi_{E_i} x_i \right) \, d\mathbb{P} \right\| = \left\| \sum_{i=1}^{n} \mathbb{P}(E_i \cap E) x_i \right\| \\ &\leq \sum_{i=1}^{n} \mathbb{P}(E_i \cap E) \|x_i\| = \int_E \|f\| \, d\mathbb{P}.\end{aligned}$$

Para el caso general cuando f es Bochner integrable, sea $(f_n)_n$ la sucesión de simples para la cual $\lim_{n\to\infty} \int_\Omega \|f - f_n\| \, d\mathbb{P} = 0$ y $\int_E f \, d\mathbb{P} = \lim_{n\to\infty} \int_\Omega f_n \, d\mathbb{P}$ para cada $E \in \mathcal{F}$.
Sea $E \in \Omega$, entonces

$$\int_E (\|f\| - \|f_n\|) \, d\mathbb{P} \leq \int_\Omega |\|f\| - \|f_n\|| \, d\mathbb{P} \leq \int_\Omega \|f - f_n\| \, d\mathbb{P},$$

pues la integral de Lebesgue es monótona. Por tanto, $\displaystyle\lim_{n\to\infty} \int_E \|f_n\| \, d\mathbb{P} = \int_E \|f\| \, d\mathbb{P}$ para cada $E \in \mathcal{F}$. Finalmente,

$$\begin{aligned}\left\| \int_E f \, d\mathbb{P} \right\| &= \left\| \lim_{n\to\infty} \int_E f_n \, d\mathbb{P} \right\| = \lim_{n\to\infty} \left\| \int_E f_n \, d\mathbb{P} \right\| \\ &\leq \lim_{n\to\infty} \int_E \|f_n\| \, d\mathbb{P} = \int_E \|f\| \, d\mathbb{P}.\end{aligned}$$

(iii) Como la serie $\sum_{n=1}^{\infty} \int_{E_n} f \, d\mathbb{P}$ está dominada término a término por la serie de números no negativos $\sum_{n=1}^{\infty} \int_{E_n} \|f\| \, d\mathbb{P} \leq \int_\Omega \|f\| \, d\mathbb{P} < \infty$, entonces $\sum_{n=1}^{\infty} \int_{E_n} f \, d\mathbb{P}$ converge absolutamente. Notemos que

$$\left\| \int_{\bigcup_{n=1}^{\infty} E_n} f \, d\mathbb{P} - \sum_{n=1}^{m} \int_{E_n} f \, d\mathbb{P} \right\| = \left\| \int_{\bigcup_{n=m+1}^{\infty} E_n} f \, d\mathbb{P} \right\|$$

pues la integral de Bochner es finitamente aditiva.
Más aún, $\lim_{m\to\infty} \mathbb{P}(\bigcup_{n=m+1}^{\infty} E_n) = 0$ y por la parte (i) obtenemos que

$$\lim_{m\to\infty} \left\| \int_{\bigcup_{n=m+1}^{\infty} E_n} f \, d\mathbb{P} \right\| = 0.$$

Por lo tanto,

$$\int_{\bigcup_{n=1}^{\infty} E_n} f \, d\mathbb{P} = \sum_{n=1}^{\infty} \int_{E_n} f \, d\mathbb{P}.$$

(iv) Si π es una partición de un conjunto $E \in \mathcal{F}$, entonces

$$\begin{aligned}\sum_{A\in\pi} \|\mathbf{F}(A)\| &= \sum_{A\in\pi} \left\| \int_A f \, d\mathbb{P} \right\| \\ &\leq \sum_{A\in\pi} \int_A \|f\| \, d\mathbb{P} = \int_E \|f\| \, d\mathbb{P}.\end{aligned}$$

Como f es Bochner integrable, $|\mathbf{F}|(E) \leq \int_E \|f\| \, d\mathbb{P} < \infty$, así $\mathbf{F}$ es de variación acotada.

Por otra parte, sea $\varepsilon > 0$ y sea $(f_n)_n$ una sucesión de simples medibles tal que $\lim_{n\to\infty} \int_\Omega \|f - f_n\| \, d\mathbb{P} = 0$. Fijemos n_0 suficientemente grande para el cual $\int_\Omega \|f - f_{n_0}\| \, d\mathbb{P} < \varepsilon$, y para cada $E \in \mathcal{F}$ escojamos π' una partición de E tales que

$$\sum_{A\in\pi'} \left\| \int_A f_{n_0} \, d\mathbb{P} \right\| = \int_E \|f_{n_0}\| \, d\mathbb{P}.$$

Tal partición siempre existe, pues como f_{n_0} es simple existen $k \in \mathbb{N}$, $x_1, \ldots, x_k \in X$ y $E_1, \ldots, E_k \in \mathcal{F}$ con $E_i \cap E_j = \emptyset$ si $i \neq j$, tal que $f_{n_0} = \sum_{i=1}^k x_i \chi_{E_i}$. Entonces podríamos tomar la partición de E dada por $\pi' = \{E_1 \cap E, \ldots, E_k \cap E\}$, la cuál satisface la igualdad deseada. En efecto,

$$\begin{aligned}\sum_{A\in\pi'} \left\| \int_A f_{n_0} \, d\mathbb{P} \right\| &= \sum_{i=1}^k \left\| \int_{E_i \cap E} f_{n_0} \, d\mathbb{P} \right\| = \sum_{i=1}^k \|x_i \cdot \mathbb{P}(E_i \cap E)\| \\ &= \int_E \sum_{i=1}^k \|x_i\| \chi_{E_i} \, d\mathbb{P} = \int_E \|f_{n_0}\| \, d\mathbb{P}.\end{aligned}$$

Tomemos ahora π una partición de E tal que refine a π' y

$$\left| |\mathbf{F}|(E) - \sum_{B\in\pi} \left\| \int_B f \, d\mathbb{P} \right\| \right| < \varepsilon.$$

Note que podemos elegir tal partición por definición de la variación de $\mathbf{F}(E)$. Además, como π refina a π', se sigue cumpliendo que

$$\int_E \|f_{n_0}\| \, d\mathbb{P} = \sum_{B\in\pi} \left\| \int_B f_{n_0} \, d\mathbb{P} \right\|.$$

Más aún,

$$\begin{aligned}\sum_{B\in\pi} \left| \left\| \int_B f_n \, d\mathbb{P} \right\| - \left\| \int_B f_{n_0} \, d\mathbb{P} \right\| \right| &\leq \sum_{B\in\pi} \left\| \int_B f_n \, d\mathbb{P} - \int_B f_{n_0} \, d\mathbb{P} \right\| \\ &\leq \sum_{B\in\pi} \int_B \|f_n - f_{n_0}\| \, d\mathbb{P} \leq \int_E \|f_n - f_{n_0}\| \, d\mathbb{P} < \varepsilon.\end{aligned}$$

Y por lo tanto,

$$\begin{aligned}
\left| |\mathbf{F}|(E) - \int_E \|f_{n_0}\| d\mathbb{P} \right| &= \left| |\mathbf{F}|(E) - \sum_{B\in\pi} \int_B \|f_{n_0}\| d\mathbb{P} \right| \\
&\leq \left| |\mathbf{F}|(E) - \sum_{B\in\pi} \left\| \int_B f_{n_0} d\mathbb{P} \right\| \right| \\
&\leq \left| |\mathbf{F}|(E) - \sum_{B\in\pi} \left\| \int_B f_n d\mathbb{P} \right\| \right| + \sum_{B\in\pi} \left| \left\| \int_B f\, d\mathbb{P} \right\| - \left\| \int_B f_{n_0} d\mathbb{P} \right\| \right| \\
&< 2\varepsilon
\end{aligned}$$

Como esto se cumple para toda n_0 suficientemente grande concluimos que

$$|\mathbf{F}|(E) = \lim_{n\to\infty} \int_E \|f_n\| d\mathbb{P} = \int_E \|f\| d\mathbb{P}.$$

■

Como consecuencia de esta última afirmación obtenemos el siguiente corolario.

Corolario 1.4.11. *Si f y g son funciones Bochner integrables tales que para cada $E \in \mathcal{F}$, $\displaystyle\int_E f\, d\mathbb{P} = \int_E g\, d\mathbb{P}$, entonces $f = g$ c.s.*

Demostración. Sea $\mathbf{F}(E) = \int_E (f-g)\, d\mathbb{P}$. Entonces $F(E) = 0$ para cada $E \in \mathcal{F}$. Entonces $|\mathbf{F}|(E) = 0$ para cada $E \in \mathcal{F}$, en particular $0 = |\mathbf{F}|(\Omega) = \int_\Omega \|f-g\|\, d\mathbb{P}$ y por tanto $\|f-g\| = 0$ c.s., lo cual implica que $f = g$ c.s. ■

El siguiente teorema exhibe una propiedad fuerte de la teoría de integración de Bochner que no tiene análogo en la teoría de integración de Lebesgue.

Teorema 1.4.12 (Hille)**.** *Sea $T\colon X \to Y$ un operador lineal cerrado, con Y un espacio de Banach. Si f es Bochner integrable con respecto a $\mathbb{P}$, entonces $T(f)$ también lo es y*

$$T\left(\int_E f\, d\mathbb{P}\right) = \int_E T(f)\, d\mathbb{P},$$

para cada $E \in \mathcal{F}$.

Demostración. Probaremos el caso cuando T es un operador lineal acotado, es decir, continuo. La prueba del caso general puede consultase en [11] (pág. 47, teo. 6).

Observe que por ser T lineal y acotado entre espacios normados existe $M > 0$ tal que $\|T(f)\| \leq M\|f\|$, luego

$$\int_\Omega \|T(f)\|\, d\mathbb{P} \leq \int_\Omega M\|f\| < \infty$$

y por lo tanto $T(f)$ es Bochner integrable (ver teorema 1.4.7). Ahora, sea $(f_n)_n$ una sucesión de funciones simples para la cual $\lim_{n\to\infty} \int_E f_n \, d\mathbb{P} = \int_E f \, d\mathbb{P}$. Primero observemos que usando la linealidad de T,

$$T\left(\int_E f_n \, d\mathbb{P}\right) = \int_E T(f_n) \, d\mathbb{P}$$

para toda función simple f_n. Por otro lado, usando la continuidad de T :

$$\begin{aligned} T\left(\int_E f \, d\mathbb{P}\right) &= T\left(\lim_{n\to\infty} \int_E f_n \, d\mathbb{P}\right) = \lim_{n\to\infty} T\left(\int_E f_n \, d\mathbb{P}\right) \\ &= \lim_{n\to\infty} \int_E T(f_n) \, d\mathbb{P} = \int_E T(f) \, d\mathbb{P}, \end{aligned}$$

donde la última igualdad se debe a que

$$\left\| \int_E T(f) \, d\mathbb{P} - \int_E T(f_n) \, d\mathbb{P} \right\| = \left\| \int_E T(f - f_n) \, d\mathbb{P} \right\| \leq \|T\| \int_E \|f - f_n\| \, d\mathbb{P} \to 0$$

cuando $n \to \infty$. ■

Corolario 1.4.13. *Si $f\colon \Omega \to X$ es Bochner integrable, entonces es Pettis integrable y las integrales coinciden.*

Demostración. La primera parte el enuciado se encuentra probada en el lema 1.3.3. Veamos entonces que las integrales coinciden. Sea x la integral de Bochner de f, es decir, $x := \int_\Omega f \, d\mathbb{P}$. Por el teorema de Hille 1.4.12, encontramos que

$$x^*(x) = x^*\left(\int_\Omega f \, d\mathbb{P}\right) = \int_\Omega x^*(f) \, d\mathbb{P}$$

para cada $x^* \in X^*$, por lo tanto x es también la integral de Pettis de f. ■

Recordamos nuevamente a los espacios «L_p» introducidos en la sección anterior: Si $1 \leq p < \infty$, el símbolo $L_p(\Omega, \mathcal{F}, \mathbb{P}; X) = L_p(\Omega, X)$ denota al conjunto de clases de equivalencia de elementos aleatorios $f\colon \Omega \to X$ que son Bochner integrables (respecto a $\mathbb{P}$) y tales que

$$\|f\|_p := (\mathbb{E}\|f\|^p)^{\frac{1}{p}} = \left(\int_E \|f\|_X^p \, d\mathbb{P}\right)^{\frac{1}{p}} < \infty.$$

Esto es, $L_p(\Omega, \mathcal{F}, \mathbb{P}; X)$ consiste de las clases de elementos aleatorios $f\colon \Omega \to X$ Bochner integrables cuyas normas $\|f\|$ pertenecen al espacio de Lebesgue $L_p(\mathbb{R})$.

Para $p = \infty$ se define el espacio $L_\infty(\Omega, \mathcal{F}, \mathbb{P}, X) = L_\infty(\Omega, X)$ como el conjunto de clases de equivalencia de funciones esencialmente acotadas $\mathbb{P}$-Bochner integrables con la norma $\|\cdot\|_\infty$ dada por

$$\|f\|_\infty := \operatorname{ess\,sup} \|f\|_X = \inf\{M > 0 \, : \|f\|_X \leq M \text{ a.e.}\}.$$

Comentario. Algunos autores, como Hoffmann-Jørgensen [15], denotan por $L_0(\Omega, \mathcal{F}, \mathbb{P}; X)$, o en breve por $L_0(\Omega; X)$, al conjunto de todos elementos aleatorios en X. En él se define la «norma»[11] $\|\cdot\|_0$ como sigue:

$$\|V\|_0 := \mathbb{E}\left\{\frac{\|V\|}{1+\|V\|}\right\} \qquad \forall\, V \in L_0(\Omega; X).$$

Esta «norma» resulta ser bastante interesante pues una sucesión $V_n \to V$ en $L_0(\Omega; X)$ si y sólo si $V_n \xrightarrow{\mathbb{P}} V$ (ver apéndice B, proposición C.0.4).

Anteriormente, cuando presentamos estos espacios en la definición 1.3.5, no mencionamos la condición «Bochner integrable», en cambio pedimos que el espacio X fuese separable. El teorema de medibilidad de Pettis justifica este cambio pues toda función Bochner integrable es fuertemente medible, y estas a su vez tienen rango separable. En cualquier caso en este texto solo estaremos interesados en los espacios separables.

1.4.3. La Propiedad de Radon-Nikodým y la esperanza condicional II

Como hemos visto hasta ahora, hemos podido extender exitosamente varias propiedades de la integral de Legesgue a la integral de Bochner, y no solo eso, para probarlo usamos en general casi los mismos métodos que en la teoría de integración de Lebesgue. Con esto en mente surge la pregunta, ¿Cuándo una medida vectorial puede verse como una integral de Bochner indefinida? Precisemos lo anterior.

Si $(\Omega, \mathcal{F}, \mathbb{P})$ es un espacio de medida finita (en particular un espacio de probabilidad) y $\mathbf{F}\colon \mathcal{F} \to X$ es una medida vectorial de la forma

$$\mathbf{F}(E) = \int_E f \, d\mathbb{P}$$

para alguna función $f\colon \Omega \to X$ Bochner integrable, entonces el teorema 1.4.10 garantiza que $\mathbf{F}$ es numerablemente aditiva, de variación acotada y absolutamente continua respecto a $\mathbb{P}$ (observe que si tomamos $A \in \mathcal{F}$ tal que $\mathbb{P}(A) = 0$, entonces $f\chi_A = 0$ en $\Omega \setminus A$, el cuál es un conjunto de probabilidad 1, luego $f\chi_A = 0$ c.s. y por tanto $\mathbf{F}(A) = \mathbf{0}$).

¿Será cierto que existe una función f Bochner integrable (la *derivada de Radon-Nikodým* de $\mathbf{F}$) tal que para cada $E \in \mathcal{F}$, $\mathbf{F}(E) = \int_E f \, d\mathbb{P}$ para cada medida vectorial $\mathbf{F}$ de variación acotada y absolutamente continua respecto a $\mathbb{P}$?

La respuesta en general es negativa.

Proposición 1.4.14. *Existe una medida vectorial con valores en c_0 de variación acotada que no tiene derivada de Radon-Nikodým.*

[11] Tal vez «seminorma» sería más apropiado, ya que $\|V\|_0 = 0$ si y sólo si $V = \mathbf{0}$ c.s.

Demostración. Sea λ la medida de Lebesgue en el intervalo $[0,1]$. Si $\mathcal{L}$ representa la σ-álgebra de subconjuntos Lebesgue medibles del $[0,1]$, para cada $E \in \mathcal{L}$ y cada $n \in \mathbb{N}$ definimos

$$\lambda_n(E) = \int_E \operatorname{sen}(2^n \pi t)\ dt,$$

y sea $\mathbf{F}\colon \mathcal{L} \to c_0$ dada por

$$\mathbf{F}(E) = (\lambda_1(E), \lambda_2(E), \ldots, \lambda_n(E), \ldots).$$

Como $E \subset [0,1]$, el lema de Riemann-Lebesgue[12] nos garantiza que

$$\begin{aligned}
\lim_{n\to\infty} \lambda_n(E) &= \lim_{n\to\infty} \int_E \operatorname{sen}(2^n \pi t)\ dt \leq \lim_{n\to\infty} \int_0^1 \operatorname{sen}((2^n\pi)t)\ dt \\
&= \lim_{k\to\infty} \int_0^1 \operatorname{sen}(kt)\ dt = 0,
\end{aligned}$$

donde $k = 2^n\pi$. Por tanto $\mathbf{F}$ es una medida vectorial definida en c_0.
Observemos además que para cada $E \in \mathcal{L}$,

$$\begin{aligned}
\|\mathbf{F}(E)\|_{c_0} &:= \sup_n |\lambda_n(E)| \leq \sup_n \left| \int_E \operatorname{sen} 2^n \pi t\, dt \right| \\
&\leq \sup_n \int_E |\operatorname{sen} 2^n\pi t|\, dt \leq \sup_n \int_E 1\, d\mathbb{P} = \lambda(E).
\end{aligned}$$

Como resultado de la desigualdad anterior tenemos que $\mathbf{F}$ es absolutamente continua respecto a λ. Resta ver que $\mathbf{F}$ es numerablemente aditiva y de variación acotada.

Sea $(E_k)_k$ una sucesión de elementos en $\mathcal{L}$ disjuntos por pares y sea $E := \bigcup_{k=1}^{\infty} E_k$, veamos que $\mathbf{F}(E) = \sum_{k=1}^{\infty} \mathbf{F}(E_k)$.

$$\begin{aligned}
\left\| \mathbf{F}(E) - \sum_{k=1}^{m} \mathbf{F}(E_k) \right\|_{c_0} &= \|(\lambda_n(E))_n - (\lambda_n(\cup_{k=1}^m E_k))_n\|_{c_0} \\
&= \|(\lambda_n(E) - \lambda_n(\cup_{k=1}^m E_k))_n\|_{c_0} = \left\| \left(\lambda_n\left(\cup_{k=m+1}^{\infty} E_k\right)\right)_n \right\|_{c_0} \\
&= \sup_n \left|\lambda_n\left(\cup_{k=m+1}^{\infty} E_k\right)\right| = \sup_n \left| \sum_{k=m+1}^{\infty} \lambda_n(E_k) \right| \\
&\leq \sup_n \sum_{k=m+1}^{\infty} |\lambda_n(E_k)| \leq \sup_n \sum_{k=m+1}^{\infty} |\lambda(E_k)| \\
&= \sum_{k=m+1}^{\infty} \lambda(E_k) = \lambda\left(\cup_{k=m+1}^{\infty} E_k\right)
\end{aligned}$$

[12] Lema de Riemann-Lebesgue: Sea f una función Riemann-integrable definida en un intervalo $[a,b]$ de la recta real, entonces $\lim_{|\alpha|\to\infty} \int_a^b f(x)e^{i\alpha x}\, dx = 0$.

pero como $\lim_{m\to\infty} \lambda(\cup_{k=m+1}^{\infty} E_k) = 0$, entonces $\mathbf{F}(E) = \sum_{k=1}^{\infty} \mathbf{F}(E_k)$ y por lo tanto $\mathbf{F}$ es numerablemente aditiva.

Además $\mathbf{F}$ es de variación acotada pues si $(E_k)_{k=1}^{n}$ es una partición finita del intervalo $[0,1]$, entonces

$$\sum_{k=1}^{n} \|\mathbf{F}(E_k)\|_\infty \leq \sum_{k=1}^{n} \lambda(E_k) = \lambda\left(\bigcup_{k=1}^{n} E_k\right) = 1$$

y por tanto $|\mathbf{F}|([0,1]) \leq 1$.

Ahora supongamos que $\mathbf{F}$ tiene derivada de Radon-Nikodým, es decir, existe $f\colon [0,1] \to c_0$ Bochner integrable tal que $\mathbf{F}(E) = \int_E f\, d\lambda$ para todo $E \in \mathcal{L}$.
Sea $f = (f_n)_n$, si denotamos por $e_n^*\colon c_0 \to \mathbb{R}$ al n-ésimo funcional coordenado en c_0 definido como $e_n^*(x) = x_n$ para cada $x = (x_n)_n \in c_0$, entonces $f_n = e_n^*(f)$ de modo que para cada n, f_n es continua y por tanto medible.
Como la variación en cada $E \in \mathcal{F}$ se calcula como $|\mathbf{F}|(E) = \int_E \|f\|_{c_0}\, d\lambda$, entonces

$$\int_E \|f\|_{c_0}\, d\lambda = \int_E \sup_n |f_n|\, d\lambda \leq 1$$

para cada $E \in \mathcal{F}$, luego $f_n\colon [0,1] \to \mathbb{R}$ es Lebesgue integrable para cada n.
Ahora notemos que para cada $m \in \mathbb{N}$,

$$\begin{aligned}\int \left\|\sum_{n=1}^{m} f_n e_n\right\|_{c_0} d\lambda &\leq \int \sum_{n=1}^{m} |f_n|\, \|e_k\|_{c_0}\, d\lambda \\ &\leq \int \sum_{n=1}^{m} |f_n|\, d\lambda = \sum_{n=1}^{m} \int |f_n|\, d\lambda < \infty\end{aligned}$$

donde $(e_n)_n$ denota a la base canónica de c_0.
Entonces el teorema 1.4.7 nos garantiza que

$$\sum_{n=1}^{m} f_n e_n := (f_1, f_2, \dots, f_{m-1}, f_m, 0, 0, \dots)$$

es una función Bochner integrable para cada m.
Por el teorema 1.4.9 de Convergencia Dominada para la integral de Bochner:

$$\begin{aligned}\mathbf{F}(E) = \int_E f\, d\lambda &= \int_E \left(\lim_{m\to\infty} \sum_{n=1}^{m} f_n e_n\right) d\lambda = \lim_{m\to\infty} \int_E \left(\sum_{n=1}^{m} f_n e_n\right) d\lambda \\ &= \lim_{m\to\infty} \sum_{n=1}^{m} \left(\int_E f_n\, d\lambda\right) e_n = \left(\int_E f_n\, d\lambda\right)_{n=1}^{\infty}\end{aligned}$$

Pero por definición $\mathbf{F}(E) = \left(\int_E \operatorname{sen}(2^n \pi t)\, dt\right)_{n=1}^{\infty}$, así que para toda n se satisface la siguiente igualdad

$$\int_E \operatorname{sen}(2^n \pi t)\, dt = \int_E f\, d\lambda,$$

y por tanto $f(t) = \operatorname{sen}(2^n \pi t)$ para casi toda $t \in [0,1]$.
Para cada n definamos los conjuntos

$$E_n = \left\{ t \in [0,1] \,:\, f_n(t) \geq \frac{1}{\sqrt{2}} \right\}.$$

Entonces $\lambda(E_n) = \frac{1}{4}$ para cada n. Más aún, por el lema de Fatou:

$$\lambda\left(\limsup_n E_n \right) \geq \limsup_n \lambda(E_n) \geq \frac{1}{4}.$$

Por último notemos que si $t \in \limsup_n E_n$, entonces $f(t) \notin c_0$ pues $f(t) \geq \frac{1}{\sqrt{2}}$ para toda n. Por lo tanto

$$\lambda\left(\{ t \in [0,1] \,:\, f(t) \in c_0 \} \right) \leq \frac{3}{4} < 1,$$

lo que contradice que f es una función con valores en c_0. ■

Se sabe que el espacio ℓ_1 cumple el teorema de Radon-Nikodým (para la prueba ver [12], pág. 181). De modo que hay espacios de Banach que satisfacen el teorema de Radon-Nikodým y otros que no, así que la validez o no de este teorema se considera como una propiedad del mismo espacio.

Definición 1.4.10. Un espacio de Banach X **tiene la Propiedad de Radon-Nikodým con respecto a** $(\Omega, \mathcal{F}, \mathbb{P})$, si y solo si, para toda medida vectorial $\mathbf{F}\colon \mathcal{F} \to X$ de variación acotada y absolutamente continua respecto a $\mathbb{P}$ existe una función Bochner integrable f tal que

$$\mathbf{F}(E) = \int_E f\, d\mathbb{P}$$

para cada $E \in \mathcal{F}$. Llamaremos a f la **derivada de Radon-Nikodým** de $\mathbf{F}$ con respecto a $\mathbb{P}$ y la denotaremos por

$$f = \frac{d\mathbf{F}}{d\mathbb{P}}.$$

Definición 1.4.11. Un espacio de Banach X **tiene la Propiedad de Radon-Nikodým (PRN)**, si y solo si X tiene la Propiedad de Radon-Nikodým respecto a cualquier espacio de medida finito.

En relación con el espacio de medida $([0,1], \mathcal{L}, \lambda)$ utilizado en el ejemplo de la proposición 1.4.14 es sabido que si X tiene la propiedad de Radon-Nikodým con respecto a dicho espacio $([0,1], \mathcal{L}, \lambda)$, entonces X tiene la propiedad de Radon-Nikodým. Esta propiedad ha sido ampliamente estudiada, ejemplos de espacios con la PRN son los reflexivos, aquellos con dual separable o los espacios de Hilbert.

Otra variante de esta definición se encuentra en [12] (ver pág. 178), y se relaciona con la integral de Pettis:

Definición 1.4.12. Un espacio de Banach X tiene la **Propiedad Débil de Radon-Nikodým** si y solo si para cada espacio de medida finito (Ω, Σ, μ) y cada medida vectorial $\mathbf{F}\colon \Sigma \to X$ de variación σ-finita, absolutamente continua respecto a μ, existe una función f Pettis integrable tal que $\mathbf{F}(E) = \int f \cdot \chi_E \, d\mu$ para cada $E \in \Sigma$. (Aquí el símbolo de integral representa a la integral de Pettis.)

Hay ciertos casos especiales en los que la derivada de Radon-Nikodým existe independientemente de la Propiedad de Radon-Nikodým (PRN). Uno de ellos es cuando el espacio de medida (Ω, Σ, μ) es atómico (ver [12], pág. 178). Otro caso especial es la esperanza condicional que estudiamos previamente en el teorema 1.3.6. Tal prueba muestra que no es necesario asumir la PRN para concluir su existencia, y no difiere en nada pedir que V sea Bochner integrable a que $V \in L_1(\Omega, \mathcal{F}, \mathbb{P}; X)$ pues hemos mostrado en el teorema 1.4.7 que tales condiciones son equivalentes. Enunciamos entonces la definición y el teorema en este nuevo contexto.

Definición 1.4.13. Sea X un espacio de Banach, $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y sea $V\colon \Omega \to X$ un elemento aleatorio Bochner integrable. Sea $\mathcal{G} \subset \mathcal{F}$ una σ-álgebra. La **esperanza condicional** de X dado $\mathcal{G}$ es el elemento aleatorio $Z\colon \Omega \to X$ único (c.s.), medible con respecto a $\mathcal{G}$, tal que

$$\mathbb{E}\left[Z\mathbb{I}_A\right] = \mathbb{E}\left[V\mathbb{I}_A\right]$$

para cada $A \in \mathcal{G}$. Y la denotamos por $Z = \mathbb{E}^{\mathcal{G}}\left[V\right]$ o $\mathbb{E}\left[V \mid \mathcal{G}\right]$.

Teorema 1.4.15 ([32])**.** *Sea X un espacio de Banach, sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, sea $V\colon \Omega \to X$ un elemento aleatorio Bochner integrable, y sea $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$. Entonces la esperanza condicional $\mathbb{E}\left[V \mid \mathcal{G}\right]$ existe.*

Demostración. Como V es Bochner integrable existe una sucesión de elementos aleatorios simples $(V_n)_n$ tal que $V_n(\omega) \to V(\omega)$ c.s.; $\int_\Omega \|V_n(\omega) - V(\omega)\| \, d\mathbb{P} \to 0$ cuando $n \to \infty$; y $\int_\Omega V_n(\omega) \, d\mathbb{P} \to \int_\Omega V(\omega) \, d\mathbb{P}$. Recordemos que $\mathbb{E}\left[V_n|\mathcal{G}\right]$ está bien definida para cada V_n y se define como $\mathbb{E}\left[V_n|\mathcal{G}\right] = \sum_{i=1}^{m_n} x_{n,i}\mathbb{E}\left[\mathbb{I}_{A_{n,i}}|\mathcal{G}\right]$. Entonces

$$\begin{aligned}
\int_\Omega \left\|\mathbb{E}\left[V_n|\mathcal{G}\right] - \mathbb{E}\left[V_m|\mathcal{G}\right]\right\| d\mathbb{P} &= \int_\Omega \left\|\mathbb{E}\left[V_n - V_m|\mathcal{G}\right]\right\| d\mathbb{P} \\
&\leq \int_\Omega \mathbb{E}\left[\|V_n - V_m\| \, |\mathcal{G}\right] \, d\mathbb{P} \qquad (1.11) \\
&= \int_\Omega \|V_n - V_m\| \, d\mathbb{P} \to 0
\end{aligned}$$

en donde la primera igualdad se debe a la linealidad de la esperanza condicional en las funciones simples, la segunda desigualdad se debe a que si V_n es simple, entonces

$$
\begin{aligned}
\|\mathbb{E}[V_n|\mathcal{G}](\omega)\| &\leq \left\| \sum_{i=1}^{m_n} x_{n,i} \mathbb{E}\left[\mathbb{I}_{A_{n,i}}|\mathcal{G}\right] \right\| \\
&\leq \sum_{i=1}^{m_n} \|x_{n,i}\| \cdot \mathbb{E}\left[\mathbb{I}_{A_{n,i}}|\mathcal{G}\right](\omega) \\
&=: \mathbb{E}\left[\|V_n\| \, |\mathcal{G}\right](\omega).
\end{aligned}
$$

La última igualdad en (1.11) es consecuencia de la definición de esperanza condicional escalar de la variable aleatoria $\|V_n - V_m\| \colon \Omega \to \mathbb{R}$. Ahora, por el teorema 3.6.1 en [27] (ver pág. 45) existe una función $Z \colon (\Omega, \mathcal{G}, \mathbb{P}; X) \to X$ Bochner integrable y por tanto fuertemente medible, único c.s., y tal que

$$
\int_\Omega \|\mathbb{E}[V_n|\mathcal{G}] - Z(\omega)\| \, d\mathbb{P} \to 0 \quad \text{cuando } n \to \infty. \tag{1.12}
$$

En consecuencia,

$$
\begin{aligned}
\left\| \int_A Z(\omega) - V(\omega) \, d\mathbb{P} \right\| &= \left\| \int_A (Z(\omega) - \mathbb{E}[V_n|\mathcal{G}]) \, d\mathbb{P} + \int_A (\mathbb{E}[V_n|\mathcal{G}] - V(\omega)) \, d\mathbb{P} \right\| \\
&\leq \int_A \|Z(\omega) - \mathbb{E}[V_n|\mathcal{G}]\| \, d\mathbb{P} + \left\| \int_A \mathbb{E}[V_n|\mathcal{G}] - V(\omega) \, d\mathbb{P} \right\| \\
&= \int_A \|Z(\omega) - \mathbb{E}[V_n|\mathcal{G}]\| \, d\mathbb{P} + \left\| \int_A V_n \, d\mathbb{P} - \int_A V(\omega) \, d\mathbb{P} \right\|
\end{aligned}
$$

que tiende a 0 conforme $n \to \infty$. Por lo tanto $\int_A Z(\omega) \, d\mathbb{P} = \int_A V(\omega) \, d\mathbb{P}$ para cada $A \in \mathcal{G}$. Así que Z satisface la definición 1.4.13. ■

Esto motiva a presentar el siguiente resultado.

Teorema 1.4.16. *Si $V \in L_1(\Omega, \mathcal{F}, \mathbb{P}; X)$ (Bochner integrable) y $\mathcal{G}$ es una sub-σ-álgebra de $\mathcal{F}$, entonces $\|\mathbb{E}[V|\mathcal{G}]\| \leq \mathbb{E}[\|V\| |\mathcal{G}]$ c.s.*

Demostración. Sea $(V_n)_n$ la sucesión de simples que converge a V como en el teorema 1.4.15. Previamente hemos mostrado que $\|\mathbb{E}[V_n|\mathcal{G}]\| \leq \mathbb{E}[\|V_n\| |\mathcal{G}]$ *c.s.* entonces, por la monotonía de la integral de Lebesgue,

$$
\int_A \|\mathbb{E}[V_n|\mathcal{G}]\| \, d\mathbb{P} \leq \int_A \mathbb{E}[\|V_n\| |\mathcal{G}] \, d\mathbb{P}.
$$

Pero por (1.12) usada en la prueba del teorema anterior

$$
\int_A \|\mathbb{E}[V_n|\mathcal{G}]\| \, d\mathbb{P} \to \int_A \|\mathbb{E}[V|\mathcal{G}]\| \, d\mathbb{P} \quad \text{cuando } n \to \infty,
$$

y también

$$
\int_A \mathbb{E}[\|V_n\| |\mathcal{G}] \, d\mathbb{P} = \int_A \|V_n\| \, d\mathbb{P} \to \int_A \|V\| \, d\mathbb{P} = \int_A \mathbb{E}[\|V\| |\mathcal{G}] \, d\mathbb{P}
$$

cuando $n \to \infty$. Por lo tanto

$$\int_A \|\mathbb{E}\left[V|\mathcal{G}\right]\| \, d\mathbb{P} \leq \int_A \mathbb{E}\left[\|V\||\mathcal{G}\right] \, d\mathbb{P}$$

para cada $A \in \mathcal{G}$. Esto concluye el resultado. ■

Teorema 1.4.17 (Desigualdad de Jensen)**.** *Sean X un espacio de Banach, $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$, y $V\colon \Omega \to X$ un elemento aleatorio Bochner integrable. Sea $C \subset X$ un subconjunto cerrado y convexo, y sea $\varphi\colon C \to \mathbb{R}$ una función convexa y continua. Si $V \in C$ c.s. y $\mathbb{E}[|\varphi(V)|] < \infty$, entonces*

$$\varphi\left(\mathbb{E}[V|\mathcal{G}]\right) \leq \mathbb{E}[\varphi(V)|\mathcal{G}].$$

Demostración. (Bosquejo) En el espacio de Banach $X' = X \oplus \mathbb{R}$ consideremos el conjunto $C' = \{(x,t) \,:\, x \in C, t \geq \varphi(x)\}$. Entonces C' es convexo pues φ es una función convexa y también es cerrado pues φ es una función continua. Así, definimos $V'\colon \Omega \to X$ como $V'(\omega) = (V(\omega), \varphi(V(\omega)))$. Entonces V' es Bochner integrable y toma valores en C', entonces $\mathbb{E}[V'|\mathcal{G}]$ existe y también toma valores en C'. Más aún,

$$\mathbb{E}[V'|\mathcal{G}] = (\mathbb{E}[V|\mathcal{G}], \mathbb{E}[\varphi(V)|\mathcal{G}]),$$

y por lo tanto $\mathbb{E}[\varphi(V)|\mathcal{G}] \geq \varphi(\mathbb{E}[V|\mathcal{G}])$. ■

1.5. Comentarios adicionales

Hemos presentado dos definiciones de integrales. Como hemos visto la integral de Pettis no usa ninguna aproximación por funciones simples, razón por la cuál algunos de los teoremas clásicos de la integral de Lebesgue no se siguen.

Por el contrario, la construcción de la integral de Bochner está fuertemente ligada con las funciones simples y la densidad de estas en $L_1(\Omega; X)$. En ambos casos la condición de separabilidad no fue un requisito; esa condición surge por otras razones que hemos expuesto en las secciones 1.1 y 1.2 de este capítulo.

Y como sucede con la integral de Lebesgue y la integral de Riemann, las de Bochner y Pettis no siempre coinciden. En el lema 1.3.3 probamos que Bochner implica Pettis, pero la implicación de regreso en general no se da. Un contraejemplo puede ser encontrado en [12] (ver pág. 182 ejem. 5.1.20). Se sabe que bajo la hipótesis de separabilidad (como en nuestro caso) los dos conceptos de integrabilidad coinciden. Más aún, Pettis observó que cualquier otra definición de integral de funciones $f\colon \Omega \to X$, que coincidiera con la de Lebesgue cuando $X = \mathbb{R}$ y con la propiedad listada en el inciso e) del teorema 1.3.1, está incluida en su definición en el sentido de que la integral de Pettis de f debe existir y ambas deben tener el mismo valor (ver [26] y [27]).

Para nuestro estudio nos es suficiente con el criterio dado en el lema 1.3.3, para más detalles sobre la integral de Pettis referimos al lector a los trabajos [11], [34] y [38].

A lo largo de este texto utilizaremos principalmente a la integral de Pettis para definir a la esperanza de un elemento aleatorio sin embargo, como posteriormente haremos notar, no usaremos ninguna propiedad particular de esta integral que no esté listada en el teorema 1.3.1, y que también cumple la integral de Bochner.

Intentamos dar dos pruebas distintas de la existencia de la esperanza condicional (que en esencia son las misma pues en ambos teoremas (1.3.6 y 1.4.15) asumimos Bochner integrabilidad del elemento aleatorio). No obstante debe observarse que la separabilidad no es necesaria siempre que se asuma Bochner integrabilidad (no es suficiente la condición de Pettis integrabilidad por si sola) como muestra la prueba del teorema 1.4.15. Y aunque la Propiedad de Radon-Nikodýn no resultó crucial para mostrar la existencia de la esperanza condicional, sí lo es para extender el siguiente teorema de Doob para martingalas en espacios de Banach:

Teorema (Doob)**.** *Sea* $(f_n)_n$ *una martingala (escalar) con respecto a* $(\mathcal{F}_n)_n$*, o más general, una submartingala (*$f_n \leq \mathbb{E}\,[f_m|\mathcal{F}_n]$ *para* $n \geq m$*). Si* $\sup_n \mathbb{E}|f_n| < \infty$, *entonces la sucesión* $(f_n)_n$ *converge casi seguramente.*

Resulta que en un espacio de Banach, se cumple el teorema de Doob si y solo si el espacio tiene la Propiedad de Radon-Nikodýn. El lector interesado puede consultar este resultado en [38] (ver pág. 135, teo. 4.3).

Capítulo 2

Leyes de los grandes números en espacios de Hilbert

Antes de estudiar leyes de grandes números en espacios normados vamos a hacerlo en un caso particular de estos: los espacios de Hilbert. A lo largo de este capítulo denotaremos por $(H, \langle\cdot,\cdot\rangle)$ al espacio de Hilbert real separable con producto interior $\langle\cdot,\cdot\rangle\colon H \times H \to \mathbb{R}$.

La geometría que nos provee el producto interior nos permitirá extender el concepto «no correlación» en H. Recordemos que en el caso clásico ($H = \mathbb{R}$), dos variables aleatorias U, W son no correlacionadas si $\mathbb{E}[U \cdot W] = \mathbb{E}[U] \cdot \mathbb{E}[W]$. Esta será la definición de elementos aleatorios no correlacionados en H cuando «$\cdot$» es un producto interior arbitario. No solamente eso, como también sucede en el caso clásico, probaremos que independencia implica no correlación.

Este resultado nos permitirá establecer la desigualdad de Kolmogorov para elementos aleatorios independientes así como una ley fuerte de grandes números que se deriva de la misma. Antes de llegar a este resultado estudiaremos el concepto de «no correlación» así como algunas leyes de grandes números para elementos aleatorios que satisfacen esta condición.

2.1. Sucesiones de elementos aleatorios no correlacionados

Sean V y Z dos elementos aleatorios en H tales que $\mathbb{E}[\|V\|^2] < \infty$ y $\mathbb{E}[\|Z\|^2] < \infty$. Si H es separable, entonces por la proposición 1.2.1 $V + Z$ y $V - Z$ también son elementos aleatorios en H, y en consecuencia $\|V + Z\|$ y $\|V - Z\|$ son variables aleatorias. Por la identidad de polarización obtenemos que

$$\langle V, Z \rangle = \frac{1}{4}\left(\|V+Z\|^2 + \|V-Z\|^2\right)$$

es una variable aleatoria. Más aún, $\mathbb{E}[\langle V, Z\rangle]$ existe pues

$$\mathbb{E}[|\langle V, Z\rangle|] \leq \mathbb{E}\left[\|V\| \cdot \|Z\|\right] \leq \left(\mathbb{E}[\|V\|^2]\right)^{\frac{1}{2}} \left(\mathbb{E}[\|Z\|^2]\right)^{\frac{1}{2}} < \infty.$$

Y ya que todo espacio de probabilidad es de medida finita, $\mathbb{E}[\|V\|^2] < \infty$ implica que $\mathbb{E}[\|V\|] < \infty$. Pero como H es completo y separable, el lema 1.3.3 garantiza que $\mathbb{E}[V]$ existe. Con esto justificamos la siguiente definición.

Definición 2.1.1. Dos elementos aleatorios V y Z en un espacio de Hilbert separable H tales que $\mathbb{E}[\|V\|^2] < \infty$ y $\mathbb{E}[\|Z\|^2] < \infty$ son **no correlacionados** si

$$\mathbb{E}[\langle V, Z\rangle] = \langle \mathbb{E}[V], \mathbb{E}[Z]\rangle.$$

Y se dicen **ortogonales** si

$$\mathbb{E}[\langle V, Z\rangle] = 0.$$

Se sigue de la definición anterior que si V y Z no están correlacionados entonces $V - \mathbb{E}[V]$ y $Z - \mathbb{E}[Z]$ son ortogonales. Más aún, podemos definir la **covarianza** de dos elementos aleatorios en H como

$$\mathrm{Cov}(V, Z) = \mathbb{E}[\langle V - \mathbb{E}[V], Z - \mathbb{E}[Z]\rangle].$$

Notemos que para cada $x \in H$, la función $f_x \colon H \to \mathbb{R}$ dada por $f_x(y) = \langle x, y\rangle = \langle y, x\rangle$ es lineal y continua, es decir $f_x \in H^*$. Además,

$$\begin{aligned}
\mathbb{E}[\langle V - \mathbb{E}[V], Z - \mathbb{E}[Z]\rangle] &= \mathbb{E}[\langle V, Z\rangle] - \mathbb{E}[\langle \mathbb{E}[V], Z\rangle] - \mathbb{E}[\langle V, \mathbb{E}[Z]\rangle] + \mathbb{E}\left[\langle \mathbb{E}[V], \mathbb{E}[Z]\rangle\right] \\
&= \mathbb{E}[\langle V, Z\rangle] - \mathbb{E}\left[f_{\mathbb{E}[V]}(Z)\right] - \mathbb{E}\left[f_{\mathbb{E}[Z]}(V)\right] + \langle \mathbb{E}[V], \mathbb{E}[Z]\rangle \\
&= \mathbb{E}[\langle V, Z\rangle] - f_{\mathbb{E}[V]}(\mathbb{E}[Z]) - f_{\mathbb{E}[Z]}(\mathbb{E}[V]) + \langle \mathbb{E}[V], \mathbb{E}[Z]\rangle \\
&= \mathbb{E}[\langle V, Z\rangle] - \langle \mathbb{E}[V], \mathbb{E}[Z]\rangle - \langle \mathbb{E}[V], \mathbb{E}[Z]\rangle + \langle \mathbb{E}[V], \mathbb{E}[Z]\rangle \\
&= \mathbb{E}[\langle V, Z\rangle] - \langle \mathbb{E}[V], \mathbb{E}[Z]\rangle.
\end{aligned}$$

Comentario. Recordemos que $\mathbb{E}[Z]$ es el elemento en H para el cuál se satiface que $f(\mathbb{E}[Z]) = \mathbb{E}[f(Z)]$ para cada $f \in H^*$, en particular satisface lo anterior para f_x (y para cada $x \in H$). Por otra parte, la igualdad anterior nos dice dos cosas. Primero, dos elementos aleatorios con esperanza nula son no correlacionados si y solo si son ortogonales. Segundo, ya que la covarianza de V y Z puede expresarse como

$$\mathrm{Cov}(V, Z) = \mathbb{E}[\langle V, Z\rangle] - \langle \mathbb{E}[V], \mathbb{E}[Z]\rangle,$$

V y Z no están correlacionados si y solo si su covarianza es cero. Y no solo eso,

$$\mathrm{Var}(V) = \mathbb{E}\left[\|V - \mathbb{E}[V]\|^2\right] = \mathbb{E}[\langle V - \mathbb{E}[V], V - \mathbb{E}[V]\rangle] = \mathrm{Cov}(V, V).$$

De esta definición se sigue que la varianza de una suma finita de elementos aleatorios no correlacionados es la suma de las varianzas. En efecto,

$$\begin{aligned}
\operatorname{Var}\left(\sum_{i=1}^{n} V_i\right) &= \mathbb{E}\left[\left\langle \sum_{i=1}^{n} V_i - \mathbb{E}\left(\sum_{i=1}^{n} V_i\right), \sum_{i=1}^{n} V_i - \mathbb{E}\left(\sum_{i=1}^{n} V_i\right)\right\rangle\right] \\
&= \mathbb{E}\left[\left\langle \sum_{i=1}^{n}(V_i - \mathbb{E}\left[V_i\right]), \sum_{j=1}^{n}(V_j - \mathbb{E}\left[V_j\right])\right\rangle\right] \\
&= \sum_{i=1}^{n} \mathbb{E}\left[\left\langle V_i - \mathbb{E}\left[V_i\right], \sum_{j=1}^{n}(V_j - \mathbb{E}\left[V_j\right])\right\rangle\right] \\
&= \sum_{i=1}^{n} \mathbb{E}\left[\sum_{j=1}^{n} \left\langle V_i - \mathbb{E}\left[V_i\right], V_j - \mathbb{E}\left[V_j\right]\right\rangle\right] \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{E}\left[\left\langle V_i - \mathbb{E}\left[V_i\right], V_j - \mathbb{E}\left[V_j\right]\right\rangle\right] \\
&= \sum_{i=1}^{n} \mathbb{E}\left[\left\langle V_i - \mathbb{E}\left[V_i\right], V_i - \mathbb{E}\left[V_i\right]\right\rangle\right] + \sum_{i\neq j} \mathbb{E}\left[\langle V_i - \mathbb{E}[V_i], V_j - \mathbb{E}[V_j]\rangle\right] \\
&= \sum_{i=1}^{n} \operatorname{Var}(V_i) + \sum_{i\neq j} \operatorname{Cov}(V_i, V_j),
\end{aligned}$$

donde el segundo sumando es cero por ser $(V_i)_{i=1}^n$ una colección finita de e.a. no correlacionados. Una primera consecuencia de esta propiedad es establecida en el siguiente resultado: el caso más simple de «ley de los grandes números».

2.1.1. Teoremas límites elementales

Teorema 2.1.1. *Si $(V_n)_n$ es una sucesión de elementos aleatorios no correlacionados en H espacio de Hilbert separable tales que*

$$\frac{1}{n^2}\sum_{k=1}^{n} \operatorname{Var}(V_k) \to 0, \tag{2.1}$$

entonces

$$\left\| \frac{1}{n}\sum_{k=1}^{n}(V_k - \mathbb{E}[V_k]) \right\| \to 0$$

en probabilidad.

Demostración. Sea $\varepsilon > 0$, por la desigualdad 1.2 de Chebyshev tenemos que

$$\begin{aligned}
\mathbb{P}\left[\left\|\frac{1}{n}\sum_{k=1}^{n}(V_k - \mathbb{E}[V_k])\right\| > \varepsilon\right] &\leq \frac{1}{\varepsilon^2}\mathbb{E}\left[\left\|\frac{1}{n}\sum_{k=1}^{n}(V_k - \mathbb{E}[V_k])\right\|^2\right] \\
&= \left(\frac{1}{\varepsilon n}\right)^2 \mathbb{E}\left[\left\langle \sum_{k=1}^{n}(V_k - \mathbb{E}[V_k]), \sum_{j=1}^{n}(V_j - \mathbb{E}[V_j])\right\rangle\right] \\
&= \left(\frac{1}{\varepsilon n}\right)^2 \sum_{k=1}^{n}\sum_{j=1}^{n} \mathbb{E}\left[\langle V_k - \mathbb{E}[V_k], V_j - \mathbb{E}[V_j]\rangle\right] \\
&= \left(\frac{1}{\varepsilon n}\right)^2 \left\{\sum_{k=1}^{n} \mathbb{E}\left[\langle V_k - \mathbb{E}[V_k], V_k - \mathbb{E}[V_k]\rangle\right]\right. \\
&\qquad \left. + \sum_{k\neq j}^{n} \mathbb{E}\left[\langle V_k - \mathbb{E}[V_k], V_j - \mathbb{E}[V_j]\rangle\right]\right\} \\
&= \left(\frac{1}{\varepsilon n}\right)^2 \sum_{k=1}^{n} \mathbb{E}\left[\|V_k - \mathbb{E}[V_k]\|^2\right] = \frac{1}{\varepsilon^2}\frac{1}{n^2}\sum_{k=1}^{n}\mathrm{Var}(V_k),
\end{aligned}$$

tomando el límite cuando $n \to \infty$ obtenemos el resultado. ■

Un corolario inmediato del teorema 2.1.1 es la ley débil de los grandes números para elementos aleatorios idénticamente distribuidos y no correlacionados, pues de (2.1) obtenemos

$$\frac{1}{n^2}\sum_{k=1}^{n}\mathrm{Var}(V_k) = \frac{1}{n}\mathrm{Var}(V_1) \to \mathbf{0}$$

cuando $n \to \infty$. Eso nos habla de lo fuerte que es la condición (2.1). A continuación estudiamos una ley con una condición más interesante y menos restrictiva que (2.1): acotamiento uniforme de las varianzas.

Teorema 2.1.2. *Si $(V_n)_n$ es una sucesión de elementos aleatorios no correlacionados en H espacio de Hilbert separable tales que para cada n, $\mathrm{Var}(V_n) \leq M$ donde M es una constante, entonces*

$$\frac{1}{n}\sum_{k=1}^{n}(V_k - \mathbb{E}[V_k]) \to \mathbf{0}$$

casi seguramente.

Demostración. Sin pérdida de generalidad supongamos que para cada número natural n, $\mathbb{E}[V_n] = \mathbf{0}$. De modo que $(V_n)_n$ no solo es una sucesión de elementos aleatorios no correlacionados sino también ortogonales. Sea $S_n = \sum_{k=1}^{n} V_k$, entonces

$$\mathbb{E}\left[\|S_n\|^2\right] = \mathrm{Var}(S_n) = \mathrm{Var}\left(\sum_{k=1}^{n} V_k\right) = \sum_{k=1}^{n}\mathrm{Var}(V_k) \leq M \cdot n,$$

y por la desigualdad de Chebyshev, para cada n

$$\mathbb{P}\left[\left\|\frac{S_n}{n}\right\| > \varepsilon\right] \leq \frac{1}{\varepsilon^2 n^2}\mathbb{E}\left[\|S_n\|^2\right] \leq \frac{M \cdot n}{\varepsilon^2 \cdot n^2} = \frac{M}{\varepsilon^2 n}.$$

Tenemos entonces que

$$\sum_{n=1}^{\infty} \mathbb{P}\left[\|S_{n^2}\| > \varepsilon n^2\right] \leq \sum_{n=1}^{\infty} \frac{M}{\varepsilon^2 n^2} = \frac{M}{\varepsilon^2} \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty.$$

Luego, por el lema 1.1.10 de Borel-Cantelli,

$$\mathbb{P}\left(\frac{\|S_{n^2}\|}{n^2} > \varepsilon \text{ para una infinidad de } n\right) = 0,$$

es decir, $\dfrac{S_{n^2}}{n^2} \to \mathbf{0}$ casi seguramente cuando $n \to \infty$. Ahora, para cada $n \geq 1$ definimos

$$D_n = \max_{n^2 \leq k < (n+1)^2} \|S_k - S_{n^2}\|.$$

Como en el intervalo $[n^2, (n+1)^2) \subset \mathbb{N}$, k toma exactamente $2n$ valores, entonces,

$$D_n^2 = \left(\max_{n^2 \leq k < (n+1)^2} \|S_k - S_{n^2}\|\right)^2 \leq 2n\|S_{(n+1)^2} - S_{n^2}\|^2 = 2n\left\|\sum_{j=n^2+1}^{(n+1)^2} V_j\right\|^2.$$

Por lo tanto,

$$\begin{aligned}
\mathbb{E}\left[D_n^2\right] &\leq 2n\mathbb{E}\left[\left\|\sum_{j=n^2+1}^{(n+1)^2} V_j\right\|^2\right] = 2n\mathbb{E}\left[\left\langle \sum_{j=n^2+1}^{(n+1)^2} V_j, \sum_{j=n^2+1}^{(n+1)^2} V_j \right\rangle\right] \\
&= 2n\mathbb{E}\left[\sum_{j=n^2+1}^{(n+1)^2} \|V_j\|^2 + \sum_{j\neq k} \langle V_j, V_k\rangle\right] \\
&= 2n \sum_{j=n^2+1}^{(n+1)^2} \mathbb{E}\left[\|V_j\|^2\right] + 2n\sum_{j\neq k} \mathbb{E}[\langle V_j, V_k\rangle] \\
&= 2n \sum_{j=n^2+1}^{(n+1)^2} \mathbb{E}\left[\|V_j\|^2\right] \\
&= 2n \sum_{j=n^2+1}^{(n+1)^2} \mathrm{Var}(V_j) \leq (2n)(2nM) = 4n^2 M.
\end{aligned}$$

Aplicando nuevamente la desigualdad de Chebyshev,

$$\mathbb{P}\left(\frac{D_n}{n^2} > \varepsilon\right) \leq \frac{1}{\varepsilon^2 n^4}\mathbb{E}\left[D_n^2\right] \leq \frac{4Mn^2}{\varepsilon^2 n^4} = \frac{4M}{\varepsilon^2}\frac{1}{n^2}.$$

Luego,

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\frac{D_n}{n^2} > \varepsilon\right) \leq \frac{4M}{\varepsilon^2} \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty.$$

Así que por Borel-Cantelli

$$\mathbb{P}\left(\frac{D_n}{n^2} > \varepsilon \text{ para una infinidad de } n\right) = 0,$$

y por tanto $\dfrac{D_n}{n^2} \to 0$ casi seguramente. Finalmente observemos que para cada $k \in \mathbb{N}$ tal que $n^2 \leq k < (n+1)^2$,

$$\begin{aligned} \|S_k\| &= \|S_{n^2} + S_k - S_{n^2}\| \leq \|S_{n^2}\| + \|S_k - S_{n^2}\| \\ &\leq \|S_{n^2}\| + \max_{n^2 \leq k < (n+1)^2} \|S_k - S_{n^2}\| \\ &= \|S_{n^2}\| + D_n. \end{aligned}$$

Como $n^2 \leq k$, entonces $\dfrac{1}{k} \leq \dfrac{1}{n^2}$ y por tanto

$$\left\| \frac{1}{k} \sum_{i=1}^{k} V_i \right\| = \left\| \frac{1}{k} S_k \right\| \leq \frac{\|S_{n^2}\|}{n^2} + \frac{D_n}{n^2},$$

lo cuál prueba el resultado. ■

2.2. Sucesiones de elementos aleatorios ortogonales

Recordemos que una familia de elementos aleatorios $(V_n)_n$ es ortogonal si cualesquiera dos elementos de la familia lo son. Diremos que la familia es **ortonormal** si además para cada $n \geq 1$ se tiene que $\mathbb{E}\left[\|V_n\|^2\right] = \mathbb{E}\left[\langle V, V\rangle\right] = 1$. Para variables aleatorias el siguiente resultado es bien conocido.

Teorema (Ley de Rademacher-Mensov). *Sea $(V_n)_n$ una sucesión de variables aleatorias ortogonales tales que* $\mathrm{Var}(V_n) < \infty$ *para cada $n \in \mathbb{N}$.*
Si $\sum_{n=1}^{\infty} n^{-2} \mathrm{Var}(V_n) \log_2^2 n < \infty$, entonces la sucesión de variables aleatorias satisface la ley fuerte de los grandes números.

Generalizaremos este teorema a cualquier espacio de Hilbert separable, y lo haremos siguiendo los mismos métodos que en el caso clásico, esto es, primero extenderemos la desigualdad de Rademacher-Mensov y posteriormente el teorema de Rademacher-Mensov (véase [29], capítulo 2).

Lema 2.2.1 (Desigualdad de Rademacher-Mensov). *Sea H un espacio de Hilbert. Si $V_1, V_2, \ldots, V_n$ son n elementos aleatorios ortonormales en H, y $c_1, c_2, \ldots, c_n$*

son números reales, entonces

$$\mathbb{E}\left[\max_{1\leq k\leq n}\left\|\sum_{j=1}^{k} c_j V_j\right\|^2\right] \leq \left(\log_2^2 4n\right)\sum_{j=1}^{n} c_j^2 \tag{2.2}$$

Demostración. Supongamos que n es de la forma $n = 2^\nu$ con $\nu \in \mathbb{Z}$. Definimos

$$\begin{aligned} S_j &= c_1V_1 + c_2V_2 + \cdots + c_jV_j \\ \phi_{\alpha\beta} &= c_{\alpha+1}V_{\alpha+1} + c_{\alpha+2}V_{\alpha+2} + \cdots + c_\beta V_\beta \end{aligned}$$

donde $\alpha = \mu 2^k$ y $\beta = \beta(\alpha) = \mu 2^k + 2^k = (\mu+1)2^k$ con $\mu = 0, 1, \ldots, 2^{\nu-k} - 1$ y $k = 0, 1, \ldots, \nu$. Hagamos la siguiente observación sobre los índices α y β.

Si $k = \nu$, entonces $\mu = 0$ luego α_1 y β_1 solo toman los valores 0 y 2^ν respectivamente.

Si $k = \nu - 1$, entonces $\mu = 0, 1$ por lo que $\alpha_2 \in \{0, 2^{\nu-1}\}$ y $\beta_2 \in \{2^{\nu-1}, 2^\nu\}$.

Si $k = \nu - 2$, entonces $\mu = 0, 1, 2, 3$ luego $\alpha_3 \in \{0, 2^{\nu-2}, 2\cdot 2^{\nu-2}, 3\cdot 2^{\nu-2}\}$ y $\beta_3 \in \{2^{\nu-2}, 2\cdot 2^{\nu-2}, 3\cdot 2^{\nu-2}, 4\cdot 2^{\nu-2} = 2^\nu\}$.

$\vdots$

Es decir, en el paso 1 tenemos al invervalo $[0, 2^\nu] \subset \mathbb{N}$ siendo α_1 y β_1 sus extremos. En el paso 2 tenemos el mismo intervalo ahora divido por la mitad generando los subintervalos $\left[0, 2^{\nu-1}\right]$ y $\left[2^{\nu-1}, 2^\nu\right]$, aquí α_2, β_2 pueden tomar los valores extremos de uno u otro intervalo. Para el paso 3 habremos dividido el intervalo $[0, 2^\nu]$ en cuatro partes iguales: $\left[0, 1\cdot 2^{\nu-2}\right]$, $\left[1\cdot 2^{\nu-2}, 2\cdot 2^{\nu-2}\right]$, $\left[2\cdot 2^{\nu-2}, 3\cdot 2^{\nu-2}\right]$ y $\left[3\cdot 2^{\nu-2}, 2^\nu\right]$. En este caso α_3 y β_3 toman por valores los extremos de alguno de estos intervalos.
Continuando este proceso, cuando $k = 0$ habremos divido el intervalo $[0, 2^\nu]$ en 2^ν partes cada una de longitud 1.

De este modo S_j puede verse como la suma de algunas $\phi_{\alpha\beta}$, así

$$S_j = \sum_i \phi_{\alpha_i\beta_i} \tag{2.3}$$

y donde los índices $\alpha_i\beta_i$ pueden elegirse de forma tal que $\beta_1 - \alpha_1 > \beta_2 - \alpha_2 > \beta_3 - \alpha_3 > \ldots$

Notemos que el número de sumandos en (2.3) es a lo más $\nu + 1$. Ahora bien, dado $j \in \mathbb{N}$ se cumple que:

$$\begin{aligned}\|S_j\|^2 = \left\| \sum_i \phi_{\alpha_i \beta_i} \right\|^2 &\leq \left(\sum_i \|\phi_{\alpha_i \beta_j}\| \right)^2 \leq \left(\sum_i 1^2 \right) \left(\sum_i \|\phi_{\alpha_i \beta_i}\|^2 \right) \\ &\leq (\nu+1) \left(\sum_i \|\phi_{\alpha_i \beta_i}\|^2 \right) \leq (\nu+1) \left(\sum_{\alpha,\beta} \|\phi_{\alpha\beta}\|^2 \right).\end{aligned}$$

Lo cuál implica que

$$\max_{1\leq j\leq n} \|S_j\|^2 \leq (\nu+1) \sum_{\alpha,\beta} \|\phi_{\alpha\beta}\|^2,$$

donde los índices α y $\beta = \beta(\alpha)$ corren sobre todos sus posibles valores. Entonces,

$$\mathbb{E}\left[\max_{1\leq j\leq n} \|S_j\|^2\right] \leq \mathbb{E}\left[(\nu+1)\sum_{\alpha,\beta}\|\phi_{\alpha\beta}\|^2\right] \leq (\nu+1)\sum_{\alpha,\beta}\mathbb{E}\left[\|\phi_{\alpha\beta}\|^2\right] \tag{2.4}$$

y donde

$$\begin{aligned}\mathbb{E}\left[\|\phi_{\alpha\beta}\|^2\right] &= \mathbb{E}\left[\langle \phi_{\alpha\beta}, \phi_{\alpha\beta}\rangle\right] \\ &= \mathbb{E}\left[\langle c_{\alpha+1}V_{\alpha+1} + \cdots + c_\beta V_\beta, c_{\alpha+1}V_{\alpha+1} + \cdots + c_\beta V_\beta\rangle\right] \\ &= \mathbb{E}\left[\sum_{i=\alpha+1}^{\beta} c_i^2\|V_i\|^2 + \sum_{i\neq j} c_i \cdot c_j \langle V_i, V_j\rangle\right] \\ &= \sum_{i=\alpha+1}^{\beta} c_i^2 \mathbb{E}\left[\|V_i\|^2\right] + \sum_{i\neq j} c_i \cdot c_j \mathbb{E}\left[\langle V_i, V_j\rangle\right] \\ &= \sum_{i=\alpha+1}^{\beta} c_i^2 \qquad \text{(pues la sucesión es ortonormal).}\end{aligned}$$

Pero los valores que toman α y β dependen de los valores que toma μ, y ésta a su vez depende k, por lo que tenemos que

$$\begin{aligned}\sum_{\alpha,\beta}\mathbb{E}\left[\|\phi_{\alpha\beta}\|^2\right] &= \sum_{\alpha,\beta}\sum_{i=\alpha+1}^{\beta} c_i^2 = \sum_{\alpha,\beta}\left[c_{\alpha+1}^2 + c_{\alpha+2}^2 + \cdots + c_\beta^2\right] \\ &= \sum_{k=0}^{\nu}\sum_{\mu=0}^{2^{\nu-k}-1} c_{\mu 2^k+1}^2 + c_{\mu 2^k+2}^2 + \cdots + c_{\mu 2^k + 2^k}^2.\end{aligned}$$

Sobre esta última expresión observemos que cuando $\mu = 0$ obtenemos los primeros 2^k sumandos, es decir, $c_1^2 + c_2^2 + \cdots + c_{2^k}^2$. Si ahora $\mu = 1$, obtenemos los siguientes 2^k sumandos, $c_{2^k+1}^2 + c_{2^k+2}^2 + \cdots + c_{2^k+2^k}^2$. Similarmente, cuando

$\mu = 2$ obtenemos los siguientes 2^k sumandos, $c^2_{2\cdot 2^k+1} + c^2_{2\cdot 2^k+2} + \cdots + c^2_{2\cdot 2^k+2^k}$. Y así sucesivamente, cuando $\mu = 2^{\nu-k} - 1$, obtenemos los últimos 2^k términos, $c^2_{2^\nu-2^k+1} + c^2_{2^\nu-2^k+2} + \cdots + c^2_{2^\nu-2^k+2^k}$, por tanto

$$\begin{aligned}\sum_{k=0}^{\nu}\sum_{\mu=0}^{2^{\nu-k}-1} c^2_{\mu 2^k+1} + c^2_{\mu 2^k+2} + \cdots + c^2_{\mu 2^k+2^k} &= \sum_{k=0}^{\nu}\sum_{j=1}^{2^\nu} c_j^2 \\ &= (\nu+1)\sum_{j=1}^{2^\nu} c_j^2.\end{aligned}$$

Retomando la desigualdad 2.4,

$$\begin{aligned}\mathbb{E}\left[\max_{1\le j\le n}\|S_j\|^2\right] &\le \mathbb{E}\left[(\nu+1)\sum_{\alpha,\beta}\|\phi_{\alpha\beta}\|^2\right] \le (\nu+1)\sum_{\alpha,\beta}\mathbb{E}\left[\|\phi_{\alpha\beta}\|^2\right] \\ &= (\nu+1)\sum_{\alpha,\beta}\sum_{i=\alpha+1}^{\beta} c_i^2 = (\nu+1)\sum_{\alpha,\beta}\left[c^2_{\alpha+1} + c^2_{\alpha+2} + \cdots + c^2_{\beta}\right] \\ &= (\nu+1)\sum_{k=0}^{\nu}\sum_{\mu=0}^{2^{\nu-k}-1} c^2_{\mu 2^k+1} + c^2_{\mu 2^k+2} + \cdots + c^2_{\mu 2^k+2^k} \\ &= (\nu+1)^2\sum_{j=1}^{2^\nu} c_j^2 = \log_2^2(2n)\sum_{j=1}^{n} c_j^2.\end{aligned}$$

Donde la última igualdad se da pues $\nu + 1 = \log_2 2^{\nu+1} = \log_2 2\cdot 2^\nu = \log_2 2\cdot n$. Por lo tanto, si $n = 2^\nu$ se tiene que

$$\mathbb{E}\left[\max_{1\le j\le n}\|S_j\|^2\right] \le (\log_2^2 2n)\sum_{j=1}^{n} c_j^2. \tag{2.5}$$

En general, cuando $2^\nu \le n < 2^{\nu+1}$ elegimos $c_{n+1} = c_{n+2} = \cdots = c_{2^\nu+1} = 0$, de modo que $S_j = S_n$ para todo $j \ge n$, y por tanto

$$\begin{aligned}\mathbb{E}\left[\max_{1\le j\le n}\|S_j\|^2\right] = \mathbb{E}\left[\max_{1\le j\le 2^{\nu+1}}\|S_j\|^2\right] &\le (\log_2^2 2\cdot 2^{\nu+1})\sum_{j=1}^{2^{\nu+1}} c_j^2 \\ &\le (\log_2^2 4\cdot 2^\nu)\sum_{j=1}^{2^{\nu+1}} c_j^2 \le (\log_2^2 4\cdot n)\sum_{j=1}^{n} c_j^2.\end{aligned}$$

■

2.2.1. Ley de los grandes números de Rademacher-Mensov

A continuación generalizamos un resultado de Rademacher-Mensov, la prueba es consecuencia de la desigualdad del lema 2.2.1.

Proposición 2.2.2 (Rademacher-Mensov)**.** *Sea $(V_n)_n$ una sucesión de elementos aleatorios ortonormales en H y sea $(c_n)_n$ una sucesión de números reales para la cuál*

$$\sum_{k=1}^{\infty} c_k^2 \log_2^2 k < \infty.$$

Entonces la serie

$$\sum_{k=1}^{\infty} c_k V_k$$

converge casi seguramente en H.

Demostración. Definamos $\vartheta_n = \sum_{k=n}^{\infty} c_k V_k$, y para cada $m \geq n$ consideremos $\vartheta_{n,m} = \sum_{k=n}^{m} c_k V_k$ su m-ésima suma parcial. Entonces,

$$\begin{aligned}
\mathbb{E}\left[\|\vartheta_{n,m}\|^2\right] &= \mathbb{E}\left[\langle \vartheta_{n,m}, \vartheta_{n,m}\rangle\right] = \mathbb{E}\left[\langle \sum_{k=n}^{m} c_k V_k, \sum_{k=n}^{m} c_k V_k\rangle\right] \\
&= \mathbb{E}\left[\langle c_n V_n + \cdots + c_m V_m, c_n V_n + \cdots + c_m V_m\rangle\right] \\
&= \mathbb{E}\left[\sum_{k=n}^{m} c_k^2 \langle V_k, V_k\rangle + \sum_{k\neq l} c_k c_l \langle V_k, V_l\rangle\right] \\
&= \sum_{k=n}^{m} c_k^2 \cdot \mathbb{E}\left[\langle V_k, V_k\rangle\right] = \sum_{k=n}^{m} c_k^2 \cdot \mathbb{E}\left[\|V_k\|^2\right] \\
&= \sum_{k=n}^{m} c_k^2.
\end{aligned}$$

Así que al tomar el límite cuando $m \to \infty$ obtenemos que

$$\mathbb{E}\left[\|\vartheta_n\|^2\right] = \sum_{k=n}^{\infty} c_k^2.$$

Denotemos por $A = \sum_{k=1}^{\infty} c_k^2 \log_2^2 k$. Como ocurre que

$$(\log_2^2 n) \sum_{k=n}^{\infty} c_k^2 = \sum_{k=n}^{\infty} c_k^2 \log_2^2 n \leq \sum_{k=n}^{\infty} c_k^2 \log_2^2 k \leq A,$$

entonces para cada $n \geq 1$

$$\mathbb{E}\left[\|\vartheta_n\|^2\right] \leq \frac{A}{\log_2^2 n}$$

y por lo tanto

$$\sum_{n=1}^{\infty} \mathbb{E}\left[\|\vartheta_{2^n}\|^2\right] \leq \sum_{n=1}^{\infty} \frac{A}{\log_2^2 2^n} = \sum_{n=1}^{\infty} \frac{A}{n^2} < \infty.$$

Del teorema C.0.2 de Beppo-Levi tenemos que la serie $\sum_{n=1}^{\infty} \|\vartheta_{2^n}\|^2$ converge c.s., lo cuál implica que $\|\vartheta_{2^n}\|^2 \to 0$ c.s. y por tanto también $\|\vartheta_{2^n}\| \to 0$ c.s. cuando $n \to \infty$.

Definiendo la siguiente sucesión de números reales

$$c'_j = \begin{cases} c_j & \text{si } j \geq 2^n \\ 0 & \text{si } j < 2^n \end{cases},$$

y aplicando la desigualdad Rademacher-Mensov (lema 2.2.1) obtenemos que

$$\begin{aligned}
\mathbb{E}\left[\max_{2^n \leq k < 2^{n+1}} \left\|\sum_{j=2^n}^{k} c_j V_j\right\|^2\right] &= \mathbb{E}\left[\max_{1 \leq k \leq 2^{n+1}-1} \left\|\sum_{j=1}^{k} c'_j V_j\right\|^2\right] \\
&\leq (\log_2^2 4 \cdot (2^{n+1} - 1)) \sum_{j=1}^{2^{n+1}-1} c_j'^2 \\
&= (\log_2^2 (2^{n+3} - 2^2)) \sum_{j=2^n}^{2^{n+1}-1} c_j^2 \\
&\leq (\log_2^2 2^{4n}) \sum_{j=2^n}^{2^{n+1}-1} c_j^2 = 16n^2 \sum_{j=2^n}^{2^{n+1}-1} c_j^2 \\
&= 16 \cdot \log_2^2 2^n \sum_{j=2^n}^{2^{n+1}-1} c_j^2 = 16 \sum_{j=2^n}^{2^{n+1}-1} c_j^2 \log_2^2 2^n \\
&\leq 16 \sum_{j=2^n}^{2^{n+1}-1} c_j^2 \log_2^2 j.
\end{aligned}$$

Lo anterior implica que

$$\begin{aligned}
\sum_{n=1}^{\infty} \mathbb{E}\left[\max_{2^n \leq k < 2^{n+1}} \left\|\sum_{j=2^n}^{k} c_j V_j\right\|^2\right] &\leq 16 \sum_{n=1}^{\infty} \sum_{j=2^n}^{2^{n+1}-1} c_j^2 \log_2^2 j \\
&\leq 16 \sum_{k=1}^{\infty} c_k^2 \log_2^2 k < \infty
\end{aligned}$$

y nuevamente, usando el teorema C.0.2 de Beppo-Levi, obtenemos que

$$\max_{2^n \leq k < 2^{n+1}} \left\| \sum_{j=2^n}^{k} c_j V_j \right\| \to 0 \quad \text{c.s. cuando } n \to \infty$$

Así que por lo anterior, dado $k \in \mathbb{N}$ tal que $2^n \leq k < 2^{n+1}$, ocurre que

$$\|\vartheta_k\| \leq \|\vartheta_k - \vartheta_{2^n}\| + \|\vartheta_{2^n}\| \leq \max_{2^n \leq k < 2^{n+1}} \left\| \sum_{j=2^n}^{k} c_j V_j \right\| + \|\vartheta_{2^n}\| \to 0 \quad \text{c.s.}$$

Lo cuál concluye la prueba de la proposición 2.2.2. ■

Para probar la extensión de la Ley de los Grandes Números de Rademacher-Mensov necesitaremos además del útil lema de Kronecker, el cual enunciamos y probamos a continuación. Cabe mencionar que la misma prueba para el caso clásico funciona para cualquier espacio normado. Así que la enunciamos de esa manera y no solo para espacios de Hilbert.

Lema 2.2.3 (Sumación por partes)**.** *Sea* $(X, \|\cdot\|)$ *un espacio normado. Sean* $(x_j)_j$ *de números reales y* $(y_k)_k$ *una sucesión en* X*. Para cada entero no negativo definimos* $X_n = \sum_{i=0}^{n} x_i$ *y* $Y_n = \sum_{i=0}^{n} y_i$*, donde* $X_0 = x_0$ *y* $Y_0 = y_0$*. Entonces,*

$$X_n Y_n = \sum_{i=0}^{n} X_i y_i + \sum_{i=1}^{n} x_i Y_{i-1}.$$

Demostración. Observemos que para cada $n \geq 1$, $x_n = X_n - X_{n-1}$ y análogamente $y_n = Y_n - Y_{n-1}$. Entonces,

$$\begin{aligned} X_n Y_n - X_0 Y_0 &= \sum_{i=1}^{n} (X_i Y_i - X_{i-1} Y_{i-1}) \\ &= \sum_{i=1}^{n} (X_i Y_i - X_i Y_{i-1} + X_i Y_{i-1} - X_{i-1} Y_{i-1}) \\ &= \sum_{i=1}^{n} X_i (Y_i - Y_{i-1}) + \sum_{i=1}^{n} (X_i - X_{i-1}) Y_{i-1} \\ &= \sum_{i=1}^{n} X_i y_i + \sum_{i=1}^{n} x_i Y_{i-1}. \end{aligned}$$

Ello implica que

$$X_n Y_n = \sum_{i=1}^{n} X_i y_i + \sum_{i=1}^{n} x_i Y_{i-1} + X_0 Y_0 = \sum_{i=0}^{n} X_i y_i + \sum_{i=1}^{n} x_i Y_{i-1}.$$

■

Lema 2.2.4 (Kronecker)**.** *Sea $(V_n)_n$ una sucesión arbitraria en X, y sea $(b_n)_n$ una sucesión de números reales de términos no negativos creciente a infinito, es decir, $0 < b_1 < b_2 < \ldots \uparrow \infty$. Supongamos que la serie $\displaystyle\sum_{k=1}^{\infty} \frac{V_k}{b_k}$ converge, entonces*

$$\lim_{n\to\infty} \frac{1}{b_n} \sum_{k=1}^{n} V_k = \mathbf{0}.$$

Demostración. Usando sumación por partes, donde $X_i = b_i$ y $y_i = \frac{V_i}{b_i}$ para cada $i \geq 1$, y si $i = 0$ definimos $X_0 = x_0 = b_0 = 0$ y $y_0 = 0$. Entonces,

$$\begin{aligned} b_n Y_n &= \sum_{i=0}^{n} b_i \frac{V_i}{b_i} + \sum_{i=1}^{n} (X_i - X_{i-1}) Y_{i-1} \\ &= \sum_{i=1}^{n} V_i + \sum_{i=1}^{n} (b_i - b_{i-1}) Y_{i-1} \\ &= \sum_{i=1}^{n} V_i + \sum_{i=0}^{n-1} (b_{i+1} - b_i) Y_i, \end{aligned}$$

es decir

$$b_n Y_n - \sum_{i=1}^{n-1} (b_{i+1} - b_i) Y_i = \sum_{i=1}^{n} V_i.$$

Ahora bien, denotemos por $Y = \lim_{n\to\infty} Y_n = \lim_{n\to\infty} \sum_{i=1}^{n} \frac{V_i}{b_i}$ (el cuál existe por hipótesis), entonces,

$$\begin{aligned} \frac{1}{b_n} \sum_{i=1}^{n} V_i = Y_n - \frac{1}{b_n} \sum_{i=1}^{n-1} (b_{i+1} - b_i) Y_i \\ &= Y_n - Y + Y - \frac{1}{b_n} \sum_{i=1}^{n-1} (b_{i+1} - b_i) Y_i \\ &= Y_n - Y + Y \frac{1}{b_n} \cdot \left(\sum_{i=0}^{n-1} (b_{i+1} - b_i) \right) - \frac{1}{b_n} \sum_{i=1}^{n-1} (b_{i+1} - b_i) Y_i \\ &= Y_n - Y + \frac{1}{b_n} \sum_{i=0}^{n-1} (b_{i+1} - b_i)(Y - Y_i). \end{aligned}$$

Sea $\varepsilon > 0$, como $Y_n \to Y$ existe $N > 0$ tal que para todo $n \geq N$ se tiene que

$\|Y_n - Y\| < \dfrac{\varepsilon}{2}$. Sea además $M = \max\limits_{0\leq j\leq N-1} \|Y - Y_i\|$. Entonces, si $n \geq N$,

$$\begin{aligned}
\left\|\frac{1}{b_n}\sum_{i=1}^{n} V_i\right\| &\leq \|Y_n - Y\| + \frac{1}{b_n}\left(\sum_{i=0}^{N-1}|b_{i+1}-b_i|\|Y - Y_i\| + \sum_{i=N}^{n-1}|b_{i+1}-b_i|\|Y-Y_i\|\right) \\
&\leq \|Y_n - Y\| + \frac{1}{b_n}\left(\sum_{i=0}^{N-1}(b_{i+1}-b_i)M + \sum_{i=N}^{n-1}(b_{i+1}-b_i)\frac{\varepsilon}{2}\right) \\
&\leq \|Y_n - Y\| + \frac{1}{b_n}M(b_N - b_1) + \frac{1}{b_n}(b_n - b_N)\frac{\varepsilon}{2} \\
&= \|Y_n - Y\| + \frac{b_N}{b_n}M + (1 - \frac{b_N}{b_n})\frac{\varepsilon}{2}.
\end{aligned}$$

Haciendo tender n a infinito obtenemos que $\left\|\dfrac{1}{b_n}\sum\limits_{i=1}^{n} V_i\right\| \leq \varepsilon$, lo cuál implica que el límite es cero. ■

Teorema 2.2.5 (Rademacher-Mensov). *Sea H un espacio de Hilbert separable. Si $(V_n)_n$ es una sucesión de elementos aleatorios ortogonales en H con* $\mathrm{Var}(V_n) < \infty$ *para cada n, y tales que*

$$\sum_{n=1}^{\infty}\frac{\mathrm{Var}(V_n)}{n^2}\log_2^2 n < \infty$$

entonces

$$\left\|\frac{1}{n}\sum_{k=1}^{n}(V_k - \mathbb{E}[V_k])\right\| \to 0$$

casi seguramente.

Demostración. Sin pérdida de generalidad podemos suponer que para cada n, $\mathbb{E}[V_n] = \mathbf{0}$, y en particular $\mathrm{Var}(V_n) = \mathbb{E}\left[\|V_n\|^2\right]$. Debido al lema de Kronecker, es suficiente probar que la serie

$$\sum_{n=1}^{\infty}\frac{V_n}{n}$$

converge casi seguramente. Con esta finalidad, para cada n definimos

$$\zeta_n = \begin{cases} \dfrac{V_n}{(\mathbb{E}[\|V_n\|^2])^{\frac{1}{2}}} & \text{si} \quad \mathrm{Var}(V_n) \neq 0 \\ 0 & \text{si} \quad \mathrm{Var}(V_n) = 0. \end{cases}$$

Entonces $(\zeta_n)_n$ es una sucesión de elementos aleatorios ortonormales. En efecto,

$$\begin{aligned}
\mathbb{E}\left[\langle \zeta_i, \zeta_j\rangle\right] &= \mathbb{E}\left[\left\langle \frac{V_i}{(\mathbb{E}\left[\|V_i\|^2\right])^{\frac{1}{2}}}, \frac{V_j}{(\mathbb{E}\left[\|V_j\|^2\right])^{\frac{1}{2}}} \right\rangle\right] \\
&= \frac{1}{(\mathbb{E}\left[\|V_i\|^2\right])^{\frac{1}{2}} \cdot (\mathbb{E}\left[\|V_j\|^2\right])^{\frac{1}{2}}} \mathbb{E}\left[\langle V_i, V_j\rangle\right] \\
&= \begin{cases} \dfrac{\mathbb{E}\left[\|V_i\|^2\right]}{\left((\mathbb{E}\left[\|V_i\|^2\right])^{\frac{1}{2}}\right)^2} & \text{si} \quad i = j \\ 0 & \text{si} \quad i \neq j \end{cases} \\
&= \begin{cases} 1 & \text{si} \quad i = j \\ 0 & \text{si} \quad i \neq j. \end{cases}
\end{aligned}$$

Así, tomando la sucesión de números reales $c_n = \dfrac{1}{n}(\mathbb{E}\left[\|V_n\|^2\right])^{\frac{1}{2}}$ obtenemos que

$$\sum_{n=1}^{\infty} c_n^2 \log_2^2 n = \sum_{n=1}^{\infty} \frac{\mathrm{Var}(V_n)}{n^2} \log_2^2 n < \infty.$$

Aplicando la proposición 2.2.2 tenemos que la serie $\displaystyle\sum_{n=1}^{\infty} c_n \zeta_n$ converge casi seguramente. Es decir,

$$\sum_{n=1}^{\infty} c_n \zeta_n = \sum_{n=1}^{\infty} \frac{(\mathbb{E}\left[\|V_n\|^2\right])^{\frac{1}{2}}}{n} \frac{V_n}{(\mathbb{E}\left[\|V_n\|^2\right])^{\frac{1}{2}}} = \sum_{n=1}^{\infty} \frac{V_n}{n}$$

converge casi seguramente en H, lo cuál prueba el resultado.

■

Comentario. No hay conflicto suponer que la sucesión $(V_n)_n$ esté formada por elementos de esperanza nula, pues de no ser así podemos «centrarlas» al sustraerles su esperanza, es decir, aplicaríamos el mismo razonamiento a la suceción de elementos aleatorios $(V_n')_n$ donde $V_n' = V_n - \mathbb{E}[V_n]$, que claramente satisface $\mathbb{E}[V_n'] = \mathbf{0}$ y $\mathrm{Var}(V_n') = \mathrm{Var}(V_n) < \infty$ para cada n.

Y no solamente eso, en tal caso, si asumimos que la sucesión es de elementos no correlacionados, entonces sería ortogonales. De modo que el teorema 2.2.5 sigue siendo válido si sustituimos la hipótesis de ortogonalidad por la de no correlación. Es decir:

Teorema 2.2.6. *Sea H un espacio de Hilbert separable. Si $(V_n)_n$ es una sucesión de elementos aleatorios no correlacionados en H con $\mathrm{Var}(V_n) < \infty$ para cada n, y tales que*

$$\sum_{n=1}^{\infty} \frac{\mathrm{Var}(V_n)}{n^2} \log_2^2 n < \infty$$

entonces la sucesión $(V_n)_n$ satisface la ley fuerte de los grandes números.

2.3. Sucesiones de elementos aleatorios independientes

Hasta ahora hemos dejado de lado el concepto de independencia para enfocarnos en estudiar los conceptos de no correlación y de ortogonalidad de elementos aleatorios. Primero porque la estructura de espacio con producto interior lo permite, y segundo porque son condiciones más sencillas de verificar que la de independencia. Esto queda plasmado en el siguiente resultado.

Lema 2.3.1. *Sea H un espacio de Hilbert separable. Si V y Z son elementos aleatorios independientes en H, con $\mathbb{E}[\|V\|^2] < \infty$ y $\mathbb{E}[\|Z\|^2] < \infty$, entonces V y Z son elementos aleatorios no correlacionados.*

Demostración. Hemos visto que $\mathbb{E}[\|V\|^2], \mathbb{E}[\|Z\|^2] < \infty$, implica que $\mathbb{E}[V]$ y $\mathbb{E}[Z]$ existen en H. Sin pérdida de generalidad supongamos que $\mathbb{E}[V] = \mathbf{0} = \mathbb{E}[Z]$, de lo contrario consideramos los elementos aleatorios «centrados» $V - \mathbb{E}[V]$ y $Z - \mathbb{E}[Z]$ cuya esperanza es $\mathbf{0}$ (el vector cero).

Como H es separable, el teorema A.1.6 nos garantiza que H tiene una base ortogonal, digamos $\{b_n : n \geq 1\}$. Ahora, dados $x, y \in H$ sabemos que

$$\langle x, y\rangle = \sum_{n=1}^{\infty} \langle x, b_n\rangle\langle y, b_n\rangle.$$

En particular, obtenemos que

$$\langle V, Z\rangle = \sum_{n=1}^{\infty} \langle V, b_n\rangle\langle Z, b_n\rangle = \lim_{m\to\infty} \sum_{n=1}^{m} \langle V, b_n\rangle\langle Z, b_n\rangle$$

Notemos que para cada $k \in \mathbb{N}$ tenemos que

$$\begin{aligned}
\left|\sum_{n=1}^{k} \langle V, b_n\rangle\langle Z, b_n\rangle\right| &\leq \sum_{n=1}^{k} |\langle V, b_n\rangle\langle Z, b_n\rangle| \\
&\leq \left(\sum_{n=1}^{k} |\langle V, b_n\rangle|^2\right)^{\frac{1}{2}} \left(\sum_{n=1}^{k} |\langle Z, b_n\rangle|^2\right)^{\frac{1}{2}} \\
&\leq \left(\sum_{n=1}^{\infty} |\langle V, b_n\rangle|^2\right)^{\frac{1}{2}} \left(\sum_{n=1}^{\infty} |\langle Z, b_n\rangle|^2\right)^{\frac{1}{2}} \\
&\leq (\|V\|^2)^{\frac{1}{2}} (\|Z\|^2)^{\frac{1}{2}} = \|V\|\|Z\|,
\end{aligned}$$

donde la segunda y última desigualdad están justificadas por las desigualdades de Hölder y Bessel respectivamente. Así, si $S_k = \sum_{n=1}^{k} \langle V, b_n\rangle\langle Z, b_n\rangle$ es la k-ésima suma parcial de $S = \sum_{n=1}^{\infty} \langle V, b_n\rangle\langle Z, b_n\rangle$, entonces $(S_k)_k$ es una sucesión de funciones tales que

I. $S_k\colon \Omega \to \mathbb{R}$ converge puntualmente a $S\colon \Omega \to \mathbb{R}$,

II. S_k está dominada por la función $\|V\|\|Z\|$ que es integrable (pues por la desigualdad de Hölder $\mathbb{E}[\|V\|\,\|Z\|] \leq \left(\mathbb{E}[\|V\|^2]\right)^{\frac{1}{2}} \left(\mathbb{E}[\|Z\|^2]\right)^{\frac{1}{2}} < \infty$).

y por lo tanto, el teorema de Convergencia Dominada nos garantiza que

$$\begin{aligned}\mathbb{E}\left[\langle V, Z\rangle\right] &= \mathbb{E}\left[\sum_{n=1}^{\infty}\langle V, b_n\rangle\langle Z, b_n\rangle\right] = \mathbb{E}\left[\lim_{m\to\infty}\sum_{n=1}^{m}\langle V, b_n\rangle\langle Z, b_n\rangle\right]\\ &= \lim_{m\to\infty}\mathbb{E}\left[\sum_{n=1}^{m}\langle V, b_n\rangle\langle Z, b_n\rangle\right] = \lim_{m\to\infty}\sum_{n=1}^{m}\mathbb{E}\left[\langle V, b_n\rangle\langle Z, b_n\rangle\right].\end{aligned}$$

En particular notemos que para cada $n \in \mathbb{N}$, y $h \in H$ la función $f_{b_n}(h) = \langle h, b_n\rangle\colon H \to \mathbb{R}$ es lineal y continua ($f_{b_n} \in H^*$), por ello también es medible. Por hipótesis V, Z son elementos aleatorios independientes, así, en virtud del lema 1.2.4 obtenemos que $\langle V, b_n\rangle, \langle Z, b_n\rangle$ son variables aleatorias independientes, y por tanto

$$\mathbb{E}\left[\langle V, b_n\rangle\langle Z, b_n\rangle\right] = \mathbb{E}\left[\langle V, b_n\rangle\right]\mathbb{E}\left[\langle Z, b_n\rangle\right].$$

En consecuencia,

$$\begin{aligned}\lim_{m\to\infty}\sum_{n=1}^{m}\mathbb{E}\left[\langle V, b_n\rangle\langle Z, b_n\rangle\right] &= \lim_{m\to\infty}\sum_{n=1}^{m}\mathbb{E}\left[\langle V, b_n\rangle\right]\mathbb{E}\left[\langle Z, b_n\rangle\right]\\ &= \lim_{m\to\infty}\sum_{n=1}^{m}\mathbb{E}\left[f_{b_n}(V)\right]\mathbb{E}\left[f_{b_n}(Z)\right]\\ &= \lim_{m\to\infty}\sum_{n=1}^{m}f_{b_n}(\mathbb{E}\left[V\right])f_{b_n}(\mathbb{E}\left[Z\right])\\ &= \lim_{m\to\infty}\sum_{n=1}^{m}\langle\mathbb{E}\left[V\right], b_n\rangle\langle\mathbb{E}\left[Z\right], b_n\rangle\\ &= \lim_{m\to\infty}\sum_{n=1}^{m}\langle\mathbf{0}, b_n\rangle\langle\mathbf{0}, b_n\rangle\\ &= \lim_{m\to\infty} 0 = 0 = \langle\mathbf{0}, \mathbf{0}\rangle = \langle\mathbb{E}\left[V\right], \mathbb{E}\left[Z\right]\rangle.\end{aligned}$$

Por lo tanto $\mathbb{E}\left[\langle V, Z\rangle\right] = \langle\mathbb{E}\left[V\right], \mathbb{E}\left[Z\right]\rangle$, es decir, V y Z son no correlacionados. ■

2.3.1. Ley de los grandes números de Kolmogorov

Proposición 2.3.2 (Desigualdad de Kolmogorov)**.** *Sea H un espacio de Hilbert separable. Si $V_1, V_2, \ldots, V_n$ son elementos aleatorios independientes en H con varianza finita y ε es cualquier número positivo, entonces*

$$\mathbb{P}\left[\max_{1\leq k\leq n}\left\|\sum_{j=1}^{k}V_j\right\| \geq \varepsilon\right] \leq \frac{1}{\varepsilon^2}\sum_{j=1}^{n}\operatorname{Var}(V_j). \tag{2.6}$$

Demostración. Sea $\varepsilon > 0$. Supongamos que $\mathbb{E}[V_j] = \mathbf{0}$, si denotamos por S_k a la k-ésima suma parcial de los elementos aleatorios V_j, $S_k = \sum_{j=1}^{k} V_j$, entonces también $\mathbb{E}[S_k] = \mathbf{0}$. Definamos los siguientes subconjuntos de Ω,

$$A = \left\{\omega : \max_{1\leq k\leq n} \|S_k\| \geq \varepsilon\right\}$$
$$A_k = \{\omega : \|S_1\| < \varepsilon, \|S_2\| < \varepsilon, \ldots, \|S_{k-1}\| < \varepsilon, \|S_k\| \geq \varepsilon\}$$

Afirmamos que $A = \displaystyle\bigcup_{k=1}^{n} A_k$.

En efecto, si $\omega \in \bigcup_{k=1}^{n} A_k$, entonces $\omega \in A_m$ para alguna $m \in \{1, 2, \ldots, n\}$, en particular $\|S_m\| \geq \varepsilon$ pero $\max\limits_{1\leq k\leq n} \|S_k\| \geq \|S_m\|$ y por lo tanto $\omega \in A$.

Por otra parte, si $\omega \in A$, entonces $\max\limits_{1\leq k\leq n} \|S_k\| \geq \varepsilon$, de modo que el conjunto $\{j : \|S_j\| \geq \varepsilon\} \neq \emptyset$. Sea $k_0 = \min\{j : \|S_j\| \geq \varepsilon\}$, entonces $\omega \in A_{k_0} \subset \displaystyle\bigcup_{k=1}^{n} A_k$.

Más aún, el conjunto A es una unión ajena de los A_k. En efecto, tomando $r \neq s$ índices podemos suponer sin perder generalidad que $r < s$. Si ocurriera que $A_r \cap A_s \neq \emptyset$ es porque existe $\omega \in A_r \cap A_s$. Luego, como $\omega \in A_s$ tenemos que $\left\|\sum_{j=1}^{p} V_j\right\| < \varepsilon$ para cada $1 \leq p < s$ en particular lo anterior es válido para $p = r$, pero como también $\omega \in A_r$, $\left\|\sum_{j=1}^{r} V_j\right\| \geq \varepsilon$ lo cuál es una contradición. Así que:

$$\begin{aligned}
\mathbb{E}\left[\|S_n\|^2 \cdot \mathbb{I}_A\right] &= \mathbb{E}\left[\|S_n\|^2 \cdot \mathbb{I}_{\bigcup_{k=1}^{n} A_k}\right] = \mathbb{E}\left[\|S_n\|^2 \cdot \sum_{k=1}^{n} \mathbb{I}_{A_k}\right] \\
&= \sum_{k=1}^{n} \mathbb{E}\left[\|S_n\|^2 \cdot \mathbb{I}_{A_k}\right] = \sum_{k=1}^{n} \mathbb{E}\left[\|S_k + S_n - S_k\|^2 \cdot \mathbb{I}_{A_k}\right] \\
&= \sum_{k=1}^{n} \mathbb{E}\left[\langle S_k + S_n - S_k, S_k + S_n - S_k\rangle \cdot \mathbb{I}_{A_k}\right] \\
&= \sum_{k=1}^{n} \mathbb{E}\left[\left(\|S_n - S_k\|^2 + \|S_k\|^2 + 2\langle S_k, S_n - S_k\rangle\right) \cdot \mathbb{I}_{A_k}\right] \\
&= \sum_{k=1}^{n} \left(\mathbb{E}\left[\|S_n - S_k\|^2 \cdot \mathbb{I}_{A_k}\right] + \mathbb{E}\left[\|S_k\|^2 \cdot \mathbb{I}_{A_k}\right] + 2\mathbb{E}\left[\langle S_k, S_n - S_k\rangle \cdot \mathbb{I}_{A_k}\right]\right)
\end{aligned}$$

Notemos que $\mathbb{E}\left[\langle S_k, S_n - S_k\rangle \cdot \mathbb{I}_{A_k}\right] = \mathbb{E}\left[\langle S_k \cdot \mathbb{I}_{A_k}, S_n - S_k\rangle\right]$, donde $S_k\mathbb{I}_{A_k}$ y $S_n - S_k$ resultan ser elementos aleatorios independientes, y como H es separable podemos aplicar el lema 2.3.1 de donde obtenemos que son $S_k\mathbb{I}_{A_k}$ y $S_n - S_k$ elementos aleatorios no correlacionados, es decir,

$$\mathbb{E}\left[\langle S_k \cdot \mathbb{I}_{A_k}, S_n - S_k\rangle\right] = \langle \mathbb{E}\left[S_k \cdot \mathbb{I}_{A_k}\right], \mathbb{E}\left[S_n - S_k\right]\rangle = 0.$$

Y por lo tanto,

$$\begin{aligned}
\sum_{j=1}^{n} \mathrm{Var}\,(V_j) = \mathrm{Var}(S_n) = \mathbb{E}\left[\|S_n\|^2\right] &\geq \mathbb{E}\left[\|S_n\|^2 \cdot \mathbb{I}_A\right] \\
&= \sum_{k=1}^{n} \left(\mathbb{E}\left[\|S_n - S_k\|^2 \cdot \mathbb{I}_{A_k}\right] + \mathbb{E}\left[\|S_k\|^2 \cdot \mathbb{I}_{A_k}\right]\right) \\
&\geq \sum_{k=1}^{n} \mathbb{E}\left[\|S_k\|^2 \cdot \mathbb{I}_{A_k}\right] \geq \sum_{k=1}^{n} \varepsilon^2 \cdot \mathbb{E}\left[\mathbb{I}_{A_k}\right] = \varepsilon^2 \sum_{k=1}^{n} \mathbb{E}\left[\mathbb{I}_{A_k}\right] \\
&= \varepsilon^2 \sum_{k=1}^{n} \mathbb{P}\left[A_k\right] = \varepsilon^2 \mathbb{P}\left[\bigcup_{k=1}^{n} A_k\right] = \varepsilon^2 \mathbb{P}\left[A\right].
\end{aligned}$$

Para el caso general, cuando la esperanza de V_j es distinta de cero, definimos la sucesión de elementos aleatorios $\eta_j = V_j - \mathbb{E}\left[V_j\right]$. Ellos cumplen tener esperanza nula y varianza finita, por lo que aplicando el resultado anterior obtenemos que dado $\varepsilon > 0$,

$$\begin{aligned}
\mathbb{P}\left[\max_{1\leq k\leq n} \|S_k - \mathbb{E}\left[S_k\right]\| \geq \varepsilon\right] = \mathbb{P}\left[\max_{1\leq k\leq n} \left\|\sum_{j=1}^{k} \eta_j\right\| \geq \varepsilon\right] &\leq \frac{1}{\varepsilon^2} \sum_{j=1}^{n} \mathrm{Var}(\eta_j) \\
&= \frac{1}{\varepsilon^2} \sum_{j=1}^{n} \mathrm{Var}(V_j - \mathbb{E}\left[V_j\right]) = \frac{1}{\varepsilon^2} \sum_{j=1}^{n} \mathrm{Var}(V_j).
\end{aligned}$$

■

Teorema 2.3.3 (Kolmogorov)**.** *Sea H un espacio de Hilbert separable. Si $(V_n)_n$ es una sucesión de elementos aleatorios independientes en H tal que*

$$\sum_{n=1}^{\infty} \frac{\mathrm{Var}(V_n)}{n^2} < \infty$$

entonces

$$\left\|\frac{1}{n} \sum_{k=1}^{n} (V_k - \mathbb{E}\left[V_k\right])\right\| \to 0 \quad c.s.$$

Demostración. Sin pérdida de generalidad supongamos que $\mathbb{E}\left[V_n\right] = 0$ para cada $n \in \mathbb{N}$ (de lo contrario consideramos a la sucesión de e.a. dados como $U_n = V_n - \mathbb{E}\left[V_n\right]$). Para cada $n \in \mathbb{N}$ sea $S_n = \sum_{k=1}^{n} V_k$ y, para cada $\varepsilon > 0$, definimos

$$A_\varepsilon = \left\{\omega \,:\, \left\|\frac{S_n(\omega)}{n}\right\| > \varepsilon \text{ para una infinidad de } n\right\}.$$

Por el lema 1.1.9, para probar el resultado basta probar que $\mathbb{P}(A_\varepsilon) = 0$ para cada $\varepsilon > 0$. Para ello definimos

$$B_{n,\varepsilon} = \left\{\omega \,:\, \left\|\frac{S_k(\omega)}{k}\right\| > \varepsilon \text{ para alguna } k \in \mathbb{N} \text{ tal que } 2^{n-1} < k \leq 2^n\right\}.$$

Entonces se tiene que

$$A_\varepsilon = \{\omega \,:\, \omega \in B_{n,\varepsilon} \text{ para una infinidad de } n\}$$

y debido al lema 1.1.10 de Borel-Cantelli, para probar que $\mathbb{P}(A_\varepsilon) = 0$ para cada $\varepsilon > 0$, basta probar que la serie $\sum\limits_{n=1}^{\infty} \mathbb{P}(B_{n.\varepsilon})$ converge para cada $\varepsilon > 0$, pero esto es consecuencia de la desigualdad de Kolmogorov. En efecto,

$$\begin{aligned}
\mathbb{P}\left(B_{n,\varepsilon}\right) &= \mathbb{P}\left[\ \left\|\frac{S_k(\omega)}{k}\right\| > \varepsilon \text{ para alguna } k \in \mathbb{N} \text{ tal que } 2^{n-1} < k \leq 2^n\right] \\
&= \mathbb{P}\left[\max_{2^{n-1}<k\leq 2^n} \left\|\frac{S_k(\omega)}{k}\right\| > \varepsilon\right] = \mathbb{P}\left[\max_{2^{n-1}<k\leq 2^n} \|S_k(\omega)\| > k\varepsilon\right] \\
&\leq \mathbb{P}\left[\max_{2^{n-1}<k\leq 2^n} \|S_k(\omega)\| > 2^{n-1}\varepsilon\right] \\
&\leq \mathbb{P}\left[\max_{1\leq k\leq 2^n} \|S_k(\omega)\| > 2^{n-1}\varepsilon\right] \leq \frac{1}{(\varepsilon 2^{n-1})^2} \sum_{k=1}^{2^n} \mathrm{Var}(V_k) \\
&= \frac{4}{\varepsilon^2 2^{2n}} \sum_{k=1}^{2^n} \mathrm{Var}(V_k)
\end{aligned}$$

y por lo tanto,

$$\begin{aligned}
\sum_{n=1}^{\infty} \mathbb{P}(B_{n,\varepsilon}) &\leq \frac{4}{\varepsilon^2} \sum_{n=1}^{\infty} \frac{1}{2^{2n}} \sum_{k=1}^{2^n} \mathrm{Var}(V_k) = \frac{4}{\varepsilon^2}\left[\frac{1}{2^2}\{\mathrm{Var}(V_1) + \mathrm{Var}(V_2)\} + \right. \\
&\quad + \frac{1}{2^4}\{\mathrm{Var}(V_1) + \mathrm{Var}(V_2) + \mathrm{Var}(V_3) + \mathrm{Var}(V_4)\} \\
&\quad \left. + \frac{1}{2^6}\{\mathrm{Var}(V_1) + \mathrm{Var}(V_2) + \mathrm{Var}(V_3) + \mathrm{Var}(V_4) + \mathrm{Var}(V_5) + \mathrm{Var}(V_6)\} + \dots\right] \\
&= \frac{4}{\varepsilon^2}\left[\mathrm{Var}(V_1)\left\{\frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + \dots\right\} + \mathrm{Var}(V_2)\left\{\frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + \dots\right\}\right. \\
&\quad \left. + \mathrm{Var}(V_3)\left\{\frac{1}{2^4} + \frac{1}{2^6} + \frac{1}{2^8} + \dots\right\} + \mathrm{Var}(V_4)\left\{\frac{1}{2^4} + \frac{1}{2^6} + \frac{1}{2^8} + \dots\right\} + \dots\right] \\
&= \frac{4}{\varepsilon^2} \sum_{k=1}^{\infty} \mathrm{Var}(V_k) \sum_{\{n\in\mathbb{N}\,:\,k\leq 2^n\}} \frac{1}{2^{2n}}.
\end{aligned}$$

Dado $k \in \mathbb{N}$, $\{n \in \mathbb{N} : k \leq 2^n\} \neq \emptyset$. Así, sea $n_0 = \min\{n \in \mathbb{N} : k \leq 2^n\}$. Entonces,

$$\sum_{\{n\in\mathbb{N}\,:\,k\leq 2^n\}} \frac{1}{2^{2n}} = \sum_{n=n_0}^{\infty} \frac{1}{2^{2n}} = \sum_{n=n_0}^{\infty} \left(\frac{1}{4}\right)^n = \frac{4}{3\cdot 2^{2n_0}} \leq \frac{4}{3k^2}$$

y por lo tanto:

$$\sum_{n=1}^{\infty} \mathbb{P}(B_{n,\varepsilon}) \leq \frac{4}{\varepsilon^2} \sum_{k=1}^{\infty} \mathrm{Var}(V_k) \frac{4}{3k^2} = \frac{16}{3\varepsilon^2} \sum_{k=1}^{\infty} \frac{\mathrm{Var}(V_k)}{k^2} < \infty.$$

■

Capítulo 3

Leyes de los grandes números en espacios normados

En el capítulo 2 vimos que pudimos extender de manera casi directa, y siguiendo casi los mismos métodos de prueba, varias leyes de grandes números como por ejemplo la de Kolmogorov (teorema 2.3.3). En definitiva tener una operación de producto interior definido en el espacio fue fundamental para tales extensiones, tan solo recordemos el lema 2.3.1 que permitió establecer una relación entre un concepto fundamental como lo es el de independencia con el de no-correlación.

Esto contrasta con el caso general ya que ninguna de las leyes de los grandes números que presentamos en el capítulo anterior (dígase teorema 2.1.1, 2.1.2, 2.2.5 o 2.3.3) se extienden de forma directa a espacios de Banach separables. Para ilustrar esto veamos el siguiente ejemplo.

Ejemplo 3.0.1. Consideremos $X = \ell_1 = \{(x_n)_n \in \mathbb{R}^\infty : \sum_{n=1}^{\infty} |x_n| < \infty\}$ dotado de la norma $\|(x_n)_n\|_{\ell_1} = \sum_{n=1}^{\infty} |x_n|$.

Denotemos por δ^n al elemento de ℓ_1 que toma el valor 1 en la n-ésima coordenada y 0 en el resto, es decir,

$$e_n = (0, \dots, 0, \underbrace{1}_{n\text{-ésimo}}, 0, \dots).$$

Tomemos $(A_n)_n$ una sucesión de variables aleatorias independientes tales que $\mathbb{P}(A_n = \pm 1) = \frac{1}{2}$ (una sucesión de Rademacher) y definamos $V_n = A_n e_n = (0, \dots, A_n, 0, \dots) = (V_{n,k})_k$. Entonces V_n es un elemento aleatorio en ℓ_1 (por la proposición 1.2.3), y no solo eso, $(V_n)_n$ es una sucesión de elementos aleatorios independientes.

Para corroborar esto último basta ver que $(f(A_n e_n))_n$ forma una sucesión de variables aleatorias independientes para cada $f \in (\ell_1)^*$. Pero como $(\ell_1)^* = \ell_\infty$, entonces dado $(f_n)_n = f \in (\ell_1)^*$,

$$f(V_n) = \sum_{k=1}^{\infty} f_k V_{n,k} = f_n A_n$$

luego $(f(V_n))_n = (f_n A_n)_n$ es una sucesión de variables aleatorias independientes para cada $f \in (\ell_1)^*$.

Observemos que $\|V_n\|_{\ell_1} = \|(0, \ldots, 0, \pm 1, 0, \ldots)\|_{\ell_1} = 1 < \infty$ (primer momento finito) y ya que $\mathbb{E}[A_n] = 0$, entonces $\mathbb{E}[V_n] = \mathbf{0}$ para cada n. Sin embargo,

$$\begin{aligned} \left\| \frac{1}{n} \sum_{k=1}^{n} V_k \right\|_{\ell_1} &= \left\| \frac{1}{n} \sum_{k=1}^{n} A_k e_k \right\|_{\ell_1} = \left\| \frac{1}{n} (A_1, A_2, \ldots, A_n, 0, \ldots) \right\|_{\ell_1} \\ &= \frac{1}{n} \sum_{k=1}^{n} |A_k| \equiv 1 \end{aligned}$$

para cada n. Por lo tanto,

$$\frac{1}{n} \sum_{k=1}^{n} V_n \nrightarrow \mathbf{0} = \mathbb{E}[V_n]$$

(leáse: no converge a $\mathbf{0}$ cuando $n \to \infty$) en ningún modo de convergencia.

El ejemplo 3.0.1 sugiere que, o bien imponemos más condiciones sobre los elementos aleatorios, o bien encontrarmos condiciones sobre el espacio que nos permitan concluir teoremas límites similares a los ya antes mencionados. Enseguida abordamos la primera de estas cuestiones, la segunda será tratada en la sección 3.3.

3.1. Ley de los grandes números de Mourier

En esta sección estudiamos la extensión de la ley fuerte de los grandes números para elementos aleatorios *independientes e idénticamente distribuidos* a espacios de Banach. Este resultado fue establecido por la matemática francesa Edith Mourier en su tesis doctoral en 1953 [24]. No solo eso, a E. Mourier también se le atribuye el inicio de la teoría de Probabilidad en espacios de Banach (véase [5]). A continuación presentamos el enunciado y prueba de tal resultado.

Teorema 3.1.1 (Mourier [24])**.** *Sea X un espacio de Banach separable y sea $(V_n)_n$ una sucesión de elementos aleatorios independientes e idénticamente distribuidos en X para los cuales $\mathbb{E}[\|V_1\|] < \infty$, entonces*

$$\left\| \frac{1}{n} \sum_{k=1}^{n} V_k - \mathbb{E}[V_1] \right\| \to 0$$

casi seguramente.

Demostración. Notemos que como $(V_n)_n$ es una sucesión de elementos aleatorios idénticamente distribuidos, $\mathbb{E}[V_n] = \mathbb{E}[V_1]$ para cada $n \geq 1$ y como por hipótesis $\mathbb{E}[\|V_1\|] < \infty$ y X separable y completo, entonces $\mathbb{E}[V_1]$ existe.

Primero supongamos que para cada n, $V_n \colon \Omega \to X$ es un e.a. discreto que toma valores $\{x_i : i \geq 1\}$ con $x_i \in X$ para cada $i \geq 1$. Dado $t \in \mathbb{N}$ definimos

$$V_k^t = \sum_{i=1}^{t} x_i \mathbb{I}_{\{V_k = x_i\}}.$$

Observemos que $\left(\mathbb{I}_{\{V_k=x_i\}}\right)_{k=1}^{\infty}$ define una sucesión de variables aleatorias independientes e idénticamente distribuidas (pues $(V_n)_n$ ya lo es). Más aún, la sucesión de variables aleatorias $\left(\mathbb{I}_{\{V_k=x_i\}}\right)_k$ satisface que para cada $k \in \mathbb{N}$,

$$\mathbb{E}\left[\mathbb{I}_{\{V_k=x_i\}}\right] = \mathbb{P}\left[\{V_k = x_i\}\right] < \infty$$

y por lo tanto la sucesión $\left(\mathbb{I}_{\{V_k=x_i\}}\right)_k$ obedece la ley fuerte de los grandes números[1], esto es,

$$\frac{1}{n}\sum_{k=1}^{n} \mathbb{I}_{\{V_k=x_i\}} \overset{\text{c.s.}}{\to} \mathbb{P}\left[\{V_1 = x_i\}\right]$$

cuando $n \to \infty$. Además,

$$\begin{aligned}
\mathbb{E}\left[V_k^t\right] &= \mathbb{E}\left[\sum_{i=1}^{t} x_i \mathbb{I}_{\{V_k=x_i\}}\right] = \sum_{i=1}^{t} \mathbb{E}\left[x_i \mathbb{I}_{\{V_k=x_i\}}\right] \\
&= \sum_{i=1}^{t} x_i \mathbb{E}\left[\mathbb{I}_{\{V_k=x_i\}}\right] = \sum_{i=1}^{t} x_i \mathbb{P}\left(\{V_k = x_i\}\right) \\
&= \sum_{i=1}^{t} x_i \mathbb{P}\left(\{V_1 = x_i\}\right) = \mathbb{E}\left[V_1^t\right]
\end{aligned}$$

donde la tercera igualdad se da por la parte d) del teorema 1.3.1 tomando el

[1] Sean $V_1, V_2, \ldots$ variables aleatorias independientes e idénticamente distribuidas con media μ. Entonces

$$\frac{1}{n}\sum_{k=1}^{n} V_k \overset{\text{c.s.}}{\to} \mu$$

cuando $n \to \infty$.

elemento aleatorio $W\colon \Omega \to X$ constante; $W(\omega) = x_i$. Por lo que, para cada t,

$$\begin{aligned}
0 \leq \left\| \frac{1}{n} \sum_{k=1}^{n} \left(V_k^t - \mathbb{E}\left[V_k^t\right] \right) \right\| &= \left\| \frac{1}{n} \sum_{k=1}^{n} \left(V_k^t - \mathbb{E}\left[V_1^t\right] \right) \right\| \\
&= \left\| \frac{1}{n} \sum_{k=1}^{n} \left(\sum_{i=1}^{t} x_i \mathbb{I}_{\{V_k = x_i\}} - \sum_{i=1}^{t} x_i \mathbb{P}\left[\{V_1 = x_i\}\right] \right) \right\| \\
&= \left\| \frac{1}{n} \sum_{k=1}^{n} \sum_{i=1}^{t} x_i \left(\mathbb{I}_{\{V_k = x_i\}} - \mathbb{P}\left[\{V_1 = x_i\}\right] \right) \right\| \\
&= \left\| \sum_{i=1}^{t} \frac{x_i}{n} \left(\sum_{k=1}^{n} \mathbb{I}_{\{V_k = x_i\}} - \mathbb{P}\left[\{V_1 = x_i\}\right] \right) \right\| \\
&\leq \sum_{i=1}^{t} \|x_i\| \left| \frac{1}{n} \sum_{k=1}^{n} \left(\mathbb{I}_{\{V_k = x_i\}} - \mathbb{P}\left[\{V_1 = x_i\}\right] \right) \right| \overset{\text{c.s.}}{\to} 0
\end{aligned}$$

cuando $n \to \infty$. Es decir, para cada t,

$$\left\| \frac{1}{n} \sum_{k=1}^{n} V_k^t - \mathbb{E}\left[V_1^t\right] \right\| \overset{\text{c.s.}}{\to} 0 \tag{3.1}$$

cuando $n \to \infty$. Ahora, para cada t definimos la sucesión $(R_k^t)_{k=1}^{\infty}$ donde

$$R_k^t = V_k - V_k^t = V_k - \sum_{i=1}^{t} x_i \mathbb{I}_{\{V_k = x_i\}}.$$

Entonces $(\|R_k^t\|)_k$ también define a una sucesión de variables aleatorias independientes e idénticamente distribuidas tales que

$$\mathbb{E}\left[\left\|R_1^t\right\|\right] = \sum_{i=t+1}^{\infty} \|x_i\| \cdot \mathbb{P}\left(\{V_1 = x_i\}\right) \leq \mathbb{E}\left[\|V_1\|\right] < \infty.$$

Así que para cada $t \in \mathbb{N}$, $(\|R_k^t\|)_k$ es una sucesión de variables aleatorias que obedece la ley fuerte de los grandes números, es decir,

$$\frac{1}{n} \sum_{k=1}^{n} \left\|R_k^t\right\| \overset{\text{c.s.}}{\to} \mathbb{E}\left[\left\|R_1^t\right\|\right] \tag{3.2}$$

cuando $n \to \infty$. Notemos además que, dado $\omega \in \Omega$,

$$\begin{aligned}
\lim_{t\to\infty} \left\|R_1^t(\omega)\right\| &= \lim_{t\to\infty} \left\|V_1(\omega) - V_1^t(\omega)\right\| \\
&= \lim_{t\to\infty} \left\| V_1(\omega) - \sum_{i=1}^{t} x_i \mathbb{I}_{\{V_1 = x_i\}}(\omega) \right\| \\
&= \left\| V_1(\omega) - \lim_{t\to\infty} \sum_{i=1}^{t} x_i \mathbb{I}_{\{V_1 = x_i\}}(\omega) \right\| = 0,
\end{aligned} \tag{3.3}$$

puesto que el segundo sumando dentro de las «barras» corre sobre todos los posibles valores de V_1. Así que $\|R_1^t\| \to 0$ (puntualmente) cuando $t \to \infty$. Y no solamente eso,

$$\begin{aligned}\left\|R_1^t(\omega)\right\| &= \left\|V_1(\omega) - \sum_{i=1}^{t} x_i \mathbb{I}_{\{V_1 = x_i\}}(\omega)\right\| \\ &= \begin{cases} 0 & \text{si} \quad V_1(\omega) = x_j \text{ con } 1 \leq j \leq t \\ \|V_1(\omega)\| & \text{si} \quad V_1(\omega) = x_j \text{ con } j > t. \end{cases}\end{aligned} \tag{3.4}$$

Por lo tanto $\|R_1^t\| \leq \|V_1\|$ para cada t. Como por hipótesis $\|V_1\|$ es integrable, el teorema de convergencia dominada garantiza que

$$\lim_{t\to\infty} \mathbb{E}\left[\left\|R_1^t\right\|\right] = \mathbb{E}\left[\lim_{t\to\infty} \left\|R_1^t\right\|\right] = 0.$$

Sea $\varepsilon > 0$, entonces existe $N_0 \in \mathbb{N}$ tal que $\mathbb{E}\left[\|R_1^t\|\right] < \dfrac{\varepsilon}{4}$ si $t \geq N_0$. Ahora, por (3.1) y (3.2), para cada entero t existen N_t y M_t conjuntos nulos para los cuales las convergencias antes citadas no se dan. Sea $S = \bigcup_{t=1}^{\infty} N_t \cup M_t$, nuevamente por (3.1) y (3.2), dado $\omega \notin S$, existe $N = N(\varepsilon, \omega)$ tal que para cada $n \geq N$

$$\left\|\frac{1}{n}\sum_{k=1}^{n} V_k^t(\omega) - \mathbb{E}\left[V_1^t(\omega)\right]\right\| < \frac{\varepsilon}{4},$$

y también

$$\left|\frac{1}{n}\sum_{k=1}^{n} \left\|R_k^t(\omega)\right\| - \mathbb{E}\left[\left\|R_1^t(\omega)\right\|\right]\right| < \frac{\varepsilon}{4}$$

lo cuál implica que

$$\frac{1}{n}\sum_{k=1}^{n} \left\|R_k^t(\omega)\right\| < \mathbb{E}\left[\left\|R_1^t(\omega)\right\|\right] + \frac{\varepsilon}{4} < \frac{\varepsilon}{2},$$

siempre que $t \geq N_0$. Luego, por la linealidad y contractividad de la integral de Pettis (ver teorema 1.3.1 incisos (a) y (f)),

$$\left\|\mathbb{E}\left[V_1^t\right] - \mathbb{E}\left[V_1\right]\right\| = \left\|\mathbb{E}\left[V_1^t - V_1\right]\right\| \leq \mathbb{E}\left[\left\|V_1^t - V_1\right\|\right] = \mathbb{E}\left[\|R_1\|\right] < \frac{\varepsilon}{4}.$$

Entonces, para cada $\omega \notin S$, $t \geq N_0$ y $n \geq N$ se tiene que

$$\begin{aligned}\left\|\frac{1}{n}\sum_{k=1}^{n} V_k(\omega) - \mathbb{E}\left[V_1\right]\right\| &= \left\|\frac{1}{n}\sum_{k=1}^{n}\left(V_k^t(\omega) - \mathbb{E}\left[V_1^t\right]\right)\right. \\ &\quad \left. + \frac{1}{n}\sum_{k=1}^{n}\left(V_k(\omega) - V_k^t(\omega)\right) + \mathbb{E}\left[V_1^t\right] - \mathbb{E}\left[V_1\right]\right\| \\ &\leq \left\|\frac{1}{n}\sum_{k=1}^{n} V_k^t(\omega) - \mathbb{E}\left[V_1^t\right]\right\| + \frac{1}{n}\sum_{k=1}^{n}\left\|R_k^t(\omega)\right\| \\ &\quad + \left\|\mathbb{E}\left[V_1^t\right] - \mathbb{E}\left[V_1\right]\right\| \\ &< \frac{\varepsilon}{4} + \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

Esto prueba el teorema Mourier para el primer caso cuando los elementos aleatorios son discretos. Enseguida haremos el caso general: por el lema 1.1.4, dado $m \in \mathbb{N}$ existe $T_m \colon X \to X$ función Borel medible que toma a lo más una cantidad numerable de valores y tal que $\|T_m(x) - x\| < \frac{1}{m}$ para cada $x \in X$. Entonces,

$$\begin{aligned}\left\|\frac{1}{n}\sum_{k=1}^{n} V_k - \mathbb{E}\left[V_k\right]\right\| &= \left\|\frac{1}{n}\sum_{k=1}^{n}(V_k - T_m(V_k))\right. \\ &\quad \left. + \frac{1}{n}\sum_{k=1}^{n}\left(T_m(V_k) - \mathbb{E}\left[T_m(V_1)\right]\right) + \mathbb{E}\left[T_m(V_1)\right] - \mathbb{E}\left[V_1\right]\right\| \\ &\leq \frac{1}{n}\sum_{k=1}^{n}\left\|V_k - T_m(V_k)\right\| + \left\|\frac{1}{n}\sum_{k=1}^{n} T_m(V_1) - \mathbb{E}\left[T_m(V_1)\right]\right\| \\ &\quad + \left\|\mathbb{E}\left[T_m(V_1)\right] - \mathbb{E}\left[V_1\right]\right\|.\end{aligned}$$

Observemos que el primer sumando del lado derecho de la desigualdad anterior satisface que

$$\frac{1}{n}\sum_{k=1}^{n}\left\|V_k - T_m(V_k)\right\| \leq \frac{1}{n}\sum_{k=1}^{n}\frac{1}{m} = \frac{1}{n}\frac{n}{m} = \frac{1}{m},$$

mientras que el tercero cumple que

$$\left\|\mathbb{E}\left[T_m(V_1)\right] - \mathbb{E}\left[V_1\right]\right\| = \left\|\mathbb{E}\left[T_m(V_1) - V_1\right]\right\| \leq \mathbb{E}\left[\left\|T_m(V_1) - V_1\right\|\right] \leq \frac{1}{m}.$$

Para el segundo sumando: como $T_m \colon X \to X$ es medible, aplicando el lema 1.2.4, obtenemos que $(T_m(V_k))_{k\geq 1}$ es una sucesión de elementos aleatorios independientes e idénticamente distribuidos en X. Además $T_m(V_k)$ es discreta para cada $k \geq 1$ y son tales que

$$\begin{aligned}\mathbb{E}\left[\|T_m(V_1)\|\right] &= \mathbb{E}\left[\|T_m(V_1) - V_1 + V_1\|\right] \\ &\leq \mathbb{E}\left[\|T_m(V_1) - V_1\|\right] + \mathbb{E}\left[\|V_1\|\right] \\ &\leq \frac{1}{m} + \mathbb{E}\left[\|V_1\|\right] < \infty.\end{aligned}$$

Así que aplicando el primer caso de la prueba obtenemos que

$$\left\| \frac{1}{n} \sum_{k=1}^{n} T_m\left(V_k\right) - \mathbb{E}\left[T_m(V_1)\right] \right\| \overset{\text{c.s.}}{\to} 0$$

cuando $n \to \infty$, y esto ocurre para cada $m \in \mathbb{N}$. Y por lo tanto existe $N_m \in \mathcal{F}$ conjunto nulo tal que la convergencia anterior no se sigue en él. Sea $S = \bigcup_{m=1}^{\infty} N_m$, así $\mathbb{P}(S) = 0$ y si $\omega \notin S$ para n suficientemente grande se sigue que

$$\left\| \frac{1}{n} \sum_{k=1}^{n} V_k - \mathbb{E}\left[V_k\right] \right\| < \frac{2}{m} + \left\| \frac{1}{n} \sum_{k=1}^{n} T_m\left(V_k\right) - \mathbb{E}\left[T_m(V_1)\right] \right\| < \frac{3}{m}.$$

Esto completa la prueba del teorema 3.1.1. ■

Comentario. Como cada espacio normado es isométricamente isomorfo a un subespacio denso de un espacio de Banach, el teorema 3.1.1 puede ser extendido a cualquier espacio normado separable (no necesariamente completo) si asumimos que $\mathbb{E}[V_1]$ existe.

El ejemplo 3.0.1 también ilustra que la hipótesis de «idéntica distribución» no puede ser sustituida por una hipótesis de acotamiento en los momentos de las variables aleatorias $\|V_n\|$, como por ejemplo acotamiento de las varianzas.

3.2. Una ley débil de los grandes números

En [35] R. L. Taylor mostró que una sucesión de elementos aleatorios idénticamente distribuidos en un espacio de Banach que tiene base de Schauder satisface la ley débil de los grandes numeros en la «topología de la norma» si y solo si satisface la ley débil de los grandes números en la «topología débil». Esto es,

$$\left\| \frac{1}{n} \sum_{k=1}^{n} V_k - \mathbb{E}\left[V_1\right] \right\| \to 0 \qquad \text{en probabilidad}$$

si y solo si

$$\left| \frac{1}{n} \sum_{k=1}^{n} f(V_k) - \mathbb{E}\left[f(V_1)\right] \right| \to 0 \qquad \text{en probabilidad,}$$

para cada $f \in X^*$. Antes de probar este resultado primero mostraremos una ligera modificación del mismo. Más precisamente, probaremos que la ley débil de los grandes números se sigue en la topología de la norma si y solo si la ley débil de los grandes números se sigue en cada coordenada de la base. Para ello vale la pena recordar algunas definiciones.

Definición 3.2.1. Sea X un espacio de Banach. Una sucesión $\{b_n : n \geq 1\}$ en X es **base de Schauder** para X si para cada $x \in X$ existe una única sucesión de escalares $\{t_n : n \geq 1\}$ tal que

$$x = \lim_{n\to\infty} \sum_{k=1}^{n} t_k b_k. \tag{3.5}$$

Diremos que **una base de Schauder** $\{b_n\}$ **es monótona** si la sucesión de números escalares $\{\|\sum_{k=1}^{n} t_k b_k\| : n \geq 1\}$ es monótona creciente para cada sucesión de escalares $\{t_n\}$.

Sin entrar mucho en detalles, cuando un espacio de Banach tiene base de Schauder $\{b_n\}$ podemos definir una sucesión $\{f_k\}$ en X^*, $f_k \colon X \to \mathbb{R}$, dada por $f_k(x) = t_k$, donde $x \in X$ y

$$x = \lim_{n\to\infty} \sum_{k=1}^{n} t_k b_k.$$

Los funcionales lineales $\{f_k\}$ son llamados **funcionales coordenados** para la base de Schauder $\{b_n\}$. La linealidad de tales funcionales es consecuencia directa de la unicidad de la representación que tiene cada $x \in X$ (ver (3.5)). Nótese que los funcionales coordenados dependen de la base y no necesariamente son continuos (ver ejemplo 3.2.2). Sin embargo, si el espacio es completo, tales funcionales resultan ser continuos (ver lema 3.2.1).

También podemos definir la siguiente sucesión de funciones lineales $\{P_n\}$ donde

$$P_n(x) = \sum_{k=1}^{n} f_k(x) b_k$$

para cada $x \in X$. Las funciones $\{P_n\}$ son llamadas **proyecciones canónicas** (u **operadores de sumas parciales**) para la base de Schauder $\{b_n\}$.

Lema 3.2.1.

a) Si X es un espacio normado con base monótona, entonces $\|P_n\| \leq 1$ para cada n, es decir, $\|P_n(x)\| \leq \|x\|$ para cada $x \in X$ y cada n.

b) Si X es un espacio de Banach con base de Schauder, entonces existe una constante positiva, $M > 0$, tal que $\|P_n\| \leq M$ para cada n, es decir, $\|P_n(x)\| \leq M\|x\|$ para cada $x \in X$ y para cada n. Nos referiremos a tal constante como ***constante de base****.*

La prueba de este lema puede ser consultada en [36] (ver página 15). Mediante este lema también se puede establecer la continuidad de los funcionales coordenados f_n pues si para cada n, $\|P_n\| \leq M$, entonces

$$\begin{aligned} |f_n(x)|\|b_n\| &= \|t_n b_n\| = \|P_n(x) - P_{n-1}(x)\| \\ &\leq (\|P_n\| + \|P_{n-1}\|)\|x\| \leq 2M\|x\|, \end{aligned}$$

y por tanto $\|f_n\| \leq \dfrac{2M}{\|b_n\|}$, es decir, f_n es continua para cada $n \in \mathbb{N}$.

Ejemplo 3.2.1. Los espacios c_0, ℓ_p y $c := \mathbb{R}^\infty$ tienen por base de Schauder a $\{e_n\}$ donde e_n es la sucesión que toma el valor 1 en la n-ésima coordenada y 0 en el resto.

Ejemplo 3.2.2. Sea X el espacio de polinomios reales definidos en el intervalo $[0,1]$, y consideremos la norma $\|x\| = \sup_{0\leq t\leq 1}|x(t)|$. El conjunto $\{b_n = t^n : n \in \mathbb{N}\}$ forma una base de Schauder para X, y los funcionales coordenados están dado por

$$f_k(x) = \left.\frac{d^k x(t)}{dt^k}\right|_{t=0} \frac{1}{k!}.$$

Para cada n definimos $q_n(t) = (1-t)^n$. Notemos que $\|q_n\| \leq 1$ para cada n y además que $|f_1(q_n)| = n$. Entonces f_1 no es continua pues $\|\frac{1}{n}q_n\| \to 0$ cuando $n \to \infty$ pero $\left|f_1\left(\frac{1}{n}q_n\right)\right| = 1$ para cada n. El lema 3.2.1 no aplica pues X no es completo.

Teorema 3.2.2 (Taylor [35])**.** *Sea X un espacio de Banach con base de Schauder $\{b_n\}$. Sea $(V_n)_n$ una sucesión de elementos aleatorios independientes e idénticamente distribuidos en X tales que $\mathbb{E}[\|V_1\|] < \infty$. Para cada funcional coordenado f_i, la ley débil de los grandes números se sigue para las variables aleatorias $(f_i(V_n))_n$ si y solo si*

$$\left\|\frac{1}{n}\sum_{k=1}^{n} V_k - \mathbb{E}[V_1]\right\| \to 0$$

en probabilidad.

Demostración. ($\Leftarrow$) La parte «si» es obvia pues convergencia en la topología de la norma implica convergencia en la topología débil, esto es:

$$\begin{aligned} \mathbb{P}\left[\left|\frac{1}{n}\sum_{k=1}^{n} f(V_k) - \mathbb{E}[f(V_1)]\right| > \varepsilon\right] &= \mathbb{P}\left[\left|\frac{1}{n}\sum_{k=1}^{n} f(V_k) - f(\mathbb{E}[V_1])\right| > \varepsilon\right] \\ &= \mathbb{P}\left[\left|f\left(\frac{1}{n}\sum_{k=1}^{n} V_k - \mathbb{E}[V_1]\right)\right| > \varepsilon\right] \\ &\leq \mathbb{P}\left[\|f\|\left\|\frac{1}{n}\sum_{k=1}^{n} V_k - \mathbb{E}[V_1]\right\| > \varepsilon\right] \\ &= \mathbb{P}\left[\left\|\frac{1}{n}\sum_{k=1}^{n} V_k - \mathbb{E}[V_1]\right\| > \frac{\varepsilon}{\|f\|}\right] \end{aligned}$$

para cada $\varepsilon > 0$ y para cada $f \in X^*$. En particular lo anterior ocurre para cada funcional coordenado f_i. Tomando el límite cuando $n \to \infty$ obtenemos que para cada i, $(f_i(V_n))_n$ obedece la ley débil de los grandes números.

($\Rightarrow$) A continuación probamos la parte «solo si». Como $\mathbb{E}\left[\|V_1\|\right] < \infty$, y X es un espacio completo separable, $\mathbb{E}\left[V_1\right]$ existe (ver el lema 1.3.3). Sin pérdida de generalidad supongamos que $\mathbb{E}\left[V_1\right] = \mathbf{0}$, de lo contrario consideramos $Z_n = V_n - \mathbb{E}\left[V_1\right]$. Sea $\varepsilon > 0$, deseamos mostrar que

$$\left\| \frac{1}{n} \sum_{k=1}^{n} V_k \right\| \to 0$$

en probabilidad, esto es, dado $\delta > 0$ existe un entero positivo $N = N(\varepsilon, \delta)$ tal que

$$\mathbb{P}\left[\left\| \frac{1}{n} \sum_{k=1}^{n} V_k \right\| > \varepsilon \right] < \delta$$

para cada $n \geq N$.

Dado $t \in \mathbb{N}$ definimos $Q_t \colon X \to X$ dada por $Q_t(x) = x - P_t(x)$, es claro que Q_t es lineal. Sea $m > 0$ la constante de la base $\{b_n\}$ para la cual $\|P_t\| < m$ para cada t. Entonces,

$$\|Q_t(x)\| = \|x - P_t(x)\| \leq \|x\| + \|P_t(x)\| \leq \|x\| \left(1 + \|P_t\|\right)$$

y por tanto $\|Q_t\| \leq 1 + \|P_t\|$ para cada t (ver apéndice A, igualdad A.4 para la definición de la norma de un operador lineal).
Ahora bien, dados n y t tenemos que

$$\begin{aligned} \frac{1}{n} \sum_{k=1}^{n} V_k &= \frac{1}{n} \sum_{k=1}^{n} \left(V_k - P_t(V_k) + P_t(V_k)\right) \\ &= \frac{1}{n} \sum_{k=1}^{n} Q_t(V_k) + \frac{1}{n} \sum_{k=1}^{n} P_t(V_k). \end{aligned} \tag{3.6}$$

Y observemos que

$$\begin{aligned} \mathbb{P}\left[\left\| \frac{1}{n} \sum_{k=1}^{n} Q_t(V_k) \right\| > \frac{\varepsilon}{2} \right] &\leq \mathbb{P}\left[\frac{1}{n} \sum_{k=1}^{n} \|Q_t(V_k)\| > \frac{\varepsilon}{2} \right] \leq \frac{2}{\varepsilon} \mathbb{E}\left[\frac{1}{n} \sum_{k=1}^{n} \|Q_t(V_k)\| \right] \\ &= \frac{2}{\varepsilon} \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}\left[\|Q_t(V_k)\|\right] = \frac{2}{\varepsilon} \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}\left[\|Q_t(V_1)\|\right] \\ &= \frac{2}{\varepsilon} \mathbb{E}\left[\|Q_t(V_1)\|\right] \end{aligned}$$

donde la segunda desigualdad es consecuencia de la desigualdad de Markov mientras que la penúltima igualdad se da por ser $(Q_t(V_k))_k$ una sucesión de elementos aleatorios idénticamente distribuidos. Y no solo eso, ya que $\{b_n\}$ es base de Schauder para X, dado V un elemento aleatorio en X, tenemos la representación

$$V = \sum_{j=1}^{\infty} f_j(V) b_j$$

y por lo tanto

$$\lim_{t\to\infty} Q_t(V_1) = V_1 - \lim_{t\to\infty} P_t(V_1) = V_1 - \lim_{t\to\infty} \sum_{j=1}^{t} f_j(V_1) b_j = \mathbf{0}.$$

De modo que $\lim_{t\to\infty} \|Q_t(V_1)\| = 0$. Como además $\|Q_t(V_1)\| \leq (1+m)\|V_1\|$ con $(m+1)\|V_1\|$ integrable, el teorema de convergencia dominada (de Lebesgue) nos asegura que

$$\lim_{t\to\infty} \mathbb{E}\left[\|Q_t(V_1)\|\right] = \mathbb{E}\left[\lim_{t\to\infty} \|Q_t(V_1)\|\right] = 0.$$

Por lo tanto, dado $\delta > 0$,

$$\mathbb{P}\left[\left\|\frac{1}{n}\sum_{k=1}^{n} Q_t(V_k)\right\| > \frac{\varepsilon}{2}\right] < \frac{\delta}{2} \tag{3.7}$$

para toda $t \geq N_0$ para alguna N_0.
Por otra parte,

$$\begin{aligned}
\mathbb{P}\left[\left\|\frac{1}{n}\sum_{k=1}^{n} P_t(V_k)\right\| > \frac{\varepsilon}{2}\right] &= \mathbb{P}\left[\left\|\frac{1}{n}\sum_{k=1}^{n}\sum_{i=1}^{t} f_i(V_k) b_i\right\| > \frac{\varepsilon}{2}\right] \\
&= \mathbb{P}\left[\left\|\frac{1}{n}\sum_{i=1}^{t}\sum_{k=1}^{n} f_i(V_k) b_i\right\| > \frac{\varepsilon}{2}\right] \\
&= \mathbb{P}\left[\left\|\frac{1}{n}\sum_{i=1}^{t} f_i\left(\sum_{k=1}^{n} V_k\right) b_i\right\| > \frac{\varepsilon}{2}\right] \\
&\leq \mathbb{P}\left[\sum_{i=1}^{t}\left| f_i\left(\frac{1}{n}\sum_{k=1}^{n} V_k\right)\right| \|b_i\| > \frac{\varepsilon}{2}\right]
\end{aligned}$$

y obsérvese que

$$\left\{\omega : \sum_{i=1}^{t}\left| f_i\left(\frac{1}{n}\sum_{k=1}^{n} V_k\right)\right| \|b_i\| > \frac{\varepsilon}{2}\right\} \subset \bigcup_{i=1}^{t}\left\{\omega : \left| f_i\left(\frac{1}{n}\sum_{k=1}^{n} V_k(\omega)\right)\right| \|b_i\| t > \frac{\varepsilon}{2}\right\}.$$

Llamemos A y B respectivamente a los conjuntos anteriores y supongamos que $A \not\subseteq B$, es decir, existe $\omega \in A$ tal que $\omega \notin B$. Pero si $\omega \notin B$, entonces para toda $i \in \{1, 2, \ldots t\}$, $\left|f_i\left(\frac{1}{n}\sum_{k=1}^{n} V_k(\omega)\right)\right| \|b_i\| \leq \frac{\varepsilon}{2t}$ lo cuál implica que $\sum_{i=1}^{t}\left|f_i\left(\frac{1}{n}\sum_{k=1}^{n} V_k(\omega)\right)\right| \|b_i\| \leq \sum_{i=1}^{t} \frac{\varepsilon}{2t} = \frac{\varepsilon}{2}$. Esto es una contradicción pues suponemos que ω sí pertenece a A. Luego,

$$\mathbb{P}\left[\sum_{i=1}^{t}\left| f_i\left(\frac{1}{n}\sum_{k=1}^{n} V_k\right)\right| \|b_i\| > \frac{\varepsilon}{2}\right] \leq \sum_{i=1}^{t} \mathbb{P}\left[\left|\frac{1}{n}\sum_{k=1}^{n} f_i\left(V_k\right)\right| > \frac{\varepsilon}{2t\,\|b_i\|}\right].$$

Pero por hipótesis para cada $i \in \mathbb{N}$, la sucesión $(f_i(V_k))_k$ satisface la ley débil de los grandes números (y $\mathbb{E}[f_i(V_1)] =: f_i(\mathbb{E}[V_1]) = f(\mathbf{0}) = 0$), entonces

$$\mathbb{P}\left[\left|\frac{1}{n}\sum_{k=1}^{n} f_i(V_k)\right| > \frac{\varepsilon}{2t\|b_i\|}\right] \to 0$$

cuando $n \to \infty$. Por lo que existe $N = N(\varepsilon, \delta)$ tal que

$$\mathbb{P}\left[\left\|\frac{1}{n}\sum_{k=1}^{n} P_t(V_k)\right\| > \frac{\varepsilon}{2}\right] < \frac{\delta}{2} \tag{3.8}$$

para cada $n \geq N$. Finalmente, usando (3.6), (3.7) y (3.8) se sigue que

$$\begin{aligned}\mathbb{P}\left[\left\|\frac{1}{n}\sum_{k=1}^{n} V_k\right\| > \varepsilon\right] &\leq \mathbb{P}\left[\left\|\frac{1}{n}\sum_{k=1}^{n} Q_t(V_k)\right\| > \frac{\varepsilon}{2}\right] + \mathbb{P}\left[\left\|\frac{1}{n}\sum_{k=1}^{n} P_t(V_k)\right\| > \frac{\varepsilon}{2}\right] \\ &< \frac{\delta}{2} + \frac{\delta}{2} = \delta\end{aligned}$$

para cada $n \geq N(\varepsilon, \delta)$. Por lo tanto

$$\left\|\frac{1}{n}\sum_{k=1}^{n} V_k\right\| \to 0$$

en probabilidad. ■

Como consecuencia de este teorema obtenemos el resultado mencionado al inicio de la sección:

Teorema 3.2.3 (Taylor [35]). *Sea X un espacio de Banach con base de Schauder $\{b_n\}$. Sea $(V_n)_n$ una sucesión de elementos aleatorios independientes e idénticamente distribuidos en X tales que $\mathbb{E}[\|V_1\|] < \infty$. Para cada $f \in X^*$, la ley débil de los grandes números se sigue para la sucesión de variables aleatorias $(f(V_n))_n$ si y solo si*

$$\left\|\frac{1}{n}\sum_{k=1}^{n} V_k - \mathbb{E}[V_1]\right\| \to 0$$

en probabilidad.

Demostración. «El regreso» es inmediato, nuevamente porque convergencia en la topología de la norma implica convergencia en la topología débil, mientras que «la ida» se hace como antes. Observe que el punto clave de esta última implicación fue separar la expresión $n^{-1}\sum_{k=1}^{n} V_k$ como sigue

$$\frac{1}{n}\sum_{k=1}^{n} V_k = \frac{1}{n}\sum_{k=1}^{n} Q_t(V_k) + \frac{1}{n}\sum_{k=1}^{n} P_t(V_k)$$

donde $\frac{1}{n}\sum_{k=1}^{n} Q_t(V_k)$ estaba acotado (en probabilidad) debido al teorema de convergencia dominada. Por otro lado, el sumando $\frac{1}{n}\sum_{k=1}^{n} P_t(V_k)$ fue controlado por la hipótesis de convergencia en probabilidad de $\frac{1}{n}\sum_{k=1}^{n} f_i(V_k)$ donde f_i es un funcional coordenado para cada $i \in \mathbb{N}$. El mismo argumento funciona en este caso, pues ahora suponemos $\frac{1}{n}\sum_{k=1}^{n} f(V_k)$ converge en probabilidad para cada $f \in X^*$, en particular esto ocurre para cada funcional coordenado f_i ya que estos pertenecen a X^*. ■

Vale la pena señalar que, mientras que la ley débil de los grandes números en la topología débil es suficiente para garantizar la ley débil de los grandes números en la topología fuerte, no lo es para garantizar la ley fuerte de los grandes números en la topología de la norma. A. Beck y B. Warren en [6] construyeron una sucesión de elementos aleatorios $(V_n)_n$ en el espacio de Banach separable c_0 tales que son

1. idénticamente distribuidas,
2. uniformemente acotadas, más precisamente, $\|V_n\| \leq 1$ para cada n,
3. $(f(V_n))_n$ satisface la ley débil de los grandes números para cada $f \in X^*$ pero no satisface la ley fuerte de los grandes números en la topología de la norma.

Por lo que con convergencia casi segura ni el teorema 3.2.2 ni el 3.2.3 son siempre posibles.

Respecto a la ley fuerte de los grandes números de Rademacher-Mensov (teo. 2.2.5) mencionamos que su extensión a espacios de Banach no es posible debido a la ausencia de una estructura multiplicativa en el espacio que permita definir el concepto de ortogonalidad. Lo que se hace en este caso es recurrir al concepto de ortogonalidad en el sentido débil, esto es, una sucesión de elementos aleatorios $(V_n)_n$ es *débilmente ortogonal*, si y sólo si, para cada $f \in X^*$, y cada $i \neq j$,

$$\mathbb{E}\left[f(V_i)f(V_j)\right] = 0.$$

Bajo ciertas condiciones (más restrictivas) se pueden obtener leyes de grandes números para sucesiones de elementos aleatorios débilmente ortogonales, sin embargo tales leyes distan de la versión dada en el teorema 2.2.5 (al respecto véase [6]).

3.3. Ley de los grandes números de Pisier y Hoffmann-Jørgensen

En esta última sección nuestro propósito es estudiar un resultado que proporciona una equivalencia entre una condición sobre la geometría del espacio llamada *tipo p* y una ley fuerte de los grandes números. Para ello necesitamos introducir algunos resultados relativos a elementos aleatorios simétricos tales como las desigualdades de Lévy, las sucesiones de Rademacher y las sumas de Rademacher, así como las importantes desigualdades de Hoffmann-Jørgensen sobre la suma de elementos aleatorios independientes.

Como antes, y al menos que se especifique lo contrario, $(X, \|\cdot\|)$ denotará un espacio de Banach separable.

3.3.1. Antecedentes

La noción de tipo (y cotipo) de un espacio de Banach aparece por primera vez en el marco de la teoría de operadores p-sumantes desarrollada a principios de los años 60. El concepto fue explícitamente introducido por B. Maurey durante el «Séminaire Maurey-Schwartz 1972/73» ([23, pág. 2], [21, pág. 270]) y de forma independiente por J. Hoffmann-Jørgensen [21, pág. 270]. Más tarde, el grupo encabezado por L. Schwartz, B. Maurey y G. Pisier, mostraría la importancia de este concepto en la teoría de espacios de Banach y la teoría de operadores. En los mismos años, Hoffmann-Jørgensen y Pisier [17] mostraron la relación entre el tipo p y la ley de los grandes números, sin embargo, la primera relación entre estos temas se remonta al origen de la Probabilidad en espacios de Banach.

En 1962, A. Beck [4] introduce la noción de B-convexidad como una forma de caracterizar a aquellos espacios de Banach en los que es válida una cierta ley de los grandes números. Un trabajo posterior de G. Pisier [28] identificaría a la clase de los espacios B-convexos con la clase de los espacios que tienen tipo p para alguna $p > 1$.[2]

En este texto no entraremos en detalles en lo que a B-convexidad se refiere, en su lugar puede consultarse el artículo de A. Beck [4] así como el de D. Giesy [13], también puede consultarse el capítulo 13 de [10]. En la bibliografía antes citada pueden encontrarse definiciones, ejemplos, propiedades, etc. de los espacios B-convexos, en particular, en [4, teo. 1, pág. 330] se puede ver la prueba del siguiente resultado:

Teorema (Beck [4])**.** *Sea X es un espacio de Banach separable. Si X es* ***B-convexo*** *y $(V_n)_n$ es una sucesión de elementos aleatorios independientes tales que $\mathbb{E}[V_n] = \mathbf{0}$ y $\mathbb{E}\left[\|V_n\|^2\right] \leq M$ para cada n y donde M es una constante positiva, entonces*

$$\left\| \frac{1}{n} \sum_{k=1}^{n} V_k \right\| \to 0$$

casi seguramente.

Adicionalmente en [4] Beck también muestra que la B-convexidad es una condición necesaria para obtener la ley fuerte de los grandes números para elementos aleatorios independientes con esperanza cero y varianza acotada. De manera más precisa, si X no es B-convexo, entonces existe una sucesión $(V_n)_n$ de elementos aleatorios independientes con $\mathbb{E}[V_n] = \mathbf{0}$ y $\mathrm{Var}(V_n) \leq M$ para cada n tal que $\left\|\frac{1}{n}\sum_{k=1}^{n} V_k\right\|$ no converge a cero de forma casi segura.

[2]Para una referencia histórica puede consultarse [23].

En resumen, en un espacio de Banach se satisface la ley fuerte de grandes números para elementos aleatorios independientes con esperanza cero y varianzas uniformemente acotadas si y solo si el espacio B-convexo (lo que sea que esto signifique). Tomando en consideración este resultado podemos preguntarnos si cambiando la condición «varianzas uniformemente acotadas» por alguna otra condición, como por ejemplo la «condición de Kolmogorov»:

$$\sum_{n=1}^{\infty} \frac{\operatorname{Var}(V_n)}{n^2} < \infty,$$

seguimos obteniendo una ley de grandes números (es decir, una extensión del teorema de Kolmogorov 2.3.3 para espacios de Banach separables del tipo B-convexos). La respuesta es en general negativa como muestra el siguiente ejemplo:

Ejemplo 3.3.1. Sea p un real tal que $1 < p < 2$, y consideremos el espacio $X = \ell_p = \{(x_n)_n \in \mathbb{R}^\infty : \sum_{n=1}^{\infty} |x_n|^p < \infty\}$ con la norma $\|(x_n)_n\|_{\ell_p} = (\sum_{n=1}^{\infty} |x_n|^p)^{\frac{1}{p}}$.[3]

Como en el ejemplo 3.0.1, e_n denota al elemento de ℓ_p que toma el valor 1 en la n-ésima coordenada y cero en el resto. Sea q un real tal que $1 - \frac{1}{p} < q < \frac{1}{2}$. Sea $(A_n)_n$ una sucesión de variables aleatorias independientes que toman los valores $\pm n^q$ con probabilidad $\frac{1}{2}$, $\mathbb{P}(A_n = \pm n^q) = \frac{1}{2}$, y definimos $V_n = A_n e_n$.

Entonces $(V_n)_n$ es una sucesión de elementos aleatorios en ℓ_p independientes tales que $\mathbb{E}[V_n] = \mathbf{0}$ y $\operatorname{Var}(V_n) = n^{2q}$ para cada n, y donde

$$\sum_{n=1}^{\infty} \frac{\operatorname{Var}(V_n)}{n^2} = \sum_{n=1}^{\infty} \frac{n^{2q}}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n^{2-2q}} < \infty,$$

pues $2 - 2q > 2 - 1 = 1$. Sin embargo, para cada n,

$$\begin{aligned}
\left\| \frac{1}{n} \sum_{k=1}^{n} V_k \right\|_{\ell_p}^{p} &= \frac{1}{n^p} \sum_{k=1}^{n} |A_k|^p = \sum_{k=1}^{n} \frac{k^{pq}}{n^p} \\
&> \sum_{k=\lfloor \frac{n}{2} \rfloor}^{n} \frac{k^{pq}}{n^p} > \sum_{k=\lfloor \frac{n}{2} \rfloor}^{n} \frac{\left(\frac{n}{2}\right)^{pq}}{n^p} \\
&= \sum_{k=\lfloor \frac{n}{2} \rfloor}^{n} \frac{n^{pq-p}}{2^{pq}} \\
&> \frac{n^{pq+1-p}}{2^{pq+1}},
\end{aligned}$$

que tiende a infinito conforme n tiende a infinito, ya que $1 - \frac{1}{p} < q$ implica que $pq + 1 - p > 0$. Aquí el símbolo $\lfloor a \rfloor$ representa la parte entera de $a \in \mathbb{R}$.

[3] ℓ_p, para $1 < p < \infty$, es un espacio B-convexo. Véase [13], ejemplo (*iii*).

La ley de los grandes números de Pisier y Hoffmann-Jørgensen usa una condición más general que la B-convexidad: el tipo p, y cubre el caso cuando imponemos una condición menos restrictiva que la condición de varianzas uniformemente acotadas dada en teorema de Beck.

3.3.2. Elementos aleatorios simétricos

Definición 3.3.1. Un elemento aleatorio $V\colon \Omega \to X$ se llama **simétrico** si V y $-V$ son idénticamente distribuidos. Y diremos que una sucesión de elementos aleatorios es simétrica si sus elementos son simétricos.

Lema 3.3.1. *Sean V y W elementos aleatorios en X independientes y p-integrables (en $L_p(\Omega; X)$) para alguna $1 \leq p < \infty$. Si V es simétrico o bien $\mathbb{E}[V] = \mathbf{0}$, entonces*

$$\mathbb{E}\left[\|W\|^p\right] \leq \mathbb{E}\left[\|V+W\|^p\right].$$

Demostración. Caso 1. Supongamos que V es simétrico. Como V y W son independientes, entonces $-V+W$ y $V+W$ son idénticamente distribuidos. Por lo tanto

$$\begin{aligned}
\left(\mathbb{E}\left[\|W\|^p\right]\right)^{\frac{1}{p}} &= \frac{1}{2}\left(\mathbb{E}\left[\|2W\|^p\right]\right)^{\frac{1}{p}} \\
&= \frac{1}{2}\left(\mathbb{E}\left[\|\left(V+W\right)+\left(W-V\right)\|^p\right]\right)^{\frac{1}{p}} \\
&\leq \frac{1}{2}\left(\left(\mathbb{E}\left[\|V+W\|^p\right]\right)^{\frac{1}{p}} + \left(\mathbb{E}\left[\|W-V\|^p\right]\right)^{\frac{1}{p}}\right) \\
&= \left(\mathbb{E}\left[\|V+W\|^p\right]\right)^{\frac{1}{p}}.
\end{aligned}$$

Caso 2. Supongamos que $\mathbb{E}[V] = \mathbf{0}$, entonces por la independencia entre V y W:

$$\begin{aligned}
\mathbb{E}\left[\|V+W\|^p\right] &= \int_X \int_X \|v+w\|^p \, d\mathbb{P}_V d\mathbb{P}_W \\
&\geq \int \left\| \int (v+w)\, d\mathbb{P}_V \right\|^p d\mathbb{P}_W \\
&= \int \left\| \int v\, d\mathbb{P}_V + w \int 1\, d\mathbb{P}_V \right\|^p d\mathbb{P}_W \\
&= \int \|\mathbb{E}[V] + w\|^p \, d\mathbb{P}_W \\
&= \int \|w\|^p \, d\mathbb{P}_W = \mathbb{E}\left[\|W\|^p\right].
\end{aligned}$$

La segunda desigualdad es consecuencia de la desigualdad de Hölder para los espacios $L_p(\Omega; X)$. En otras palabras, probamos que bajo las hipótesis de simetría o de esperanza nula en V,

$$\frac{1}{2}\|2W\|_p \leq \|V+W\|_p.$$

■

Otra forma equivalente de dar la definición 3.3.1 es como sigue: V es simétrica si tiene la misma distribución que εV, donde ε es una variable aleatoria Rademacher independiente de V. En efecto, supongamos que ε es una variable aleatoria Rademacher independiente de V, y sea $B \in \mathcal{B}(X)$, si $\mathbb{P}(V \in B) = \mathbb{P}(-V \in B)$ entonces

$$\begin{aligned}\mathbb{P}(\varepsilon V \in B) &= \mathbb{P}(\varepsilon V \in B, \varepsilon = +1) + \mathbb{P}(\varepsilon V \in B, \varepsilon = -1)\\ &= \mathbb{P}(V \in B, \varepsilon = +1) + \mathbb{P}(-V \in B, \varepsilon = -1)\\ &= \mathbb{P}(V \in B)\frac{1}{2} + \mathbb{P}(-V \in B)\frac{1}{2} = \mathbb{P}(V \in B).\end{aligned}$$

Por otro lado, si suponemos que $\mathbb{P}_{\varepsilon V} = \mathbb{P}_V$,

$$\begin{aligned}\mathbb{P}(V \in B) &= \mathbb{P}(\varepsilon V \in B)\\ &= \mathbb{P}(\varepsilon V \in B \,|\, \varepsilon = +1)\,\mathbb{P}(\varepsilon = +1) + \mathbb{P}(\varepsilon V \in B \,|\, \varepsilon = -1)\,\mathbb{P}(\varepsilon = -1)\\ &= \mathbb{P}(V \in B)\frac{1}{2} + \mathbb{P}(-V \in B)\frac{1}{2},\end{aligned}$$

por lo tanto $\mathbb{P}(V \in B) = \mathbb{P}(-V \in B)$ para cada $B \in \mathcal{B}(X)$.

Esta simple observación da paso a la técnica llamada *aleatorización* (o también *simetrización*) la cuál es una de las herramientas más útiles en la teoría de probabilidad en espacios de Banach y que permite aplicar desigualdades para *sumas de Rademacher* a sumas de variables aleatorias simétricas e independientes (véase [21] pág. 47).

Comentamos además que para cada elemento aleatorio V, existe una forma canónica de generar un elemento aleatorio que se «aproxima» a V: considere $V^* = V - V'$ donde V' es una copia independiente de V, es decir, $\mathbb{P}_V = \mathbb{P}_{V'}$ y V y V' son independientes. El elemento aleatorio V^* es llamado **simetrización** de V. Si $(V_n)_n$ es una suceción de elementos aleatorios, $(V_n^*)_n$ se llama *simetrización* de $(V_n)_n$ si $V_n^* = V_n - V_n'$ donde $(V_n)_n$ y $(V_n')_n$ son independientes y equidistribuidas. La simetrización de $(V_n)_n$ siempre existe al menos en el espacio producto $(\Omega \times \Omega, \mathcal{F} \otimes \mathcal{F}, \mathbb{P} \times \mathbb{P})$ (véase [15, pág. 161]).

Lema 3.3.2. *Sean $V_1, \dots, V_n$ elementos aleatorios independientes y simétricos en X. Entonces para cada elección $\varepsilon_1, \dots, \varepsilon_n \in \{-1, +1\}$ los elementos aleatorios en X^n, $(V_1, \dots, V_n)$ y $(\varepsilon_1 V_1, \dots, \varepsilon_n V_n)$, son idénticamente distribuidos.*

Demostración. Note que si $(V_i)_{i=1}^n$ son independientes entonces $(\pm V_i)_{i=1}^n$ también lo son, entonces el resultado se sigue de la observación hecha después de la definición 1.2.5. Es decir,

$$\begin{aligned}\mathbb{P}_{(V_1,\dots,V_n)} &= \mathbb{P}_{V_1} \times \cdots \times \mathbb{P}_{V_n}\\ &= \mathbb{P}_{\pm V_1} \times \cdots \times \mathbb{P}_{\pm V_n} \qquad \text{(por la simetría)}\\ &= \mathbb{P}_{(\pm V_1,\dots,\pm V_n)}\\ &= \mathbb{P}_{(\varepsilon_1 V_1,\dots,\varepsilon_n V_n)}.\end{aligned}$$

■

El resultado anterior se extiende a cualquier sucesión de signos $(\varepsilon_n)_n$ y cualquier sucesión de elementos aleatorios $(V_n)_n$.

La definición 3.3.1 de *sucesión simétrica* puede ser reformulada como sigue: una sucesión de elementos aleatorios $(V_n)_n$ es simétrica si, para cada elección de signos ± 1, $(\pm V_n)_n$ tiene la misma ley de distribución que $(V_n)_n$ (esto es, para cada $N \in \mathbb{N}, (\pm V_1, \ldots, \pm V_N)$ tiene la misma ley que $(V_1, \ldots, V_N)$ en X^N).

De acuerdo con [21], esta definición es equivalente a decir que $(V_n)_n$ tiene la misma distribución que $(\varepsilon_n V_n)_n$ donde $(\varepsilon_n)_n$ es una sucesión de Rademacher independiente de (V_n).

Las sumas parciales de una sucesión de elementos aleatorios simétricos satisfacen ciertas desigualdades importantes que se conocen como **desigualdades de Lévy.**

Proposición 3.3.3 (Lévy [21])**.** *Sea $(V_i)_i$ una sucesión independiente y/o simétrica de elementos aleatorios en X. Para cada k, sea $S_k = \sum_{i=1}^{k} V_i$, entonces para cada entero n y cada $\varepsilon > 0$,*

$$\mathbb{P}\left[\max_{1\le k\le n} \|S_k\| > \varepsilon\right] \le 2\mathbb{P}\left[\|S_n\| > \varepsilon\right] \tag{3.9}$$

y

$$\mathbb{P}\left[\max_{1\le i\le n} \|V_i\| > \varepsilon\right] \le 2\mathbb{P}\left[\|S_n\| > \varepsilon\right]. \tag{3.10}$$

Si además $(S_k)_k$ converge casi seguramente a S, entonces

$$\mathbb{P}\left[\sup_{k} \|S_k\| > \varepsilon\right] \le 2\mathbb{P}\left[\|S\| > \varepsilon\right]. \tag{3.11}$$

Demostración. Las pruebas de estas desigualdades son similares a la prueba de la desigualdad de Kolmogorov que dimos en 2.3.2. Para ilustrar esto hagamos (3.9). Dado $n \in \mathbb{N}$ definamos

$$\begin{aligned} A &= \left\{\max_{1\le k\le n} \|S_k\| > \varepsilon\right\} \\ A_k &= \{\|S_1\| \le \varepsilon, \ldots, \|S_{k-1}\| \le \varepsilon, \|S_k\| > \varepsilon\} \qquad \text{para } k = 1, \ldots, n \end{aligned}$$

entonces $\biguplus_{k=1}^{n} A_k = A$ (unión disjunta). Ahora, para cada $k \le n$ definamos $S_n^{(k)} = S_k - V_{k+1} - \cdots - V_n = 2S_k - S_n$ y sean

$$B_k = A_k \cap \{\|S_n\| > \varepsilon\} \quad \text{y} \quad C_k = A_k \cap \{\|S_n^{(k)}\| > \varepsilon\}.$$

Notemos que si $\omega \in A_k$, entonces

$$\|S_n(\omega) + S_n^{(k)}\| = 2\|S_k(\omega)\| > 2\varepsilon$$

lo que significa que o bien $\|S_n(\omega)\| > \varepsilon$ o bien $\|S_n^{(k)}\| > \varepsilon$, es decir, $A_k \subset B_k \cup C_k$, y por lo tanto

$$A_k = B_k \cup C_k = [A_k \cap \{\|S_n\| > \varepsilon\}] \cup \left[A_k \cap \{\|S_n^{(k)}\| > \varepsilon\}\right].$$

Por otra parte, del lema 3.3.2, obtenemos que los elementos aleatorios en X^n, $V = (V_1, \ldots, V_n)$ y $V^{(k)} = (V_1, \ldots, V_k, -V_{k+1}, \ldots, -V_n)$ tienen la misma distribución y como resultado

$$\mathbb{P}(B_k) = \mathbb{P}\left(A_k \cap \{\|S_n\| > \varepsilon\}\right) = \mathbb{P}\left(A_k \cap \{\|S_n^{(k)}\| > \varepsilon\}\right) = \mathbb{P}(C_k).$$

En efecto, si tomamos $W_k \colon X^k \to \mathbb{R}$ (para $k \leq n$) tal que $W_k(x_1, \ldots, x_k) = \left\|\sum_{j=1}^k x_k\right\|$, entonces W_k es continua (por ser composición de funciones continuas, la suma y la norma) por lo tanto

$$D_k = \bigcap_{j=1}^{k-1} \{W_j(x_1, \ldots, x_j) \leq \varepsilon\} \cap \{W_k(x_1, \ldots, x_k) > \varepsilon, W_n(x_1, \ldots, x_n) > \varepsilon\}$$

define a un boreliano de X^n, i.e. $D_k \in \mathcal{B}(X^n)$. Entonces,

$$\begin{aligned}
\mathbb{P}(B_k) &= \mathbb{P}\left(A_k \cap \{\|S_n\| > \varepsilon\}\right) \\
&= \mathbb{P}\left(V^{-1}(D_k)\right) \\
&= \mathbb{P}_V\left(D_k\right) \\
&= \mathbb{P}_{V^{(k)}}\left(D_k\right) \\
&= \mathbb{P}\left((V^{(k)})^{-1}(D_k)\right) \\
&= \mathbb{P}\left(A_k \cap \{\|2S_k - S_n\| > \varepsilon\}\right) = \mathbb{P}(C_k).
\end{aligned}$$

Así que

$$\begin{aligned}
\mathbb{P}(A_k) &= \mathbb{P}\left((A_k \cap \{\|S_n\| > \varepsilon\}) \cup \left(A_k \cap \{\|S_n^{(k)}\| > \varepsilon\}\right)\right) \\
&= \mathbb{P}\left(A_k \cap \{\|S_n\| > \varepsilon\}\right) + \mathbb{P}\left(A_k \cap \{\|S_n^{(k)}\| > \varepsilon\}\right) \\
&= 2\mathbb{P}\left(A_k \cap \{\|S_n\| > \varepsilon\}\right)
\end{aligned}$$

y por lo tanto

$$\begin{aligned}
\mathbb{P}(A) = \sum_{k=1}^n \mathbb{P}(A_k) &\leq 2\sum_{k=1}^n \mathbb{P}\left(A_k \cap \{\|S_n\| > \varepsilon\}\right) \\
&= 2\mathbb{P}\left(\left(\bigcup_{k=1}^n A_k\right) \cap \{\|S_n\| > \varepsilon\}\right) \\
&= 2\mathbb{P}\left(A \cap \{\|S_n\| > \varepsilon\}\right) \\
&\leq 2\mathbb{P}\left(\{\|S_n\| > \varepsilon\}\right).
\end{aligned}$$

La prueba de (3.10) es totalmente análoga y se obtiene de sustituir

$$\begin{aligned}
A_k &= \{\|V_1\| \leq \varepsilon, \ldots, \|V_{k-1}\| \leq \varepsilon, \|V_k\| > \varepsilon\} \\
S_n^{(k)} &= -V_1 - \cdots - V_{k-1} + V_k - V_{k+1} - \cdots - V_n
\end{aligned}$$

en la prueba de (3.9).

Resta probar la última desigualdad: (3.11). Sea $\Omega_0 \in \mathcal{F}$ de probabilidad 1, $\mathbb{P}(\Omega_0) = 1$, tal que $\sum_{i=1}^{\infty} V_i$ está definido (i.e. existe en X), tal conjunto existe pues suponemos que $(S_k)_k$ converge casi seguramente. Denotaremos por S a tal límite y por $M = \sup_k \|S_k(\omega)\|$.
Para cada $\varepsilon > 0$ definamos

$$\begin{aligned} A &= \left\{\omega \in \Omega_0 \,:\, \sup_k \|S_k(\omega)\| > \varepsilon\right\} = \{\omega \in \Omega_0 \,:\, M(\omega) > \varepsilon\}, \\ B &= \{\omega \in \Omega_0 \,:\, \|S(\omega)\| > \varepsilon\}, \end{aligned}$$

y como antes, dividamos el conjunto A en los conjuntos disjuntos (pero ahora en una cantidad numerable de ellos):

$$A_k = \{\omega \in \Omega_0 \,:\, \|S_1\| \leq \varepsilon, \ldots, \|S_{k-1}\| \leq \varepsilon, \|S_k\| > \varepsilon\}$$

para cada $k \in \mathbb{N}$, es decir, $\biguplus_{k=1}^{\infty} A_m = A$.
Así, si $\omega \in A_k$, entonces al menos uno de lo siguientes elementos aleatorios

$$\begin{aligned} S(\omega) &= S_k(\omega) + (V_{k+1}(\omega) + V_{k+2}(\omega) + \ldots) \\ S'(\omega) &= S_k(\omega) - (V_{k+1}(\omega) + V_{k+2}(\omega) + \ldots) = 2S_k - S \end{aligned}$$

se encuentra fuera de de $B[\mathbf{0}, \varepsilon]$. Esto es,

$$A_k \subset [A_k \cap \{\|S\| > \varepsilon\}] \cup [A_k \cap \{\|S'\| > \varepsilon\}],$$

con lo que concluimos que se da la igualdad entre estos conjuntos.

Pero como los V_i son simétricos, S y S' tienen la misma distribución así que los «vectores» $U = (V_1, \ldots, V_k, S)$ y $U^{(k)} = (V_1, \ldots, V_k, S')$ también tienen la misma distribución, esto implica que los conjuntos $M_k = A_k \cap \{\|S\| > \varepsilon\}$ y $N_k = A_k \cap \{\|S'\| > \varepsilon\}$ ocurren con la misma probilidad. Es decir,

$$\mathbb{P}\left(A_k \cap \{\|S\| > \varepsilon\}\right) = \mathbb{P}\left(A_k \cap \{\|S'\| > \varepsilon\}\right).$$

Para verificar esto, como en la prueba de la desigualdad (3.9), consideremos la función que suma $W_m(x_1, \ldots, x_m) = \sum_{j=1}^{m} x_j$ y definamos

$$\begin{aligned} D_k = \bigcap_{j=1}^{k-1} &\{\|W_j(x_1, \ldots, x_j)\| \leq \varepsilon\} \cap \{\|W_k(x_1, \ldots, x_k)\| > \varepsilon\} \\ &\cap \{\|W_{k+1}(x_1, \ldots, x_{k+1}) - W_k(x_1, \ldots, x_k)\| > \varepsilon\}. \end{aligned}$$

Claramente $D_k \in \mathcal{B}(X^{k+1})$. Luego,

$$\begin{aligned} \mathbb{P}(M_k) &= \mathbb{P}\left(A_k \cap \{\|S\| > \varepsilon\}\right) \\ &= \mathbb{P}\left(U^{-1}(D_k)\right) \\ &= \mathbb{P}_U\left(D_k\right) \\ &= \mathbb{P}_{U^{(k)}}\left(D_k\right) \\ &= \mathbb{P}\left((U^{(k)})^{-1}(D_k)\right) \\ &= \mathbb{P}\left(A_k \cap \{\|2S_k - S\| > \varepsilon\}\right) = \mathbb{P}(N_k). \end{aligned}$$

Por lo tanto

$$\mathbb{P}(A_k) = \mathbb{P}(M_k \cup N_k) \leq 2\mathbb{P}(M_k) = 2\mathbb{P}(A_k \cap \{\|S\| > \varepsilon\}) = 2\mathbb{P}(A_k \cap B).$$

Pero

$$\begin{aligned}
\mathbb{P}(B \cap A) &= \mathbb{P}\left(B \cap \left(\bigcup_{k=1}^{\infty} A_k\right)\right) = \mathbb{P}\left(\bigcup_{k=1}^{\infty} B \cap A_k\right) \\
&= \sum_{k=1}^{\infty} \mathbb{P}(B \cap A_k) \geq \frac{1}{2}\sum_{k=1}^{\infty} \mathbb{P}(A_k) = \frac{1}{2}\mathbb{P}\left(\bigcup_{k=1}^{\infty} A_k\right) = \frac{1}{2}\mathbb{P}(A).
\end{aligned}$$

Lo que implica que

$$\mathbb{P}\left(\|S\| > \varepsilon\right) \geq \frac{1}{2}\mathbb{P}\left(\sup_k \|S_k\| > \varepsilon\right).$$

■

Estas desigualdades a su vez son importantes pues implican el célebre teorema de Itô y Nisio (el cual es la generalización del teorema de Lévy a espacios de Banach) que enunciamos a continuación[4]:

Teorema 3.3.4 (Itô-Nisio [19]). *Sea $(V_i)_i$ una sucesión de elementos aleatorios independientes en X. Para cada n, denotamos por μ_n a la distribución de la n-ésima suma parcial $S_n = \sum_{i=1}^{n} V_n$, y sea $S\colon \Omega \to X$ un elemento aleatorio. Son equivalentes:*

1. $\lim_{n\to\infty} S_n = S$ *converge c.s.,*
2. $\lim_{n\to\infty} S_n = S$ *converge en probabilidad,*
3. $(\mu_n)_n$ *converge (en la métrica de Prohorov).*

Por la separabilidad del espacio X, la condición 3 es equivalente a **convergencia débil en medida** también llama **convergencia en distribución**, y esto significa que

$$\lim_{n\to\infty} \int_X f(x)\, d\mu_n = \int_X f(x)\, d\mu$$

para cada función $f\colon X \to \mathbb{R}$ continua y acotada. No estudiaremos en este texto este modo de convergencia, referimos al lector a [8] y [21] para un desarrollo detallado de este concepto.

La prueba del teorema anterior no es trivial. Primero hemos de mencionar que (1) ⇒ (2) ⇒ (3) es inmediato pues convergencia casi segura implica convergencia en probabilidad y esta a su vez implica convergencia en distribución.

[4] El enunciado del teorema 3.3.4 tiene una primera versión para una *sucesión de elementos aleatorios simétricos*, tanto el enunciando como la prueba de esta versión puede ser encontradas en [21, pág. 48, teo. 2.4].

La implicación (3) $\Rightarrow$ (2) requiere definir varios conceptos adicionales como el de medida de probabilidad *firme* o *tensa* así como el de *función característica* o *transformada de fourier* de una medida de probabilidad. La prueba de este teorema (particularmente de esta última implicación), así como las definiciones de lo antes mencionado se encuentran en [19] (ver teo. 3.1, pág. 37).

En todo caso solo nos interesa establecer la equivalencia entre (1) y (2).

Teorema 3.3.5 (Itô-Nisio [9])**.** *Sea $(V_n)_n$ una sucesión de elementos aleatorios independientes. Si $\sum_{n=1}^{\infty} V_n$ converge en probabilidad, entonces $\sum_{n=1}^{\infty} V_n$ converge casi seguramente.*

Demostración. Comencemos la prueba haciendo la siguiente observación: Si $(f_n)_n$ es una sucesión de funciones de Ω en X, y denotamos por $C = \{\omega : (f_n(\omega))_n \text{ converge}\}$, entonces

$$\begin{aligned} C &= \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n+1}^{\infty} \left\{\omega : \|f_n(\omega) - f_k(\omega)\| \leq \frac{1}{m}\right\} \\ &= \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n+1}^{\infty} \left\{\omega : \max_{n<j\leq k} \|f_n(\omega) - f_j(\omega)\| \leq \frac{1}{m}\right\}. \end{aligned}$$

Estas igualdades son inmediatas pues la sucesión converge es si es de Cauchy (recuerde que el espacio X es completo). Como consecuencia

$$\mathbb{P}(C) = \lim_{m\to\infty} \lim_{n\to\infty} \lim_{k\to\infty} \mathbb{P}\left\{\max_{n<j\leq k} \|f_n - f_j\| \leq \frac{1}{m}\right\}.$$

Esto último proporciona otra manera de mostrar la convergencia casi segura de la sucesión $(f_n)_n$ siempre y cuando el límite anterior sea igual a 1.

Sea $S_n = \sum_{j=1}^{n} V_n$, entonces como $(S_n)_n$ converge en probabilidad tenemos que para cada $\varepsilon > 0$, con $0 < \varepsilon < 1$, existe $M \in \mathbb{N}$ tal que si $n > m > M$:

$$\mathbb{P}\left(\|S_{m,n}\| > \varepsilon\right) < \varepsilon \tag{3.12}$$

donde

$$S_{m,n} = S_n - S_m = \sum_{j=m+1}^{n} V_j.$$

Para cada entero k con $m < k \leq n$ definamos los conjuntos medibles

$$\begin{aligned} A_k &= \left\{\max_{m<j\leq k-1} \|S_{m,j}\| \leq 2\varepsilon; \|S_{m,k}\| > 2\varepsilon; \|S_{k,n}\| \leq \varepsilon\right\} \\ A &= \{\|S_{m,n}\| > \varepsilon\}, \end{aligned}$$

entonces $(A_k)_{k=m+1}^{n}$ forma una colección de eventos ajenos tales que $\bigcup_{k=m+1}^{n} A_k \subset A$ y en consecuencia

$$\sum_{k=m+1}^{n} \mathbb{P}\left\{\max_{m<j\leq k-1} \|S_{m,j}\| \leq 2\varepsilon; \|S_{m,k}\| > 2\varepsilon; \|S_{k,n}\| \leq \varepsilon\right\} \leq \mathbb{P}\left\{\|S_{m,n}\| > \varepsilon\right\}.$$

Pero como por hipótesis $(V_n)_n$ es una sucesión de elementos aleatorios independientes, y $\max_{m<j\le k-1}\|S_{m,j}\|$ es un elemento aleatorio medible respecto a $\sigma(V_{m+1},\ldots,V_k)$ mientras que $\|S_{m,k}\|$ lo es respecto a $\sigma(V_{k+1},\ldots,V_n)$, entonces

$$\mathbb{P}\left\{\max_{m<j\le k-1}\|S_{m,j}\|\le 2\varepsilon; \|S_{m,k}\|>2\varepsilon; \|S_{k,n}\|\le\varepsilon\right\} =$$
$$\mathbb{P}\left\{\max_{m<j\le k-1}\|S_{m,j}\|\le 2\varepsilon; \|S_{m,k}\|>2\varepsilon\right\}\mathbb{P}\left\{\|S_{k,n}\|\le\varepsilon\right\}.$$

Ahora observemos que si

$$A'_k=\left\{\max_{m<j\le k-1}\|S_{m,j}\|\le 2\varepsilon; \|S_{m,k}\|>2\varepsilon\right\}$$

para $m<k\le n$, entonces $(A'_k)_{k=m+1}^n$ es también una sucesión de eventos disjuntos que cumple

$$\bigcup_{k=m+1}^n A'_k=\left\{\max_{m<k\le n}\|S_{m,j}\|>2\varepsilon\right\}$$

(recuerde la prueba de la desigualdad de Kolmogorov en la proposición 2.3.2). Por lo tanto,

$$\begin{aligned}
\mathbb{P}\left\{\|S_{m,n}\|>\varepsilon\right\} &\ge \sum_{k=m+1}^n \mathbb{P}\left(A'_k\right)\mathbb{P}\left\{\|S_{k,n}\|\le\varepsilon\right\}\\
&\ge \sum_{k=m+1}^n \mathbb{P}\left(A'_k\right)\min_{m<k\le n}\mathbb{P}\left\{\|S_{k,n}\|\le\varepsilon\right\}\\
&= \mathbb{P}\left(\bigcup_{k=m+1}^n A'_k\right)\min_{m<k\le n}\mathbb{P}\left\{\|S_{k,n}\|\le\varepsilon\right\}\\
&= \mathbb{P}\left\{\max_{m<k\le n}\|S_{m,j}\|>2\varepsilon\right\}\min_{m<k\le n}\mathbb{P}\left\{\|S_{k,n}\|\le\varepsilon\right\}.
\end{aligned}$$

Pero por (3.12) tenemos que

$$\min_{m<k\le n}\mathbb{P}\left\{\|S_{k,n}\|\le\varepsilon\right\}>1-\varepsilon$$

siempre $m>M$, y por lo tanto para $n>m>M$

$$\mathbb{P}\left\{\max_{m<k\le n}\|S_{m,j}\|>2\varepsilon\right\}\le\frac{1}{1-\varepsilon}\mathbb{P}\left\{\|S_{m,n}\|>\varepsilon\right\}<\frac{\varepsilon}{1-\varepsilon}. \tag{3.13}$$

Tomando los límites cuando n,m tienden a infinito y ε tendiendo a cero, y haciendo uso de la observación mencionada al principio de la demostración, obtenemos que el límite en (3.13) es igual a cero, por lo que concluimos que $(S_n)_n$ converge casi seguramente. ■

Para variables aleatorias el siguiente resultado es bien conocido.

Lema 3.3.6. *Sea* $0 < p < \infty$, *y sea* $V : \Omega \to \mathbb{R}$ *una variable aleatoria tal que* $\mathbb{P}(V \geq 0) = 1$, *entonces*

$$\mathbb{E}\left[X^p\right] = \int_0^\infty p \cdot t^{p-1} \mathbb{P}\left(V > t\right)\, dt.$$

Teorema 3.3.7 (Hoffmann-Jørgensen [15])**.** *Sea* $(V_n)_n$ *una sucesión de elementos aleatorios independientes en* X. *Sea* $S_n = \sum_{k=1}^n V_k$. *Supongamos que existe un elemento aleatorio* S *tal que* $S_n \xrightarrow{c.s} S$. *Sea* $0 < p < \infty$. *Los siguientes enunciados son equivalentes:*

1. $S_n \to S$ *en* $L_p(\Omega; X)$,
2. $\sup_n \|S_n\| \in L_p(\mathbb{R})$,
3. $\sup_n \|V_n\| \in L_p(\mathbb{R})$.

Demostración. Probaremos el teorema para el caso cuando los elementos aleatorios son además simétricos, la prueba del caso general pueden ser encontrada tanto en [3] como en [15].

Observése que por las desigualdades de Lévy (teorema 3.3.3) así como por el lema 3.3.6, (1) implica (2) y (3) (ilustraremos esto a continuación). Por otra parte, (2) implica a (1) debido al teorema de Convergencia Dominada, por lo que basta probar que (3) implica (2).

(3) $\Rightarrow$ (2) : Hagamos $M = \sup_n \|S_n\|$ y $N = \sup_n \|V_n\|$, y sean

$$R_n(t) = \mathbb{P}\left(\|S_n\| \geq t\right); \qquad R(t) = \mathbb{P}\left(M \geq t\right); \qquad Q(t) = \mathbb{P}\left(N \geq t\right).$$

Primero se probará que para cada $t, s \geq 0$ y $n \geq 1$,

$$R_n(2t+s) \leq Q(s) + 4R_n^2(t). \tag{3.14}$$

Sea τ el tiempo de paro definido por

$$\tau = \operatorname{ínf}\{n \geq 1 \,:\, \|S_n\| \geq t\}$$

donde $\operatorname{ínf}(\emptyset) = \infty$. Entonces $\|S_n\| \geq 2t + s$ implica que $\tau \leq n$, y por tanto

$$R_n(2t+s) = \sum_{j=1}^n \mathbb{P}\left(\|S_n\| \geq 2t+s, \tau = j\right).$$

Si $\tau = j$ y $\|S_n\| \geq 2t + s$, entonces $\|S_{j-1}\| < t$ y

$$\|S_n - S_j\| \geq \|S_n\| - \|S_{j-1}\| - \|X_j\| \geq (2t+s) - t - \|X_j\| \geq t + s - N.$$

Además, note que $\{\tau = j\}$ es un evento medible respecto a la σ-álgebra $\sigma(V_1, \ldots, V_j)$ mientras que $\{\|S_n - S_j\| \geq t\}$ lo es respecto a la σ-álgebra $\sigma(V_{j+1}, \ldots, V_n)$, y por lo tanto dichos eventos son independientes. Entonces,

$$\begin{aligned}
\mathbb{P}(\tau = j, \|S_n\| \geq 2t+s) &\leq \mathbb{P}(\tau = j, \|S_n - S_j\| \geq t + s - N) \\
&\leq \mathbb{P}(\tau = j, N \geq s) + \mathbb{P}(\tau = j, \|S_n - S_j\| \geq t) \\
&= \mathbb{P}(\tau = j, N \geq s) + +\mathbb{P}(\tau = j)\mathbb{P}(\|S_n - S_j\| \geq t).
\end{aligned}$$

Así que al sumar sobre todos los posibles valores de j obtenemos que

$$R_n(2t+s) \leq Q(s) + \sum_{j=1}^{n} \mathbb{P}(\tau = j)\mathbb{P}(\|S_n - S_j\| \geq t).$$

Ahora, sean $Y_1 = S_n - S_j$ y $Y_2 = S_j$; entonces Y_1 y Y_2 son elementos elementos aleatorios simétricos e independientes tales que $Y_1 + Y_2 = S_n$, así que por el teorema 3.3.3 tenemos que

$$\mathbb{P}(\|S_n - S_j\| \geq t) \leq 2\mathbb{P}(\|S_n\| \geq t) = 2R_n(t),$$

y como $\bigcup_{j=1}^{n}\{\tau = j\} = \{\text{máx}_{1\leq j\leq n} \|S_j\| \geq t\}$ (ver prueba de la desigualdad de Kolmogorov) entonces

$$\begin{aligned}
R_n(2t+s) &\leq Q(s) + 2R_n(t)\sum_{j=1}^{n} \mathbb{P}(\tau = j) \\
&= Q(s) + 2R_n(t)\mathbb{P}\left(\max_{1\leq j\leq n} \|S_j\| \geq t\right) \\
&\leq Q(s) + 4R_n^2(t),
\end{aligned}$$

donde la última desigualdad de obtiene de aplicar otra vez el teorema 3.3.3 (desigualdad de Lévy). Esto prueba 3.14. Observe que aplicando nuevamente el teorema 3.3.3 encontramos que $R_n(t) \leq R(t) \leq 2\sup_n R_n(t)$, luego

$$R(2t+s) \leq 2Q(s) + 8R^2(t) \quad \text{para cada}\, t, s \geq 0.$$

Ahora bien, queremos mostrar que $\mathbb{E}\left[M^p\right] < \infty$, pero debido al lema 3.3.6 esto es equivalente a probar que la integral indefinida $\int_0^\infty px^{p-1}R(x)\,dx$ es finita. Para ello: escojamos $t_0 > 0$ tal que $R(t_0) \leq \frac{1}{16\cdot 3^p}$. Si $A > 3t_0$, tenemos que

$$\begin{aligned}
\int_0^A px^{p-1}R(x)\,dx &= 3^p \int_0^{A/3} px^{p-1}R(3x)\,dx \\
&\leq 3^p \cdot 2\int_0^{A/3} px^{p-1}Q(x)\,dx + 3^p \cdot 8\int_0^{A/3} px^{p-1}R^2(t)\,dx \\
&\leq 2\cdot 3^p \cdot E\left[N^p\right] + 3^p \cdot 8\int_0^{t_0} px^{p-1}\,dx + 3^p \cdot 8\int_{t_0}^{A/3} px^{p-1}R(t_0)R(x)\,dx \\
&\leq C + \frac{1}{2}\int_0^A px^{p-1}R(x)\,dx
\end{aligned}$$

donde $C = 2\cdot 3^p \cdot E\left[N^p\right] + 8\cdot 3^p \cdot t_0^p < \infty$, y por lo tanto,

$$\int_0^A px^{p-1}R(x)\,dx \leq 2C \qquad \text{para cada}\, A > 3t_0.$$

Haciendo A tender a infinito encontramos que $M \in L_p(\mathbb{R})$. ■

3.3.3. Desigualdades de Hoffmann-Jørgensen

A continuación presentamos brevemente las ideas de lo que Hoffmann-Jørgensen denominó como **principio de comparación** (ver [15]). Comenzamos con $(V_n)_n$ una sucesión de elementos aleatorios independientes, y $(\xi_n)_n$ y $(\eta_n)_n$ dos sucesiones de variables aleatorias independientes. La idea es estudiar el comportamiento de la serie $\sum_{n=1}^{\infty} \eta_n X_n$ en términos del comportamiento de la serie $\sum_{n=1}^{\infty} \xi_n X_n$. Primero consideramos el caso cuando η_n no es variable aleatoria y $\xi_n \equiv 1$.

Lema 3.3.8 (Hoffmann-Jørgensen)**.** *Sean $V_1, V_2, \ldots, V_n$ elementos aleatorios independientes, p-integrables con $1 \leq p < \infty$. Si $\mathbb{E}[V_j] = \mathbf{0}$ para cada $j = 1, 2, \ldots, n$, entonces*

$$\left(\mathbb{E} \left\| \sum_{j=1}^{n} a_j V_j \right\|^p \right)^{\frac{1}{p}} \leq 2 \max_{1 \leq j \leq n} |a_j| \left(\mathbb{E} \left\| \sum_{j=1}^{n} V_j \right\|^p \right)^{\frac{1}{p}} \tag{3.15}$$

para cada $a_1, a_2, \ldots, a_n \in \mathbb{R}$.

Si V_j es simétrico para cada $j = 1, 2, \ldots, n$, entonces

$$\left(\mathbb{E} \left\| \sum_{j=1}^{n} a_j V_j \right\|^p \right)^{\frac{1}{p}} \leq \max_{1 \leq j \leq n} |a_j| \left(\mathbb{E} \left\| \sum_{j=1}^{n} V_j \right\|^p \right)^{\frac{1}{p}} \tag{3.16}$$

para cada $a_1, a_2, \ldots, a_n \in \mathbb{R}$.

Demostración. Sea $\sigma \subset \{1, 2, \ldots, n\}$, entonces por el lema 3.3.1 tenemos que

$$\mathbb{E} \left\| \sum_{j \in \sigma} V_j \right\|^p \leq \mathbb{E} \left\| \sum_{j=1}^{n} V_j \right\|^p.$$

Primero supongamos que $a_j = \pm 1$ para cada j, y definamos $\sigma_+ = \{j : a_j = +1\}$ y $\sigma_- = \{j : a_j = -1\}$; entonces

$$\begin{aligned}
\left(\mathbb{E} \left\| \sum_{j=1}^{n} a_j V_j \right\|^p \right)^{\frac{1}{p}} &= \left(\mathbb{E} \left\| \sum_{j \in \sigma_+} a_j V_j + \sum_{j \in \sigma_-} a_j V_j \right\|^p \right)^{\frac{1}{p}} \\
&\leq \left(\mathbb{E} \left\| \sum_{j \in \sigma_+} a_j V_j \right\|^p \right)^{\frac{1}{p}} + \left(\mathbb{E} \left\| \sum_{j \in \sigma_-} a_j V_j \right\|^p \right)^{\frac{1}{p}} \\
&= \left(\mathbb{E} \left\| \sum_{j \in \sigma_+} V_j \right\|^p \right)^{\frac{1}{p}} + \left(\mathbb{E} \left\| \sum_{j \in \sigma_-} V_j \right\|^p \right)^{\frac{1}{p}} \\
&\leq 2 \left(\mathbb{E} \left\| \sum_{j=1}^{n} V_j \right\|^p \right)^{\frac{1}{p}}.
\end{aligned}$$

Esto prueba (3.15) para este primer caso.

Ahora, sea $K_n = \{(a_1, \ldots, a_n) \in \mathbb{R}^n : |a_1| \leq 1, \ldots, |a_n| \leq 1\}$, entonces K_n es un subconjunto compacto y convexo de $\mathbb{R}^n$, y cuyos puntos extremos son los 2^n puntos de la forma $(\pm 1, \pm 1, \ldots, \pm 1)$. Entonces por el teorema de Carathéodory (ver por ejemplo teorema 6.28 en [10]) dado $a = (a_1, \ldots, a_n) \in K_n$, a se puede representar como una combinación convexa de a lo más $n+1$ puntos extremos. Es decir, podemos encontrar $\lambda_m \geq 0$ con $m = 1, \ldots, n+1$ tales que

$$a_j = \sum_{m=1}^{n+1} \lambda_m a_{jm}, \quad \sum_{m=1}^{n+1} \lambda_m = 1,$$

donde $a_{jm} = \pm 1$. Entonces encontramos que

$$\begin{aligned}
\left(\mathbb{E} \left\| \sum_{j=1}^{n} a_j V_j \right\|^p \right)^{\frac{1}{p}} &= \left(\mathbb{E} \left\| \sum_{j=1}^{n} \left(\sum_{m=1}^{n+1} \lambda_m a_{jm} \right) V_j \right\|^p \right)^{\frac{1}{p}} \\
&= \left(\mathbb{E} \left\| \sum_{m=1}^{n+1} \lambda_m \sum_{j=1}^{n} a_{jm} V_j \right\|^p \right)^{\frac{1}{p}} \\
&\leq \sum_{m=1}^{n+1} \left(\mathbb{E} \left\| \lambda_m \sum_{j=1}^{n} a_{jm} V_j \right\|^p \right)^{\frac{1}{p}} \qquad \text{(desigualdad de Minkowski)} \\
&= \sum_{m=1}^{n+1} \lambda_m \left(\mathbb{E} \left\| \sum_{j=1}^{n} a_{jm} V_j \right\|^p \right)^{\frac{1}{p}} \\
&\leq \sum_{m=1}^{n+1} \lambda_m \left\{ 2 \left(\mathbb{E} \left\| \sum_{j=1}^{n} V_j \right\|^p \right)^{\frac{1}{p}} \right\} \qquad \text{(por el caso 1)} \\
&= 2 \left(\mathbb{E} \left\| \sum_{j=1}^{n} V_j \right\|^p \right)^{\frac{1}{p}}.
\end{aligned}$$

Por lo tanto, si escogemos a los escalares $a'_i = \dfrac{a_i}{\max\limits_{1 \leq j \leq n} |a_j|}$ obtenemos el resultado, pues $|a'_i| \leq 1$ para cada i y por el caso anterior:

$$\frac{1}{\max\limits_{1 \leq j \leq n} |a_j|} \left(\mathbb{E} \left\| \sum_{j=1}^{n} a_j V_j \right\|^p \right)^{\frac{1}{p}} \leq 2 \left(\mathbb{E} \left\| \sum_{j=1}^{n} V_j \right\|^p \right)^{\frac{1}{p}}.$$

La prueba de (3.16) es similar. Como V_j es simétrico para cada $j = 1, 2, \ldots, n$, entonces

$$\mathbb{E} \left\| \sum_{j=1}^{n} a_j V_j \right\|^p = \mathbb{E} \left\| \sum_{j=1}^{n} V_j \right\|^p$$

siempre que $a_j = \pm 1$. La última parte de (3.16) se concluye como en (3.15). ■

Corolario 3.3.9 (Hoffmann-Jørgensen)**.** *Sean $V_1, V_2, \ldots, V_n$ elementos aleatorios en independientes, p-integrables con $1 \leq p < \infty$ y $\mathbb{E}[V_j] = \mathbf{0}$ para cada $j = 1, 2, \ldots, n$. Si $\eta_1, \eta_2, \ldots, \eta_n$ son variables aleatorias p-integrables, tales que $(V_1, V_2, \ldots, V_n)$ y $(\eta_1, \eta_2, \ldots, \eta_n)$ son independientes, entonces*

$$2^{-p}\mathbb{E}\left[\min_{1\leq j\leq n} |\eta_j|^p\right] \mathbb{E}\left\|\sum_{j=1}^{n} V_j\right\|^p \leq \mathbb{E}\left\|\sum_{j=1}^{n} \eta_j V_j\right\|^p \leq 2^p \mathbb{E}\left[\max_{1\leq j\leq n} |\eta_j|^p\right] \mathbb{E}\left\|\sum_{j=1}^{n} V_j\right\|^p.$$

Demostración. Por el lema 3.3.8:

$$\mathbb{E}\left\|\sum_{j=1}^{n} a_j V_j\right\|^p \leq 2^p \max_{1\leq j\leq n} |a_j|^p \mathbb{E}\left\|\sum_{j=1}^{n} V_j\right\|^p$$

para cada $a_1, \ldots, a_n \in \mathbb{R}$. Y si ahora volvemos a aplicar el lema 3.3.8 a la los elementos aleatorios $W_j = a_j V_j$ y a los escalares $b_j = \frac{1}{a_j}$ para cada $j = 1, \ldots, n$ obtenemos:

$$\mathbb{E}\left\|\sum_{j=1}^{n} b_j W_j\right\|^p \leq 2^p \max_{1\leq j\leq n} |b_j|^p \mathbb{E}\left\|\sum_{j=1}^{n} W_j\right\|^p$$

o equivalentemente

$$\mathbb{E}\left\|\sum_{j=1}^{n} V_j\right\|^p \leq 2^p \max_{1\leq j\leq n} |b_j|^p \, \mathbb{E}\left\|\sum_{j=1}^{n} a_j V_j\right\|^p,$$

pero como $\max\limits_{1\leq j\leq n} |b_j|^p = \dfrac{1}{\min\limits_{1\leq j\leq n} \left|b_j^{-1}\right|^p}$ y $b_j^{-1} = a_j$, entonces

$$2^{-p} \min_{1\leq j\leq n} |a_j|^p \, \mathbb{E}\left\|\sum_{j=1}^{n} V_j\right\|^p \leq \mathbb{E}\left\|\sum_{j=1}^{n} a_j V_j\right\|^p.$$

En resumen, probamos que

$$2^{-p} \min_{1\leq j\leq n} |a_j|^p \, \mathbb{E}\left\|\sum_{j=1}^{n} V_j\right\|^p \leq \mathbb{E}\left\|\sum_{j=1}^{n} a_j V_j\right\|^p \leq 2^p \max_{1\leq j\leq n} |a_j|^p \mathbb{E}\left\|\sum_{j=1}^{n} V_j\right\|^p$$

para cada sucesión de escalares $a_1, \ldots, a_n$.

En particular, si tomamos a las variables aleatorias $\eta_1, \ldots, \eta_n$ y calculamos la esperanza respecto a la distribución conjunta de $(\eta_1, \ldots, \eta_n)$ en la desigualdad anterior, entonces por la hipótesis de independencia de la familia $\{\eta_j\}_{j=1}^n$ y entre los vectores $(\eta_1, \eta_2, \ldots, \eta_n)$ y $(V_1, V_2, \ldots, V_n)$ tenemos que

$$2^{-p}\mathbb{E}\left[\min_{1\leq j\leq n} |\eta_j|^p\right]\mathbb{E}\left\|\sum_{j=1}^n V_j\right\|^p \leq \mathbb{E}\left\|\sum_{j=1}^n \eta_j V_j\right\|^p \leq 2^p\mathbb{E}\left[\max_{1\leq j\leq n} |\eta_j|^p\right]\mathbb{E}\left\|\sum_{j=1}^n V_j\right\|^p.$$

Con esto último tratamos de decir que como

$$\mathbb{E}\left\|\sum_{j=1}^n V_j\right\|^p = \int_X \left\|\sum_{j=1}^n V_j\right\|^p d\mathbb{P}_{(V_1,\ldots,V_n)} \tag{3.17}$$

y $(V_1, \ldots, V_n)$ es independiente de $(\eta_1, \eta_2, \ldots, \eta_n)$, cuando integramos respecto a la distribución conjunta de $(\eta_1, \eta_2, \ldots, \eta_n)$, la expresión en (3.17) sale de esta integral. Este mismo proceso se repite en los otros términos en desigualdad anterior. ■

Teorema 3.3.10 (Hoffmann-Jørgensen). *Sean $(\eta_j)_j$ y $(\xi_j)_j$ dos sucesiones de variables aleatorias independientes, p-integrables con $1 \leq p < \infty$. Sea $(V_j)_j$ una sucesión de elementos aleatorios independientes y p-integrables tales que*

$$N = \sup_j |\eta_j| \in L_p(\Omega; \mathbb{R}), \tag{3.18}$$

$$a = \inf_j \mathbb{E}|\xi_j| > 0, \tag{3.19}$$

$$(V_j)_j \text{ y } (\xi_j)_j \text{ son independientes, y } (V_j)_j \text{ y } (\eta_j)_j \text{ son independientes,} \tag{3.20}$$

$$\mathbb{E}\left[\eta_j V_j\right] = \mathbb{E}\left[\xi_j V_j\right] = 0 \text{ para cada } j \geq 1. \tag{3.21}$$

Entonces

$$\mathbb{E}\left\|\sum_{j=1}^n \eta_j V_j\right\|^p \leq \left(\frac{8}{a}\right)^p \mathbb{E}\left[N^p\right]\mathbb{E}\left\|\sum_{j=1}^n \xi_j V_j\right\|^p$$

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

Demostración. Sea $(\varepsilon_j)_j$ una sucesión de Rademacher que es independiente de las tres sucesiones antes mencionadas, $((V_j)_j, (\xi_j)_j, (\eta_j)_j)$. Por el corolario 3.3.9 (primera desigualdad), aplicado a $U_j = \eta_j V_j$ y $\nu_j = \varepsilon_j$ tenemos

$$2^{-p}\mathbb{E}\left[\min_{1\leq j\leq n} |\nu_j|^p\right]\mathbb{E}\left\|\sum_{j=1}^n U_j\right\|^p \leq \mathbb{E}\left\|\sum_{j=1}^n \nu_j U_j\right\|^p,$$

es decir,

$$\mathbb{E}\left\|\sum_{j=1}^n \eta_j V_j\right\|^p \leq 2^p\mathbb{E}\left\|\sum_{j=1}^n \eta_j \varepsilon_j V_j\right\|^p.$$

Pero como $\varepsilon_j V_j$ es simétrica (pues ε_j de Rademacher) y (η_j) es independiente de $(\varepsilon_j V_j)_j$ (pues ya lo era de tanto de $(\varepsilon_j)_j$ como de $(X_j)_j$), encontramos que por el corolario 3.3.9 (segunda desigualdad):

$$\begin{aligned} 2^p \mathbb{E} \left\| \sum_{j=1}^{n} \eta_j \varepsilon_j V_j \right\|^p &\leq 2^p \left(2^p \mathbb{E} \left[\max_{1 \leq j \leq n} |\eta_j|^p \right] \mathbb{E} \left\| \sum_{j=1}^{n} \varepsilon_j V_j \right\|^p \right) \\ &\leq 4^p \mathbb{E} \left[\sup_j |\eta_j|^p \right] \mathbb{E} \left\| \sum_{j=1}^{n} \varepsilon_j V_j \right\|^p \\ &= 4^p \mathbb{E} \left[N^p \right] \mathbb{E} \left\| \sum_{j=1}^{n} \varepsilon_j V_j \right\|^p . \end{aligned}$$

Tomemos los escalares $a_j = \mathbb{E}|\xi_j| > 0$ y denotemos por $b_0 = \dfrac{1}{\min\limits_{1 \leq j \leq n} a_j^p}$, entonces

$$\begin{aligned} 2^p \mathbb{E} \left\| \sum_{j=1}^{n} \eta_j \varepsilon_j V_j \right\|^p &\leq 4^p \mathbb{E} \left[N^p \right] b_0 b_0^{-1} \mathbb{E} \left\| \sum_{j=1}^{n} \varepsilon_j V_j \right\|^p \\ &\leq 4^p \mathbb{E} \left[N^p \right] \frac{1}{\min\limits_{1 \leq j \leq n} a_j^p} \mathbb{E} \left\| \sum_{j=1}^{n} a_j \varepsilon_j V_j \right\|^p \\ &\leq \left(\frac{4}{a} \right)^p \mathbb{E} \left[N^p \right] \mathbb{E} \left\| \sum_{j=1}^{n} a_j \varepsilon_j V_j \right\|^p . \end{aligned}$$

Sea $\mathcal{G}$ la σ-álgebra generada por $(\varepsilon_1, \ldots, \varepsilon_n, V_1, \ldots, V_n)$, y definamos

$$\varepsilon_j^* = \begin{cases} -\varepsilon_j & \text{si} \quad \xi_j < 0, \\ \varepsilon_j & \text{si} \quad \xi_j \geq 0. \end{cases}$$

Utilizando algunas propiedades de la esperanza condicional enlistadas en el teorema 1.3.7 tenemos que:

$$\begin{aligned} \left\| \sum_{j=1}^{n} a_j \varepsilon_j V_j \right\|^p = \left\| \sum_{j=1}^{n} \mathbb{E} \left(|\xi_j| \right) \varepsilon_j V_j \right\|^p &= \left\| \mathbb{E} \left(\sum_{j=1}^{n} |\xi_j| \varepsilon_j V_j \,\middle|\, \mathcal{G} \right) \right\|^p \\ &\leq \mathbb{E} \left(\left\| \sum_{j=1}^{n} |\xi_j| \varepsilon_j V_j \right\|^p \,\middle|\, \mathcal{G} \right) \quad \text{(desigualdad de Jensen)} \\ &= \mathbb{E} \left(\left\| \sum_{j=1}^{n} \xi_j \varepsilon_j^* V_j \right\|^p \,\middle|\, \mathcal{G} \right) \end{aligned}$$

y por lo tanto

$$\mathbb{E}\left\|\sum_{j=1}^{n}\eta_j V_j\right\|^p \leq \left(\frac{4}{a}\right)^p \mathbb{E}\left[N^p\right]\mathbb{E}\left\|\sum_{j=1}^{n}\xi_j\varepsilon_j^* V_j\right\|^p.$$

Se sigue de la definición de ε_j^* que la sucesión $(\varepsilon_j^*)_j$ es independiente tanto de $(\xi_j)_j$ como de $(X_j)_j$, y nuevamente por el corolario 3.3.9:

$$\mathbb{E}\left\|\sum_{j=1}^{n}\varepsilon_j^*\xi_j V_j\right\|^p \leq 2^p\mathbb{E}\left[\max_{1\leq j\leq n}|\varepsilon_j^*|\right]\mathbb{E}\left\|\sum_{j=1}^{n}\xi_j V_j\right\|^p$$

y por lo tanto

$$\mathbb{E}\left\|\sum_{j=1}^{n}\eta_j V_j\right\|^p \leq \left(\frac{8}{a}\right)^p \mathbb{E}\left[N^p\right]\mathbb{E}\left\|\sum_{j=1}^{n}\xi_j V_j\right\|^p.$$

■

3.3.4. Tipo p de un espacio de Banach

En esta sección introduciremos la definición del *tipo* de un espacio de Banach, aunque hay varias formas de presentarla (dependiendo del enfoque a estudiar). En este texto seguimos la dada en el artículo [17] de J. Hoffmann-Jørgensen y G. Pisier.

Sea $(\varepsilon_n)_n$ una sucesión de Rademacher y denotemos por $X^\infty = \Pi_{n=1}^\infty X$, definamos

$$C(X) = \left\{(x_n)_n \in X^\infty \;:\; \sum_{n=1}^{\infty}\varepsilon_n x_n \text{ converge en probabilidad}\right\}.$$

Estos espacios fueron introducidos y estudiados por J. Hoffmann-Jørgensen en [15] y también en [16]. Debe observarse que $(\varepsilon_n x_n)_n$ es una sucesión de elementos aleatorios independientes y simétricos. Queremos mostrar que podemos cambiar «*convergencia en probabilidad*» por «*convergencia en* $L_p(\Omega; X)$» en la definición de $C(X)$. Con esta finalidad presentamos lo siguientes dos resultados:

Lema 3.3.11. *Si la serie* $S = \sum_{n=1}^\infty \varepsilon_n x_n$ *converge casi seguramente y satisface que*

$$\mathbb{P}\left(\|S\| > r\right) < \frac{\alpha}{2} \tag{3.22}$$

para alguna $r > 0$ *y* $\alpha > 0$, *entonces*

$$\mathbb{P}\left(\|S\| > 2r\right) < \alpha^2. \tag{3.23}$$

Demostración. Como por hipótesis $\sum_{n=1}^{\infty} \varepsilon_n x_n$ converge c.s. fijemos $\Omega_0 \in \mathcal{F}$ el conjunto de probabilidad 1, $\mathbb{P}(\Omega_0) = 1$, donde la serie anterior converge. Si $m \in \mathbb{N}$, sea $S_m = \sum_{n=1}^{m} \varepsilon_n x_n$ y $\lim_{m\to\infty} S_m = \sum_{n=1}^{\infty} \varepsilon_n x_n = S$. Dado $r > 0$, consideramos los subconjuntos de Ω_0

$$A = \left\{ \sup_m \|S_m\| > r \right\}, \quad B = \{\|S\| > r\}, \quad C = \{\|S\| > 2r\}.$$

Como lo hemos estado haciendo anteriormente, particionamos A en una cantitad numberable de eventos (disjuntos)

$$A_m = \{\|S_1\| \leq r, \ldots, \|S_{m-1}\| \leq r, \|S_m\| > r\}$$

para cada $m \in \mathbb{N}$, es decir, $\biguplus_{m=1}^{\infty} A_m = A$. También definimos

$$C_m = \{\|S - S_{m-1}\| > r\} = \left\{ \left\| \sum_{n=1}^{\infty} \varepsilon_n x_n - \sum_{n=1}^{m-1} \varepsilon_n x_n \right\| > r \right\} = \left\{ \left\| \sum_{n=m}^{\infty} \varepsilon_n x_n \right\| > r \right\}.$$

Note que si $\omega \in A_m \cap C$, entonces

$$\|S - S_{m-1}\| \geq \|S\| - \|S_{m-1}\| > 2r - r = r$$

por lo que $A_m \cap C \subset C_m$. Pero además, como $C \subset B \subset A$ y A es unión disjunta de los A_m, entonces

$$\mathbb{P}(C) = \mathbb{P}(C \cap A) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} (A_m \cap C) \right) = \sum_{m=1}^{\infty} \mathbb{P}(A_m \cap C) \leq \sum_{m=1}^{\infty} \mathbb{P}(A_m \cap C_m).$$

Lo siguiente que haremos será mostrar que A_m y C_m son eventos independientes. Para ello notemos que

(1) A_m solo depende de $\varepsilon_1, \ldots, \varepsilon_m$;

(2) C_m solo depende de $\varepsilon_m \varepsilon_{m+1}, \varepsilon_m \varepsilon_{m+2}, \ldots$;

(3) Las variables aleatorias $\varepsilon_1, \ldots, \varepsilon_m, \varepsilon_m \varepsilon_{m+1}, \varepsilon_m \varepsilon_{m+2}, \ldots$ son independientes.

A continuación mostraremos (2): como $\mathbb{P}(\varepsilon_m^2 = 1) = 1$, y

$$|\varepsilon_m| \left\| \sum_{n=m}^{\infty} \varepsilon_n x_n \right\| = \left\| \varepsilon_m \cdot \sum_{n=m}^{\infty} \varepsilon_n x_n \right\| = \left\| x_m + \sum_{n=m+1}^{\infty} \varepsilon_m \varepsilon_n x_n \right\|$$

tenemos que

$$C_m = \left\{ \left\| x_m + \sum_{n=m+1}^{\infty} \varepsilon_m \varepsilon_n x_n \right\| > r \right\}$$

solo depende de las variables $\varepsilon_m \varepsilon_{m+1}, \varepsilon_m \varepsilon_{m+2}, \ldots$.

Para (3) consideremos $\delta_n = \pm 1$ para cada $n \in \mathbb{N}$. Usando la independencia de la sucesión $(\varepsilon_n)_n$ así como la propiedad $\delta_n^2 = 1$ tenemos que

$$\begin{aligned}
&\mathbb{P}(\varepsilon_1 = \delta_1, \ldots, \varepsilon_m = \delta_m, \varepsilon_m\varepsilon_{m+1} = \delta_{m+1}, \ldots, \varepsilon_m\varepsilon_k = \delta_k) \\
&= \mathbb{P}(\varepsilon_1 = \delta_1, \ldots, \varepsilon_m = \delta_m, \varepsilon_{m+1} = \delta_m\delta_{m+1}, \ldots, \varepsilon_k = \delta_m\delta_k) \\
&= \mathbb{P}(\varepsilon_1 = \delta_1) \cdot \ \ldots \ \cdot \mathbb{P}(\varepsilon_m = \delta_m) \cdot \mathbb{P}(\varepsilon_{m+1} = \delta_m\delta_{m+1}) \cdot \ \ldots \ \cdot \mathbb{P}(\varepsilon_k = \delta_m\delta_k) \\
&= \frac{1}{2^k}
\end{aligned}$$

Por otra parte, si $m + 1 \leq j \leq k$:

$$\begin{aligned}
&\mathbb{P}(\varepsilon_m\varepsilon_j = \delta_j) \\
&= \mathbb{P}(\varepsilon_m\varepsilon_j = \delta_j \,|\, \varepsilon_m = +1)\mathbb{P}(\varepsilon_m = +1) + \mathbb{P}(\varepsilon_m\varepsilon_j = \delta_j \,|\, \varepsilon_m = -1)\mathbb{P}(\varepsilon_m = -1) \\
&= \mathbb{P}(\varepsilon_j = \delta_j)\frac{1}{2} + \mathbb{P}(\varepsilon_j = -\delta_j)\frac{1}{2} \\
&= \frac{1}{2}
\end{aligned}$$

y entonces

$$\mathbb{P}(\varepsilon_1 = \delta_1) \cdot \ \ldots \ \cdot \mathbb{P}(\varepsilon_m = \delta_m) \cdot \mathbb{P}(\varepsilon_m\varepsilon_{m+1} = \delta_{m+1}) \cdot \ \ldots \ \cdot \mathbb{P}(\varepsilon_m\varepsilon_k = \delta_k) = \frac{1}{2^k}.$$

Esto es,

$$\begin{aligned}
&\mathbb{P}(\varepsilon_1 = \delta_1, \ldots, \varepsilon_m = \delta_m, \varepsilon_m\varepsilon_{m+1} = \delta_{m+1}, \ldots, \varepsilon_m\varepsilon_k = \delta_k) \\
&= \mathbb{P}(\varepsilon_1 = \delta_1) \cdot \ \ldots \ \cdot \mathbb{P}(\varepsilon_m = \delta_m) \cdot \mathbb{P}(\varepsilon_m\varepsilon_{m+1} = \delta_{m+1}) \cdot \ \ldots \ \cdot \mathbb{P}(\varepsilon_m\varepsilon_k = \delta_k)
\end{aligned}$$

para toda $k \in \mathbb{N}$. Por lo tanto, la sucesión de variables aleatorias

$$\varepsilon_1, \ldots, \varepsilon_m, \varepsilon_m\varepsilon_{m+1}, \varepsilon_m\varepsilon_{m+2}, \ldots$$

es independiente. Así, como $A_m \in \sigma(\varepsilon_1, \ldots, \varepsilon_m)$ y $C_m \in \sigma(\varepsilon_m\varepsilon_{m+1}, \varepsilon_m\varepsilon_{m+1}, \ldots)$ entonces A_m y C_m son independientes, luego

$$\mathbb{P}(A_m \cap C_m) = \mathbb{P}(A_m)\mathbb{P}(C_m).$$

En resumen,

$$\begin{aligned}
\mathbb{P}(C) &\leq \sum_{m=1}^{\infty} \mathbb{P}(A_m \cap C_m) = \sum_{m=1}^{\infty} \mathbb{P}(A_m)\mathbb{P}(C_m) \leq \sum_{m=1}^{\infty} \mathbb{P}(A_m) \sup_m \mathbb{P}(C_m) \\
&= \mathbb{P}(A) \sup_m \mathbb{P}(C_m),
\end{aligned}$$

así que si probamos que $\mathbb{P}(C_m) \leq 2\mathbb{P}(B)$ para cada m ya habremos acabado. En efecto, de la desigualdad de Lévy (3.11) en el teorema 3.3.3 obtenemos que $\mathbb{P}(A) \leq 2\mathbb{P}(B) < 2\dfrac{\alpha}{2}$ y en dado caso que $\mathbb{P}(C_m) \leq 2\mathbb{P}(B)$:

$$\mathbb{P}(C) \leq 2\mathbb{P}(A)\mathbb{P}(B) < 2\alpha \cdot \frac{\alpha}{2} = \alpha^2.$$

Con este propósito introducimos los siguientes conjuntos:

$$C_{m,\delta}^{+} = \left(\bigcap_{j<m} \{\varepsilon_j = \delta_j\} \right) \cap C_m \cap \{\|S\| > r\}$$

$$C_{m,\delta}^{-} = \left(\bigcap_{j<m} \{\varepsilon_j = \delta_j\} \right) \cap C_m \cap \{\|S - 2S_{m-1}\| > r\}.$$

Afirmamos que estos dos conjuntos tienen la misma probabilidad. En efecto, ya que S y $S - 2S_{m-1}$ tienen la misma distribución, los elementos $U = (\varepsilon_1 x_1, \ldots, \varepsilon_{m-1} x_{m-1}, S)$ y $V = (\varepsilon_1 x_1, \ldots, \varepsilon_{m-1} x_{m-1}, S - 2S_{m-1})$ también tienen la misma distribución. Así, si consideramos E_m el subconjunto de X^n (y $n \geq m$) cuyos elementos son de la forma $(t_1 x_1, t_2 x_2, \ldots, t_n x_n)$ donde $t_j \in \mathbb{R}$ y $x_j \in X$), tales que:

1. $t_j = \delta_j$ para cada $1 \leq j \leq m-1$,
2. $\left\| \sum_{j=m}^{n} t_j x_j \right\| > r$,
3. $\left\| \sum_{j=1}^{k-1} t_j x_j - \sum_{j=m}^{n} t_j x_j \right\| > r$.

Entonces,

$$\begin{aligned}
\mathbb{P}\left(C_{m,\delta}^{+}\right) &= \mathbb{P}\left(U^{-1}(E_m)\right) \\
&= \mathbb{P}_U\left(E_m\right) \\
&= \mathbb{P}_V\left(E_m\right) \\
&= \mathbb{P}\left(V^{-1}(E_m)\right) = \mathbb{P}\left(C_{m,\delta}^{-}\right).
\end{aligned}$$

Ahora, observemos que podemos escribir

$$C_{m,\delta}^{+} = \left(\bigcap_{j<m} \{\varepsilon_j = \delta_j\} \right) \cap C_m \cap B,$$

además, si $\omega \in C_m$, entonces $\|S(\omega) - S_{m-1}(\omega)\| > r$, lo cual implica que

$$\begin{aligned}
2r &< 2\|S(\omega) - S_{m-1}(\omega)\| \\
&= \|S(\omega) + (S(\omega) - 2S_{m-1}(\omega))\| \\
&\leq \|S(\omega)\| + \|S(\omega) - 2S_{m-1}(\omega)\|
\end{aligned}$$

por lo que $\|S(\omega)\| > r$ o bien $\|S(\omega) - 2S_{m-1}(\omega)\| > r$, y entonces

$$C_m \subset \{\|S\| > r\} \cup \{\|S - 2S_{m-1}\| > r\}.$$

Por lo tanto

$$C_m \cap \left(\bigcap_{j<m} \{\varepsilon_j = \delta_j\} \right) = C^+_{m,\delta} \cup C^-_{m,\delta}.$$

Finalmente,

$$\begin{aligned}
\mathbb{P}(C_m) &= \sum_{\delta_1,\ldots,\delta_{m-1}=\pm 1} \mathbb{P}\left(C_m \cap \bigcap_{j<m} \{\varepsilon_j = \delta_j\} \right) \\
&= \sum_{\delta_1,\ldots,\delta_{m-1}=\pm 1} \mathbb{P}\left(C^+_{m,\delta} \cup C^-_{m,\delta} \right) \\
&\leq 2 \cdot \sum_{\delta_1,\ldots,\delta_{m-1}=\pm 1} \mathbb{P}\left(C^+_{m,\delta} \right) \\
&\leq 2 \cdot \sum_{\delta_1,\ldots,\delta_{m-1}=\pm 1} \mathbb{P}\left(B \cap \bigcap_{j<m} \{\varepsilon_j = \delta_j\} \right) \\
&\leq 2\mathbb{P}(B).
\end{aligned}$$

Así que $\sup_m \mathbb{P}(C_m) \leq 2\mathbb{P}(B)$ por lo que la desigualdad (3.23) queda demostrada. ■

Siguiendo la notación del lema anterior enunciamos el siguiente resultado.

Lema 3.3.12. *Si la serie $\sum_{n=1}^{\infty} \varepsilon_n x_n$ converge casi seguramente, entonces $\|S\|$ pertenece a $L_p(\Omega; \mathbb{R})$ para cada $1 \leq p < \infty$.*

Demostración. Nuevamente, ya que la serie $\sum_{n=1}^{\infty} \varepsilon_n x_n$ converge c.s. en lo sucesivo consideraremos que estamos trabajando en un conjunto $\Omega_0 \in \mathcal{F}$ de probabilidad 1 donde la serie antes mencionada converge.

Ahora, sea $\varphi(t) \colon [0, \infty) \to [0, \infty)$ una función continua y creciente. Notemos que $\|S\|$ es una variable aleatoria y así también lo es $\varphi(\|S\|)$ (pues φ es continua y por tanto medible). Entonces su esperanza está dada por la integral de Riemann-Stieltjes:

$$\mathbb{E}\left[\varphi(Z)\right] = \int_0^\infty \varphi(t)\, dF_Z(t)$$

donde $Z = \|S\|$ y $F_Z(t) = \mathbb{P}\left(Z \leq t\right)$ es la función de distribución de Z, así que

$$\mathbb{E}\left[\varphi(\|S\|)\right] = -\int_0^\infty \varphi(t)\, dp(t) \tag{3.24}$$

donde $p(t) = \mathbb{P}(\|S\| > t)$. Entonces escojamos $r > 0$ tal que $p(t) = \mathbb{P}(\|S\| > r) < \frac{\alpha}{2}$ donde $\alpha < \frac{1}{2}$, entonces por el lema anterior aplicado recursivamente encontramos que

$$p(r) < \frac{1}{4} 2\alpha,\ p(2r) < \frac{1}{4}(2\alpha)^2, \ldots,\ p(2^n r) < \frac{1}{4}(2\alpha)^{2n}.$$

Luego
$$-\int_{2^n r}^{2^{n+1}r} \varphi(t)\, dp(t) < \frac{1}{4}(2\alpha)^{2^n}\varphi(2^{n+1}r).$$
Por lo que la integral en (3.24) converge si la serie
$$\sum_{n=1}^{\infty}(2\alpha)^{2^n}\varphi(2^{n+1}r) < \infty,$$
en particular esto ocurre cuando $\varphi(t) = t^p$, para cualquier $1 \leq p < \infty$. En efecto, ya que $\alpha < \frac{1}{2}$ entonces $\beta = 2\alpha < 1$, luego la serie
$$\sum_{n=1}^{\infty}\beta^{2^n}\left(2^{n+1}r\right)^p = 2^p r^p \sum_{n=1}^{\infty}\beta^{2^n}2^{np},$$
converge debido a la prueba del cociente. ■

Denotemos por $S_n = \sum_{j=1}^{n}\varepsilon_j x_j$, entonces $S_n \overset{c.s.}{\to} S$, y por el lema anterior sabemos que $\|S\| \in L_p(\mathbb{R})$. Sea $M = \sup_n \|S_n\|$, usando la desigualdad 3.11 (de Lévy) tenemos que $\mathbb{P}(M \geq t) \leq 2\mathbb{P}(\|S\| \geq t)$, luego
$$\int_0^A px^{p-1}\mathbb{P}(M > t)\, dt \leq 2\int_0^A px^{p-1}\mathbb{P}(\|S\| > t) \leq 2\mathbb{E}\left[\|S\|^p\right] < \infty$$
para cada $A > 0$, y por lo tanto $M = \sup_n \|S_n\| \in L_p(\mathbb{R})$. De acuerdo al teorema 3.3.7 esto es equivalente a la convergencia de la serie $S = \sum_{n=1}^{\infty}\varepsilon_n x_n$ en $L_p(\Omega; X)$.

Siendo este el caso podemos tomar $\|\cdot\|_p$ la norma usual en $L_p(\Omega; X)$ y definir una norma (para $p \in \mathbb{R}_+$) en $C(X)$ dada por
$$|x|_p = \left\|\sum_{n=1}^{\infty}\varepsilon_n x_n\right\|_p := \left(\mathbb{E}\left\|\sum_{n=1}^{\infty}\varepsilon_n x_n\right\|^p\right)^{\frac{1}{p}}$$
para cada $x = (x_n)_n \in C(X)$. Se puede probar que $(C(X), |\cdot|_p)$ es de Fréchet para $p \in \mathbb{R}_+$ (es de Banach para $p \geq 1$) y que en dado caso las topologías inducidas por la norma $|\cdot|_p$ son más fuertes que la topología producto en $C(X)$, y no solo eso, se puede ver que todas estas $|\cdot|_p$-topologías son equivalentes, esto es:

Lema 3.3.13. *Sea $(\varepsilon_n)_n$ una sucesión de Rademacher, si definimos el conjunto L dado por*
$$L = \left\{\sum_{n=1}^{\infty}\varepsilon_n x_n \;:\; (x_n)_n \in C(X)\right\}$$
entonces $L \subset L_p(\Omega; X)$ y todas las L_p-topologías coinciden en L, donde $p \in \mathbb{R}_+$.

Las L_p-topologías en L son las topologías inducidas por las normas $\|\cdot\|_p$ en L, luego, para probar que las L_p-topologías son equivalentes hay que probar dichas normas son equivalentes, esto es, dados $0 < p, q < \infty$ hay que exhibir una constante $C > 0$ tal que

$$\|R\|_q \leq C\|R\|_p,$$

donde $R \in L$ es una serie de Rademacher de la forma $\sum_n \varepsilon_n x_n$. Pero este resultado es conocido como desigualdad de Kahane [21, pág. 100, teorema 4.7], el cual probamos en la proposición C.1.2.

Diremos que un espacio de Banach, X, es de *Rademacher de tipo* p, con $p \in [1, 2]$, si $\ell_p(X) \subset C(X)$.[5] Esto es, si para cada sucesión en $(x_n)_n \subset X$ la convergencia de la serie $\sum_{n=1}^{\infty} \|x_n\|^p$ implica la convergencia casi segura de la serie $\sum_{n=1}^{\infty} \varepsilon_n x_n$, donde $(\varepsilon_n)_n$ es una sucesión de Rademacher.

Cuando el espacio X es de tipo p la inclusión de $\ell(X)$ en $C(X)$ tiene gráfica cerrada, luego es continua. Esto proporciona la siguiente definición:

Definición 3.3.2. Sea $1 \leq p \leq 2$. Diremos que un espacio de Banach X es **Rademacher de tipo** p, o en breve de **tipo** p, si existe una constante $0 < C < \infty$ tal que

$$\mathbb{E}\left\|\sum_{n=1}^{m} \varepsilon_n x_n\right\|^p \leq C \sum_{n=1}^{m} \|x_n\|^p$$

para cada $m \geq 1$ y para cada $x_1, x_2, \ldots, x_m \in X$.

Comentario. En términos del espacio $C(X)$ la definición 3.3.2 es como sigue [17]: X es de tipo p y existe una constante $0 < C < \infty$ tal que

$$\mathbb{E}\left\|\sum_{n=1}^{\infty} \varepsilon_n x_n\right\|^p \leq C \sum_{n=1}^{\infty} \|x_n\|^p$$

para cada $(x_n)_n \in C(X)$. La equivalencia se obtiene de aplicar el teorema de Itô-Nisio 3.3.4 (que establece que la convergencia en probabilidad y casi segura de una serie de elementos aleatorios independientes son equivalentes) y del lema de Fatou.

No vamos ahondar en esta definición pero comentamos que la restricción en los valores de p se debe a que todo espacio de Banach es de tipo para $p \in (0, 1]$, mientras que el único espacio de tipo p, para $p > 2$, es el trivial: $\{\mathbf{0}\}$.

Lo anterior es consecuencia de las *desigualdades de Kahane-Khinchin*, al respecto puede consultarse [1] (ver pág. 138), o bien [10] (ver pág. 218).

En el siguiente teorema probaremos que un espacio satisface la ley fuerte de los grandes números si y solo si el espacio es de tipo p, y aunque esta pareciera

[5] $\ell_p(X) = \left\{(x_n)_n \in X^{\infty} : \sum_{n=1}^{\infty} \|x_n\|^p < \infty\right\}$.

ser una condición extraña en realidad no lo es tanto. Fue G. Pisier quién probó en [28] que un espacio X es B-convexo si y solo si X es de tipo p para algún $p > 1$ (recordemos que B-convexidad es una condición equivalente a una ley fuerte de grandes números). Mostrar esto queda fuera del alcance de esta tesis, para una mejor referencia aparte de la ya antes mencionada recomendamos ver [10] los capítulos 11, 12 y 13 (especialmente este último).

En lo sucesivo las sucesiones de elementos aleatorios que consideremos serán tales que $\mathbb{E}[V_n] = \mathbf{0}$ para cada n, esto es, las variables aleatorias están centradas. Diremos que una sucesión de elementos aleatorios $(V_n)_n$ satisface la **condición de Chung** si

$$\sum_{n=1}^{\infty} \frac{1}{n^p} \mathbb{E}\left[\|V_n\|^p\right] < \infty. \tag{3.25}$$

Teorema 3.3.14 (Høffmann-Jørgensen y Pisier)**.** *Sea $1 \leq p \leq 2$, entonces los siguientes cuatro enunciados son equivalentes:*

1. *X es de tipo p.*

2. *Existe $C > 0$ tal que*

$$\mathbb{E}\left\|\sum_{j=1}^{n} V_j\right\|^p \leq C \sum_{j=1}^{n} \mathbb{E}\|V_j\|^p$$

   para cada elección $V_1, V_2, \ldots, V_n$ de elementos aleatorios independientes con media nula y p-ésimo momento finito.

3. *La ley fuerte de los grandes números,*

$$\frac{1}{n} \sum_{k=1}^{n} V_k \overset{c.s.}{\to} \mathbf{0},$$

   se sigue para toda sucesión $(V_n)_n$ de elementos aleatorios independientes con esperanza nula que satisface la condición (3.25) de Chung.

4. *Si $\sum_{n=1}^{\infty} n^{-p}\|x_n\|^p < \infty$, donde $(x_n)_n \subset X$, y $(\varepsilon_n)_n$ es una sucesión de Rademacher, entonces*

$$\frac{1}{n} \sum_{k=1}^{n} \varepsilon_k x_k \overset{\mathbb{P}}{\to} \mathbf{0}.$$

Demostración. (1) $\Rightarrow$ (2). Como X es de tipo p, existe una constante positiva, $B > 0$, tal que

$$\mathbb{E}\left\|\sum_{j=1}^{n} \varepsilon_j x_j\right\|^p \leq B \sum_{j=1}^{n} \|x_j\|^p \tag{3.26}$$

para cada $x_1, x_2, \ldots, x_n \in X$ y para cada $n \in \mathbb{N}$. Consideremos $V_1, \ldots, V_n$ n elementos aleatorios como (2), y además independientes de la suceción $(\varepsilon_n)_n$ de Rademacher. Entonces, por el teorema de Fubini,

$$\mathbb{E}\left\|\sum_{j=1}^{n} \varepsilon_j V_j\right\|^p \leq B \sum_{j=1}^{n} \mathbb{E}\|V_j\|^p.$$

Ahora observemos que se satisfacen las hipótesis del teorema 3.3.10 donde la sucesión $(\eta_j)_j \equiv (1)_j$ (la sucesión constante igual a 1), por lo tanto

$$\mathbb{E}\left\|\sum_{j=1}^{n} V_j\right\|^p \leq 8^p \mathbb{E}\left\|\sum_{j=1}^{n} \varepsilon_j V_j\right\|^p.$$

Luego la condición (2) se sigue para $C = 8^p B$.

(2) $\Rightarrow$ (3). Sea $(V_n)_n$ una sucesión de elementos aleatorios centrados y p-integrables que satisfacen la condición (3.25) de Chung, esto es, la serie

$$\sum_{n=1}^{\infty} \frac{1}{n^p} \mathbb{E}\left[\|V_n\|^p\right] < \infty.$$

Veamos que $n^{-1} \sum_{k=1}^{n} V_k \to \mathbf{0}$ c.s.

Denotemos por $S_n = \sum_{j=1}^{n} j^{-p} \mathbb{E}\left[\|V_j\|^p\right]$, entonces la sucesión de sumas parciales $(S_n)_n$ es de Cauchy. Así, dado $\varepsilon > 0$ existe $N > 0$ tal que

$$\sum_{j=m+1}^{n} \frac{1}{j^p} \mathbb{E}\left[\|V_j\|^p\right] < \varepsilon$$

para cada $n > m \geq N$. Afirmamos que la serie $\sum_{n=1}^{\infty} n^{-1} V_n$ converge en $L_p(\Omega; X)$. En efecto, sea $T_n = \sum_{j=1}^{n} j^{-1} V_j$ y tomemos m y n índices tales que $n > m$. Como suponemos que la condición (2) se da, entonces existe $C > 0$ tal que

$$\mathbb{E}\left[\|T_n - T_m\|^p\right] = \mathbb{E}\left\|\sum_{j=m+1}^{n} \frac{V_j}{j}\right\|^p \leq C \sum_{j=m+1}^{n} \mathbb{E}\left\|\frac{V_j}{j}\right\|^p < C\varepsilon$$

para cada $n \geq m \geq N$. Por lo tanto la sucesión $(T_n)_n$ es de Cauchy en $L_p(\Omega; X)$ así que la serie $\sum_{n=1}^{\infty} n^{-1} V_n$ converge en $L_p(\Omega; X)$ (al ser este un espacio completo), pero como convergencia en L_p (o media p-ésima) implica convergencia en probabilidad (ver lema 1.1.7), entonces $\sum_{n=1}^{\infty} n^{-1} V_n$ converge en probabilidad. Luego, por el teorema 3.3.5 de Itô-Nisio, la serie $\sum_{n=1}^{\infty} n^{-1} V_n$ converge casi seguramente. El lema 2.2.4 de Kronecker concluye que:

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} V_k = \mathbf{0}$$

casi seguramente.

(3) $\Rightarrow$ (4). Sea $(x_n)_n$ una sucesión en X tal que la serie $\sum_{n=1}^{\infty} n^{-p}\|x_n\|^p < \infty$, y sea $(\varepsilon)_j$ sucesión de Rademacher. Para cada $n \in \mathbb{N}$ definimos $V_n : \Omega \to X$ como $V_n = \varepsilon_n x_n$, la proposición 1.2.3 garantiza que V_n es un elemento aleatorio en X para cada n. Más aún, por el inciso (d) del teorema 1.3.1 , $\mathbb{E}[V_n] = \mathbb{E}[\varepsilon_n x_n] = \mathbb{E}[\varepsilon_n] \cdot x_n = \mathbf{0}$, además la sucesión $(V_n)_n$ es independiente por el lema 1.2.7 y satisface la condición de Chung pues

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{\mathbb{E}\left[\|V_n\|^p\right]}{n^p} &= \sum_{n=1}^{\infty} \frac{\mathbb{E}\left[|\varepsilon_n|^p \cdot \|x_n\|^p\right]}{n^p} \\ &\leq \sum_{n=1}^{\infty} \frac{\|x_n\|^p \cdot \mathbb{E}\left[|\varepsilon_n|^p\right]}{n^p} \leq \sum_{n=1}^{\infty} \frac{\|x_n\|^p}{n^p} < \infty.\end{aligned}$$

Por lo tanto $(V_n)_n = (\varepsilon_n x_n)_n$ obedece la ley fuerte de los grandes números, es decir

$$\frac{1}{n}\sum_{k=1}^{n} \varepsilon_k x_k \to \mathbf{0}$$

casi seguramente, y por tanto también en probabilidad.

(4) $\Rightarrow$ (1). Sea $(x_n)_n$ una sucesión en X para la cuál

$$\sum_{n=1}^{\infty} \frac{1}{n^p}\|x_n\|^p < \infty.$$

Sea $(\varepsilon_n)_n$ una sucesión de Rademacher, entonces, por la hipótesis (4),

$$\frac{1}{n}\sum_{j=1}^{n} \varepsilon_j x_j \to \mathbf{0} \quad \text{en probabilidad.}$$

Por lo observado en el lema 3.3.13 (es decir, por la equivalencias de los momentos de las sumas de Rademacher (teorema C.1.2)) y del lema C.1.3 (véase observación C.1.1):

$$\frac{1}{n}\sum_{j=1}^{n} \varepsilon_j x_j \to \mathbf{0} \quad \text{en } L_p(\Omega; X).$$

Definamos los siguientes espacios:

$$G = \left\{ (x_j)_j \in X^{\infty} \;:\; \sum_{j=1}^{\infty} j^{-p}\|x_j\|^p < \infty \right\}, \qquad |x|_G = \left(\sum_{j=1}^{\infty} j^{-p}\|x_j\|^p \right)^{\frac{1}{p}}$$

$$F = \left\{ (x_j)_j \in X^{\infty} \;:\; \sup_n \left\| \frac{1}{n}\sum_{j=1}^{n} \varepsilon_j x_j \right\|_p < \infty \right\}, \; |x|_F = \sup_n \left\| \frac{1}{n}\sum_{j=1}^{n} \varepsilon_j x_j \right\|_p .$$

Entonces $(G, |\cdot|_G)$ y $(F, |\cdot|_F)$ son ambos espacios de Banach, y por lo anterior tenemos que $G \subset F$. Así, si consideramos $i\colon G \to F$ la inyección canónica (i.e. $i(x) = x$ para cada $x = (x_n)_n \in G$), obtenemos una función lineal entre dos espacios de Banach que tiene gráfica cerrada: Si $(x_m)_{m=1}^\infty$ es una sucesión en G que converge x, y $(i(x_m))_{m=1}^\infty = (x_m)_{m=1}^\infty$ converge a y en F, entonces $x = y$. Para ello note que convergencia en la norma $|\cdot|_G$ implica que las componentes $(x_{m,j})_{m=1}^\infty$ convergen a x_j (en X) mientras que convergencia en la norma $|\cdot|_F$ implica la convergencia de las componentes $(x_{m,j})_{m=1}^\infty$ a y_j (en X), por lo que $x_j = y_j$ para cada $j \in \mathbb{N}$. Entonces, por el teorema A.2.3 de la gráfica cerrada, i es continua, de modo que existe una constante $C > 0$ tal que $|i(x)|_F \leq C|x|_G$ para cada $x \in G$. Es decir,

$$\sup_n \frac{1}{n} \left(\mathbb{E} \left\| \sum_{j=1}^n \varepsilon_j x_j \right\|^p \right)^{\frac{1}{p}} \leq C \left(\sum_{j=1}^\infty \frac{1}{j^p} \|x_j\|^p \right)^{\frac{1}{p}}$$

para cada $x_1, x_2, \ldots \in X$. Entonces,

$$\frac{1}{n} \left(\mathbb{E} \left\| \sum_{j=1}^n \varepsilon_j x_j \right\|^p \right)^{\frac{1}{p}} \leq C \left(\sum_{j=1}^n \frac{1}{j^p} \|x_j\|^p \right)^{\frac{1}{p}}$$

para cada $x_1, \ldots, x_n \in X$ y para cada $n \in \mathbb{N}$, o equivalentemente,

$$\mathbb{E} \left\| \sum_{j=1}^n \varepsilon_j x_j \right\|^p \leq C^p n^p \sum_{j=1}^n \frac{\|x_j\|^p}{j^p}. \tag{3.27}$$

Así, sean $x_1, x_2, \ldots, x_n \in X$ y definamos

$$y_j = \begin{cases} 0 & \text{si} \quad 1 \leq j \leq N, \\ x_{j-N} & \text{si} \quad N < j \leq N + n \end{cases}$$

donde N es entero positivo. Entonces

$$\begin{aligned} \mathbb{E} \left\| \sum_{j=1}^n \varepsilon_j x_j \right\|^p &= \mathbb{E} \left\| \sum_{j=1}^{N+n} \varepsilon_j y_j \right\|^p \\ &\leq C^p (N+n)^p \sum_{j=N+1}^{N+n} \frac{\|x_{j-N}\|^p}{j^p} \qquad \text{(por 3.27)} \\ &\leq C^p (N+n)^p \sum_{j=N+1}^{N+n} \frac{\|x_{j-N}\|^p}{(N+1)^p} \\ &= C^p \left(\frac{N+n}{N+1} \right)^p \sum_{j=1}^n \|x_j\|^p \end{aligned}$$

para cada $N \geq 1$, por lo que al tomar el límite cuando N tiende a infinito

$$\mathbb{E}\left\|\sum_{j=1}^{n} \varepsilon_j x_j\right\|^p \leq K \sum_{j=1}^{n} \|x_j\|^p$$

para cada $x_1, \ldots, x_n \in X$ y para cada $n \in \mathbb{N}$, es decir, X es de tipo p. ■

Observación 3.3.1. El teorema 3.3.14 que acabamos de demostrar no es aplicable a la sucesión de elementos aleatorios que dimos en el ejemplo 3.3.1, pues como $q > 1 - \frac{1}{p}$,

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}\|V_n\|^p}{n^p} = \sum_{n=1}^{\infty} \frac{n^{qp}}{p} > \sum_{n=1}^{\infty} \frac{n^{p-1}}{n^p} = \sum_{n=1}^{\infty} \frac{1}{n} = \infty,$$

es decir, no se satisface la condición de Chung.

Apéndice A

Análisis funcional

Tanto en este capítulo, como en el siguiente, describimos definiciones y resultados que se asumen por ser bien conocidos, o bien debido a que son de naturaleza auxiliar como para merecer la separación del texto principal.

A.1. Sobre espacios métricos, normados y de Hilbert

Definición A.1.1. Un conjunto M no vacío es un **espacio métrico** si existe una función $d\colon M \times M \to [0, \infty)$ con las siguientes propiedades:

1. $d(x, y) \geq 0$ para cada $(x, y) \in M \times M$ y
   $d(x, y) = 0$, si y sólo si $x = y$
2. $d(x, y) = d(y, x)$
3. $d(x, y) \leq d(x, z) + d(z, y)$ para todo $x, y, z \in M$.

Al número real $d(x, y)$ se le llama **distancia** de x a y y la función d es llamada **métrica**. Denotaremos por el símbolo (M, d) al espacio métrico M con la métrica d.

Definición A.1.2. Una sucesión $(x_n)_{n\in\mathbb{N}}$ es un espacio métrico (M, d) es una **sucesión de Cauchy** si para todo $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que $d(x_n, x_m) < \varepsilon$ si $n, m \geq N$.

Definición A.1.3. Diremos que un espacio métrico (M, d) es **completo** si toda sucesión de Cauchy en M converge a un elemento en M.

Definición A.1.4. Un **espacio vectorial** sobre un campo $\mathbb{F}$ es un conjunto X no vacío, dotado de dos operaciones; la operación *suma* $+\colon X \times X \to X$ que a cada para de elementos $(\mathbf{u}, \mathbf{v}) \in X \times X$ le asigna el elemento $\mathbf{w} \in X$ tal que

$\mathbf{w} = \mathbf{u} + \mathbf{v}$, y la operación *multiplicación por escalar* $\cdot\colon \mathbb{F} \times X \to X$ que a cada $\mathbf{v} \in X$ y $t \in \mathbb{F}$ le asigna el elemento $t\mathbf{v} \in X$, tales operaciones satisfacen que para cada $\mathbf{u}$,$\mathbf{v}$,$\mathbf{w} \in X$ y $s, t \in \mathbb{F}$

1. $\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$
2. $\mathbf{u}+ (\mathbf{v}+\mathbf{w}) = (\mathbf{u}+ \mathbf{v})+\mathbf{w}$
3. Existe un elemento $\mathbf{0} \in X$, tal que $\mathbf{v} + \mathbf{0} = \mathbf{v}$
4. Para cada $\mathbf{v} \in X$, existe un elemento $\hat{\mathbf{v}} \in X$, tal que $\mathbf{v} + \hat{\mathbf{v}} = \mathbf{0}$
5. $1\mathbf{v} = \mathbf{v}$
6. $s(t\mathbf{v}) = (st)\mathbf{v}$
7. $(s + t)\mathbf{v} = s\mathbf{v} + t\mathbf{v}$
8. $s(\mathbf{u} + \mathbf{v}) = s\mathbf{u} + s\mathbf{v}$

Nos referiremos como vectores (o a veces simplemente puntos) a los elementos de X. Llamaremos *neutro aditivo*, o de acuerdo a lo anterior, *vector cero* al elemento $\mathbf{0}$ descrito en 3 y haremos la convención de denotarlo con el número cero en negritas.

A lo largo de este trabajo convendremos que todos los espacios vectoriales con los cuales trabajaremos serán reales, es decir, $\mathbb{F} = \mathbb{R}$.

Definición A.1.5. Sea X un espacio vectorial. Una función $\rho\colon X \to \mathbb{R}$ es una **seminorma** en X si satisface las siguientes propiedades:

1. $\rho(ax) = |a|\rho(x)$ para todo $a \in \mathbb{R}, x \in X$
2. $\rho(x + y) \leq \rho(x) + \rho(y)$ para todo $x, y \in X$

De la propiedad 1 tenemos que $\rho(0) = \rho(0 \cdot x) = 0 \cdot \rho(x) = 0$, y haciendo uso de la propiedad 2 tenemos que $0 = \rho(x + (-1)x) \leq \rho(x) + \rho((-1)x) = \rho(x) + |-1|\rho(x) = 2\rho(x)$ y por lo tanto, $\rho(x) \geq 0$ para cada $x \in X$.

Definición A.1.6. Sea X un espacio vectorial. Diremos que una función $\|\cdot\|_X\colon X \to \mathbb{R}$ es una **norma** en X si $\|\cdot\|_X$ es una seminorma tal que si $\|x\| = 0$, entonces $x = 0$.

En tal caso se dice que X es un espacio normado con norma $\|\cdot\|_X$ y se denota por $(X, \|\cdot\|_X)$. A menudo se usaremos la notación $(X, \|\cdot\|)$ si no hay confusión sobre la norma que estamos considerando en el espacio X.
Nótese que todo espacio normado es un espacio métrico, donde la metrica está dada por la norma de la siguiente manera: $d(x, y) := \|x - y\|$.

Definición A.1.7. Un espacio normado se llama **espacio de Banach** si es completo (respecto a la distancia inducida por la norma).

Definición A.1.8. Sea H un espacio vectorial. Una función $\langle\cdot,\cdot\rangle\colon H\times H\to\mathbb{F}$, donde $\mathbb{F}=\mathbb{R},\mathbb{C}$, es un **producto interior** en H si para cada $x,y,z\in X$ y cada $\alpha\in\mathbb{F}$ se verifica lo siguiente:

1. $\langle x+y,z\rangle=\langle x,z\rangle+\langle y,z\rangle$;
2. $\langle ax,y\rangle=a\langle x,y\rangle$;
3. $\overline{\langle x,y\rangle}=\langle y,x\rangle$;
4. $\langle x,x\rangle>0$ si $x\neq 0$

En tal caso decimos que H es un espacio con producto interior y se denota por $(H,\langle\cdot\rangle)$. Note que el producto interior en H induce una norma (en H) dada por $\|x\|:=\sqrt{\langle x,x\rangle}$ para cada $x\in H$.

Definición A.1.9. Un espacio con producto interior $(H,\langle\cdot,\cdot\rangle)$ se llama **espacio de Hilbert** si es un espacio de Banach respecto a la norma dada anteriormente.

La barra sobre el producto interior de x e y, escrito en 3, representa el complejo conjudado de ese número. Sin embargo, durante este trabajo asumiremos que todos los espacios de Hilbert con lo que trabajaaremos serán reales. De manera que dicha propiedad 3 ahora se traduce en decir que el producto interior es conmutativo.

Si x e y denotan elementos en un espacio de Hilbert entonces los siguientes resultados son válidos.

Proposición A.1.1 (Desigualdad de Cauchy-Schwarz)**.**

$$|\langle x,y\rangle|\leq\|x\|\cdot\|y\|$$

Proposición A.1.2 (Desigualdad del triángulo)**.**

$$\|x+y\|\leq\|x\|+\|y\|$$

En particular si x e y son ortogonales, es decir, $\langle x,y\rangle=0$, entonces

$$\|x+y\|^2=\|x\|^2+\|y\|^2$$

Proposición A.1.3. *Sea $(H,\langle\cdot,\cdot\rangle)$ un espacio de Hilbert. El producto intetrior $\langle\cdot,\cdot\rangle$ es una función continua de $H\times H$ en $\mathbb{R}$.*

Demostración. Para cada $x_1,y_1,x_2,y_2\in H$ la desiguadal de Cauchy-Schwarz implica que

$$\begin{aligned}|\langle x_1,y_1\rangle-\langle x_2,y_2\rangle|&=|\langle x_1-x_2,y_1\rangle-\langle x_2,y_1-y_2\rangle|\\&\leq\|y_1\|\|x_1-x_2\|+\|x_2\|\|y_1-y_2\|.\end{aligned}$$

lo cual a su vez implica la continuidad del producto interior, pues si tomamos $x_n\to x$ y $y_n\to y$ cuando $n\to\infty$, entonces

$$|\langle x_n,y_n\rangle-\langle x,y\rangle|\leq\|y_n\|\|x_n-x\|+\|x\|\|y_n-y\|$$

que tiende a cero pues $\|y_n\|$ es acotada al ser $(y_n)_n$ convergente. ∎

Proposición A.1.4. *Si $(x_n)_n$ es una base, entonces para cada $y \in H$ existe una sucesión $(c_n)_n$ de números reales para la cual*

$$y = \lim_n \sum_{k=1}^{n} c_k x_k \tag{A.1}$$

Es decir,

$$\|y - (c_1x_1 + c_2x_2 + \cdots + c_nx_n)\| \to 0 \quad (\textit{cuando } n \to \infty)$$

Si (x_n) es una base ortonormal entonces la serie en (A.1) es llamada *expansión de Fourier* de y.

Proposición A.1.5 (Desigualdad de Bessel)**.** *Sea $(x_n)_n$ una sucesión ortonormal y sea $c_n = \langle y, x_n\rangle$, donde y es un elemento en H espacio de Hilbert. Entonces*

$$\sum_{k=1}^{\infty} c_k^2 \leq \|y\|^2$$

En particular si la sucesión (x_n) es una base, entonces

$$\sum_{k=1}^{\infty} c_k^2 = \|y\|^2.$$

Teorema A.1.6. *Sea H de Hilbert. H tiene una base ortonormal, si y sólo si H es separable.*

A.2. Espacios vectoriales topológicos

Definición A.2.1. Decimos que X es un **espacio vectorial topológico** o **espacio topológico lineal** si es un espacio vectorial real provisto de una topología que hace continua a la operaciones suma de vectores y multiplicación por escalar.

Para entender más a fondo la definición anterior vamos a requerir de los siguiente conceptos.

Sea (X, τ) un espacio topológico, se dice que $\mathcal{B} \subset \tau$ es una base la topología τ si para cada $U \in \tau$ y cada $x \in U$, existe $V \in \mathcal{B}$ tal que $x \in V \subset U$.
Por ejemplo, en un espacio métrico las bolas abiertas forman una base para la topología usual.

Ahora, dado un punto $x \in X$, diremos que V_x es una vecindad de x si existe $U \in \tau$ tal que $x \in U \subset V_x$, y diremos que $\mathcal{V}_x \subset \mathcal{N}(x)$ es una base de vecindades de x (base local en x) si para cada $V_x \in \mathcal{N}(x)$ existe $U_x \in \mathcal{V}_x$ tal que $U_x \subset V_x$, donde $\mathcal{N}(x)$ denota al conjunto de todas las vecindades de x.
Por ejemplo, en un espacio métrico las bolas cerradas con centro en un punto x

forman una base de vecindades de x.

Ahora bien, recordemos que una función $f\colon (Y,\tau_Y) \to (X,\tau_X)$ es continua un punto $y_0 \in Y$ si para cada $U_{f(y_0)} \in \tau_X$ (ó $\mathcal{V}_{f(y_0)}$), existe $V_{y_0} \in \tau_Y$ (ó $\mathcal{V}_{y_0}$) tal que si $y \in V_{y_0}$, entonces $f(y) \in U_{f(y_0)}$.

Por otro lado, sabemos que $U_1 \times U_2$ con $U_1 \in \tau_Y$ y $U_2 \in \tau_X$ forman una base para la topología producto de $Y \times X$. Entonces,

1. La operación **suma**

$$\begin{aligned} +\colon X \times X &\to X \\ (x,y) &\mapsto x+y, \end{aligned}$$

   es continua en un punto (x_0, y_0) si para cada abierto V de X tal que $x_0 + y_0 \in V$ existen $U_1, U_2 \in \tau$ tal que $(x_0,y_0) \in U_1 \times U_2$ y si $(x,y) \in U_1 \times U_2$, entonces $x + y \in V$.

2. La operación **producto por escalar**

$$\begin{aligned} \cdot\colon \mathbb{F} \times X &\to X \\ (\lambda,x) &\mapsto \lambda x, \end{aligned}$$

   es continua en (λ_0, x_0) si para cada abierto V de X tal que $\lambda_0 x_0 \in V$ existe $\delta > 0$ y $U \in \tau$ tal que $(\lambda_0, x_0) \in B(\lambda_0,\delta) \times U$ y si $(\lambda, x) \in B(\lambda_0,\delta) \times U$, entonces $\lambda x \in U$.,

Si en particular, la topología que estamos considerando es inducida por una métrica, la continuidad de las operaciones suma y producto por escalar están caracterizadas de la siguietne forma: Dado $\varepsilon > 0$ existe $\delta > 0$ tal que

1. si $d(x,x_0) < \delta$, $d(y,y_0) < \delta$ implica que $d(x+y, x_0+y_0) < \varepsilon$
2. si $|\lambda - \lambda_0| < \delta$, $d(x,x_0) < \delta$ implica que $d(\lambda x, \lambda_0 x_0) < \varepsilon$.

Como consecuencia de esto llegamos a siguiente resultado:

Proposición A.2.1. *Todo espacio normado es un espacio vectorial topológico.*

Demostración.

1. **Continuidad de la suma.** Sea $\varepsilon > 0$, entonces elegimos $\delta := \dfrac{\varepsilon}{2}$ tal que $\|x - x_0\|$, $\|y - y_0\| < \delta$, entonces

$$\|(x+y) - (x_0+y_0)\| \leq \|x - x_0\| + \|y - y_0\| < 2\delta = \varepsilon.$$

2. **Continuidad del producto por escalar.** Sea $\varepsilon > 0$, y notemos lo siguiente

$$\begin{aligned} \|\lambda x - \lambda_0 x_0\| &= \|(\lambda - \lambda_0)(x - x_0) + \lambda_0(x - x_0) + (\lambda - \lambda_0)x_0\| \\ &\leq |\lambda - \lambda_0| \, \|x - x_0\| + |\lambda_0| \, \|x - x_0\| + |\lambda - \lambda_0 \, \|x_0\| \end{aligned} \tag{A.2}$$

entonces, escogiendo $\delta > 0$, con $\delta < \text{mín}\left\{1, \frac{\varepsilon}{3}, \frac{\varepsilon}{3(|\lambda_0|+1)}, \frac{\varepsilon}{3(\|x_0\|+1)}\right\}$ y tal que $|\lambda - \lambda_0| < \delta$ y $\|x - x_0\| < \delta$, otenemos de la desigualdad A.2 que

$$\begin{aligned} \|\lambda x - \lambda_0 x_0\| &\leq \delta^2 + |\lambda_0|\delta + \delta\,\|x_0\| \\ &< \frac{\varepsilon}{3}\cdot 1 + |\lambda_0|\cdot\frac{\varepsilon}{3(|\lambda_0|+1)} + \frac{\varepsilon}{3(\|x_0\|+1)}\|x_0\| \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned} \tag{A.3}$$

■

Hemos de notar que en la prueba de la proposición anterior no usamos que si "$\|x\| = 0$, *entonces* $x = 0$", por lo que esta prueba permanece siendo válida para espacios seminormados.

El siguiente teorema provee una forma muy útil de caracterizar a las funciones lineales continuas entre dos espacios normados arbitrarios.

Teorema A.2.2. *Sean $(X, \|\cdot\|_X)$ y $(Y, \|\cdot\|_Y)$ dos espacios normados y sea $T\colon X \to Y$ una función lineal. Son equivalentes*

1. *T es Lipschitiz continua*
2. *T es es uniformemente continua*
3. *T es continua en X*
4. *T es continua en 0*
5. *Si (x_n) es una sucesión en X tal que $x_n \to \mathbf{0}$, entonces $T(x_n) \to \mathbf{0}$*
6. *T es acotada*
7. *Existe $M > 0$ tal que $\|T(x)\|_Y \leq M\|x\|_X$, para todo $x \in X$.*
8. $\sup\limits_{\|x\|_X \leq 1} \|T(x)\|_Y < \infty$

La prueba de este teorema puede encontrase en [22] a lo largo del capítulo 4.

No es el propósito de esta tesis estudiar las propiedades del espacio de funciones lineales continuas de un espacio normado a otro, pero si mencionamos que a este espacio se le denota por $B(X,Y)$, y de acuerdo al teorema anterior, $T \in B(X,Y)$, si y sólo si, existe $M > 0$ tal que $\|T(x)\| \leq M\|x\|$ para cada $x \in X$. Tal espacio resulta ser un espacio vectorial, y no sólo eso, también resulta ser un espacio normado donde la norma esta dada como sigue:

$$\|T\| := \text{ínf}\,\{M > 0 \,:\, \|T(x)\| \leq M\|x\|\}\,. \tag{A.4}$$

Más aún, se puede mostrar que tal norma se puede calcular como

$$\|T\| = \text{mín}\{M > 0 \,:\, \|T(x)\| \leq M\|x\|\} = \sup_{x\in B[0,1]} \|T(x)\| = \sup_{x\in B(0,1)} \|T(x)\|,$$

y si además el espacio $X \neq \{\mathbf{0}\}$, también se tiene que

$$\sup_{x\in B(0,1)} \|T(x)\| = \sup_{\|x\|=1} \|T(x)\| = \sup_{x\in X} \frac{\|T(x)\|}{\|x\|}.$$

Observación A.2.1. De esto último queremos destacar lo siguiente:

1. Si $T \in B(X,Y)$, entonces para cada $x \in X$

$$\|T(x)\| \leq \|T\|\,\|x\|.$$

2. Si Y es de Banach, entonces $B(X,Y)$ es también un espacio de Banach con respecto a la norma definida anteriormente.

Un caso particularmente importante es cuando $Y = \mathbb{R}$, en tal caso llamamos a la función T **funcional lineal**.

Definición A.2.2. Sea X un espacio vectorial topológico, definimos el **espacio dual topológico** de X como

$$X^* = \{f \colon X \to \mathbb{R} \mid f \text{ es lineal y continua}\}$$

Análogamente, al espacio formado por las funcionales lineales en X le llamaremos **espacio dual algebraico** y lo denotaremos por $X^\dagger$.
Como consecuencia de lo anteriormente discutido obtenemos que X^* es un espacio de Banach respecto a la norma definida en (A.4).

Los siguientes resultados son ampliamente conocidos y pueden ser encontrados a lo largo del texto [30].

Teorema A.2.3 (Gráfica Cerrada)**.** *Supongamos que* $(X, \|\cdot\|)$ *y* $(Y, \|\cdot\|)$ *son espacios de Banach y* $T \colon X \to Y$ *es lineal.* T *es continua si y sólo si* $Graf(T) = \{(x, T(x)) \,:\, x \in X\}$ *es cerrada.*

Teorema A.2.4 (Extensión de Hahn-Banach)**.** *Sea* X *un espacio vectorial,* $S \subset X$ *un subespacio,* $\rho \colon X \to \mathbb{R}$ *un funcional sublineal y* $f \in S^\dagger$ *lineal tal que* $f(x) \leq \rho(x)$ *para cada* $x \in S$*. Entonces existe* $F \in X^\dagger$ *tal que* $F\restriction_S = f$ *y* $F(x) \leq \rho(x)$ *para cada* $x \in X$*.*

Corolario A.2.5 (Hahn-Banach)**.** *Sea* $(X, \|\cdot\|)$ *un espacio normado y* $x_0 \neq 0$*. Entonces existe* $f \in X^*$ *tal que* $f(x_0) = \|x_0\|$ *y* $\|f\| = 1$*.*

Definición A.2.3. Se dice que un espacio vectorial topológico X **tiene suficientes funcionales continuas** si dados $x \neq y$ en X, existe $f \in X^*$ tal que $f(x) \neq f(y)$.

Como consecuencia del corolario A.2.5 tenemos el siguiente corolario.

Corolario A.2.6. *Todo espacio normado tiene suficientes funcionales continuas.*

Demostración. En efecto, sean $x, y \in X$ con $x \neq y$ entonces $x - y \neq \mathbf{0}$. Por el corolario A.2.5 existe $f \in X^*$ tal que $f(x-y) = \|x-y\| \neq 0$ y entonces $f(x) - f(y) \neq 0$, por lo tanto $f(x) \neq f(y)$. ■

Teorema A.2.7. *Sean X e Y espacios de Banach, y sea D un subespacio denso de X. Estonces todo operador lineal continuo definido en D, $T\colon D \to Y$, tiene una única extensión lineal y continua en todo X.*

Apéndice B

Teoría de la medida

B.1. Sigma álgebra de Borel

Definición B.1.1. Una colección $\mathcal{S}$ de subconjuntos de un conjunto X es una σ-álgebra si

1. $\emptyset, X \in \mathcal{S}$,

2. $A \in \mathcal{S}$, entonces $X \setminus A \in \mathcal{S}$,

3. $A_1, A_2, \ldots, A_n, \ldots, \in \mathcal{S}$, entonces $\displaystyle\bigcup_{i=1}^{\infty} A_i \in \mathcal{S}$.

Como $\displaystyle\bigcap_i A_i = X \setminus \left(\bigcup_i (X \setminus A_i) \right)$, entonces $\mathcal{S}$ es también cerrada bajo intersecciones numerables. No sólo eso, si $\{\mathcal{S}_\alpha : \alpha \in \Lambda\}$ es una familia de σ-álgebras de subconjuntos de X, entonces $\displaystyle\bigcap_{\alpha \in \Lambda} \mathcal{S}_\alpha$ es también una σ-álgebra. La prueba de esta afirmación es inmediata de la definición (B.1.1).
Una vez observado esto ya podemos definir la σ-álgebra generada por una colección de subconjuntos de X.

Definición B.1.2. Sea $\mathcal{R}$ una colección de subconjuntos de X. La sigma-álgebra generada por $\mathcal{R}$, que denotamos por $\sigma(\mathcal{R})$, es la colección

$$\sigma(\mathcal{R}) = \bigcap \{\mathcal{S} : \mathcal{S} \text{ es una } \sigma\text{-álgebra de subconjuntos de } X \text{ y } \mathcal{R} \subset \mathcal{S}\}$$

Una consecuencia inmediata de esta definición es que $\sigma(\mathcal{R})$ es la mínima σ-álgebra que contiene a $\mathcal{R}$. En efecto, si $\mathcal{S}'$ es otra σ-álgebra que contiene a $\mathcal{R}$ entonces $\sigma(\mathcal{R}) \subset \mathcal{S}'$ (pues ella ya es uno de los «intersectandos»), así tenemos que $\mathcal{R} \subset \sigma(\mathcal{R}) \subset \mathcal{S}'$.

Las siguientes dos propiedades resultan especialmente útiles para probar contenciones entre σ-álgebras generadas, pues nos dicen que basta verificar las contenciones entre los conjuntos generadores para concluir las conteciones de sus respectivas la σ-álgebras generadas.

Lema B.1.1.

1. *Si $\mathcal{R}_1$ y $\mathcal{R}_2$ son dos colecciones de subconjuntos de X tales que $\mathcal{R}_1 \subset \mathcal{R}_2$, entonces $\sigma(\mathcal{R}_1) \subset \sigma(\mathcal{R}_2)$.*

2. *Si $\mathcal{F}$ es una σ-álgebra, entonces $\sigma(\mathcal{F}) = \mathcal{F}$.*

Demostración.

1. Tenemos que $\mathcal{R}_1 \subset \mathcal{R}_2 \subset \sigma(\mathcal{R}_2)$, de modo que $\sigma(\mathcal{R}_2)$ es una σ-álgebra que contine a $\mathcal{R}_1$, por ser $\sigma(\mathcal{R}_1)$ la mínima σ-álgebra que contiene a $\mathcal{R}_1$ concluimos que $\sigma(\mathcal{R}_1) \subset \sigma(\mathcal{R}_2)$.

2. Por definición ocurre que $\mathcal{F} \subset \sigma(\mathcal{F})$. Para la otra contención, observemos que $\mathcal{F}$ es en particular una σ-álgebra que contiene a $\mathcal{F}$ por lo tanto $\sigma(\mathcal{F}) \subset \mathcal{F}$.

∎

Por **espacio de Borel** entenderemos a la pareja $(X, \mathcal{S})$ donde X es un conjunto (no vacío) y $\mathcal{S}$ es una σ-álgebra de subconjuntos de X. En particular estaremos interesados cuando la σ-álgebra $\mathcal{S}$ es la σ-álgebra de Borel de X para X un espacio métrico.

Definición B.1.3. Sea M un espacio métrico. La **σ-álgebra de Borel** de M, denotada por $\mathcal{B}(M)$, es la σ-álgebra generada por la familia de todos los subconjuntos abiertos de M. A los elementos de $\mathcal{B}(M)$ se les llaman *borelianos*.

Si U es un abierto en $\mathcal{B}(M)$, entonces $M \setminus U$ es un cerrado y también pertenece a $\mathcal{B}(M)$. Los siguientes resultados tienen como propósito probar que $\mathcal{B}(M)$ es también generada por los conjuntos cerrados de M.

Definición B.1.4. Sea M un espacio métrico. Un subconjunto A de M es G_δ si existe una sucesión de conjuntos abiertos (U_n) tal que $A = \bigcap_{n=1}^{\infty} U_n$. El complemento de un conjunto G_δ recibe el nombre de F_σ.

Proposición B.1.2. *Sea (M, d) un espacio métrico. Para cada $A \subset M$ (fijo) definimos $f(p) := d(p, A) = \inf\{d(p, a) : a \in A\}$. Entonces para cada $p, q \in A$*

$$|f(p) - f(q)| \leq d(p, q).$$

En particular, f es una función uniformemente continua.

Demostración. Sea $a \in A$, por la desigualdad del triángulo y la definición de $d(p, A)$ tenemos que

$$d(p, A) \leq d(p, a) \leq d(p, q) + d(q, a).$$

Esto significa que para toda $a \in A$,

$$d(p, A) - d(p, q) \leq d(q, a)$$

de modo que $d(p, A) - d(p, q)$ es una cota inferior del conjunto $\{d(q, a) : a \in A\}$ y por lo tanto

$$d(p, A) - d(p, q) \leq d(q, A).$$

Con lo que encontramos que

$$d(p, A) \leq d(p, q) + d(q, A).$$

Análogamente se tiene que

$$d(q, A) \leq d(p, q) + d(p, A).$$

De estas dos últimas desigualdades podermos concluir que $|f(p) - f(q)| \leq d(p, q)$. ■

Corolario B.1.3. *En un espacio métrico M todo subconjunto cerrado es G_δ y todo subconjunto abierto es F_σ.*

Demostración. Sea C un subconjunto cerrado de M. Entonces

$$\begin{aligned} C &= \{x : d(x, G) = 0\} \\ &= \bigcap_{n=1}^{\infty} \left\{ x : d(x, G) < \frac{1}{n} \right\}. \end{aligned}$$

Por la proposición anterior sabemos que $f(x) = d(x, G)$ es una función continua, por lo tanto $f^{-1}[(-\infty, \frac{1}{n})] = \left\{x : d(x, G) < \frac{1}{n}\right\}$ es un conjunto abierto en M. La segunda parte de la afirmación es cierta por definición de conjunto F_σ. ■

Corolario B.1.4. *En un espacio métrico (M, d) la σ-álgebra de Borel está generada por los conjuntos cerrados.*

Demostración. Si $\mathcal{A}$ y $\mathcal{C}$ representan a la clase de los abiertos y los cerrados en M respectivamente, el corolario anterior nos garantiza que $\mathcal{A} \subset \sigma(\mathcal{C})$ y en virtud de minimalidad de la σ-álgebra $\sigma(\mathcal{A})$, obtenemos que $\sigma(\mathcal{A}) = \mathcal{B}(M) \subset \sigma(\mathcal{C})$. Así también $\mathcal{C} \subset \mathcal{B}(M)$ y por lo tanto $\sigma(\mathcal{C}) = \mathcal{B}(M)$. ■

Dicho de otra forma, $\mathcal{B}(M)$ es la mínima familia de subconjuntos de M que contiene a todos los subconjuntos abiertos y cerrados de M, y que es cerrada bajo intersecciones y uniones numerables.

Recordemos que un espacio métrico M es **separable** contiene un subconjunto denso y numerable, esto es, existen $x_1, x_2, \ldots$ en M tal que $\overline{\{x_1, x_2, \ldots\}} = M$ (donde $\overline{A}$ denota la cerradura topológica de A en X).

Probaremos que bajo la hipótesis de separabilidad la σ-álgebra $\mathcal{B}(M)$ está generada por las bolas abiertas (o cerradas).

Lema B.1.5. *Si M es un espacio métrico separable, entonces $\mathcal{B}(M)$ es igual a la σ-álgebra generada por las bolas abiertas (cerradas) de M.*

Demostración. Denotemos por $\mathcal{B}$ a la σ-álgebra generada por las bolas abiertas (o cerradas) en M. Entonces, según lo antes visto $\mathcal{B} \subset \mathcal{B}(X)$. Para la otra contención hagamos notar dos cosas.

1. Todo abierto es unión numerable de bolas abiertas. Sea $D = \{d_n : n \geq 1\}$ el denso numerable en M, y sea U un abierto en M. Entonces para cada $x \in U$ y existe $r > 0$, con r racional tal que la bola abierta con centro en x y de radio r, $B(x,r)$, se queda contenida en U. Como D es denso en M, y en particular para $B(x, \frac{1}{3})$, se tiene que $B(x, \frac{1}{3}) \cap D \neq \emptyset$, por lo que existe $y_x \in D$, tal que $d(x, y_x) < \frac{r}{3} < \frac{r}{2}$. Entonces $x \in B(y_x, \frac{r}{2}) \subset B(x,r) \subset U$.

Notemos que hay a lo más una cantidad numerable de bolas del tipo $B\left(y_x, \frac{r}{2}\right)$, pues la colección de bolas con centro en D y radio *racional* tiene a lo más la cardinalidad de $\mathbb{N} \times \mathbb{N}$ que a su vez tiene la misma cardinalidad de $\mathbb{N}$. Por lo tanto,

$$U = \bigcup_{x \in U} B\left(y_x, \frac{r}{2}\right)$$

es una unión a lo más numerable.

2. Todo abierto es unión numerable de bolas cerradas. Por el caso 1 todo abierto es unión numerable de bolas abiertas, pero estas a su vez son uniones numerables de bolas cerradas, es decir, en general si $x \in M$ y $r > 0$ (no necesariamente racional) se tiene que

$$B(x,r) = \bigcup_{n=n_r}^{\infty} B[x, r - 1/n]$$

donde n_r es un entero positivo para el cuál $r - \frac{1}{n} > 0$. Y como la unión numerable de una unión numerable sigue siendo numerable podemos concluir que todo abierto es unión numerable de bolas cerradas.

Con esto podemos concluir que dado un abierto U de M, $U \in \mathcal{B}$ y por lo tanto $\mathcal{B}(M) \subset \mathcal{B}$ probando así la igualdad de las σ-álgebras. ∎

Definición B.1.5. Sea $(X, \mathcal{B}(X))$ y $(Y, \mathcal{B}(Y))$ dos espacios de Borel y sea $f : X \to Y$ una función. Diremos que f es una función **Borel medible**, o simplemente **medible**, si para cada $E \in \mathcal{B}(Y)$,

$$f^{-1}(E) = \{x \in X : f(x) \in E\} \in \mathcal{B}(X).$$

B.2. Sigma álgebra producto

Definición B.2.1. Sean $(X, \mathcal{B}(X)), (Y, \mathcal{B}(Y))$ dos espacios de Borel, supongamos además que X, Y son métricos. Por **rectángulo de Borel** en $X \times Y$, entendemos a los conjuntos de la forma $A \times B$ donde $A \in \mathcal{B}(X)$ y $B \in \mathcal{B}(Y)$.

Definimos la **σ-álgebra producto** en $X \times Y$, y la denotamos por $\mathcal{B}(X) \otimes \mathcal{B}(Y)$, como la σ-álgebra generada por la familia de rectángulos de Borel. Esto es,

$$\mathcal{B}(X) \otimes \mathcal{B}(Y) = \sigma(\{A \times B \,:\, A \in \mathcal{B}(X), B \in \mathcal{B}(Y)\}).$$

Llamaremos **espacio producto** a la pareja $(X \times Y, \mathcal{B}(X) \otimes \mathcal{B}(Y))$, por lo que una función $f \colon X \times Y \to Z$ será medible si la imagen inversa de borelianos de Z es un elemento de la σ-álgebra producto.

En el espacio producto también podemos considerar la σ-álgebra de Borel $\mathcal{B}(X \times Y) = \sigma(\{U \,:\, U \text{ es abierto en } X \times Y\})$.

Es claro de la definición que $\mathcal{B}(X) \otimes \mathcal{B}(Y) \subset \mathcal{B}(X \times Y)$ (pues producto de abiertos es abierto), la otra contención no se da en general. Daremos un contraejemplo de esto último y además probaremos que bajo la hipótesis de separabilidad de X e Y se da la igualdad.

Proposición B.2.1. *Sean X, Y espacios métricos separables. Entonces $\mathcal{B}(X \times Y) = \mathcal{B}(X) \otimes \mathcal{B}(Y)$.*

Demostración. Es suficiente probar que todo abierto U en $X \times Y$ pertenece a $\mathcal{B}(X) \otimes \mathcal{B}(Y)$. Como Y es separable, Y tiene una base numerable $\{V_n\}$. Entonces U puede representarse como unión numerable de conjuntos de la forma $U_\alpha \times V_n$ donde U_α es abierto en X. Para cada $n \geq 1$ fijo, sea W_n la unión de todos subconjuntos U_α tales que $U_\alpha \times V_n \subset U$. Entonces $U = \bigcup_{n=1}^{\infty} (W_n \times V_n) \in \mathcal{B}(X) \otimes \mathcal{B}(Y)$. ■

Para ver que no se da la ingualdad de las σ-álgebras vamos a necesitar de los siguiente lemas.

Lema B.2.2. *Sea $\mathcal{F}$ una colección de subconjuntos de un conjunto X y consideremos la σ-algebra generada por la familia $\mathcal{F}$, $\sigma(\mathcal{F})$. Si $A \in \sigma(\mathcal{F})$, entonces existe una subfamilia numerable $\mathcal{F}'$ de $\mathcal{F}$ tal que $A \in \sigma(\mathcal{F}')$.*

Demostración. Sea $\mathcal{A} = \bigcup \sigma(\mathcal{E})$ donde $\mathcal{E}$ varía sobre todas las subfamilias numerables de la familia $\mathcal{F}$. Veamos que $\mathcal{A}$ es una σ-algebra:

1. Es claro que tanto $\emptyset$ como X pertenecen a $\mathcal{A}$, pues $\sigma(\mathcal{E})$ les contiene para cada $\mathcal{E} \subset \mathcal{F}$ numerable.

2. Si $A \in \mathcal{A}$, entonces $A \in \sigma(\mathcal{E})$ para alguna $\mathcal{E} \subset \mathcal{F}$ numerable, y como $\sigma(\mathcal{E})$ es una σ-álgebra tenemos que $X \setminus A \in \sigma(\mathcal{E})$, por lo que $X \setminus A \in \mathcal{A}$.

3. Sea $(A_n)_n$ una sucesión en $\mathcal{A}$, entonces $A_n \in \mathcal{A}$ para cada $n \geq 1$, y por tanto para cada $n \geq 1$ existe $\mathcal{E}_n$ subfamilia numerable de $\mathcal{F}$ tal que $A_n \in$

$\sigma(\mathcal{E}_n)$. Tomemos $\mathcal{E}' = \bigcup_{n=1}^{\infty} \mathcal{E}_n$, entonces $\mathcal{E}'$ es una subfamilia numerable de $\mathcal{F}$ y por tanto $\sigma(\mathcal{E}') \subset \mathcal{A}$. Además, puesto que para cada $n \geq 1$, $\mathcal{E}_n \subset \mathcal{E}' \subset \sigma(\mathcal{E}')$ tenemos que $\sigma(\mathcal{E}_n) \subset \sigma(\mathcal{E}')$, luego $\bigcup_{n=1}^{\infty} \sigma(\mathcal{E}_n) \subset \sigma(\mathcal{E}')$. Así, $\bigcup_{n=1}^{\infty} A_n \in \bigcup_{n=1}^{\infty} \sigma(\mathcal{E}_n) \subset \mathcal{A}$. Lo cual prueba que $\mathcal{A}$ es cerrado bajo uniones numerables.

Por lo tanto $\mathcal{A}$ es una σ-álgebra. Más aún, $\mathcal{F} \subset \mathcal{A}$, pues si $B \in \mathcal{F}$ entonces $\{B\} \subset \mathcal{F}$ donde $\{B\}$ es numerable y por tanto $\{B\} \subset \sigma(\{B\}) \subset \mathcal{A}$, con lo cual $B \in \mathcal{A}$. Por otra parte, si $B \in \mathcal{A}$ entonces $B \in \sigma(\mathcal{E})$ para alguna subfamilia numerable $\mathcal{E}$ de $\mathcal{F}$, pero como $\mathcal{E} \subset \mathcal{F} \subset \sigma(\mathcal{F})$ entonces $\sigma(\mathcal{E}) \subset \sigma(\mathcal{F})$. Por lo tanto $B \in \sigma(\mathcal{F})$. Hemos probado que

$$\mathcal{F} \subset \mathcal{A} \subset \sigma(\mathcal{F})$$

y por tanto $\mathcal{A} = \sigma(\mathcal{F})$. Podemos concluir que si $C \in \sigma(\mathcal{F})$, entonces existe una subfamilia numerable $\mathcal{F}'$ de $\mathcal{F}$ tal que $C \in \mathcal{F}'$. ■

Lema B.2.3. *Sea X un conjunto con cardinalidad mayor que $\mathfrak{c}$. Dada una colección numerable $\{U_n\}$ de subconjuntos de X, existen $x \neq y$ en X tales que para cada $n \geq 1$ se cumple que $x, y \in U_n$ o bien $x, y \notin U_n$, en tal caso decimos que la familia $\{U_n\}$ no distingue a x y y. Más aún, la colección Σ de todos los subconjuntos de X que no distingen a x e y es una σ-álgebra.*

Demostración. Supongamos que dados $x \neq y$ en X existe U_n tal que $x \in U_n$ pero $y \notin U_n$ (o bien $y \in U_n$ pero $x \notin U_n$). Definimos $h\colon X \to \{0,1\}^{\mathbb{N}}$ como $h(x) = (a_n)_n$ donde $a_n = 1$ si $x \in U_n$ y $a_n = 0$ si $x \notin U_n$. Entonces h es inyectiva y por tanto, $\operatorname{card}(X) \leq \operatorname{card}(\{0,1\}^{\mathbb{N}}) = \mathfrak{c}$, lo que contradice la hipótesis sobre la cardinalidad de X. Por tanto deben existir $x, y \in X$ con $x \neq y$ tales que para cada $n \geq 1, x, y \in U_n$ o $x, y \notin U_n$.
Ahora veamos que colección Σ de subconjuntos de X que no distingue a x y y es una σ-álgebra.

1. Es claro que $\emptyset$, X pertenecen a Σ.

2. Sea $A \in \Sigma$, entonces $x, y \in A$ o bien $x, y \notin A$. Si $x, y \in A$, tenemos que $x, y \notin X \setminus A$ y por tanto $X \setminus A \in \Sigma$. Por otro lado, si $x, y \notin A$ significa que $x, y \in X \setminus A$. Ello implica que $X \setminus A \in \Sigma$.

3. Sea $(A_n)_n$ una sucesión en Σ, entonces $A_n \in \Sigma$ para cada $n \geq 1$. Si $x \in \bigcup_{n=1}^{\infty} A_n$, entonces $x \in A_{n_0}$ para algún $n_0 \geq 1$, lo cual implica que $x, y \in A_{n_0}$ y por tanto $x, y \in \bigcup_{n=1}^{\infty} A_n$, así $\bigcup_{n=1}^{\infty} A_n \in \Sigma$. Por otro lado, si $x \notin$

$\bigcup_{n=1}^{\infty} A_n$ implica que $x, y \notin A_n$ para todo $n \geq 1$ y por tanto, $x, y \notin \bigcup_{n=1}^{\infty} A_n$ y entonces $\bigcup_{n=1}^{\infty} A_n \in \Sigma$.

■

Proposición B.2.4. *Sea* (X, d) *un espacio métrico con cardinalidad mayor que* $\mathfrak{c}$. *Entonces* $\mathcal{B}(X) \otimes \mathcal{B}(X) \subsetneqq \mathcal{B}(X \times X)$.

Demostración. Es claro que $\mathcal{B}(X) \otimes \mathcal{B}(X) \subset \mathcal{B}(X \times X)$. Veamos que $\mathcal{B}(X \times X)$ no está contenida en $\mathcal{B}(X) \otimes \mathcal{B}(X)$.

Supongamos lo contrario. Consideremos $D = \{(x, x) : x \in X\}$ el conjunto diagonal en $X \times X$, y notemos que al ser $d \colon X \to [0, \infty)$ una función continua y $\{0\}$ un cerrado en $[0, \infty)$ entonces $d^{-1}(\{0\}) = D$ es cerrado en $X \times X$ y por tanto D pertenece a $\mathcal{B}(X \times X)$. Afirmamos que $D \notin \mathcal{B}(X) \otimes \mathcal{B}(X)$.

Supongamos que $D \in \mathcal{B}(X) \otimes \mathcal{B}(X)$. Como $\mathcal{B}(X) \otimes \mathcal{B}(X)$ es la σ-álgebra generada por la familia de rectángulos de Borel $\{A \times B : A, B \in \mathcal{B}(X)\}$, por el lema B.2.2 existe una subfamilia numerable de rectángulos medibles $\{A_n \times B_n\}_{n \geq 1}$ tal que $D \in \sigma(\{A_n \times B_n\})$. Ahora, por el lema B.2.3 existen $x, y \in X$ tales que $x \neq y$ y la colección numerable $\{A_n, B_n : n \geq 1\}$ de subconjuntos de X no distingue a x y y. Entonces $\{A_n \times B_n : n \geq 1\}$ es una colección numerable de subconjuntos de $X \times X$ que no distingue a los elementos (x, x) y (x, y) de $X \times X$.

Denotemos por Σ a la σ-álgebra de subconjuntos de $X \times X$ que no disguen a (x, x) y (x, y). Entonces $\{A_n \times B_n\} \subset \Sigma$ y por tanto $\sigma(\{A_n \times B_n\}) \subset \Sigma$. Como $D \in \sigma(\{A_n \times B_n\})$ tenemos que $D \in \Sigma$, de modo que D es un subconjunto de $X \times X$ que no distingue a (x, x) y (x, y), pero esto no es posible, ya que $(x, x) \in D$ pero $(x, y) \notin D$.

Por lo tanto $D \notin \mathcal{B}(X) \otimes \mathcal{B}(X)$, así, la σ-álgebra $\mathcal{B}(X) \otimes \mathcal{B}(X)$ está contenida propiamente en $\mathcal{B}(X \times X)$. ■

Apéndice C

Resultados varios

Teorema C.0.1 (Convergencia Dominada de Lebesgue [29]). *Si $f_1, f_2, \dots$ es una sucesión de funciones integrables y existe una función integrable $b(x)$ tal que $|f_n| \leq b$ para cada $n = 1, 2, \dots$ entonces*

$$\lim_{n\to\infty} \int f_n \, d\mathbb{P} = \int \lim_{n\to\infty} f_n \, d\mathbb{P}$$

siempre que $\lim\limits_{n\to\infty} f_n$ exista casi seguramente.

Teorema C.0.2 (Beppo-Levi [29]). *Si $f_1, f_2, \dots$ es una sucesión de funciones integrables y no negativas tales que $\sum_{n=1}^{\infty} \int f_n < \infty$, entonces la serie $\sum_{n=1}^{\infty} f_n$ es convergente casi seguramente y*

$$\int \sum_{n=1}^{\infty} f_n = \sum_{n=1}^{\infty} \int f_n.$$

Lema C.0.3. *Sea $(V_n)_n$ es una sucesión de elementos aleatorios en un espacio de Banach separable X, supongamos que $V_n \xrightarrow{\mathbb{P}} V$ y $\|V_n\| \leq \|W\|$ para cada n donde $W \in L_p(\Omega; X)$, entonces $V_n \xrightarrow{L_p} V$ (en p-media).*

Demostración. Como $\mathbb{E}\left[\|V_n\|^p\right] \leq \mathbb{E}\left[\|W\|^p\right] < \infty$, entonces $V_n \in L_p(\Omega; X)$ para cada n. Sea $\varepsilon > 0$, entonces

$$\begin{aligned} \{\|V\| > \|W\| + \varepsilon\} &\subset \{\|V\| > \|V_n\| + \varepsilon\} \\ &\subset \{\|V\| - \|V_n\| > \varepsilon\} \\ &\subset \{\|V - V_n\| > \varepsilon\} \end{aligned}$$

para cada n, lo cual implica que

$$\mathbb{P}\left(\|V\| > \|W\| + \varepsilon\right) \leq \lim_{n\to\infty} \mathbb{P}\left(\|V - V_n\| > \varepsilon\right) = 0.$$

Pero como esto es válido para cada $\varepsilon > 0$, en particular, tomando $\varepsilon = \frac{1}{k}$ con $k \in \mathbb{N}$ obtenemos

$$0 \leq \mathbb{P}\left(\|V\| > \|W\|\right) \leq \lim_{k\to\infty} \mathbb{P}\left(\|V\| > \|W\| + \frac{1}{k}\right) = 0.$$

Es decir, $\|V\| \leq \|W\|$ c.s. y en consecuencia $V \in L_p(\Omega; X)$.

Ahora supongamos que V_n no converge a V en p-media (en L_p), es decir, existe $\varepsilon > 0$ tal que para cada n,

$$\mathbb{E}\left[\|V_n - V\|^p\right] \geq \varepsilon. \tag{C.1}$$

Así que existe $(V_{n_k})_k$ una subsucesión tal que $\mathbb{E}\left[\|V_{n_k} - V\|^p\right] \geq \varepsilon$, por otra parte sabemos $V_n \xrightarrow{\mathbb{P}} V$ por lo que también $V_{n_k} \xrightarrow{\mathbb{P}} V$, y entonces por el lema 1.1.13 esta subsucesión admite una subsucesión $(V_{n_{k_j}})_j$ tal que $V_{n_{k_j}} \xrightarrow{c.s.} V$, es decir, $V_{n_{k_j}} - V \xrightarrow{c.s.} \mathbf{0}$. Por la primera parte de la demostración $\|V_{n_{k_j}} - V\| \leq 2\|W\|$, así que por el teorema de convergencia dominada de Lebesgue

$$\lim_{j\to\infty} \mathbb{E}\left[\|V_{n_{k_j}} - V\|^p\right] = \mathbb{E}\left[\lim_{j\to\infty} \|V_{n_{k_j}} - V\|^p\right] = 0$$

lo cual contradice (C.1). Por lo tanto V_n sí converge a V en p-media. ■

Proposición C.0.4. *Sean $V_1, V_2, \ldots, V$ elementos aleatorios en un espacio de Banach separable. Entonces, $V_n \xrightarrow{\mathbb{P}} V$ si y solo si*

$$\lim_{n\to\infty} \mathbb{E}\left\{\frac{\|V_n - V\|}{1 + \|V_n - V\|}\right\} = 0.$$

La prueba de esta proposición es análoga a la que se tiene para variables aleatorias.

Demostración. Sin pérdida de generalidad supongamos que $V = \mathbf{0}$. Queremos mostrar que $V_n \xrightarrow{\mathbb{P}} V$ si y sólo si $\lim_{n\to\infty} \mathbb{E}\left\{\frac{\|V_n\|}{1+\|V_n\|}\right\} = 0$. Primero supongamos que $V_n \xrightarrow{\mathbb{P}} 0$. Sea $\varepsilon > 0$, entonces $\lim_{n\to\infty} \mathbb{P}\left(\|V_n\| > \varepsilon\right) = 0$. Notemos que

$$\frac{\|V_n\|}{1 + \|V_n\|} \leq \frac{\|V_n\|}{1 + \|V_n\|}\mathbb{I}_{\{\|V_n\| > \varepsilon\}} + \varepsilon \mathbb{I}_{\{\|V_n\| \leq \varepsilon\}} \leq \mathbb{I}_{\{\|V_n\| > \varepsilon\}} + \varepsilon.$$

Y por tanto,

$$\mathbb{E}\left\{\frac{\|V_n\|}{1 + \|V_n\|}\right\} \leq \mathbb{P}\left(\|V_n\| > \varepsilon\right) + \varepsilon.$$

Esto implica que

$$\lim_{n\to\infty} \mathbb{E}\left\{\frac{\|V_n\|}{1 + \|V_n\|}\right\} \leq \varepsilon$$

para toda $\varepsilon > 0$. El resultado se obtiene al hacer tender ε a cero.

Ahora supongamos que $\lim_{n\to\infty} \mathbb{E}\left\{\frac{\|V_n\|}{1+\|V_n\|}\right\} = 0$. Como la función $f(x) = \frac{x}{1+x}$ es estrictamente creciente en el intervalo $[0,\infty)$, tenemos que

$$\frac{\varepsilon}{1+\varepsilon}\mathbb{I}_{\{\|V_n\|>\varepsilon\}} \leq \frac{\|V_n\|}{1+\|V_n\|}\mathbb{I}_{\{\|V_n\|>\varepsilon\}} \leq \frac{\|V_n\|}{1+\|V_n\|}$$

donde $\varepsilon > 0$ es fijo (pero arbitrario). Por lo que tomando esperanzas y sacando límites,

$$\frac{\varepsilon}{1+\varepsilon}\lim_{n\to\infty} \mathbb{P}\left(\|V_n\| > \varepsilon\right) \leq \lim_{n\to\infty} \mathbb{E}\left\{\frac{\|V_n\|}{1+\|V_n\|}\right\} = 0.$$

Como $\varepsilon > 0$ es fijo, $\lim_{n\to\infty} \mathbb{P}\left(\|V_n\| > \varepsilon\right)$. ■

Recordemos que si $L_0(\Omega, X)$ denota al conjunto de elementos aleatorios en X, y en él definimos la métrica

$$\|V\|_0 = \mathbb{E}\left\{\frac{\|V\|}{1+\|V\|}\right\}$$

para cada $V \in L_0(\Omega, X)$, entonces la proposición anterior muestra que la L_0-topología (la inducida por la métrica $\|\cdot\|_0$) induce la convergencia en probabilidad.

C.1. Equivalencia de los momentos de sumas de Rademacher

El propósito de esta sección será discutir el lema 3.3.13 sobre la equivalencia de las L_p-topologías en L. Recordemos que este espacio, L, es el espacio de las series de Rademacher de la forma $\sum_{k=1}^{\infty} \varepsilon_k x_k$, que convergen casi seguramente. Por esta razón nos fijaremos en las sumas de Rademacher $\sum_{k=1}^{n} \varepsilon_k x_k$. Puede consultarse el caso general en [21] (ver teorema 4.7).

Recordemos también que las L_p-topologías en L son las topologías inducidas por las p-normas en L, por lo que en nuestro caso, lo que tenemos que probar es las p-normas son equivalentes en el conjunto de sumas de Rademacher, o dicho de otro modo, los momentos de las sumas de Rademacher son equivalentes, esto es, dados $0 < p, q < \infty$ debemos mostrar que existe $C > 0$ tal que

$$\left\|\sum_{k=1}^{n} \varepsilon_k x_k\right\|_q \leq C \left\|\sum_{k=1}^{n} \varepsilon_k x_k\right\|_p$$

para cada $n \in \mathbb{N}$ y cada $x_1, \ldots, x_n \in X$.

Para ello necesitaremos del siguiente lema, el cuál ya fue implícitamente probado en la prueba del lema 3.3.11.

Lema C.1.1. *Consideremos la suma de Rademacher $\sum_{k=1}^{n} \varepsilon_k x_k$ en el espacio de Banach X. Entonces, para cada $a > 0$ y $n \geq 1$,*

$$\mathbb{P}\left(\left\|\sum_{k=1}^{n} \varepsilon_k x_k\right\| \geq 2a\right) \leq 4\left(\mathbb{P}\left(\left\|\sum_{k=1}^{n} \varepsilon_k x_k\right\| \geq a\right)\right)^2.$$

Teorema C.1.2 (Desigualdad de Kahane)**.** *Dados* $0 < p, q < \infty$, *existe una constante* $K_{p,q} > 0$ *tal que si* X *es un espacio de Banach y* $x_k \in X$, *con* $k \geq 1$, *entonces para cada* n,

$$\left(\mathbb{E}\left\|\sum_{k=1}^{n} \varepsilon_k x_k\right\|^q\right)^{\frac{1}{q}} \leq K_{p,q} \cdot \left(\mathbb{E}\left\|\sum_{k=1}^{n} \varepsilon_k x_k\right\|^p\right)^{\frac{1}{p}}.$$

Demostración. Supongamos que $0 < p < q < \infty$, cuando $p \geq q$ la afirmación es inmediata con $K_{p,q} = 1$ (basta usar la desigualdad de Jensen). Hagamos

$$M = \left(\mathbb{E}\left\|\sum_{k=1}^{n} \varepsilon_k x_k\right\|^p\right)^{\frac{1}{p}}.$$

Si $M = 0$ la desigualdad es trivial y se verifica con cualquier constante positiva, digamos $K_{p,q} = 1$. Así, supongamos que $M > 0$, y sea $u_k = \frac{x_k}{M}$, con $1 \leq k \leq n$, para el cual

$$\mathbb{E}\left\|\sum_{k=1}^{n} \varepsilon_k u_k\right\|^p = \mathbb{E}\left\|\sum_{k=1}^{n} \varepsilon_k \frac{x_k}{M}\right\|^p = 1. \tag{C.2}$$

Por la desigualdad 1.2 de Chebyshev:

$$\mathbb{P}\left(\left\|\sum_{k=1}^{n} \varepsilon_k u_k\right\| \geq 8^{\frac{1}{p}}\right) = \mathbb{P}\left(\left\|\sum_{k=1}^{n} \varepsilon_k u_k\right\|^p \geq 8\right) \leq \frac{1}{8}\mathbb{E}\left(\left\|\sum_{k=1}^{n} \varepsilon_k u_k\right\|^p\right) = \frac{1}{8}.$$

Supongamos, por inducción, que para $m \geq 0$ tenemos que

$$\mathbb{P}\left(\left\|\sum_{k=1}^{n} \varepsilon_k u_k\right\| \geq 2^m \cdot 8^{\frac{1}{p}}\right) \leq \frac{1}{4} \cdot 2^{-2^m}. \tag{C.3}$$

Aplicando el lema C.1.1 y C.3,

$$\mathbb{P}\left(\left\|\sum_{k=1}^{n} \varepsilon_k u_k\right\| \geq 2^{m+1} \cdot 8^{\frac{1}{p}}\right) \leq 4\left(\mathbb{P}\left(\left\|\sum_{k=1}^{n} \varepsilon_k u_k\right\| \geq 2^m \cdot 8^{\frac{1}{p}}\right)\right)^2 \leq \frac{1}{4} \cdot 2^{-2^{m+1}},$$

esto muestra que C.3 se cumple para cada $m \geq 0$.

Por otra parte, sabemos que por el lema 3.3.6,

$$\mathbb{E}\left\|\sum_{k=1}^{n} \varepsilon_k u_k\right\|^q = \int_0^\infty q t^{q-1} \mathbb{P}\left(\left\|\sum_{k=1}^{n} \varepsilon_k u_k\right\| \geq t\right) dt. \tag{C.4}$$

Para cada $t \geq 0$ definamos

$$f(t) = q t^{q-1} \mathbb{P}\left(\left\|\sum_{k=1}^{n} \varepsilon_k u_k\right\| \geq t\right),$$

y hagamos $\alpha_0 = 0$, $\alpha_m = 2^{m-1} \cdot 8^{\frac{1}{p}}$ para cada $m = 1, 2, \ldots$.

Usando C.3 y C.4,

$$\begin{aligned}
\mathbb{E}\left\|\sum_{k=1}^{n} \varepsilon_k u_k\right\|^q &= \int_0^\infty f(t)\,dt \\
&= \int_0^{\alpha_1} f(t)\,dt + \sum_{m=0}^{\infty} \int_{\alpha_{m+1}}^{\alpha_{m+2}} f(t)\,dt \\
&\leq \int_0^{\alpha_1} qt^{q-1}\,dt + \sum_{m=0}^{\infty} \int_{\alpha_{m+1}}^{\alpha_{m+2}} qt^{q-1}\mathbb{P}\left(\left\|\sum_{k=1}^{n} \varepsilon_k u_k\right\| \geq \alpha_{m+1}\right) dt \\
&\leq \alpha_1^q + \sum_{m=0}^{\infty} \int_{\alpha_{m+1}}^{\alpha_{m+2}} qt^{q-1} \cdot \frac{1}{4} \cdot 2^{-2^m}\,dt \\
&\leq 8^{\frac{q}{p}} + \frac{1}{4} \sum_{m=0}^{\infty} 2^{-2^m} \left(\alpha_{m+2}^q - \alpha_{m+1}^q\right) < \infty.
\end{aligned}$$

Tomando $K_{p,q}^q := 8^{\frac{q}{p}} + \frac{1}{4}\sum_{m=0}^{\infty} 2^{-2^m} \left(\alpha_{m+2}^q - \alpha_{m+1}^q\right)$ tenemos que

$$\left(\mathbb{E}\left\|\sum_{k=1}^{n} \varepsilon_k u_k\right\|^q\right)^{\frac{1}{q}} \leq K_{p,q}$$

y por C.2,

$$\left(\mathbb{E}\left\|\sum_{k=1}^{n} \varepsilon_k u_k\right\|^q\right)^{\frac{1}{q}} \leq K_{p,q} \cdot \left(\mathbb{E}\left\|\sum_{k=1}^{n} \varepsilon_k u_k\right\|^p\right)^{\frac{1}{p}}.$$

Finalmente, como $u_k = \frac{x_k}{M}$,

$$\left(\mathbb{E}\left\|\sum_{k=1}^{n} \varepsilon_k x_k\right\|^q\right)^{\frac{1}{q}} \leq K_{p,q} \cdot \left(\mathbb{E}\left\|\sum_{k=1}^{n} \varepsilon_k x_k\right\|^p\right)^{\frac{1}{p}}.$$

■

Lema C.1.3. *Sea V una variable aleatoria positiva tal que para algún $q > p > 0$ y alguna constante positiva C,*

$$\|V\|_q \leq C\|V\|_p.$$

Entonces, si $\varepsilon > 0$ es tal que $\mathbb{P}(V > \varepsilon) \leq (2C^p)^{q/(p-q)}$, se tiene que

$$\|V\|_p \leq 2^{\frac{1}{p}}\varepsilon \qquad y \qquad \|V\|_q \leq 2^{\frac{1}{p}} C\varepsilon.$$

Demostración. Como $q > p$ tomemos $p' = \frac{q}{p}$ y $q' = \frac{p'}{p'-1} = \frac{q}{q-p}$, entonces por la desigualdad de Hölder:

$$\begin{aligned}
\mathbb{E}\left[V^p \mathbb{I}_{\{V>\varepsilon\}}\right] &\leq \left(\mathbb{E}\left[(V^p)^{p'}\right]\right)^{1/p'} \left(\mathbb{E}\left[\mathbb{I}_{\{v>\varepsilon\}}\right]\right)^{1/q'} \\
&\leq \left(\mathbb{E}\left[V^q\right]\right)^{p/q} \left(\mathbb{P}(V > \varepsilon)\right)^{(p-q)/q} = \|V\|_q^p \left(\mathbb{P}(V > \varepsilon)\right)^{1-p/q}.
\end{aligned}$$

Luego,

$$\begin{aligned}
\mathbb{E}\left[V^p\right] &= \int_{\{V\leq\varepsilon\}} V^p \, d\mathbb{P} + \int_{\{V>\varepsilon\}} V^p \, d\mathbb{P} \\
&\leq \varepsilon^p + \int_{\{V>\varepsilon\}} f^p \, d\mathbb{P} \\
&\leq \varepsilon^p + \|V\|_q^p \left(\mathbb{P}(V>\varepsilon)\right)^{1-p/q} \\
&\leq \varepsilon^p + C^p \|V\|_p^p \cdot \frac{1}{2C^p} = \varepsilon^p + \frac{\|V\|_p^p}{2}
\end{aligned}$$

y por lo tanto, $\|V\|_p^p = \mathbb{E}\left[V^p\right] \leq 2\varepsilon^p$. ∎

Observación C.1.1. Notemos que como consecuencia de este lema, si $(V_n)_n$ es una sucesión de variables aleatorias (o elementos aleatorios) tales que para algún $q > p > 0$ y $C > 0$, $\|V_n\|_q \leq C\|V_n\|_p$ para cada n, y si $(V_n)_n$ converge en probabilidad a alguna variable aleatoria V, entonces, como $\sup_n \|V_n\|_q < \infty$ por el lema C.1.3, $(V_n)_n$ también converge a V en $L_{q'}$ para toda $q' < q$.

En particular, y debido a la desigualdad C.1.2 de Kahane, esto último ocurre para la sucesión que consideramos en la implicación (4) $\Rightarrow$ (1) del teorema 3.3.14.

Bibliografía

[1] F. Albiac and N. J. Kalton, *Topics in Banach space theory*, Springer, 2006.

[2] M. Á. G. Álvarez, *Introducción a la teoría de la probabilidad II. Segundo curso*, Fondo de Cultura Economica, 2015.

[3] A. Araujo and E. Giné, *The central limit theorem for real and Banach valued random variables*, John Wiley & Sons, 1980.

[4] A. Beck, *A convexity condition in banach spaces and the strong law of large numbers*, Proceedings of the American Mathematical Society, 13 (1962), pp. 329–334.

[5] A. Beck, *Probability in Banach Spaces: Proceedings of the First International Conference on Probability in Banach Spaces, 20 - 26 July 1975, Oberwolfach*, Lecture Notes in Mathematics, Springer Berlin Heidelberg, 2006.

[6] A. Beck and P. Warren, *Strong laws of large numbers for weakly orthogonal sequences of banach space-valued random variables*, Ann. Probab., 2 (1974), pp. 918–925.

[7] P. Billingsley, *Probability and Measure*, Wiley Series in Probability and Statistics, Wiley, 1995.

[8] ——, *Convergence of probability measures*, Wiley Series in Probability and Statistics, Wiley, 1999.

[9] K. L. Chung, *A course in probability theory*, Probability and mathematical statistics, Academic press, 2001.

[10] J. Diestel, H. Jarchow, and A. Tonge, *Absolutely summing operators*, vol. 43, Cambridge university press, 1995.

[11] J. Diestel and J. Uhl, *Vector Measures*, Mathematical surveys and monographs, American Press, 1977.

[12] G. A. Edgar and L. Sucheston, *Stopping times and directed processes*, vol. 47, Cambridge University Press, 1992.

[13] D. P. Giesy, *On a convexity condition in normed linear spaces*, Transactions of the American Mathematical Society, 125 (1966), pp. 114–146.

[14] U. Grenander, *Probabilities on algebraic structures*, Courier Corporation, 2008.

[15] J. Hoffmann-Jørgensen, *Sums of independent banach space valued random variables*, Studia Mathematica, 52 (1974), pp. 159–186.

[16] J. Hoffmann-Jørgensen, *Probability in banach space*, in Ecole d'été de probabilités de Saint-Flour VI-1976, Springer, 1977, pp. 1–186.

[17] J. Hoffmann-Jorgensen and G. Pisier, *The law of large numbers and the central limit theorem in banach spaces*, Ann. Probab., 4 (1976), pp. 587–599.

[18] T. Hytönen, J. van Neerven, M. Veraar, and L. Weis, *Analysis in Banach Spaces: Volume II: Probabilistic Methods and Operator Theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics, Springer International Publishing, 2018.

[19] K. Itô and M. Nisio, *On the convergence of sums of independent banach space valued random variables*, Osaka J. Math., 5 (1968), pp. 35–48.

[20] J.-P. Kahane, *Some random series of functions*, vol. 5, Cambridge University Press, 1993.

[21] M. Ledoux and M. Talagrand, *Probability in Banach spaces: isoperimetry and processes*, Springer, 1991.

[22] I. J. Maddox, *Elements of functional analysis*, Cambridge University Press, 1988.

[23] B. Maurey, *Type, cotype and k-convexity*, Handbook of the geometry of Banach spaces, 2 (2003), pp. 1299–1332.

[24] E. Mourier, *Eléments aléatoires dans un espace de banach*, Ann. Inst. H. Poincaré, 13 (1953), pp. 161–244.

[25] W. Padgett and R. Taylor, *Laws of large numbers for normed linear spaces and certain Fréchet spaces*, Springer, 1973.

[26] B. J. Pettis, *On integration in vector spaces*, Transactions of the American Mathematical Society, 44 (1938), pp. 277–304.

[27] R. S. Phillips and E. Hille, *Functional analysis and semi-groups*, American Mathematical Society, 1957.

[28] G. Pisier, *Sur les espaces qui ne contiennent pas de l_n^1 uniformément*, Séminaire Analyse fonctionnelle (dit"Maurey-Schwartz"), 1973 (1974), pp. 1–19.

[29] P. Révész, *The laws of large numbers*, American Press, 1968.

[30] W. Rudin, *Functional analysis. International series in pure and applied mathematics*, McGraw-Hill, Inc., New York, 1991.

[31] V. I. Rybakov, *On conditional mathematical expectations for functions integrable in the pettis sense*, Mathematical notes of the Academy of Sciences of the USSR, 10 (1971), pp. 764–767.

[32] F. S. Scalora, *Abstract martingale convergence theorems.*, Pacific J. Math., 11 (1961), pp. 347–374.

[33] K. Sedor, *The Law of Large Numbers and its Applications*, 2015. `https://www.lakeheadu.ca/sites/default/files/uploads/77/images/Sedor%20Kelly.pdf` Fecha de consulta: 04/06/2018.

[34] M. Talagrand, *Pettis integral and measure theory*, no. 307, American Mathematical Soc., 1984.

[35] R. L. Taylor, *Weak laws of large numbers in normed linear spaces*, Ann. Math. Statist., 43 (1972), pp. 1267–1274.

[36] ——, *Stochastic convergence of weighted sums of random elements in linear spaces*, Springer, 1978.

[37] G. E. F. Thomas, *On some negative properties of the pettis integral*, Lecture notes, (1976), p. 131.

[38] N. Vakhania, V. Tarieladze, and S. Chobanyan, *Probability distributions on Banach spaces*, vol. 14, Springer Science & Business Media, 1987.