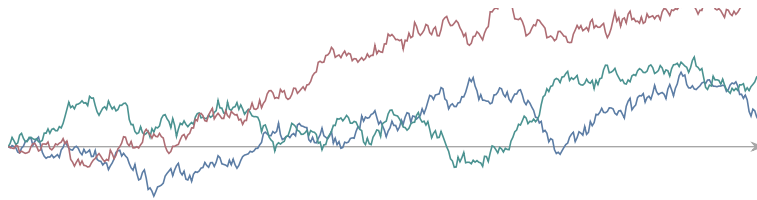


Procesos Estocásticos

Notas de clase

Martingalas, procesos de Poisson, movimiento Browniano
y cálculo estocástico



Marcelo Gallardo

`marcelo.gallardo@pucp.edu.pe`

Índice general

Prefacio	viii
Notación y convenciones	xi
I Martingalas a tiempo discreto	1
1. Esperanza condicional	2
1.1. Preliminares de teoría de la medida	2
1.2. Existencia y unicidad	4
1.3. Propiedades básicas	5
1.4. Teoremas de convergencia condicionales	8
1.5. La desigualdad de Jensen	10
2. Martingalas: definiciones y ejemplos	12
2.1. Filtraciones, procesos adaptados y predecibles	12
2.2. Martingalas, submartingalas y supermartingalas	13
2.3. Funciones convexas de martingalas	14
2.4. Ejemplos fundamentales	15
2.5. La transformada de martingalas	16
2.6. La descomposición de Doob	18
3. Convergencia de martingalas	20
3.1. Upcrossings	20
3.2. El teorema de convergencia casi segura	22
3.3. Ejemplo: procesos de ramificación	24

4. Integrabilidad uniforme y convergencia en L^1	27
4.1. Modos de convergencia	27
4.2. Integrabilidad uniforme	28
4.3. El teorema de Vitali	31
4.4. Martingalas cerradas: el último elemento	32
5. Tiempos de parada y desigualdades de Doob	35
5.1. Definiciones	35
5.2. El teorema de muestreo opcional	36
5.2.1. Un contraejemplo instructivo	38
5.2.2. El teorema general	38
5.3. Las desigualdades maximales de Doob	42
II Aplicaciones en finanzas	45
6. Mercados financieros a tiempo discreto	46
6.1. El mercado	46
6.2. Arbitraje	49
6.3. Probabilidades martingala equivalentes	50
6.4. El primer teorema fundamental	52
7. Valoración de derivados y completitud	54
7.1. Contratos	54
7.2. Precio justo	55
7.3. Cobertura y completitud	57
8. Opciones americanas y la envolvente de Snell	61
8.1. Contratos americanos	61
8.2. La envolvente de Snell	62
8.3. El tiempo óptimo de ejercicio	64
8.4. Existencia de sobrecobertura	65
8.5. La call americana no se ejerce antes	66

III	Procesos a tiempo continuo	67
9.	Trayectorias càdlàg y tiempos de parada continuos	68
9.1.	El espacio $\mathbb{D}_S$	68
9.2.	Medibilidad y la filtración canónica	70
9.3.	Tiempos de parada y tiempos opcionales	72
9.4.	Procesos medibles, adaptados y progresivos	75
10.	El proceso de Poisson	78
10.1.	La distribución de Poisson	78
10.2.	Medidas aleatorias de Poisson	79
10.3.	Primera construcción del proceso de Poisson	82
10.4.	Propiedades de las trayectorias	85
10.5.	Segunda construcción: tiempos de espera exponenciales	87
11.	Las propiedades de Markov	90
11.1.	La familia $(\mathbb{P}_x)_{x \in \mathbb{Z}_+}$	90
11.2.	La propiedad de Markov débil	91
11.3.	La propiedad de Markov fuerte	93
11.4.	Los tiempos de espera son exponenciales i.i.d.	95
12.	Vectores y procesos Gaussianos	97
12.1.	Variables Gaussianas	97
12.2.	Vectores Gaussianos	99
12.3.	Espacios y procesos Gaussianos	100
12.4.	Medidas Gaussianas (ruido blanco)	102
12.5.	Ejercicios resueltos	105
13.	El movimiento Browniano	108
13.1.	Motivación: procesos de Lévy	108
13.2.	El pre-movimiento Browniano	109
13.2.1.	La integral de Wiener	112
13.3.	La continuidad de las trayectorias	112

13.4. La medida de Wiener	114
13.5. La propiedad de Markov y la ley 0–1 de Blumenthal	116
13.6. La propiedad de Markov fuerte y el principio de reflexión	119
13.7. Variación cuadrática y no diferenciabilidad	122
13.8. Ejercicios resueltos	123
 IV Cálculo estocástico	 128
 14. Martingalas a tiempo continuo y variación cuadrática	 129
14.1. Martingalas a tiempo continuo	129
14.2. Variación cuadrática	132
14.3. Semimartingalas continuas	133
 15. La integral de Itô	 135
15.1. Procesos elementales	135
15.2. Extensión por isometría	137
15.3. Localización	139
15.4. Integración respecto de semimartingalas	140
15.5. Procesos de Itô	141
 16. La fórmula de Itô	 143
16.1. El enunciado	143
16.2. Demostración	144
16.3. Versiones multidimensionales	145
16.4. Ejemplos de cálculo	146
16.5. La caracterización de Lévy	148
 17. Ecuaciones diferenciales estocásticas	 150
17.1. Existencia y unicidad	150
17.2. Ejemplos	152
17.3. Difusiones, generador y Feynman–Kac	153
 18. Girsanov, representación de martingalas y Black–Scholes	 156

18.1. El teorema de Girsanov	156
18.2. Representación de martingalas	158
18.3. El modelo de Black–Scholes	159
18.4. La fórmula de Black–Scholes	161
18.5. Del modelo binomial a Black–Scholes	163

Apéndices **167**

A. Resultados de teoría de la medida **167**

A.1. Clases de conjuntos y unicidad de medidas	167
A.2. Integración	168
A.3. Radon–Nikodym	169
A.4. Convergencia y funciones características	169
A.5. Lema de Grönwall	170

B. Síntesis: formulario **171**

B.1. Esperanza condicional	171
B.2. Martingalas a tiempo discreto	171
B.3. Mercados a tiempo discreto	172
B.4. Procesos a tiempo continuo	173
B.5. Cálculo estocástico	174

C. Ejercicios adicionales **176**

Bibliografía **180**

Prefacio

Estas notas reúnen, corrigen y completan el material del curso de *Procesos Estocásticos* dictado en la Pontificia Universidad Católica del Perú durante el semestre 2023–2. El texto original estaba repartido en seis documentos independientes —escritos, en su mayor parte, en tiempo real durante las clases— que se solapaban parcialmente y que usaban convenciones de notación distintas. Esta edición los unifica en un solo hilo argumental.

Qué contiene

El curso recorre, en cuatro etapas, el camino que va de la esperanza condicional hasta la fórmula de Black–Scholes:

- **Parte I — Martingalas a tiempo discreto.** La esperanza condicional como objeto central, las martingalas y sus teoremas de convergencia, la integrabilidad uniforme, los tiempos de parada y las desigualdades de Doob. Es el núcleo técnico del curso.
- **Parte II — Aplicaciones en finanzas.** Mercados a tiempo discreto con Ω finito: arbitraje, probabilidades martingala equivalentes, los dos teoremas fundamentales, valoración y cobertura de derivados, y opciones americanas vía la envolvente de Snell. Toda la Parte II es, literalmente, la Parte I traducida al lenguaje del mercado.
- **Parte III — Procesos a tiempo continuo.** El espacio de trayectorias càdlàg, el proceso de Poisson (construido de dos maneras distintas), las propiedades de Markov débil y fuerte, los procesos Gaussianos y el movimiento Browniano.
- **Parte IV — Cálculo estocástico.** *Material nuevo, no cubierto en las notas originales:* martingalas a tiempo continuo y variación cuadrática, la construcción de la integral de Itô, la fórmula de Itô, ecuaciones diferenciales estocásticas, el teorema de Girsanov y, como cierre, la deducción de la fórmula de Black–Scholes —que es la Parte II llevada al límite continuo.

Qué se cambió respecto de las notas originales

1. **Notación unificada.** Las notas originales alternaban entre $\mathcal{G}_n$ y $\mathcal{F}_n$ para las filtraciones, entre $\mathcal{L}^p$ y L^p , entre Q y $\mathbb{Q}$ para la probabilidad de riesgo neutro, entre ϑ_s y θ_s para los operadores de traslación, y usaban σ simultáneamente como tiempo de parada, como σ -álgebra generada y como volatilidad. El [cuadro de notación](#) fija una única convención, respetada en todo el texto.
2. **Errores matemáticos corregidos.** Se corrigieron, entre otros: la hipótesis de los teoremas de convergencia (es $\sup_n \mathbb{E}[|X_n|] < \infty$, no $\mathbb{E}[|X_n|] < \infty$ para cada n , que es parte de la definición); la caracterización de la independencia de un vector Gaussiano (matriz de covarianzas *diagonal*, no la identidad); la suma de Gaussianas (requiere independencia); la recursión de la envolvente de Snell; el enunciado del lema $\mathcal{A} \subsetneq B_{\mathbb{Q}}$; varios signos en desarrollos de $\mathbb{E}[(X_t - X_s)^2]$ y de $\mathbb{E}[|X_{t+h} - X_t|^2]$; y la dirección de varias inclusiones y complementos. Las correcciones no se señalan en el texto: simplemente el enunciado que aparece es el correcto.
3. **Demostraciones completadas.** Varias pruebas estaban esbozadas o interrumpidas. Se completaron, en particular, el teorema de muestreo opcional, la desigualdad L^p de Doob, la construcción de la medida aleatoria de Poisson y la propiedad de Markov fuerte.
4. **Figuras rehechas en TikZ.** Todas las figuras (antes imágenes rasterizadas) se redibujaron vectorialmente, y se añadieron varias nuevas: trayectorias simuladas, el árbol binomial, diagramas de pago, el principio de reflexión, la construcción de la integral de Itô.
5. **Redundancias eliminadas.** El material que aparecía dos o tres veces en documentos distintos (esperanza condicional, martingalas, mercados) se fusionó conservando la versión más completa de cada enunciado y de cada demostración.

Cómo leer estas notas

Los capítulos 1–5 son prerrequisito de todo lo demás. Un lector interesado únicamente en finanzas puede pasar de la Parte I directamente a la Parte II y luego al [capítulo 18](#). Un lector interesado en el cálculo estocástico puede leer la Parte I, saltar la Parte II, y seguir por [capítulo 12](#), [13](#) y la Parte IV; el [capítulo 10](#) sobre el proceso de Poisson es independiente de la Parte IV, pero el [capítulo 9](#) (trayectorias càdlàg y tiempos opcionales)

sí se usa después.

Referencias

El tratamiento de martingalas a tiempo discreto sigue de cerca a Durrett y a Williams; la Parte II, a Lambertson–Lapeyre y Shreve; la Parte III y IV, a Le Gall y a Revuz–Yor. La bibliografía completa está al final.

Lima, 2023–2

Notación y convenciones

A lo largo de todo el texto se usan las convenciones siguientes. Salvo mención explícita, $(\Omega, \mathcal{F}, \mathbb{P})$ denota un espacio de probabilidad fijo.

Espacios y σ -álgebras

$(\Omega, \mathcal{F}, \mathbb{P})$	espacio de probabilidad.
$\mathcal{G}, \mathcal{H}$	sub- σ -álgebras genéricas de $\mathcal{F}$.
$(\mathcal{F}_n)_{n \in \mathbb{Z}_+}, (\mathcal{F}_t)_{t \geq 0}$	filtraciones (discreta y continua).
$\mathcal{F}_\infty$	$\sigma(\bigcup_n \mathcal{F}_n)$.
$\mathcal{F}_{t-}, \mathcal{F}_{t+}$	$\sigma(\bigcup_{s < t} \mathcal{F}_s)$ y $\bigcap_{s > t} \mathcal{F}_s$.
$\mathcal{F}_\tau$	σ -álgebra de los sucesos anteriores al tiempo de parada τ .
$\sigma(Z), \sigma(Z, W)$	σ -álgebra generada; $\sigma(Z) \vee \sigma(W) = \sigma(\sigma(Z) \cup \sigma(W))$.
“ $Z \in \mathcal{G}$ ”	abuso cómodo de lenguaje por “ Z es $\mathcal{G}$ -medible”.
$\mathcal{B}_E$	borelianos de E ; escribimos $\mathcal{B}_{\mathbb{R}}, \mathcal{B}_{\mathbb{R}^d}$.
$L^p(\Omega, \mathcal{F}, \mathbb{P})$	$\{X : \mathbb{E}[X ^p] < \infty\}$; siempre L^p , nunca $\mathcal{L}^p$.
$\mathbb{D}_S, \mathcal{C}_S$	trayectorias càdlàg (resp. continuas) de $[0, \infty)$ en S .
X_t	proceso canónico $X_t(\alpha) = \alpha(t)$ sobre $\mathbb{D}_S$ o $\mathcal{C}_S$.
θ_t	operador de traslación: $(\theta_t \alpha)(s) = \alpha(t + s)$. <i>(En las partes III y IV, θ designa también la velocidad de reversión de un Ornstein–Uhlenbeck y el precio de mercado del riesgo; el contexto lo distingue sin ambigüedad.)</i>

Probabilidad

$\mathbb{P}, \mathbb{E}$	probabilidad y esperanza de referencia.
$\mathbb{Q}, \mathbb{E}_{\mathbb{Q}}$	probabilidad martingala equivalente (PME) y su esperanza.
$\mathbb{P}_x, \mathbb{E}_x$	ley del proceso iniciado en x , y su esperanza.
$\mathbb{W}$	medida de Wiener sobre $\mathcal{C}_{\mathbb{R}}$.
$\mathbb{E}[X \mathcal{G}]$	esperanza condicional.
$\mathbb{E}[X; A]$	abreviatura de $\mathbb{E}[X \mathbf{1}_A] = \int_A X d\mathbb{P}$.
$\mathbf{1}_A$	función indicadora de A .
c. s.	casi seguramente (con probabilidad 1).
$X \perp \mathcal{G}$	$\sigma(X)$ y $\mathcal{G}$ son independientes.
$\mathcal{N}(m, \sigma^2), \text{ Poiss}(\lambda), \text{Exp}(\lambda)$	leyes normal, de Poisson y exponencial.
$\xrightarrow[\text{ley}]{\text{c. s.}}, \xrightarrow{\mathbb{P}}, \xrightarrow{L^p}$	convergencia c. s., en probabilidad, en L^p , en ley.

Procesos

M_n, X_n	martingala / proceso genérico a tiempo discreto.
τ, ρ	tiempos de parada (siempre $\tau \leq \rho$ cuando se comparan).
X^τ	proceso detenido, $X_n^\tau = X_{n \wedge \tau}$.
Γ_n	proceso predecible (Γ_n es $\mathcal{F}_{n-1}$ -medible); estrategia.
$(\Gamma \cdot M)_n$	$\sum_{k=1}^n \Gamma_k (M_k - M_{k-1})$, también escrito $\int_0^n \Gamma dM$.
$U_n^{\alpha\beta}$	número de <i>upcrossings</i> de $[\alpha, \beta]$ hasta el instante n .
N_t	proceso de Poisson; ξ_n tiempos entre saltos, $J_n = \xi_1 + \dots + \xi_n$.
B_t, W_t	movimientos Brownianos.

Tres convenciones que conviene tener presentes

1. **“Positivo” significa ≥ 0 .** Cuando decimos que una submartingala es positiva queremos decir $X_n \geq 0$; si hace falta excluir el cero lo diremos explícitamente.
2. **$\inf \emptyset = +\infty$.** Todo tiempo de parada definido como un ínfimo toma el valor $+\infty$ cuando el conjunto es vacío; por eso los tiempos de parada toman valores en $\overline{\mathbb{Z}}_+$ o en $[0, \infty]$.
3. **Igualdades entre esperanzas condicionales.** $\mathbb{E}[X \mid \mathcal{G}]$ es una *clase de equivalencia*: toda igualdad o desigualdad que la involucre vale c. s., aunque no lo escribamos.

Parte I

Martingalas a tiempo discreto

CAPÍTULO 1

Esperanza condicional

Toda la teoría de martingalas descansa sobre un único objeto: la esperanza condicional respecto de una σ -álgebra. La idea es sencilla de enunciar — $\mathbb{E}[X | \mathcal{G}]$ es la *mejor predicción de X con la información contenida en $\mathcal{G}$* — pero su construcción exige algo de teoría de la medida. Este capítulo la desarrolla y establece las propiedades que se usarán sin comentario en el resto del texto.

1.1 Preliminares de teoría de la medida

Fijemos un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ y una variable aleatoria $X : \Omega \rightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, esto es, una función medible. Dada una variable aleatoria Z escribimos

$$\sigma(Z) = \{Z^{-1}(A) : A \in \mathcal{B}_{\mathbb{R}}\} = \{\{Z \in A\} : A \in \mathcal{B}_{\mathbb{R}}\},$$

y, análogamente,

$$\sigma(Z, W) = \{(Z, W)^{-1}(A) : A \in \mathcal{B}_{\mathbb{R}^2}\} = \underbrace{\sigma(\sigma(Z) \cup \sigma(W))}_{\text{la menor } \sigma\text{-álgebra que las contiene}} =: \sigma(Z) \vee \sigma(W).$$

Notación. Dada $\mathcal{G} \subset \mathcal{F}$ una sub- σ -álgebra, si $\sigma(Z) \subset \mathcal{G}$ entonces con la información de $\mathcal{G}$ podemos responder toda pregunta sobre Z . En ese caso escribimos informalmente $Z \in \mathcal{G}$ para decir que Z es $\mathcal{G}$ -medible.

Observación. Si $Z, W \in \mathcal{G}$, entonces $Z + W \in \mathcal{G}$, $ZW \in \mathcal{G}$ y, más generalmente, $\varphi(Z, W) \in \mathcal{G}$ para toda φ boreliana.

Recordemos el siguiente resultado de teoría de la medida, que usaremos repetidamente para manipular límites de sucesiones de variables aleatorias.

Proposición 1.1. Sean $f_n : X \rightarrow \overline{\mathbb{R}}$, $n \in \mathbb{N}$, funciones medibles. Entonces

$$\sup_{n \in \mathbb{N}} f_n, \quad \inf_{n \in \mathbb{N}} f_n, \quad \limsup_n f_n, \quad \liminf_n f_n$$

son medibles.

Demostración. Escribiendo $\varphi = \sup_n f_n$ y $\psi = \inf_n f_n$, para todo $a \in \mathbb{R}$,

$$\begin{aligned}\varphi^{-1}([-\infty, a]) &= \left\{x : \sup_n f_n(x) \leq a\right\} = \bigcap_{n \in \mathbb{N}} \underbrace{\{x : f_n(x) \leq a\}}_{\in \mathcal{F}}, \\ \psi^{-1}([a, +\infty]) &= \left\{x : \inf_n f_n(x) \geq a\right\} = \bigcap_{n \in \mathbb{N}} \underbrace{\{x : f_n(x) \geq a\}}_{\in \mathcal{F}}.\end{aligned}$$

Como los intervalos $[-\infty, a]$ (resp. $[a, +\infty]$) generan $\mathcal{B}_{\mathbb{R}}$, esto prueba que φ y ψ son medibles. Luego,

$$\limsup_n f_n(x) = \inf_{k \geq 1} \sup_{n \geq k} f_n(x) \stackrel{(*)}{=} \lim_{k \rightarrow \infty} \sup_{n \geq k} f_n(x),$$

donde $(*)$ vale porque $k \mapsto \sup_{n \geq k} f_n(x)$ es *decreciente*. Ya sabemos que $\varphi_k := \sup_{n \geq k} f_n$ es medible para todo k ; en particular $\inf_{k \geq 1} \varphi_k = \limsup_n f_n$ es medible. El caso del $\liminf$ es análogo. ■

Corolario 1.2. Si $Z_1, Z_2, \dots$ son $\mathcal{G}$ -medibles, entonces $\limsup_n Z_n$ y $\liminf_n Z_n$ son $\mathcal{G}$ -medibles.

El siguiente lema es la razón por la que la esperanza condicional está bien definida *como clase de equivalencia*.

Lema 1.3. Sean Z y W variables aleatorias integrables. Entonces

$$Z \leq W \text{ c. s.} \iff \int_A Z d\mathbb{P} \leq \int_A W d\mathbb{P} \text{ para todo } A \in \mathcal{F}.$$

Demostración. $(\Rightarrow)$ Si $Z \leq W$ c.s., la monotonía de la integral da $\int_A Z \leq \int_A W$ para todo $A \in \mathcal{F}$.

$(\Leftarrow)$ Supongamos que $\int_A Z \leq \int_A W$ para todo A pero que $\mathbb{P}\{Z > W\} > 0$. Como $\{Z > W\} = \bigcup_{k \geq 1} A_k$ con $A_k = \{Z - W \geq 1/k\}$, existe k con $\mathbb{P}(A_k) > 0$, y entonces

$$\int_{A_k} Z - \int_{A_k} W = \int_{A_k} (Z - W) d\mathbb{P} \geq \frac{1}{k} \mathbb{P}(A_k) > 0,$$

lo cual contradice la hipótesis. ■

Corolario 1.4. Sean Z y W integrables. Entonces

$$Z = W \text{ c. s.} \iff \int_A Z = \int_A W \text{ para todo } A \in \mathcal{F}.$$

De hecho, basta verificar la igualdad para $A \in \sigma(Z, W)$.

1.2 Existencia y unicidad

Teorema 1.5 (Existencia). *Sea X integrable y $\mathcal{G} \subset \mathcal{F}$ una sub- σ -álgebra. Entonces existe una variable aleatoria Z tal que*

(I) Z es $\mathcal{G}$ -medible e integrable;

(II) $\int_A Z d\mathbb{P} = \int_A X d\mathbb{P}$ para todo $A \in \mathcal{G}$.

Además, si W es integrable, $\mathcal{G}$ -medible y cumple $\int_A W = \int_A X$ para todo $A \in \mathcal{G}$, entonces $Z = W$ c. s..

Demostración. Existencia. Supongamos primero $X \geq 0$ y consideremos la restricción $\mathbb{P}|_{\mathcal{G}}$. Definamos

$$\mu(A) = \int_A X d\mathbb{P}, \quad A \in \mathcal{G}.$$

Esto define una medida finita sobre $(\Omega, \mathcal{G})$ que es absolutamente continua respecto de $\mathbb{P}|_{\mathcal{G}}$: si $\mathbb{P}(A) = 0$ entonces $\mu(A) = 0$. Por el teorema de Radon–Nikodym ([teorema A.14](#)) existe $Z : \Omega \rightarrow [0, \infty)$ $\mathcal{G}$ -medible con

$$\underbrace{\mu(A)}_{\int_A X d\mathbb{P}} = \int_A Z d\mathbb{P}, \quad A \in \mathcal{G}.$$

En el caso general escribimos $X = X^+ - X^-$, obtenemos Z^+ asociada a X^+ y Z^- asociada a X^- , y tomamos $Z = Z^+ - Z^-$; nótese que $Z^\pm$ son integrables porque $\mathbb{E}[Z^\pm] = \mathbb{E}[X^\pm] < \infty$.

Unicidad. Si Z y W son ambas $\mathcal{G}$ -medibles e integrables y $\int_A Z = \int_A W$ para todo $A \in \mathcal{G}$, el [corolario 1.4](#) aplicado sobre el espacio $(\Omega, \mathcal{G}, \mathbb{P}|_{\mathcal{G}})$ da $Z = W$ c. s.. ■

Definición 1.6 (Esperanza condicional). Sea X integrable y $\mathcal{G} \subset \mathcal{F}$. Llamamos $\mathbb{E}[X | \mathcal{G}]$ a la clase de todas las variables aleatorias Z que son $\mathcal{G}$ -medibles, integrables y verifican

$$\int_A Z d\mathbb{P} = \int_A X d\mathbb{P}, \quad A \in \mathcal{G}.$$

Cada representante de esa clase es una *versión* de $\mathbb{E}[X | \mathcal{G}]$.

Observación. $\mathbb{E}[X | \mathcal{G}]$ es una *variable aleatoria*, no un número. Escribimos $\mathbb{E}[X | Y]$ como abreviatura de $\mathbb{E}[X | \sigma(Y)]$; en tal caso, por el lema de factorización de Doob–Dynkin, existe una función boreliana φ con

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

Observación. Si $X \perp \mathcal{G}$ —es decir, si $\sigma(X)$ y $\mathcal{G}$ son independientes— entonces $\mathbb{E}[X | \mathcal{G}] = \mathbb{E}[X]$. En efecto, la constante $\mathbb{E}[X]$ es $\mathcal{G}$ -medible y para $A \in \mathcal{G}$, $\int_A X = \mathbb{E}[X \mathbf{1}_A] = \mathbb{E}[X] \mathbb{P}(A) = \int_A \mathbb{E}[X]$.

Observación. Si $X \in \mathcal{G}$ entonces $\mathbb{E}[X | \mathcal{G}] = X$; el candidato $Z = X$ satisface trivialmente ambas condiciones.

Teorema 1.7 (Localización). Si $X_1 = X_2$ sobre $B \in \mathcal{G}$, entonces $\mathbb{E}[X_1 | \mathcal{G}] = \mathbb{E}[X_2 | \mathcal{G}]$ c. s. sobre B .

Demostración. Para $A \in \mathcal{G}$ con $A \subset B$ se tiene $A \in \mathcal{G}$ y $\int_A \mathbb{E}[X_1 | \mathcal{G}] = \int_A X_1 = \int_A X_2 = \int_A \mathbb{E}[X_2 | \mathcal{G}]$. Como $\mathbb{E}[X_1 | \mathcal{G}] - \mathbb{E}[X_2 | \mathcal{G}]$ es $\mathcal{G}$ -medible y su integral se anula sobre todo $A \in \mathcal{G}$ contenido en B , el corolario 1.4 aplicado sobre $(B, \mathcal{G} \cap B)$ concluye. ■

Ejemplo 1.8 (Caso con densidad conjunta). Supongamos que (X, Y) tiene densidad conjunta $f(x, y)$, es decir $\mathbb{P}\{(X, Y) \in B\} = \int_B f(x, y) dx dy$ para todo $B \in \mathcal{B}_{\mathbb{R}^2}$, y que $f_Y(y) = \int f(x, y) dx > 0$ para todo y . Si $\mathbb{E}[|g(X)|] < \infty$, entonces $\mathbb{E}[g(X) | Y] = h(Y)$ con

$$h(y) = \frac{\int g(x) f(x, y) dx}{\int f(x, y) dx}.$$

Esta es la fórmula que se ve en un primer curso; obsérvese que es un *caso particular* de la definición, no la definición.

1.3 Propiedades básicas

Proposición 1.9. Sean $X, Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ y $\mathcal{G} \subset \mathcal{F}$. Entonces:

- (a) Si X es $\mathcal{G}$ -medible, $\mathbb{E}[X | \mathcal{G}] = X$.
- (b) (Linealidad.) $X \mapsto \mathbb{E}[X | \mathcal{G}]$ es lineal en L^1 : para $\alpha, \alpha' \in \mathbb{R}$, $\mathbb{E}[\alpha X + \alpha' Y | \mathcal{G}] = \alpha \mathbb{E}[X | \mathcal{G}] + \alpha' \mathbb{E}[Y | \mathcal{G}]$.
- (c) (Monotonía.) Si $X \leq Y$ c. s., entonces $\mathbb{E}[X | \mathcal{G}] \leq \mathbb{E}[Y | \mathcal{G}]$ c. s..

(d) (Propiedad de la torre.) Si $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$,

$$\mathbb{E}[\mathbb{E}[X | \mathcal{G}] | \mathcal{H}] = \mathbb{E}[X | \mathcal{H}] \quad c. s.$$

(e) $|\mathbb{E}[X | \mathcal{G}]| \leq \mathbb{E}[|X| | \mathcal{G}]$ c. s..

(f) (Esperanza total.) $\mathbb{E}[\mathbb{E}[X | \mathcal{G}]] = \mathbb{E}[X]$.

Demostración. (a) Es la definición: si X es $\mathcal{G}$ -medible, trivialmente $\int_A X = \int_A X$ para todo $A \in \mathcal{G}$.

(b) Sean $X, X' \in L^1$ y $\alpha, \alpha' \in \mathbb{R}$. La variable $\alpha\mathbb{E}[X | \mathcal{G}] + \alpha'\mathbb{E}[X' | \mathcal{G}]$ es $\mathcal{G}$ -medible, integrable y para $A \in \mathcal{G}$

$$\int_A (\alpha\mathbb{E}[X | \mathcal{G}] + \alpha'\mathbb{E}[X' | \mathcal{G}]) = \alpha \int_A X + \alpha' \int_A X' = \int_A (\alpha X + \alpha' X'),$$

de modo que cumple la propiedad característica de $\mathbb{E}[\alpha X + \alpha' X' | \mathcal{G}]$.

(c) Para todo $A \in \mathcal{G}$, $\int_A \mathbb{E}[X | \mathcal{G}] = \int_A X \leq \int_A Y = \int_A \mathbb{E}[Y | \mathcal{G}]$, y el [lema 1.3](#) aplicado sobre $(\Omega, \mathcal{G}, \mathbb{P}|_{\mathcal{G}})$ da la conclusión.

(d) La variable $\mathbb{E}[X | \mathcal{H}]$ es $\mathcal{H}$ -medible e integrable. Para $A \in \mathcal{H} \subset \mathcal{G}$,

$$\int_A \mathbb{E}[X | \mathcal{G}] = \int_A X = \int_A \mathbb{E}[X | \mathcal{H}],$$

donde la primera igualdad usa $A \in \mathcal{G}$ y la segunda $A \in \mathcal{H}$. Luego $\mathbb{E}[X | \mathcal{H}]$ cumple la propiedad que define a $\mathbb{E}[\mathbb{E}[X | \mathcal{G}] | \mathcal{H}]$. Conviene leer esta propiedad como un enunciado sobre *información*: condicionar dos veces, quedándose al final con la información más pobre, es lo mismo que condicionar una sola vez a la información más pobre.

(e) Por (b) y (c) aplicados a $X = X^+ - X^-$,

$$|\mathbb{E}[X | \mathcal{G}]| = |\mathbb{E}[X^+ | \mathcal{G}] - \mathbb{E}[X^- | \mathcal{G}]| \leq \mathbb{E}[X^+ | \mathcal{G}] + \mathbb{E}[X^- | \mathcal{G}] = \mathbb{E}[|X| | \mathcal{G}].$$

(f) Tomar $A = \Omega \in \mathcal{G}$ en la definición. ■

Teorema 1.10 (Sacar lo conocido). Sea X una variable aleatoria $\mathcal{G}$ -medible y Y tal que $\mathbb{E}[|Y|] < \infty$ y $\mathbb{E}[|XY|] < \infty$. Entonces

$$\mathbb{E}[XY | \mathcal{G}] = X \mathbb{E}[Y | \mathcal{G}] \quad c. s.$$

En particular, el resultado vale si X es $\mathcal{G}$ -medible y acotada, o si $X, Y \geq 0$ y X es $\mathcal{G}$ -medible.

Demostración. Por el argumento estándar de máquina, basta probarlo para $X = \mathbf{1}_B$ con $B \in \mathcal{G}$, extenderlo por linealidad a X simple $\mathcal{G}$ -medible, pasar al límite monótono para $X \geq 0$ $\mathcal{G}$ -medible, y finalmente descomponer $X = X^+ - X^-$, $Y = Y^+ - Y^-$.

Para $X = \mathbf{1}_B$ y $A \in \mathcal{G}$, como $A \cap B \in \mathcal{G}$,

$$\int_A \mathbf{1}_B \mathbb{E}[Y | \mathcal{G}] = \int_{A \cap B} \mathbb{E}[Y | \mathcal{G}] = \int_{A \cap B} Y = \int_A \mathbf{1}_B Y,$$

y $\mathbf{1}_B \mathbb{E}[Y | \mathcal{G}]$ es $\mathcal{G}$ -medible; esto es exactamente la propiedad que caracteriza a $\mathbb{E}[\mathbf{1}_B Y | \mathcal{G}]$. ■

Teorema 1.11 (Interpretación L^2 : mejor predictor). *Supongamos $\mathbb{E}[X^2] < \infty$. Entonces $\mathbb{E}[X | \mathcal{G}]$ es la única variable $\mathcal{G}$ -medible $Y \in L^2$ que minimiza el error cuadrático medio $\mathbb{E}[(X - Y)^2]$. Es decir, $\mathbb{E}[\cdot | \mathcal{G}]$ es la proyección ortogonal de $L^2(\Omega, \mathcal{F}, \mathbb{P})$ sobre el subespacio cerrado $L^2(\Omega, \mathcal{G}, \mathbb{P})$.*

Demostración. Sea $Z = \mathbb{E}[X | \mathcal{G}]$; por Jensen (lema 1.14) $Z \in L^2$. Para $Y \in L^2(\Omega, \mathcal{G}, \mathbb{P})$,

$$\mathbb{E}[(X - Y)^2] = \mathbb{E}[(X - Z)^2] + 2\mathbb{E}[(X - Z)(Z - Y)] + \mathbb{E}[(Z - Y)^2].$$

El término cruzado se anula: $Z - Y$ es $\mathcal{G}$ -medible y acotada por truncamiento, luego por el teorema 1.10 y la torre

$$\mathbb{E}[(X - Z)(Z - Y)] = \mathbb{E}[\mathbb{E}[(X - Z)(Z - Y) | \mathcal{G}]] = \mathbb{E}\left[(Z - Y) \underbrace{\mathbb{E}[X - Z | \mathcal{G}]}_{=0}\right] = 0.$$

Por lo tanto $\mathbb{E}[(X - Y)^2] = \mathbb{E}[(X - Z)^2] + \mathbb{E}[(Z - Y)^2] \geq \mathbb{E}[(X - Z)^2]$, con igualdad si y solo si $Y = Z$ c. s.. ■

La idea en una frase

$\mathbb{E}[X | \mathcal{G}]$ es la variable aleatoria que “mejor imita” a X usando solo la información de $\mathcal{G}$: es $\mathcal{G}$ -medible y tiene las mismas integrales que X sobre todos los sucesos de $\mathcal{G}$. En L^2 eso es literalmente una proyección ortogonal.

1.4 Teoremas de convergencia condicionales

Los tres teoremas de convergencia de la integral de Lebesgue tienen su versión condicional. En lo que sigue, “no negativa” significa que la variable toma valores en $[0, \infty]$; para tales variables $\mathbb{E}[X | \mathcal{G}]$ se define por el mismo procedimiento (Radon–Nikodym para medidas σ -finitas), y toma valores en $[0, \infty]$.

Teorema 1.12. *Sea X una variable aleatoria con valores en $[0, \infty]$. Existe una variable aleatoria $\mathcal{G}$ -medible con valores en $[0, \infty]$, denotada $\mathbb{E}[X | \mathcal{G}]$, tal que para toda $Z \geq 0$ $\mathcal{G}$ -medible*

$$\mathbb{E}[XZ] = \mathbb{E}[\mathbb{E}[X | \mathcal{G}] Z].$$

Además $\mathbb{E}[X | \mathcal{G}]$ es única salvo en un conjunto de probabilidad cero.

Proposición 1.13 (Propiedades para variables no negativas). *Sean X, X' no negativas y $a, b \geq 0$. Entonces:*

- (a) $\mathbb{E}[aX + bX' | \mathcal{G}] = a\mathbb{E}[X | \mathcal{G}] + b\mathbb{E}[X' | \mathcal{G}]$.
- (b) Si X es $\mathcal{G}$ -medible, $\mathbb{E}[X | \mathcal{G}] = X$.
- (c) $\mathbb{E}[\mathbb{E}[X | \mathcal{G}]] = \mathbb{E}[X]$.
- (d) (Convergencia monótona condicional.) Si $X_n \uparrow X$ con $X_n \geq 0$, entonces $\mathbb{E}[X_n | \mathcal{G}] \uparrow \mathbb{E}[X | \mathcal{G}]$ c. s.. En consecuencia, para (Y_n) no negativas,

$$\mathbb{E}\left[\sum_{n \in \mathbb{N}} Y_n \mid \mathcal{G}\right] = \sum_{n \in \mathbb{N}} \mathbb{E}[Y_n | \mathcal{G}].$$

- (e) (Fatou condicional.) Si $X_n \geq 0$,

$$\mathbb{E}\left[\liminf_n X_n \mid \mathcal{G}\right] \leq \liminf_n \mathbb{E}[X_n | \mathcal{G}] \quad \text{c. s.}$$

- (f) (Convergencia dominada condicional.) Si $X_n \rightarrow X$ c. s. y existe $Z \geq 0$ integrable con $|X_n| \leq Z$ c. s. para todo n , entonces $\mathbb{E}[X_n | \mathcal{G}] \rightarrow \mathbb{E}[X | \mathcal{G}]$ c. s. y en L^1 .
- (g) (Jensen.) Si $\varphi : \mathbb{R} \rightarrow \mathbb{R}_+$ es convexa y $X \in L^1$, entonces $\mathbb{E}[\varphi(X) | \mathcal{G}] \geq \varphi(\mathbb{E}[X | \mathcal{G}])$.

Demostración. Nos concentramos en los teoremas de convergencia; (a)–(c) son inmediatos de la definición.

(d) *Convergencia monótona.* Sea $X_n \uparrow X$. Para todo $A \in \mathcal{G}$,

$$0 \leq X_n \mathbf{1}_A \uparrow X \mathbf{1}_A,$$

de modo que, por el teorema de convergencia monótona usual,

$$\underbrace{\int_A X_n \mathbf{1}_A}_{=\int_A \mathbb{E}[X_n | \mathcal{G}]} \longrightarrow \underbrace{\int_A X \mathbf{1}_A}_{=\int_A \mathbb{E}[X | \mathcal{G}]} . \quad (1.1)$$

Por otro lado, poniendo $Z_n := \mathbb{E}[X_n | \mathcal{G}]$, la monotonía (proposición 1.9(c)) da $Z_1 \leq Z_2 \leq \dots$ c. s., de modo que $\Omega^* := \bigcap_{k \geq 1} \{Z_k \leq Z_{k+1}\}$ cumple $\mathbb{P}(\Omega^*) = 1$. Podemos entonces definir

$$\eta = \begin{cases} \lim_n Z_n, & \text{si } \omega \in \Omega^*, \\ 0, & \text{en caso contrario,} \end{cases}$$

que es $\mathcal{G}$ -medible con valores en $[0, \infty]$. Para $A \in \mathcal{G}$ se tiene $Z_n \mathbf{1}_A \uparrow \eta \mathbf{1}_A$ y, nuevamente por convergencia monótona, $\int Z_n \mathbf{1}_A \rightarrow \int \eta \mathbf{1}_A$. Comparando con (1.1),

$$\int_A \mathbb{E}[X | \mathcal{G}] = \int_A \eta \quad \text{para todo } A \in \mathcal{G},$$

y por lo tanto $\mathbb{E}[X | \mathcal{G}] = \eta = \lim_n \mathbb{E}[X_n | \mathcal{G}]$ c. s..

(e) *Fatou.* Aplicando (d) a $\inf_{n \geq k} X_n \uparrow \liminf_n X_n$ y la monotonía,

$$\begin{aligned} \mathbb{E}[\liminf_n X_n | \mathcal{G}] &= \mathbb{E}\left[\lim_k \inf_{n \geq k} X_n \middle| \mathcal{G}\right] = \lim_k \mathbb{E}\left[\inf_{n \geq k} X_n \middle| \mathcal{G}\right] \\ &\leq \lim_k \inf_{n \geq k} \mathbb{E}[X_n | \mathcal{G}] = \liminf_n \mathbb{E}[X_n | \mathcal{G}]. \end{aligned}$$

(f) *Convergencia dominada.* Como $X_n \rightarrow X$ c. s. y $|X_n| \leq Z$, también $|X| \leq Z$ y X es integrable. Las variables $Z \pm X_n$ son no negativas, de modo que por (e),

$$\begin{aligned} \mathbb{E}[Z - X | \mathcal{G}] &= \mathbb{E}[\liminf_n (Z - X_n) | \mathcal{G}] \leq \mathbb{E}[Z | \mathcal{G}] - \limsup_n \mathbb{E}[X_n | \mathcal{G}], \\ \mathbb{E}[Z + X | \mathcal{G}] &= \mathbb{E}[\liminf_n (Z + X_n) | \mathcal{G}] \leq \mathbb{E}[Z | \mathcal{G}] + \liminf_n \mathbb{E}[X_n | \mathcal{G}]. \end{aligned}$$

Cancelando $\mathbb{E}[Z | \mathcal{G}]$ (que es finita c. s. pues Z es integrable) y combinando,

$$\mathbb{E}[X | \mathcal{G}] \leq \liminf_n \mathbb{E}[X_n | \mathcal{G}] \leq \limsup_n \mathbb{E}[X_n | \mathcal{G}] \leq \mathbb{E}[X | \mathcal{G}],$$

es decir $\mathbb{E}[X_n | \mathcal{G}] \rightarrow \mathbb{E}[X | \mathcal{G}]$ c. s.. Para la convergencia en L^1 , sea

$$\mathcal{U}_N = \sup\{|X_n - X_m| : n, m \geq N\}.$$

Como $X_n \rightarrow X$ c. s., la sucesión (X_n) es de Cauchy c. s. y $\mathcal{U}_N \downarrow 0$ c. s., con $0 \leq \mathcal{U}_N \leq 2Z$ integrable. Por convergencia monótona (aplicada a $2Z - \mathcal{U}_N \uparrow 2Z$), $\mathbb{E}[\mathcal{U}_N | \mathcal{G}] \rightarrow 0$ c. s. y, tomando esperanzas, $\mathbb{E}[\mathcal{U}_N] \rightarrow 0$. Como

$$|\mathbb{E}[X_n | \mathcal{G}] - \mathbb{E}[X | \mathcal{G}]| \leq \mathbb{E}[|X_n - X| | \mathcal{G}] \leq \mathbb{E}[\mathcal{U}_N | \mathcal{G}] \quad (n \geq N),$$

tomando esperanzas se obtiene $\mathbb{E}[|\mathbb{E}[X_n | \mathcal{G}] - \mathbb{E}[X | \mathcal{G}]|] \leq \mathbb{E}[\mathcal{U}_N] \rightarrow 0$. ■

Observación. El teorema de convergencia monótona condicional sigue siendo válido para $X_n \downarrow X$ siempre que X_1 sea integrable.

1.5 La desigualdad de Jensen

Lema 1.14 (Desigualdad de Jensen condicional). *Sea X integrable, $\mathcal{G} \subset \mathcal{F}$ una sub- σ -álgebra y $\varphi : \mathbb{R} \rightarrow (-\infty, +\infty]$ convexa¹ tal que $\varphi(X)$ es integrable o $\varphi \geq 0$. Entonces*

$$\varphi(\mathbb{E}[X | \mathcal{G}]) \leq \mathbb{E}[\varphi(X) | \mathcal{G}] \quad \text{c. s.}$$

Demostración. La clave es que una función convexa es el supremo de las rectas afines que la minoran, y que ese supremo puede tomarse sobre una familia numerable. Definamos

$$\Lambda_\varphi = \{T_{a,b} : T_{a,b}(x) = ax + b \leq \varphi(x) \ \forall x \in \mathbb{R}, \ a, b \in \mathbb{Q}\},$$

que es numerable, no vacío, y verifica

$$\varphi(x) = \left(\sup_{T \in \Lambda_\varphi} T \right)(x), \quad x \in \mathbb{R}.$$

Para cada $T = T_{a,b} \in \Lambda_\varphi$ se tiene $\varphi(X) \geq T(X)$, de modo que, por monotonía y linealidad,

$$\underbrace{\mathbb{E}[\varphi(X) | \mathcal{G}]}_{=: W} \geq \mathbb{E}[T(X) | \mathcal{G}] = a\mathbb{E}[X | \mathcal{G}] + b = T\left(\underbrace{\mathbb{E}[X | \mathcal{G}]}_{=: Z}\right) \quad \text{c. s.}$$

¹Siempre es posible extender una función convexa definida en un intervalo a toda la recta asignándole el valor $+\infty$ fuera de él.

Es decir, para cada T existe $\Omega_T \subset \Omega$ con $\mathbb{P}(\Omega_T) = 1$ y $W(\omega) \geq T(Z(\omega))$ para $\omega \in \Omega_T$. Como Λ_φ es numerable,

$$\Omega^* = \bigcap_{T \in \Lambda_\varphi} \Omega_T \quad \text{cumple} \quad \mathbb{P}(\Omega^*) = 1,$$

y sobre Ω^* vale $W(\omega) \geq T(Z(\omega))$ para todo $T \in \Lambda_\varphi$ simultáneamente. Tomando supremo en T ,

$$W(\omega) \geq \sup_{T \in \Lambda_\varphi} T(Z(\omega)) = \varphi(Z(\omega)), \quad \omega \in \Omega^*. \quad \blacksquare$$

Por qué la numerabilidad importa

La desigualdad $W \geq T(Z)$ vale c. s. *para cada T fijo*, con un conjunto excepcional que depende de T . Si Λ_φ fuera no numerable, la unión de esos conjuntos excepcionales podría tener probabilidad positiva —o ni siquiera ser medible— y el argumento se caería. Este tipo de cuidado reaparecerá varias veces (por ejemplo, al probar el teorema de convergencia de martingalas en el [capítulo 3](#), donde el supremo se toma sobre pares de racionales).

Martingalas: definiciones y ejemplos

Una martingala modela un juego *justo*: conocida toda la historia hasta el instante n , la mejor predicción del capital futuro es el capital actual. Esta idea, formalizada con la esperanza condicional del capítulo anterior, resulta ser sorprendentemente potente: de ella saldrán teoremas de convergencia, desigualdades maximales y, en la [parte II](#), la noción misma de precio justo.

2.1 Filtraciones, procesos adaptados y predecibles

Definición 2.1 (Filtración). Una *filtración* sobre $(\Omega, \mathcal{F}, \mathbb{P})$ es una familia creciente de sub- σ -álgebras

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}.$$

El parámetro $n \in \mathbb{Z}_+$ se interpreta como el tiempo, y $\mathcal{F}_n$ como la información disponible en el instante n . Al par $((\Omega, \mathcal{F}, \mathbb{P}), (\mathcal{F}_n)_{n \in \mathbb{Z}_+})$ lo llamamos *espacio de probabilidad filtrado*.

Definición 2.2. Un proceso $(X_n)_{n \in \mathbb{Z}_+}$ es

- *adaptado* a $(\mathcal{F}_n)$ si X_n es $\mathcal{F}_n$ -medible para todo n ;
- *predecible* si X_0 es $\mathcal{F}_0$ -medible y X_n es $\mathcal{F}_{n-1}$ -medible para todo $n \geq 1$.

Adaptado significa “el valor de X_n se conoce en el instante n ”; predecible, “se conoce ya en el instante $n - 1$ ”. La distinción será crucial: una estrategia de inversión debe ser predecible (se decide *antes* de observar el movimiento del mercado), mientras que el valor de la cartera es meramente adaptado.

Ejemplo 2.3 (Filtración canónica). Si $(X_n)_{n \in \mathbb{Z}_+}$ es un proceso aleatorio, para cada n definimos

$$\mathcal{F}_n^X = \sigma(X_0, X_1, \dots, X_n),$$

la menor σ -álgebra que hace medibles a $X_0, \dots, X_n$. Entonces $(\mathcal{F}_n^X)_{n \in \mathbb{Z}_+}$ es una filtración, la *filtración canónica* de (X_n) : es la más pequeña a la que (X_n) está adaptado.

Ejemplo 2.4 (Filtración diádica). Sea $\Omega = [0, 1)$, $\mathcal{F}$ los borelianos de $[0, 1)$ y $\mathbb{P}$ la medida de Lebesgue. Para cada $n \in \mathbb{Z}_+$ definimos

$$\mathcal{F}_n = \sigma \left(\left[\frac{i-1}{2^n}, \frac{i}{2^n} \right) : i \in \{1, 2, \dots, 2^n\} \right).$$

Entonces $(\mathcal{F}_n)_{n \in \mathbb{Z}_+}$ es una filtración, conocida como *filtración diádica*: $\mathcal{F}_n$ codifica el conocimiento de los primeros n dígitos binarios de ω .

2.2 Martingalas, submartingalas y supermartingalas

Definición 2.5 (Martingala). Sea $((\Omega, \mathcal{F}, \mathbb{P}), (\mathcal{F}_n)_{n \in \mathbb{Z}_+})$ un espacio filtrado y $M_n : \Omega \rightarrow \mathbb{R}$. Decimos que $(M_n, \mathcal{F}_n)_{n \in \mathbb{Z}_+}$ es una *martingala* si

- (I) (M_n) está adaptado a $(\mathcal{F}_n)$;
- (II) $\mathbb{E}[|M_n|] < \infty$ para todo $n \in \mathbb{Z}_+$;
- (III) para todo $0 \leq n \leq m$,

$$\mathbb{E}[M_m | \mathcal{F}_n] = M_n \quad \text{c. s.}$$

Si en (iii) se reemplaza $=$ por $\geq$ se trata de una *submartingala*; si se reemplaza por $\leq$, de una *supermartingala*.

Un truco mnemotécnico

La terminología es contraintuitiva: una *submartingala* *crece* en media ($\mathbb{E}[X_n] \leq \mathbb{E}[X_{n+1}]$) y una *supermartingala* *decrece*. El nombre viene de la teoría del potencial, no del comportamiento del proceso. Para recordarlo: “la vida de un jugador es supermartingala”.

Observación 2.6 (Basta un paso). Para verificar (iii) basta probar

$$M_n = \mathbb{E}[M_{n+1} | \mathcal{F}_n] \quad \text{para todo } n.$$

En efecto, aplicando la propiedad de la torre reiteradamente,

$$\mathbb{E}[M_m | \mathcal{F}_n] = \mathbb{E} \left[\underbrace{\dots \mathbb{E} \left[\underbrace{\mathbb{E}[M_m | \mathcal{F}_{m-1}] | \mathcal{F}_{m-2}}_{=M_{m-1}} \dots}_{=M_{m-2}} \middle| \mathcal{F}_n \right] = M_n.$$

Lo mismo vale para sub- y supermartingalas, usando además la monotonía.

Observación 2.7 (Formulación integral). La condición $M_n = \mathbb{E}[M_m | \mathcal{F}_n]$ equivale a

$$\mathbb{E}[M_n \mathbf{1}_A] = \mathbb{E}[M_m \mathbf{1}_A] \quad \text{para todo } A \in \mathcal{F}_n,$$

es decir, $\mathbb{E}[(M_m - M_n) \mathbf{1}_A] = 0$ para todo $A \in \mathcal{F}_n$. En el caso submartingala la igualdad se reemplaza por

$$\int_A M_n \leq \int_A M_m \quad \text{para todo } A \in \mathcal{F}_n,$$

y en el caso supermartingala por $\geq$. Esta formulación, sin esperanzas condicionales, es a menudo la más cómoda para las demostraciones.

Proposición 2.8 (Reglas de cálculo). Sean (X_n) , (Y_n) procesos adaptados e integrables. Entonces:

- (a) (X_n) es supermartingala $\iff (-X_n)$ es submartingala.
- (b) (X_n) es martingala $\iff$ es simultáneamente sub- y supermartingala.
- (c) Si (X_n) e (Y_n) son submartingalas (resp. supermartingalas), $(X_n + Y_n)$ lo es.
- (d) Si (M_n) y (N_n) son martingalas, $\alpha M_n + \beta N_n$ lo es para cualesquiera $\alpha, \beta \in \mathbb{R}$.
- (e) Si X es $\mathcal{F}_0$ -medible e integrable, $M_n \equiv X$ es martingala.
- (f) Si (X_n) es submartingala entonces $n \mapsto \mathbb{E}[X_n]$ es creciente; si es martingala, $\mathbb{E}[X_n] = \mathbb{E}[X_0]$ para todo n .

Demostración. (a)–(e) son consecuencias inmediatas de la linealidad y la monotonía de la esperanza condicional. Para (f), tomar esperanzas en (iii) y usar $\mathbb{E}[\mathbb{E}[X_m | \mathcal{F}_n]] = \mathbb{E}[X_m]$. ■

2.3 Funciones convexas de martingalas

Las dos proposiciones siguientes son la fuente principal de submartingalas.

Proposición 2.9. Sea (M_n) una martingala y $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ convexa con $\mathbb{E}[|\varphi(M_n)|] < \infty$ para todo n . Entonces $X_n = \varphi(M_n)$ es una submartingala.

Demostración. Por la propiedad de martingala y la desigualdad de Jensen condicional (lema 1.14), para $n \leq m$,

$$\underbrace{\varphi(M_n)}_{=X_n} = \varphi(\mathbb{E}[M_m | \mathcal{F}_n]) \leq \underbrace{\mathbb{E}[\varphi(M_m) | \mathcal{F}_n]}_{=\mathbb{E}[X_m | \mathcal{F}_n]}. \quad \blacksquare$$

Proposición 2.10. Sea (X_n) una submartingala y $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ convexa y creciente con $\mathbb{E}[|\varphi(X_n)|] < \infty$ para todo n . Entonces $Y_n = \varphi(X_n)$ es una submartingala.

Demostración. Por ser (X_n) submartingala, $X_n \leq \mathbb{E}[X_m | \mathcal{F}_n]$. Como φ es creciente,

$$\varphi(X_n) \leq \varphi(\mathbb{E}[X_m | \mathcal{F}_n]),$$

y aplicando Jensen al miembro derecho,

$$\varphi(X_n) \leq \varphi(\mathbb{E}[X_m | \mathcal{F}_n]) \leq \mathbb{E}[\varphi(X_m) | \mathcal{F}_n]. \quad \blacksquare$$

Ejemplo 2.11. Consecuencias inmediatas que usaremos muchas veces:

- Si (M_n) es martingala, $(|M_n|)$ es submartingala ($\varphi(x) = |x|$).
- Si (M_n) es martingala con $M_n \in L^p$, $p \geq 1$, entonces $(|M_n|^p)$ es submartingala.
- Si (X_n) es submartingala y $\alpha \in \mathbb{R}$, entonces $((X_n - \alpha)^+)$ es submartingala, pues $x \mapsto (x - \alpha)^+$ es convexa y creciente. Este caso es la clave del lema de *upcrossings* del capítulo 3.

2.4 Ejemplos fundamentales

Ejemplo 2.12 (Sumas de variables centradas). Sean $\xi_1, \xi_2, \dots$ variables aleatorias independientes, integrables y de media cero. Definamos

$$S_0 = 0, \quad S_n = \sum_{j=1}^n \xi_j, \quad \mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n).$$

Entonces $(S_n, \mathcal{F}_n)$ es una martingala. En efecto, para $A \in \mathcal{F}_n$ y $m > n$, como $\mathbf{1}_A$ es $\mathcal{F}_n$ -medible y $\xi_{n+1}, \dots, \xi_m$ son independientes de $\mathcal{F}_n$,

$$\mathbb{E}[(S_m - S_n)\mathbf{1}_A] = \mathbb{E}\left[\sum_{j=n+1}^m \xi_j \mathbf{1}_A\right] = \mathbb{E}\left[\sum_{j=n+1}^m \xi_j\right] \cdot \mathbb{E}[\mathbf{1}_A] = 0 \cdot \mathbb{P}(A) = 0.$$

Por la observación 2.7 concluimos.

Ejemplo 2.13 (Productos de variables de media uno). Sean $\xi_1, \xi_2, \dots$ independientes, integrables, con $\mathbb{E}[\xi_j] = 1$ para todo j . Definiendo

$$M_0 = 1, \quad M_n = \prod_{j=1}^n \xi_j, \quad \mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n),$$

$(M_n, \mathcal{F}_n)$ es una martingala: $\mathbb{E}[M_{n+1} | \mathcal{F}_n] = M_n \mathbb{E}[\xi_{n+1} | \mathcal{F}_n] = M_n \mathbb{E}[\xi_{n+1}] = M_n$, donde se usó el [teorema 1.10](#). Si además $\xi_j \geq 0$, la martingala es positiva; este es el prototipo de las *densidades* y reaparecerá en el teorema de Girsanov ([capítulo 18](#)).

Ejemplo 2.14 (Martingala cerrada). Sea Z integrable y $(\mathcal{F}_n)$ una filtración cualquiera. Entonces

$$M_n = \mathbb{E}[Z | \mathcal{F}_n]$$

es una martingala, por la propiedad de la torre. Veremos en el [capítulo 4](#) que toda martingala uniformemente integrable es de esta forma.

Ejemplo 2.15 (Apuestas de un jugador). Un jugador apuesta repetidamente a cara o cruz con una moneda equilibrada, ganando o perdiendo una unidad. Su fortuna S_n es la martingala del [ejemplo 2.12](#). Si la moneda está sesgada en su contra, $\mathbb{E}[\xi_j] < 0$ y (S_n) es supermartingala: por muy ingeniosa que sea su estrategia, no podrá convertirla en martingala. Este es el contenido preciso de la [proposición 2.17](#) de más abajo.

2.5 La transformada de martingalas

Introducimos ahora el análogo discreto de la integral estocástica. Es una construcción elemental, pero toda la Parte IV consistirá en darle sentido a su versión continua.

Definición 2.16 (Transformada de martingalas). Sean (M_n) un proceso adaptado y $(\Gamma_n)_{n \geq 1}$ un proceso predecible. La *transformada de M por Γ* es el proceso

$$(\Gamma \cdot M)_0 = 0, \quad (\Gamma \cdot M)_n = \sum_{k=1}^n \Gamma_k (M_k - M_{k-1}), \quad n \geq 1.$$

También escribimos $\int_0^n \Gamma dM := (\Gamma \cdot M)_n$.

La interpretación es transparente: si M_k es el precio de un activo y Γ_k el número de unidades que el inversor decidió tener durante el período $(k-1, k]$ —decisión tomada en $k-1$, de ahí la predecibilidad—, entonces $(\Gamma \cdot M)_n$ es la ganancia acumulada hasta n .

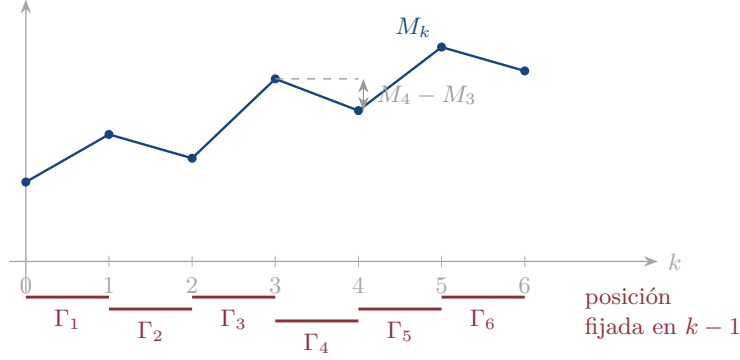


Figura 2.1: La transformada $(\Gamma \cdot M)_n = \sum_{k \leq n} \Gamma_k (M_k - M_{k-1})$. La posición Γ_k se mantiene constante durante el intervalo $(k-1, k]$ y se decide con la información de $\mathcal{F}_{k-1}$: el proceso (Γ_n) es predecible, no meramente adaptado.

Proposición 2.17 (Estabilidad de la transformada). Sea $((\Omega, \mathcal{F}, \mathbb{P}), (\mathcal{F}_n))$ un espacio filtrado y (Γ_n) predecible y acotado. Entonces:

- (a) Si (M_n) es martingala, $(\Gamma \cdot M)$ es martingala.
- (b) Si (X_n) es submartingala (resp. supermartingala) y además $\Gamma_n \geq 0$, entonces $(\Gamma \cdot X)$ es submartingala (resp. supermartingala).

Demostración. El proceso $(\Gamma \cdot M)$ es adaptado e integrable (pues Γ es acotado y $M_k \in L^1$). Para el incremento, como Γ_n es $\mathcal{F}_{n-1}$ -medible y acotada, el teorema 1.10 da

$$\begin{aligned} \mathbb{E}[(\Gamma \cdot M)_n - (\Gamma \cdot M)_{n-1} \mid \mathcal{F}_{n-1}] &= \mathbb{E}[\Gamma_n (M_n - M_{n-1}) \mid \mathcal{F}_{n-1}] \\ &= \Gamma_n \mathbb{E}[M_n - M_{n-1} \mid \mathcal{F}_{n-1}]. \end{aligned}$$

Si (M_n) es martingala, el último factor es 0 y se obtiene (a). Si (X_n) es submartingala el factor es ≥ 0 ; multiplicado por $\Gamma_n \geq 0$ sigue siendo ≥ 0 , lo que da (b). El caso supermartingala es análogo. ■

Lectura financiera de la proposición

Con una martingala no hay estrategia *predecible* que produzca ganancia esperada positiva: $\mathbb{E}[(\Gamma \cdot M)_n] = 0$ cualquiera sea Γ . Es la formulación matemática de “no se puede ganarle a un juego justo”, y será el corazón del primer teorema fundamental de las finanzas (teorema 6.15). Nótese que la predecibilidad es esencial: si se permitiera Γ_k $\mathcal{F}_k$ -medible, bastaría

tomar $\Gamma_k = \text{sgn}(M_k - M_{k-1})$ para ganar siempre.

2.6 La descomposición de Doob

Toda submartingala se descompone, de manera única, en una martingala más una parte creciente predecible. Es el análogo discreto —y elemental— del teorema de Doob–Meyer, y lo usaremos para las opciones americanas (capítulo 8).

Teorema 2.18 (Descomposición de Doob). *Sea $(X_n)_{n \in \mathbb{Z}_+}$ una submartingala. Entonces existe una única descomposición*

$$X_n = M_n + A_n, \quad n \in \mathbb{Z}_+,$$

donde (M_n) es una martingala con $M_0 = X_0$ y (A_n) es un proceso predecible, creciente, con $A_0 = 0$.

Demostración. Existencia. Definamos $A_0 = 0$ y, recursivamente,

$$A_n - A_{n-1} := \mathbb{E}[X_n - X_{n-1} \mid \mathcal{F}_{n-1}] \geq 0, \quad n \geq 1,$$

donde la desigualdad es la propiedad de submartingala. Por construcción A_n es $\mathcal{F}_{n-1}$ -medible, es decir predecible, y creciente. Poniendo $M_n = X_n - A_n$ se tiene $M_0 = X_0$ y

$$\mathbb{E}[M_n - M_{n-1} \mid \mathcal{F}_{n-1}] = \mathbb{E}[X_n - X_{n-1} \mid \mathcal{F}_{n-1}] - (A_n - A_{n-1}) = 0,$$

usando que $A_n - A_{n-1}$ es $\mathcal{F}_{n-1}$ -medible. Luego (M_n) es martingala.

Unicidad. Si $X_n = M_n + A_n = \tilde{M}_n + \tilde{A}_n$ con ambas descomposiciones del tipo indicado, entonces $D_n := A_n - \tilde{A}_n = \tilde{M}_n - M_n$ es a la vez predecible y martingala, con $D_0 = 0$. Por predecibilidad, $D_n = \mathbb{E}[D_n \mid \mathcal{F}_{n-1}] = D_{n-1}$, luego $D_n \equiv D_0 = 0$. ■

Observación. Para supermartingalas se obtiene $X_n = M_n - A_n$ con (A_n) predecible creciente y $A_0 = 0$: basta aplicar el teorema a $(-X_n)$.

Ejemplo 2.19 (El compensador de M^2). Sea (M_n) una martingala con $M_n \in L^2$. Por el ejemplo 2.11, (M_n^2) es una submartingala; su descomposición de Doob es $M_n^2 = N_n + \langle M \rangle_n$ con

$$\langle M \rangle_n = \sum_{k=1}^n \mathbb{E}[(M_k - M_{k-1})^2 \mid \mathcal{F}_{k-1}],$$

pues $\mathbb{E}[M_k^2 - M_{k-1}^2 \mid \mathcal{F}_{k-1}] = \mathbb{E}[(M_k - M_{k-1})^2 \mid \mathcal{F}_{k-1}]$ (desarrollar el cuadrado y usar la propiedad de martingala). El proceso $\langle M \rangle$ es la *variación cuadrática predecible* de M ; su análogo continuo, construido en el [capítulo 14](#), es el objeto que hace funcionar toda la integración estocástica.

CAPÍTULO 3

Convergencia de martingalas

El teorema central de la teoría es que una submartingala acotada en L^1 converge casi seguramente. La demostración es un pequeño prodigio: se reduce a controlar cuántas veces el proceso puede subir de un nivel α a un nivel $\beta > \alpha$, y ese control se obtiene aplicando la [proposición 2.17](#) a una estrategia de compraventa absolutamente explícita.

3.1 Upcrossings

Sea (X_n) un proceso y $\alpha < \beta$ dos niveles. Un *upcrossing* (cruce ascendente) del intervalo $[\alpha, \beta]$ es un tramo de la trayectoria que empieza por debajo de α y termina por encima de β . Denotamos por $U_n^{\alpha\beta}$ el número de upcrossings completados por $X_0, X_1, \dots, X_n$.

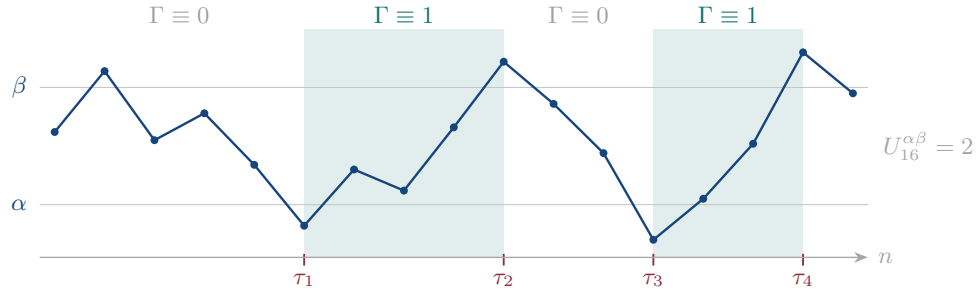


Figura 3.1: Dos upcrossings del intervalo $[\alpha, \beta]$. La estrategia Γ del [lema 3.1](#) vale 1 exactamente en las zonas sombreadas (“comprar al bajar de α , vender al subir de β ”) y 0 fuera de ellas. Cada upcrossing completado aporta al menos $\beta - \alpha$ a la ganancia ($\Gamma \cdot X$).

Lema 3.1 (Desigualdad de upcrossings, caso normalizado). *Sea (X_n) una submartingala positiva y U_n el número de upcrossings de $[0, 1]$ hasta el instante n . Entonces*

$$\mathbb{E}[U_n] \leq \mathbb{E}[X_n - X_0].$$

Demostración. Definimos los tiempos de parada

$$\begin{aligned}\tau_0 &= 0, \\ \tau_1 &= \min\{n \geq \tau_0 : X_n \leq 0\}, \\ \tau_2 &= \min\{n \geq \tau_1 : X_n \geq 1\}, \\ \tau_3 &= \min\{n \geq \tau_2 : X_n \leq 0\},\end{aligned}$$

y así sucesivamente (con $\min \emptyset = +\infty$). Definimos ahora el proceso predecible

$$\Gamma_{n+1} = \begin{cases} 0, & \text{si } \tau_0 \leq n < \tau_1, \\ 1, & \text{si } \tau_1 \leq n < \tau_2, \\ 0, & \text{si } \tau_2 \leq n < \tau_3, \\ \vdots & \text{y así sucesivamente,} \end{cases}$$

que es predecible porque para decidir el valor de Γ_{n+1} solo hace falta saber en qué tramo estamos en el instante n , y eso es $\mathcal{F}_n$ -medible. Como cada upcrossing completo aporta al menos 1 a la suma $\sum \Gamma_k(X_k - X_{k-1})$ y los tramos con $\Gamma = 1$ incompletos aportan una cantidad ≥ -0 (pues $X \geq 0$ y el tramo empieza en un valor ≤ 0 , luego $X_{\tau_1} = 0$ y el proceso es positivo), obtenemos

$$U_n \leq (\Gamma \cdot X)_n \quad \text{para todo } n.$$

Por otro lado, descomponiendo la ganancia total,

$$X_n - X_0 = (\Gamma \cdot X)_n + ((1 - \Gamma) \cdot X)_n,$$

y tomando esperanzas,

$$\mathbb{E}[X_n - X_0] = \mathbb{E}[(\Gamma \cdot X)_n] + \mathbb{E}[(1 - \Gamma) \cdot X)_n].$$

Ahora bien, $1 - \Gamma$ es predecible, acotado y positivo, de modo que por la [proposición 2.17\(b\)](#), $((1 - \Gamma) \cdot X)$ es una submartingala que arranca en 0; por lo tanto

$$0 = \mathbb{E}[(1 - \Gamma) \cdot X)_0] \leq \mathbb{E}[(1 - \Gamma) \cdot X)_n].$$

Combinando las tres relaciones,

$$\mathbb{E}[U_n] \leq \mathbb{E}[(\Gamma \cdot X)_n] \leq \mathbb{E}[X_n - X_0]. \quad \blacksquare$$

Proposición 3.2 (Desigualdad de upcrossings de Doob). *Sea (X_n) una submartingala y $\alpha < \beta$ reales. Entonces*

$$(\beta - \alpha) \mathbb{E}[U_n^{\alpha\beta}] \leq \mathbb{E}[(X_n - \alpha)^+] - \mathbb{E}[(X_0 - \alpha)^+], \quad (3.1)$$

o, equivalentemente,

$$\mathbb{E}[U_n^{\alpha\beta}] \leq \mathbb{E}\left[\left(\frac{X_n - \alpha}{\beta - \alpha}\right)^+\right] - \mathbb{E}\left[\left(\frac{X_0 - \alpha}{\beta - \alpha}\right)^+\right]. \quad (3.2)$$

Demostración. Aplicar el lema 3.1 a

$$\tilde{X}_n = \left(\frac{X_n - \alpha}{\beta - \alpha}\right)^+,$$

que es una submartingala positiva por el ejemplo 2.11 ($x \mapsto ((x - \alpha)/(\beta - \alpha))^+$ es convexa y creciente) y cuyos upcrossings de $[0, 1]$ coinciden exactamente con los upcrossings de $[\alpha, \beta]$ de (X_n) . ■

3.2 El teorema de convergencia casi segura

La desigualdad (3.1) permite acotar el número *total* de upcrossings, y de ahí sale la convergencia. Sea

$$0 \leq U_n^{\alpha\beta} \uparrow U^{\alpha\beta} := \lim_n U_n^{\alpha\beta}.$$

Por el lema de Fatou y (3.2),

$$\mathbb{E}[U^{\alpha\beta}] \leq \liminf_n \mathbb{E}[U_n^{\alpha\beta}] \leq \frac{1}{\beta - \alpha} \liminf_n \mathbb{E}[(X_n - \alpha)^+].$$

Como $(x - \alpha)^+ \leq x^+ + |\alpha|$, basta acotar uniformemente $\mathbb{E}[X_n^+]$:

$$\mathbb{E}[U^{\alpha\beta}] \leq \frac{1}{\beta - \alpha} \left(\sup_n \mathbb{E}[X_n^+] + |\alpha| \right). \quad (3.3)$$

Por lo tanto, si $\sup_n \mathbb{E}[X_n^+] < \infty$ entonces $\mathbb{E}[U^{\alpha\beta}] < \infty$ y, en particular, $\mathbb{P}\{U^{\alpha\beta} < \infty\} = 1$.

Teorema 3.3 (Convergencia de submartingalas). *Sea (X_n) una submartingala con*

$$\sup_n \mathbb{E}[X_n^+] < \infty.$$

Entonces existe una variable aleatoria $Z : \Omega \rightarrow \mathbb{R}$ integrable tal que $X_n \xrightarrow{c.s.} Z$.

Atención a la hipótesis

La condición es $\sup_n \mathbb{E}[X_n^+] < \infty$, con el supremo *dentro*. Pedir solamente $\mathbb{E}[X_n^+] < \infty$ para cada n no dice nada: eso ya está garantizado por la integrabilidad, que forma parte de la definición de submartingala.

Demostración. Denotemos $\tilde{Z} = \liminf_n X_n : \Omega \rightarrow [-\infty, +\infty]$.

Paso 1: el límite existe c. s. en $[-\infty, \infty]$. Por (3.3), para cada par de racionales $\alpha < \beta$ se tiene $U^{\alpha\beta} < \infty$ c. s.. Como $\mathbb{Q}^2$ es numerable,

$$\Omega^* = \bigcap_{\substack{\alpha, \beta \in \mathbb{Q} \\ \alpha < \beta}} \{U^{\alpha\beta} < \infty\} \quad \text{cumple} \quad \mathbb{P}(\Omega^*) = 1.$$

Fijemos $\omega \in \Omega^*$ y veamos que $\liminf_n X_n(\omega) = \limsup_n X_n(\omega)$. Si no fuera así, existirían racionales $\alpha < \beta$ con

$$\liminf_n X_n(\omega) < \alpha < \beta < \limsup_n X_n(\omega),$$

y entonces la trayectoria tomaría valores $< \alpha$ y valores $> \beta$ infinitas veces alternadamente, es decir $U^{\alpha\beta}(\omega) = \infty$: contradicción. Por lo tanto

$$\Omega^* \subset \{\tilde{Z} = \liminf_n X_n = \limsup_n X_n\},$$

esto es, $X_n \rightarrow \tilde{Z}$ c. s. en $[-\infty, +\infty]$.

Paso 2: el límite es finito e integrable. La función $x \mapsto x^+$ es continua de $\overline{\mathbb{R}}$ en $\overline{\mathbb{R}}$ (con las convenciones $(+\infty)^+ = +\infty$ y $(-\infty)^+ = 0$), de modo que $0 \leq X_n^+ \rightarrow \tilde{Z}^+$ c. s.. Por Fatou,

$$\mathbb{E}[\tilde{Z}^+] \leq \liminf_n \mathbb{E}[X_n^+] \leq \sup_n \mathbb{E}[X_n^+] < \infty.$$

Para la parte negativa usamos que $\mathbb{E}[X_n] \geq \mathbb{E}[X_0]$ (submartingala):

$$\mathbb{E}[X_n^-] = \mathbb{E}[X_n^+] - \mathbb{E}[X_n] \leq \mathbb{E}[X_n^+] - \mathbb{E}[X_0] \leq \sup_n \mathbb{E}[X_n^+] - \mathbb{E}[X_0] =: C',$$

uniformemente en n . Como $x \mapsto x^-$ también es continua, $X_n^- \rightarrow \tilde{Z}^-$ c. s. y, por Fatou, $\mathbb{E}[\tilde{Z}^-] \leq \liminf_n \mathbb{E}[X_n^-] \leq C' < \infty$. En particular $\mathbb{P}\{|\tilde{Z}| < \infty\} = 1$. Definiendo

$$Z = \begin{cases} \tilde{Z}, & \text{sobre } \Omega^* \cap \{|\tilde{Z}| < \infty\}, \\ 0, & \text{en caso contrario,} \end{cases}$$

obtenemos una variable real, integrable, con $X_n \xrightarrow{\text{c. s.}} Z$. ■

Los tres corolarios siguientes son las formas en que el teorema se usa en la práctica.

Corolario 3.4 (Martingalas acotadas en L^1). *Si (M_n) es una martingala con $\sup_n \mathbb{E}[|M_n|] < \infty$, entonces $M_n \xrightarrow{c.s.} Z$ con Z integrable.*

Demostración. Una martingala es en particular submartingala y $\mathbb{E}[M_n^+] \leq \mathbb{E}[|M_n|]$. ■

Corolario 3.5. *Si (Y_n) es una supermartingala con $\sup_n \mathbb{E}[Y_n^-] < \infty$, entonces $Y_n \xrightarrow{c.s.} Z$ con Z integrable.*

Corolario 3.6 (Supermartingalas positivas). *Si (Y_n) es una supermartingala con $Y_n \geq 0$, entonces $Y_n \xrightarrow{c.s.} Z$ con $Z \geq 0$ integrable y $\mathbb{E}[Z] \leq \mathbb{E}[Y_0]$.*

Demostración. $Y_n^- \equiv 0$, así que el corolario 3.5 da la convergencia. Por Fatou, $\mathbb{E}[Z] \leq \liminf_n \mathbb{E}[Y_n] \leq \mathbb{E}[Y_0]$. ■

La convergencia c. s. no basta

Estos teoremas dan convergencia *casi segura*, no en L^1 : en general $\mathbb{E}[Z] \neq \lim_n \mathbb{E}[X_n]$. El ejemplo canónico es el proceso de ramificación crítico de la sección siguiente, donde $M_n \rightarrow 0$ c. s. pero $\mathbb{E}[M_n] = 1$ para todo n . Determinar cuándo hay convergencia en L^1 es exactamente el problema del capítulo 4.

3.3 Ejemplo: procesos de ramificación

Consideremos una población en la que cada individuo, independientemente, produce un número aleatorio de descendientes con ley μ sobre $\{0, 1, 2, \dots\}$. Formalmente, sea

$$\{\xi_{nj} : n, j \in \mathbb{N}\}$$

una familia independiente con $\xi_{nj} \sim \mu$, y definamos

$$Z_0 = 1, \quad Z_n = \sum_{j=1}^{Z_{n-1}} \xi_{nj} \quad (n \geq 1), \quad \text{con la convención } \sum_{j=1}^0 (\cdot) = 0.$$

Así $\{Z_n = 0\} \uparrow \{\text{extinción}\}$. Sea

$$m = \sum_{k=0}^{\infty} k \mu(\{k\}) = \mathbb{E}[\xi_{11}] \in (0, \infty)$$

el número medio de hijos por individuo, y tomemos la filtración

$$\mathcal{F}_0 = \{\emptyset, \Omega\}, \quad \mathcal{F}_n = \sigma\{\xi_{kj} : 1 \leq k \leq n, j \in \mathbb{N}\}, \quad n \geq 1.$$

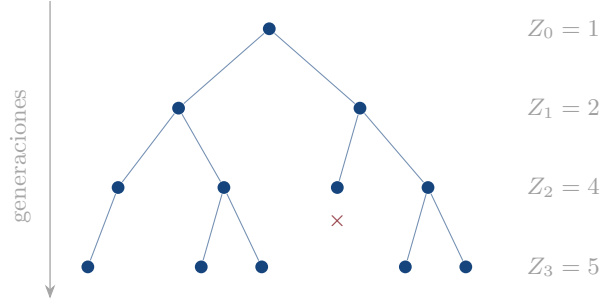


Figura 3.2: Tres generaciones de un proceso de ramificación. El individuo marcado con $\times$ no tuvo descendencia. El proceso normalizado $M_n = Z_n/m^n$ es una martingala.

Proposición 3.7. *El proceso*

$$M_n = \frac{Z_n}{m^n}, \quad n \in \mathbb{Z}_+,$$

es una martingala respecto de $(\mathcal{F}_n)$.

Demostración. Z_n es $\mathcal{F}_n$ -medible y $\sigma\{\xi_{nj} : j \in \mathbb{N}\}$ es independiente de $\mathcal{F}_{n-1}$. Para $A \in \mathcal{F}_{n-1}$, descomponiendo según el valor de Z_{n-1} ,

$$\begin{aligned} \int_A M_n &= \frac{1}{m^n} \sum_{k=0}^{\infty} \int_{A \cap \{Z_{n-1}=k\}} Z_n \\ &= \frac{1}{m^n} \sum_{k=0}^{\infty} \mathbb{E}[(\xi_{n1} + \dots + \xi_{nk}) \mathbf{1}_{A \cap \{Z_{n-1}=k\}}] \\ &= \frac{1}{m^n} \sum_{k=0}^{\infty} k m \mathbb{P}(A \cap \{Z_{n-1} = k\}) \quad (\text{independencia}) \\ &= \frac{1}{m^{n-1}} \sum_{k=0}^{\infty} \int_{A \cap \{Z_{n-1}=k\}} Z_{n-1} = \int_A \frac{Z_{n-1}}{m^{n-1}} = \int_A M_{n-1}. \end{aligned}$$

Por la [observación 2.7](#), (M_n) es martingala. ■

Como $M_n \geq 0$, el [corolario 3.6](#) da $M_n \xrightarrow{\text{c.s.}} M_\infty$ con M_∞ integrable. El caso crítico es el más interesante.

Proposición 3.8 (Extinción en el caso crítico). *Supongamos $m = 1$ y $\mu(\{1\}) < 1$. Entonces $Z_n = M_n \xrightarrow{\text{c.s.}} 0$; es decir, la población se extingue con probabilidad 1.*

Demostración. Con $m = 1$ tenemos $M_n = Z_n$, martingala positiva, luego $Z_n \xrightarrow{\text{c.s.}} Z_\infty$ con Z_∞ finita c.s.. Como Z_n toma valores enteros y converge, la sucesión $(Z_n(\omega))$ debe ser *eventualmente constante* para casi todo ω . Basta entonces ver que $\mathbb{P}\{Z_\infty = u\} = 0$ para todo entero $u \geq 1$. Escribamos

$$\{Z_\infty = u\} = \bigcup_{r=1}^{\infty} E_r, \quad E_r = \bigcap_{n \geq r} \{Z_n = u\},$$

y notemos que $\{Z_r = u, \dots, Z_k = u\} \downarrow E_r$ cuando $k \uparrow \infty$. Ahora, condicionando sucesivamente y usando la independencia,

$$\begin{aligned} \mathbb{P}\{Z_r = u, Z_{r+1} = u, \dots, Z_{r+i} = u\} &= \mathbb{P}\{Z_r = u, \dots, Z_{r+i-1} = u\} \cdot \underbrace{\mathbb{P}\{\xi_{r+i,1} + \dots + \xi_{r+i,u} = u\}}_{=: \lambda} \\ &\vdots \\ &= \mathbb{P}\{Z_r = u\} \lambda^i. \end{aligned}$$

La clave es que $\lambda < 1$: si fuera $\lambda = 1$, la suma de $u \geq 1$ copias independientes de μ valdría u con probabilidad 1, lo que con $\mathbb{E}[\xi] = 1$ obliga a $\mu(\{1\}) = 1$, contra la hipótesis. Haciendo $i \rightarrow \infty$ obtenemos $\mathbb{P}(E_r) = 0$ para todo r , luego $\mathbb{P}\{Z_\infty = u\} = 0$ y, como $u \geq 1$ era arbitrario, $\mathbb{P}\{Z_\infty > 0\} = 0$. ■

Observación. Este es el ejemplo prometido: $Z_n \xrightarrow{\text{c.s.}} 0$ pero $\mathbb{E}[Z_n] = 1$ para todo n . La martingala converge c.s. pero no en L^1 , y no es uniformemente integrable. La masa “se escapa al infinito”: condicionado a no extinguirse —suceso de probabilidad que tiende a cero— el proceso es enorme.

CAPÍTULO 4

Integrabilidad uniforme y convergencia en L^1

El capítulo 3 nos dio convergencia casi segura. Pero para las aplicaciones —y en particular para poder escribir $X_n = \mathbb{E}[X_\infty | \mathcal{F}_n]$ — necesitamos convergencia en L^1 . La condición exacta que hace falta es la *integrabilidad uniforme*: una forma cuantitativa de decir que ninguna parte de la masa de la familia se escapa al infinito.

4.1 Modos de convergencia

Definición 4.1. Sea $(X_n)_{n \in \mathbb{N}}$ una sucesión de variables aleatorias con valores en $\mathbb{R}^d$.

- (a) $X_n \xrightarrow{\text{c.s.}} X$ (*casi seguramente*) si $\mathbb{P}(\{\omega : X_n(\omega) \rightarrow X(\omega)\}) = 1$.
- (b) Para $p \in [1, \infty)$ con $\mathbb{E}[|X|^p] < \infty$ y $\mathbb{E}[|X_n|^p] < \infty$, $X_n \xrightarrow{L^p} X$ si $\lim_n \mathbb{E}[|X_n - X|^p] = 0$.
- (c) $X_n \xrightarrow{\mathbb{P}} X$ (*en probabilidad*) si para todo $\varepsilon > 0$, $\lim_n \mathbb{P}\{|X_n - X| > \varepsilon\} = 0$.
- (d) $X_n \xrightarrow{\text{ley}} X$ (*en ley*) si $\mathbb{E}[f(X_n)] \rightarrow \mathbb{E}[f(X)]$ para toda f continua y acotada.

Las relaciones entre estos modos se resumen en el diagrama siguiente. La única implicación que *no* es automática —y que este capítulo se encarga de caracterizar— es la que va de la convergencia en probabilidad a la convergencia en L^1 .

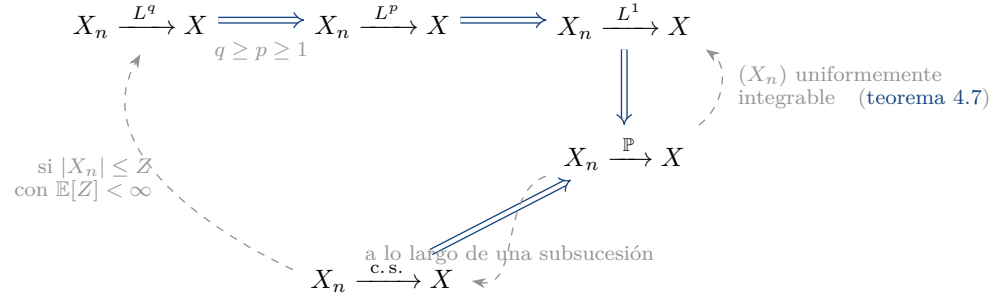


Figura 4.1: Relación entre los distintos tipos de convergencia. Las flechas dobles son implicaciones incondicionales; las punteadas requieren la hipótesis indicada.

Recordemos que el teorema de convergencia dominada dice precisamente

$$\left. \begin{array}{l} X_n \xrightarrow{\text{c.s.}} X \\ |X_n| \leq Z \text{ con } Z \text{ integrable} \end{array} \right\} \implies X_n \xrightarrow{L^1} X.$$

El objetivo de este capítulo es debilitar la dominación.

4.2 Integrabilidad uniforme

Definición 4.2 (Integrabilidad uniforme). Una familia $(X_\lambda)_{\lambda \in \Lambda}$ de variables aleatorias es *uniformemente integrable* (u. i.) si

$$\lim_{t \rightarrow \infty} \sup_{\lambda \in \Lambda} \int_{\{|X_\lambda| > t\}} |X_\lambda| d\mathbb{P} = 0,$$

lo que también se escribe $\lim_{t \rightarrow \infty} \left(\sup_{\lambda \in \Lambda} \mathbb{E}[|X_\lambda| ; |X_\lambda| > t] \right) = 0$.

Observación 4.3 (Ejemplos inmediatos). (a) Si Λ es finito y cada X_λ es integrable, la familia es u. i.

(b) Si $|X_\lambda| \leq |Y_\lambda|$ para todo λ y (Y_λ) es u. i., entonces (X_λ) es u. i.

(c) Si $|X_\lambda| \leq Z$ con Z integrable, para todo λ , entonces (X_λ) es u. i. (dominación $\Rightarrow$ u. i.).

(d) Si (X_λ) e (Y_λ) son u. i., entonces $(X_\lambda + Y_\lambda)$ es u. i.

Proposición 4.4 (Caracterización). $(X_\lambda)_{\lambda \in \Lambda}$ es uniformemente integrable si y solo si se cumplen las dos condiciones siguientes:

- (I) (Acotación en L^1 .) Existe C con $\mathbb{E}[|X_\lambda|] \leq C$ para todo $\lambda \in \Lambda$.
 (II) (Continuidad absoluta uniforme.) Para todo $\varepsilon > 0$ existe $\delta > 0$ tal que

$$\mathbb{P}(A) < \delta \implies \int_A |X_\lambda| d\mathbb{P} < \varepsilon \quad \text{para todo } \lambda \in \Lambda.$$

Demostración. ($\Rightarrow$) Supongamos (X_λ) u.i. Como

$$\lim_{t \rightarrow \infty} \left\{ \sup_\lambda \int_{\{|X_\lambda| > t\}} |X_\lambda| \right\} = 0,$$

podemos elegir t^* con $\sup_\lambda \int_{\{|X_\lambda| > t^*\}} |X_\lambda| \leq 1$. Entonces, para todo λ ,

$$\mathbb{E}[|X_\lambda|] = \underbrace{\int_{\{|X_\lambda| \leq t^*\}} |X_\lambda|}_{\leq t^*} + \underbrace{\int_{\{|X_\lambda| > t^*\}} |X_\lambda|}_{\leq 1} \leq 1 + t^*,$$

lo que prueba (i). Para (ii), fijado $\varepsilon > 0$ elegimos t con

$$\int_{\{|X_\lambda| > t\}} |X_\lambda| < \frac{\varepsilon}{2} \quad \text{para todo } \lambda,$$

y tomamos $\delta = \varepsilon/(2t)$. Entonces, para $\mathbb{P}(A) < \delta$,

$$\begin{aligned} \int_A |X_\lambda| &= \int_{A \cap \{|X_\lambda| > t\}} |X_\lambda| + \int_{A \cap \{|X_\lambda| \leq t\}} |X_\lambda| \\ &\leq \int_{\{|X_\lambda| > t\}} |X_\lambda| + t \mathbb{P}(A) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

($\Leftarrow$) Supongamos (i) y (ii), y fijemos $\varepsilon > 0$. Sea $\delta > 0$ el dado por (ii). Por la desigualdad de Markov y (i),

$$\mathbb{P}\{|X_\lambda| > t\} \leq \frac{\mathbb{E}[|X_\lambda|]}{t} \leq \frac{C}{t} \quad \text{para todo } \lambda.$$

Eligiendo $t_0 = C/\delta$ aseguramos que, para $t > t_0$ y todo λ ,

$$\mathbb{P}\{|X_\lambda| > t\} < \delta \implies \int_{\{|X_\lambda| > t\}} |X_\lambda| < \varepsilon,$$

usando (ii) con $A = \{|X_\lambda| > t\}$. Como ε era arbitrario, la familia es u.i. ■

Observación. Si (X_n) es u.i. y $\mathbb{P}(A_n) \rightarrow 0$, entonces $\int_{A_n} |X_n| \rightarrow 0$: es la condición (ii) de la proposición anterior.

Las dos condiciones siguientes son las que más se usan para *verificar* la integrabilidad uniforme.

Lema 4.5 (Acotación en L^p , $p > 1$). *Si existe $p > 1$ y $C > 0$ tales que $\mathbb{E}[|X_\lambda|^p] \leq C$ para todo $\lambda \in \Lambda$, entonces (X_λ) es uniformemente integrable.*

Demostración. Sobre $\{|X_\lambda| > t\}$ se tiene $1 < (|X_\lambda|/t)^{p-1}$, de modo que

$$\int_{\{|X_\lambda| > t\}} |X_\lambda| \leq \int_{\{|X_\lambda| > t\}} |X_\lambda| \left(\frac{|X_\lambda|}{t} \right)^{p-1} = \frac{1}{t^{p-1}} \int_{\{|X_\lambda| > t\}} |X_\lambda|^p \leq \frac{C}{t^{p-1}} \xrightarrow{t \rightarrow \infty} 0,$$

uniformemente en λ . ■

El caso $p = 1$ está excluido

La acotación en L^1 no implica integrabilidad uniforme: tomando $X_n = n\mathbf{1}_{[0,1/n]}$ sobre $([0,1], \text{Leb})$ se tiene $\mathbb{E}[X_n] = 1$ para todo n , pero $\int_{\{X_n > t\}} X_n = 1$ para todo $t < n$. Esta es exactamente la situación del proceso de ramificación crítico.

Proposición 4.6 (Esperanzas condicionales de una variable fija). *Sea $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ y $(\mathcal{G}_\alpha)_{\alpha \in \Lambda}$ una familia arbitraria de sub- σ -álgebras de $\mathcal{F}$. Entonces la familia*

$$Z_\alpha = \mathbb{E}[X | \mathcal{G}_\alpha], \quad \alpha \in \Lambda,$$

es uniformemente integrable.

Demostración. Sea $\varepsilon > 0$. Como X es integrable, existe $\delta > 0$ tal que $\mathbb{P}(A) < \delta$ implica $\mathbb{E}[|X|; A] < \varepsilon$ (continuidad absoluta de la integral). Sea $M > 0$ con $\mathbb{E}[|X|]/M < \delta$.

Por Jensen condicional, $|Z_\alpha| \leq \mathbb{E}[|X| | \mathcal{G}_\alpha]$, luego $\{|Z_\alpha| > M\} \subset \{\mathbb{E}[|X| | \mathcal{G}_\alpha] > M\} =: A_\alpha$, que es un suceso de $\mathcal{G}_\alpha$. Entonces

$$\mathbb{E}[|Z_\alpha|; |Z_\alpha| > M] \leq \mathbb{E}[\mathbb{E}[|X| | \mathcal{G}_\alpha]; A_\alpha] = \mathbb{E}[|X|; A_\alpha],$$

donde la última igualdad es la definición de esperanza condicional aplicada a $A_\alpha \in \mathcal{G}_\alpha$. Finalmente, por Markov,

$$\mathbb{P}(A_\alpha) \leq \frac{\mathbb{E}[\mathbb{E}[|X| | \mathcal{G}_\alpha]]}{M} = \frac{\mathbb{E}[|X|]}{M} < \delta,$$

de modo que $\mathbb{E}[|X|; A_\alpha] < \varepsilon$ para todo α . ■

4.3 El teorema de Vitali

Teorema 4.7 (Vitali). Sea (X_n) una sucesión con $\mathbb{E}[|X_n|] < \infty$ y supongamos que $X_n \xrightarrow{\mathbb{P}} X$. Son equivalentes:

- (I) (X_n) es uniformemente integrable;
- (II) X es integrable y $X_n \xrightarrow{L^1} X$;
- (III) X es integrable y $\mathbb{E}[|X_n|] \rightarrow \mathbb{E}[|X|] < \infty$.

Demostración. (i) $\Rightarrow$ (ii). Por la [proposición 4.4](#) existe C con $\mathbb{E}[|X_n|] \leq C$ para todo n . Como $X_n \xrightarrow{\mathbb{P}} X$, existe una subsucesión $X_{n_k} \xrightarrow{\text{c.s.}} X$; entonces $|X_{n_k}| \rightarrow |X|$ c.s. y, por Fatou,

$$\mathbb{E}[|X|] \leq \liminf_k \mathbb{E}[|X_{n_k}|] \leq C < \infty,$$

de modo que X es integrable. La familia $(|X_n - X|)_n$ es u.i., pues $|X_n - X| \leq |X_n| + |X|$ y la suma de familias u.i. es u.i. ([observación 4.3\(d\)](#), con $\{X\}$ u.i. por ser finita). Sea $\varepsilon > 0$; escribimos

$$\mathbb{E}[|X_n - X|] = \int_{\{|X_n - X| > \varepsilon\}} |X_n - X| + \int_{\{|X_n - X| \leq \varepsilon\}} |X_n - X| \leq \int_{A_n} |X_n - X| + \varepsilon,$$

con $A_n = \{|X_n - X| > \varepsilon\}$. Como $X_n \xrightarrow{\mathbb{P}} X$ se tiene $\mathbb{P}(A_n) \rightarrow 0$, y como $(|X_n - X|)$ es u.i., la condición (ii) de la [proposición 4.4](#) da $\int_{A_n} |X_n - X| \rightarrow 0$. Por lo tanto

$$0 \leq \limsup_n \mathbb{E}[|X_n - X|] \leq \varepsilon,$$

y siendo $\varepsilon > 0$ arbitrario, $\mathbb{E}[|X_n - X|] \rightarrow 0$.

(ii) $\Rightarrow$ (iii). Inmediato de $|\mathbb{E}[|X_n|] - \mathbb{E}[|X|]| \leq \mathbb{E}[|X_n - X|]$.

(iii) $\Rightarrow$ (i). Para $s > 0$ escribamos $\varphi_s(x) = (|x| - s)^+$, de modo que $|x| - \varphi_s(x) = \min(|x|, s)$ es una función continua y acotada por s . Como $X_n \xrightarrow{\mathbb{P}} X$, también $\min(|X_n|, s) \xrightarrow{\mathbb{P}} \min(|X|, s)$ y, por convergencia dominada (la dominante es la constante s),

$$\mathbb{E}[\min(|X_n|, s)] \xrightarrow{n \rightarrow \infty} \mathbb{E}[\min(|X|, s)].$$

Restando esto de la hipótesis $\mathbb{E}[|X_n|] \rightarrow \mathbb{E}[|X|]$ obtenemos

$$\mathbb{E}[\varphi_s(X_n)] \xrightarrow{n \rightarrow \infty} \mathbb{E}[\varphi_s(X)] \quad \text{para cada } s > 0 \text{ fijo.} \quad (4.1)$$

Por otro lado, para $t = 2s$,

$$\int_{\{|X_n|>t\}} |X_n| = \int_{\{|X_n|>t\}} (|X_n|-t) + t \mathbb{P}\{|X_n| > t\} \leq \mathbb{E}[\varphi_t(X_n)] + t \mathbb{P}\{|X_n| > t\},$$

y como sobre $\{|X_n| > t\} = \{|X_n| > 2s\}$ se tiene $\varphi_s(X_n) \geq s = t/2$, la desigualdad de Markov da $t \mathbb{P}\{|X_n| > t\} \leq 2 \mathbb{E}[\varphi_s(X_n)]$. Como además $\varphi_t \leq \varphi_s$, concluimos

$$\int_{\{|X_n|>2s\}} |X_n| \leq 3 \mathbb{E}[\varphi_s(X_n)]. \quad (4.2)$$

Ahora fijemos $\varepsilon > 0$. Como $X \in L^1$ y $\varphi_s(X) \downarrow 0$ cuando $s \uparrow \infty$, por convergencia dominada existe s_0 con $\mathbb{E}[\varphi_{s_0}(X)] < \varepsilon/6$; por (4.1) existe N tal que $\mathbb{E}[\varphi_{s_0}(X_n)] < \varepsilon/3$ para $n \geq N$, y entonces (4.2) da $\int_{\{|X_n|>2s_0\}} |X_n| < \varepsilon$ para $n \geq N$. Las finitas variables $X_1, \dots, X_{N-1}$ forman una familia u.i. por sí solas (observación 4.3(a)), de modo que agrandando t si hace falta se obtiene la cota para todos los índices a la vez. ■

Corolario 4.8. *Para una submartingala (X_n) que converge c. s., son equivalentes:*

- (I) (X_n) es uniformemente integrable;
- (II) (X_n) converge en L^1 .

Demostración. Consecuencia del teorema 4.7, ya que la convergencia c. s. y la convergencia en L^1 implican ambas convergencia en probabilidad. ■

4.4 Martingalas cerradas: el último elemento

Definición 4.9 (Último elemento). Decimos que una (sub/super)martingala $(X_n)_{n \in \mathbb{Z}_+}$ tiene último elemento, o que está cerrada, si existen $X_\infty \in L^1$ y una σ -álgebra $\mathcal{F}_\infty \subset \mathcal{F}$ que contiene a todas las $\mathcal{F}_n$, tales que $(X_n, \mathcal{F}_n)_{n \in \mathbb{Z}_+ \cup \{\infty\}}$ sigue siendo (sub/super)martingala. Explícitamente: X_∞ es $\mathcal{F}_\infty$ -medible y, para todo n ,

$$\mathbb{E}[X_\infty | \mathcal{F}_n] \begin{cases} = X_n & \text{(martingala),} \\ \geq X_n & \text{(submartingala),} \\ \leq X_n & \text{(supermartingala).} \end{cases}$$

Lema 4.10. Si $X_n \xrightarrow{L^1} X$, entonces $\mathbb{E}[X_n; A] \rightarrow \mathbb{E}[X; A]$ para todo $A \in \mathcal{F}$.

Demostración. $|\mathbb{E}[X_n; A] - \mathbb{E}[X; A]| \leq \mathbb{E}[|X_n - X|] \rightarrow 0$. ■

Lema 4.11. *Si una martingala (X_n) converge en L^1 hacia X , entonces $X_n = \mathbb{E}[X | \mathcal{F}_n]$ para todo n .*

Demostración. Por la propiedad de martingala, para $m > n$ y $A \in \mathcal{F}_n$,

$$\mathbb{E}[X_m; A] = \mathbb{E}[X_n; A].$$

Por el [lema 4.10](#), el miembro izquierdo tiende a $\mathbb{E}[X; A]$ cuando $m \rightarrow \infty$, mientras que el derecho no depende de m . Luego $\mathbb{E}[X_n; A] = \mathbb{E}[X; A]$ para todo $A \in \mathcal{F}_n$, que es justamente $X_n = \mathbb{E}[X | \mathcal{F}_n]$. ■

Teorema 4.12 (Caracterización de las martingalas cerradas). *Sea (X_n) una martingala. Son equivalentes:*

- (I) (X_n) tiene último elemento;
- (II) (X_n) es uniformemente integrable;
- (III) (X_n) converge c. s. y en L^1 hacia una variable X_∞ , y $X_n = \mathbb{E}[X_\infty | \mathcal{F}_n]$ para todo n .

Demostración. (ii) $\Rightarrow$ (iii). Por la [proposición 4.4\(i\)](#) se tiene $\sup_n \mathbb{E}[|X_n|] \leq C < \infty$, de modo que el [corolario 3.4](#) da $X_n \xrightarrow{\text{c.s.}} X_\infty$ con X_∞ integrable. Como (X_n) es u. i. y la convergencia c. s. implica convergencia en probabilidad, el [teorema 4.7](#) da $X_n \xrightarrow{L^1} X_\infty$, y el [lema 4.11](#) da $X_n = \mathbb{E}[X_\infty | \mathcal{F}_n]$.

(iii) $\Rightarrow$ (i). Basta tomar $\mathcal{F}_\infty = \sigma(\bigcup_n \mathcal{F}_n)$; la variable X_∞ , que es límite c. s. de variables $\mathcal{F}_\infty$ -medibles, es $\mathcal{F}_\infty$ -medible.

(i) $\Rightarrow$ (ii). Si (X_n) tiene último elemento, entonces $X_n = \mathbb{E}[X_\infty | \mathcal{F}_n]$ con $X_\infty \in L^1$ fija, y la [proposición 4.6](#) dice exactamente que tal familia es uniformemente integrable. ■

Teorema 4.13 (Submartingalas positivas). *Sea (X_n) una submartingala positiva.*

- (a) Si (X_n) es uniformemente integrable, tiene último elemento.
- (b) Recíprocamente, si (X_n) tiene último elemento, es uniformemente integrable.

Demostración. (a) Si (X_n) es u.i., entonces $\mathbb{E}[X_n^+] \leq \mathbb{E}[|X_n|] \leq C$ para todo n , y el [teorema 3.3](#) da $X_n \xrightarrow{\text{c.s.}} X_\infty$ con X_∞ integrable; por Vitali, $X_n \xrightarrow{L^1} X_\infty$. Sea $n \leq m$ y $A \in \mathcal{F}_n$; por ser submartingala,

$$\int_A X_n \leq \int_A X_m.$$

Haciendo $m \rightarrow \infty$ y usando el [lema 4.10](#), $\int_A X_n \leq \int_A X_\infty$ para todo $A \in \mathcal{F}_n$, es decir $X_n \leq \mathbb{E}[X_\infty | \mathcal{F}_n]$.

(b) Supongamos $X_n \geq 0$ con último elemento X_∞ . Como $\{X_n > t\} \in \mathcal{F}_n$, la propiedad de submartingala cerrada da

$$\int_{\{X_n > t\}} X_n \leq \int_{\{X_n > t\}} X_\infty.$$

Fijemos $\varepsilon > 0$ y sea $\delta > 0$ tal que $\mathbb{P}(A) < \delta$ implica $\int_A X_\infty < \varepsilon$. Por Markov y la monotonía de las esperanzas,

$$\mathbb{P}\{X_n > t\} \leq \frac{\mathbb{E}[X_n]}{t} \leq \frac{\mathbb{E}[X_\infty]}{t}.$$

Tomando $t^* = \mathbb{E}[X_\infty] / \delta$ se obtiene, para $t > t^*$ y todo n ,

$$\int_{\{X_n > t\}} X_n \leq \int_{\{X_n > t\}} X_\infty < \varepsilon,$$

lo que prueba la integrabilidad uniforme. ■

Resumen del capítulo

Para una martingala (M_n) :

Hipótesis	Conclusión
$\sup_n \mathbb{E}[M_n] < \infty$	$M_n \xrightarrow{\text{c.s.}} M_\infty$ integrable
(M_n) uniformemente integrable	además $M_n \xrightarrow{L^1} M_\infty$ y $M_n = \mathbb{E}[M_\infty \mathcal{F}_n]$
$\sup_n \mathbb{E}[M_n ^p] < \infty$, $p > 1$	además $M_n \xrightarrow{L^p} M_\infty$ (corolario 5.17)

La primera hipótesis no implica la segunda (proceso de ramificación crítico); la tercera sí implica la segunda ([lema 4.5](#)).

Tiempos de parada y desigualdades de Doob

Un tiempo de parada es una regla para decidir *cuándo detenerse* que solo usa información pasada. El teorema de muestreo opcional dice que detener una martingala en un tiempo de parada no cambia su esperanza —con hipótesis que hay que cuidar, como muestra el ejemplo del jugador terco de la [sección 5.2.1](#)— y las desigualdades maximales de Doob controlan el máximo de la trayectoria por su valor final.

5.1 Definiciones

Definición 5.1 (Tiempo de parada). Dado $(\Omega, \mathcal{F})$ con una filtración $(\mathcal{F}_n)_{n \in \mathbb{Z}_+}$, una variable

$$\tau : \Omega \longrightarrow \bar{\mathbb{Z}}_+ = \mathbb{Z}_+ \cup \{\infty\}$$

es un *tiempo de parada* si

$$\{\tau \leq n\} \in \mathcal{F}_n \quad \text{para todo } n \in \mathbb{Z}_+.$$

Equivalentemente, $\{\tau = n\} \in \mathcal{F}_n$ para todo n , o bien $\{\tau > n\} \in \mathcal{F}_n$ para todo n .

Definición 5.2 (σ -álgebra de los sucesos anteriores a τ). Dado un tiempo de parada τ , definimos

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \leq n\} \in \mathcal{F}_n \text{ para todo } n \in \mathbb{Z}_+\}.$$

Ejercicio 5.3. Verifique las propiedades siguientes.

- (a) $\mathcal{F}_\tau$ es una σ -álgebra, y τ es $\mathcal{F}_\tau$ -medible (es decir, $\sigma(\tau) \subset \mathcal{F}_\tau$).
- (b) Si $\tau \equiv n$ es constante, $\mathcal{F}_\tau = \mathcal{F}_n$.
- (c) Si $\tau \leq \rho$ son tiempos de parada, $\mathcal{F}_\tau \subset \mathcal{F}_\rho$.
- (d) $\tau \wedge \rho$ y $\tau \vee \rho$ son tiempos de parada, y $\mathcal{F}_{\tau \wedge \rho} = \mathcal{F}_\tau \cap \mathcal{F}_\rho$.

(e) Si (X_n) es adaptado, $X_\tau \mathbf{1}_{\{\tau < \infty\}}$ es $\mathcal{F}_\tau$ -medible.

Indicación. Para (c): si $A \in \mathcal{F}_\tau$, entonces $A \cap \{\rho \leq n\} = (A \cap \{\tau \leq n\}) \cap \{\rho \leq n\} \in \mathcal{F}_n$, donde se usó $\{\rho \leq n\} \subset \{\tau \leq n\}$. Para (e): para D boreliano, $\{X_\tau \in D, \tau < \infty\} \cap \{\tau \leq n\} = \bigcup_{k \leq n} (\{X_k \in D\} \cap \{\tau = k\}) \in \mathcal{F}_n$. ■

Definición 5.4 (Proceso detenido). Dados τ y un proceso (X_n) , el *proceso detenido* X^τ es

$$X_n^\tau(\omega) = X_{n \wedge \tau(\omega)}(\omega), \quad n \in \mathbb{Z}_+.$$

Proposición 5.5 (Detener preserva el carácter). Sea (X_n) una (sub/super)martingala y τ un tiempo de parada respecto de $(\mathcal{F}_n)$. Entonces $(X_n^\tau, \mathcal{F}_n)$ es una (sub/super)martingala. En particular, para una martingala, $\mathbb{E}[X_{n \wedge \tau}] = \mathbb{E}[X_0]$ para todo n .

Demostración. Definamos $H_k = \mathbf{1}_{\{\tau > k-1\}} = \mathbf{1}_{\{\tau \geq k\}}$ para $k \geq 1$. Como $\{\tau > k-1\} = \{\tau \leq k-1\}^c \in \mathcal{F}_{k-1}$, el proceso (H_k) es predecible, acotado y positivo. Además

$$X_n^\tau = X_0 + (H \cdot X)_n = X_0 + \sum_{k=1}^n \mathbf{1}_{\{\tau \geq k\}}(X_k - X_{k-1}),$$

pues la suma telescópica exactamente hasta el índice $n \wedge \tau$. La [proposición 2.17](#) concluye. La última afirmación se obtiene tomando esperanzas y usando $X_0^\tau = X_0$. ■

5.2 El teorema de muestreo opcional

Empecemos por el caso de horizonte finito, donde no hace falta ninguna hipótesis adicional.

Lema 5.6 (Horizonte finito). Sea $X_0, X_1, \dots, X_m$ una submartingala y $\tau : \Omega \rightarrow \{0, 1, \dots, m\}$ un tiempo de parada. Entonces X_τ es integrable y

$$\mathbb{E}[X_0] \leq \mathbb{E}[X_\tau] \leq \mathbb{E}[X_m].$$

Demostración. La integrabilidad es clara: $|X_\tau| \leq \sum_{k=0}^m |X_k|$.

Primera desigualdad. Por la [proposición 5.5](#), (X_n^τ) es submartingala, luego

$$\underbrace{\mathbb{E}[X_0^\tau]}_{=\mathbb{E}[X_0]} \leq \underbrace{\mathbb{E}[X_m^\tau]}_{=\mathbb{E}[X_\tau]},$$

donde se usó que $\tau \leq m$ implica $X_m^\tau = X_{m \wedge \tau} = X_\tau$.

Segunda desigualdad. Con H como en la [proposición 5.5](#), descomponemos

$$X_n = \underbrace{X_0 + (H \cdot X)_n}_{=X_n^\tau} + \underbrace{((1-H) \cdot X)_n}_{=:Y_n},$$

y (Y_n) es submartingala con $Y_0 = 0$ porque $1-H \geq 0$ es predecible y acotado. Evaluando en $n = m$ y tomando esperanzas,

$$\mathbb{E}[X_m] = \mathbb{E}[X_\tau] + \mathbb{E}[Y_m] \geq \mathbb{E}[X_\tau] + \mathbb{E}[Y_0] = \mathbb{E}[X_\tau]. \quad \blacksquare$$

Proposición 5.7 (Caracterización por tiempos de parada). Sean $X_0, X_1, \dots, X_m$ variables aleatorias integrables y adaptadas. Son equivalentes:

- (I) $(X_n)_{n \leq m}$ es submartingala;
- (II) para todo par de tiempos de parada $\tau, \rho : \Omega \rightarrow \{0, 1, \dots, m\}$ con $\tau \leq \rho$ se tiene $\mathbb{E}[X_\tau] \leq \mathbb{E}[X_\rho]$.

La versión con igualdad caracteriza a las martingalas: $(X_n)_{n \leq m}$ es martingala si y solo si $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$ para todo tiempo de parada τ con valores en $\{0, \dots, m\}$.

Demostración. (i) $\Rightarrow$ (ii). Sea $\tilde{X}_n = X_n^\rho$, que es submartingala por la [proposición 5.5](#). Aplicando el [lema 5.6](#) a $\tilde{X}$ con el tiempo de parada τ ,

$$\mathbb{E}[\underbrace{\tilde{X}_\tau}_{=X_{\tau \wedge \rho} = X_\tau}] \leq \mathbb{E}[\underbrace{\tilde{X}_m}_{=X_\rho}].$$

(ii) $\Rightarrow$ (i). Fijemos $n \in \{0, \dots, m-1\}$ y $A \in \mathcal{F}_n$. La variable

$$\tau = \begin{cases} n, & \text{sobre } A, \\ n+1, & \text{sobre } A^c, \end{cases}$$

es un tiempo de parada (pues $A \in \mathcal{F}_n$), y $\tau \leq \rho \equiv n+1$. Entonces

$$\begin{aligned} \mathbb{E}[X_\tau] &= \mathbb{E}[X_n \mathbf{1}_A] + \mathbb{E}[X_{n+1} \mathbf{1}_{A^c}], \\ \mathbb{E}[X_\rho] &= \mathbb{E}[X_{n+1} \mathbf{1}_A] + \mathbb{E}[X_{n+1} \mathbf{1}_{A^c}], \end{aligned}$$

y restando, $\mathbb{E}[X_\tau] \leq \mathbb{E}[X_\rho]$ equivale a $\mathbb{E}[X_n \mathbf{1}_A] \leq \mathbb{E}[X_{n+1} \mathbf{1}_A]$ para todo $A \in \mathcal{F}_n$, que es la propiedad de submartingala ([observación 2.7](#)). $\blacksquare$

5.2.1 Un contraejemplo instructivo

Ejemplo 5.8 (El jugador terco). Sean (ξ_n) independientes con $\mathbb{P}\{\xi_n = 1\} = \mathbb{P}\{\xi_n = -1\} = \frac{1}{2}$, y

$$S_0 = 0, \quad S_n = \sum_{j=1}^n \xi_j, \quad \mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n).$$

(S_n) es una martingala (ejemplo 2.12). Consideremos

$$\tau = \min\{n \geq 0 : S_n = 1\},$$

la estrategia “jugar hasta ganar una unidad”. Entonces τ es un tiempo de parada y $\mathbb{P}\{\tau < \infty\} = 1$.

En efecto, (S_n^τ) es una martingala con $S_n^\tau \leq 1$ para todo n , de modo que $\mathbb{E}[(S_n^\tau)^+] \leq 1$ y el teorema 3.3 da $S_n^\tau \xrightarrow{\text{c.s.}} Z$ con Z integrable. Pero sobre $\{\tau = \infty\}$ se tiene $S_n^\tau = S_n$, y el paseo aleatorio simple *no* converge (oscila entre $\pm\infty$); luego

$$\mathbb{P}\{\tau = \infty, S_n^\tau \rightarrow Z\} = 0 \implies \mathbb{P}\{\tau = \infty\} = 0.$$

Sin embargo,

$$\underbrace{\mathbb{E}[S_0]}_{=0} \neq \underbrace{\mathbb{E}[S_\tau]}_{=1}.$$

Moraleja

El teorema de muestreo opcional *falla* sin hipótesis adicionales cuando el horizonte no es acotado. Lo que ocurre aquí es que $\mathbb{E}[\tau] = \infty$: la estrategia funciona, pero puede requerir un tiempo arbitrariamente largo y una deuda arbitrariamente grande. Las hipótesis que sí funcionan son: horizonte acotado (lema 5.6), τ acotado, o existencia de último elemento (teorema 5.9).

5.2.2 El teorema general

Teorema 5.9 (Muestreo opcional de Doob). Sea $(X_n, \mathcal{F}_n)_{n \in \mathbb{Z}_+}$ una (sub/super)martingala con último elemento X_∞ . Entonces, para cualquier par de tiempos de parada $\tau \leq \rho$, las variables X_τ y X_ρ son integrables y

$$\mathbb{E}[X_0] \bowtie \mathbb{E}[X_\tau] \bowtie \mathbb{E}[X_\rho] \bowtie \mathbb{E}[X_\infty],$$

donde $\bowtie$ es $=$ para martingalas, $\leq$ para submartingalas y $\geq$ para supermartingalas.

Teorema 5.10 (Versión con condicionamiento). *En las hipótesis del teorema 5.9,*

$$\int_A X_\tau d\mathbb{P} \preceq \int_A X_\rho d\mathbb{P} \quad \text{para todo } A \in \mathcal{F}_\tau,$$

es decir, $X_\tau \preceq \mathbb{E}[X_\rho | \mathcal{F}_\tau]$. En particular, para una martingala, $\mathbb{E}[X_\rho | \mathcal{F}_\tau] = X_\tau$ c. s..

Demostración del teorema 5.9. Caso 1: (X_n) es martingala. Por el teorema 4.12, $X_n = \mathbb{E}[X_\infty | \mathcal{F}_n]$ para todo $n \in \mathbb{Z}_+$. Por Jensen, $(|X_n|, \mathcal{F}_n)_{n \in \mathbb{Z}_+}$ es una submartingala cerrada por $|X_\infty|$, de modo que $|X_n| \leq \mathbb{E}[|X_\infty| | \mathcal{F}_n]$ y entonces

$$\mathbb{E}[|X_\tau|] = \sum_{n \in \mathbb{Z}_+} \int_{\{\tau=n\}} |X_n| \leq \sum_{n \in \mathbb{Z}_+} \int_{\{\tau=n\}} |X_\infty| = \mathbb{E}[|X_\infty|] < \infty,$$

donde se usó $\{\tau = n\} \in \mathcal{F}_n$. Luego X_τ es integrable. Con la misma descomposición y el teorema de convergencia dominada (la dominante es $|X_\infty|$),

$$\mathbb{E}[X_\tau] = \sum_{n \in \mathbb{Z}_+} \int_{\{\tau=n\}} X_n = \sum_{n \in \mathbb{Z}_+} \int_{\{\tau=n\}} X_\infty = \mathbb{E}[X_\infty],$$

donde en el paso central se usó de nuevo $\{\tau = n\} \in \mathcal{F}_n$ y $X_n = \mathbb{E}[X_\infty | \mathcal{F}_n]$. Como el argumento vale para cualquier tiempo de parada, en particular para $\tau \equiv 0$ y para ρ , obtenemos $\mathbb{E}[X_0] = \mathbb{E}[X_\tau] = \mathbb{E}[X_\rho] = \mathbb{E}[X_\infty]$.

Caso 2: supermartingalas positivas con $X_\infty = 0$. Sea $X_n \geq 0$ supermartingala con último elemento $X_\infty = 0$, y $\tau \leq \rho$ tiempos de parada. Fijemos $j \in \mathbb{Z}_+$. Para $k \geq j$, sobre el suceso $\{\tau = j, \rho \geq k\} \in \mathcal{F}_k$ la propiedad de supermartingala da

$$\int_{\{\tau=j, \rho \geq k\}} X_k \geq \int_{\{\tau=j, \rho \geq k+1\}} X_{k+1} + \int_{\{\tau=j, \rho=k\}} X_\rho,$$

pues $\{\tau = j, \rho \geq k\}$ se parte en $\{\rho \geq k+1\}$ y $\{\rho = k\}$, y sobre el primero $\int X_k \geq \int X_{k+1}$ mientras que sobre el segundo $X_k = X_\rho$. Reordenando,

$$\int_{\{\tau=j, \rho \geq k\}} X_k - \int_{\{\tau=j, \rho \geq k+1\}} X_{k+1} \geq \int_{\{\tau=j, \rho=k\}} X_\rho.$$

Sumando en $k = j, j+1, \dots, m$ (suma telescópica) se obtiene, para $m > j$,

$$\underbrace{\int_{\{\tau=j\}} X_j}_{= \int_{\{\tau=j\}} X_\tau} - \underbrace{\int_{\{\tau=j, \rho \geq m+1\}} X_{m+1}}_{\geq 0} \geq \int_{\{\tau=j, j \leq \rho \leq m\}} X_\rho,$$

y por lo tanto $\int_{\{\tau=j\}} X_\tau \geq \int_{\{\tau=j, \rho \leq m\}} X_\rho$. Haciendo $m \rightarrow \infty$ y usando convergencia monótona ($X_\rho \geq 0$),

$$\int_{\{\tau=j\}} X_\tau \geq \int_{\{\tau=j, \rho < \infty\}} X_\rho \geq 0.$$

Sumando ahora en $j = 0, 1, 2, \dots$,

$$\int_{\{\tau < \infty\}} X_\tau \geq \int_{\{\tau < \infty, \rho < \infty\}} X_\rho.$$

Como $X_\infty = 0$, ambos miembros son $\mathbb{E}[X_\tau]$ y $\mathbb{E}[X_\rho]$ respectivamente, de donde $\mathbb{E}[X_\tau] \geq \mathbb{E}[X_\rho]$. Tomando $\tau \equiv 0$ resulta además $\infty > \mathbb{E}[X_0] \geq \mathbb{E}[X_\rho]$.

Caso general. Sea (X_n) una supermartingala con último elemento X_∞ . Definiendo

$$M_n = \mathbb{E}[X_\infty | \mathcal{F}_n], \quad n \in \mathbb{Z}_+,$$

se tiene que (M_n) es una martingala cerrada (Caso 1) y que

$$X_n = M_n + \underbrace{(X_n - M_n)}_{=: Y_n},$$

donde (Y_n) es una supermartingala positiva —pues $X_n \geq \mathbb{E}[X_\infty | \mathcal{F}_n] = M_n$ — con último elemento $Y_\infty = X_\infty - X_\infty = 0$. Aplicando el Caso 1 a M y el Caso 2 a Y y sumando se concluye. El caso de submartingalas se obtiene cambiando el signo. ■

Demostración del teorema 5.10. Fijemos $A \in \mathcal{F}_\tau$ y definamos

$$\gamma = \begin{cases} \tau, & \text{sobre } A, \\ \rho, & \text{sobre } A^c. \end{cases}$$

Entonces γ es un tiempo de parada: para $n \in \mathbb{Z}_+$,

$$\{\gamma \leq n\} = (A \cap \{\tau \leq n\}) \cup (A^c \cap \{\rho \leq n\}),$$

y el primer conjunto está en $\mathcal{F}_n$ porque $A \in \mathcal{F}_\tau$, mientras que el segundo lo está porque $A^c \cap \{\rho \leq n\} = \{\rho \leq n\} \setminus (A \cap \{\tau \leq n\})$, usando $\{\rho \leq n\} \subset \{\tau \leq n\}$. Como $\gamma \leq \rho$, el teorema 5.9 (caso submartingala) da $\mathbb{E}[X_\gamma] \leq \mathbb{E}[X_\rho]$, o sea

$$\int_A X_\gamma + \int_{A^c} X_\gamma \leq \int_A X_\rho + \int_{A^c} X_\rho.$$

Como $\gamma = \rho$ sobre A^c , los términos sobre A^c se cancelan y queda

$$\underbrace{\int_A X_\gamma}_{=\int_A X_\tau} \leq \int_A X_\rho. \quad \blacksquare$$

Corolario 5.11 (Casos particulares de uso frecuente). (a) Si (X_n) es una martingala uniformemente integrable, entonces para todo tiempo de parada τ ,

$$\mathbb{E}[X_0] = \mathbb{E}[X_\tau] = \mathbb{E}[X_\infty],$$

donde $X_n \rightarrow X_\infty$ c. s. y en L^1 .

(b) Si (X_n) es una submartingala uniformemente integrable, para todo τ , $\mathbb{E}[X_0] \leq \mathbb{E}[X_\tau] \leq \mathbb{E}[X_\infty]$.

(c) Si (Z_n) es una supermartingala positiva, entonces siempre tiene último elemento —a saber, $Z_\infty = \lim_n Z_n$, que existe y es integrable por el corolario 3.6— y por lo tanto, para todo tiempo de parada τ ,

$$\mathbb{E}[Z_0] \geq \mathbb{E}[Z_\tau] \geq \mathbb{E}[Z_\infty].$$

Demostración. Solo (c) requiere comentario. Por el corolario 3.6, $Z_n \xrightarrow{\text{c. s.}} Z_\infty$ con $Z_\infty \geq 0$ integrable. Por el lema de Fatou condicional, para n fijo,

$$\mathbb{E}[Z_\infty | \mathcal{F}_n] = \mathbb{E}[\liminf_m Z_m | \mathcal{F}_n] \leq \liminf_m \mathbb{E}[Z_m | \mathcal{F}_n] \leq Z_n,$$

de modo que $(Z_n, \mathcal{F}_n)_{n \in \mathbb{Z}_+}$ es una supermartingala con último elemento Z_∞ ; el teorema 5.9 concluye. ■

Una supermartingala positiva no tiene por qué ser u. i.

Es tentador —y falso— concluir de (c) que toda supermartingala positiva es uniformemente integrable. El contraejemplo es el proceso de ramificación crítico de la proposición 3.8: Z_n es una martingala positiva (en particular supermartingala) con $\mathbb{E}[Z_n] = 1$ para todo n , pero $Z_n \xrightarrow{\text{c. s.}} 0$, de modo que $\mathbb{E}[Z_\infty] = 0 < 1 = \mathbb{E}[Z_0]$ y la convergencia no es en L^1 . Nótese que la desigualdad $\mathbb{E}[Z_0] \geq \mathbb{E}[Z_\infty]$ de (c) es *estricta* en ese caso, y que eso es precisamente lo que distingue a (c) del enunciado de (a) para martingalas u. i.

La diferencia con el teorema 4.13(b) está en la dirección de la desigualdad: para una submartingala positiva cerrada se tiene $X_n \leq \mathbb{E}[X_\infty | \mathcal{F}_n]$, y el último elemento *domina*, lo que da la integrabilidad uniforme. Para una supermartingala el último elemento está dominado, y el argumento no se puede invertir.

5.3 Las desigualdades maximales de Doob

Teorema 5.12 (Desigualdad maximal, L^1). Sea (X_n) una submartingala positiva. Para todo $\lambda > 0$ y todo $N \in \mathbb{Z}_+$,

$$\mathbb{P}\left\{\max_{0 \leq n \leq N} X_n \geq \lambda\right\} \leq \frac{\mathbb{E}[X_N]}{\lambda}.$$

Demostración. Definamos

$$\tau = \min\{0 \leq n \leq N : X_n \geq \lambda\} \wedge N,$$

que es un tiempo de parada acotado, y observemos que

$$\left\{\max_{0 \leq n \leq N} X_n \geq \lambda\right\} = \{X_\tau \geq \lambda\} =: A \in \mathcal{F}_\tau.$$

Sobre A se tiene $X_\tau \geq \lambda$, de modo que

$$\lambda \mathbb{P}(A) \leq \int_A X_\tau.$$

Como $\tau \leq N$ son tiempos de parada acotados y $(X_n)_{n \leq N}$ es submartingala, el [teorema 5.10](#) (aplicable con horizonte finito, sin necesidad de último elemento) da

$$\int_A X_\tau \leq \int_A X_N \leq \int_\Omega X_N = \mathbb{E}[X_N],$$

usando en la última desigualdad que $X_N \geq 0$. Combinando, $\lambda \mathbb{P}(A) \leq \mathbb{E}[X_N]$. ■

Teorema 5.13 (Desigualdad maximal para martingalas). Sea (M_n) una martingala. Para todo $\lambda > 0$,

$$\mathbb{P}\left\{\max_{0 \leq n \leq N} |M_n| > \lambda\right\} \leq \frac{\mathbb{E}[|M_N|]}{\lambda}.$$

Demostración. $(|M_n|)$ es una submartingala positiva ([proposición 2.9](#)); aplicar el [teorema 5.12](#). ■

Ejercicio 5.14. Sean (M_n) y $(\widetilde{M}_n)$ dos martingalas respecto de la misma filtración. Pruebe que

$$\mathbb{P}\left\{\max_{0 \leq n \leq N} |M_n - \widetilde{M}_n| > \lambda\right\} \leq \frac{\mathbb{E}[|M_N - \widetilde{M}_N|]}{\lambda}.$$

Indicación: $M - \widetilde{M}$ es una martingala.

Teorema 5.15 (Desigualdad L^p de Doob). *Sea (X_n) una submartingala positiva y*

$$X_N^* = \max_{0 \leq n \leq N} X_n.$$

Si $p > 1$ y q es su exponente conjugado ($1/p + 1/q = 1$, es decir $q = p/(p-1)$), entonces

$$(\mathbb{E}[(X_N^*)^p])^{1/p} \leq q (\mathbb{E}[X_N^p])^{1/p}.$$

Demostración. Recordemos que para $Z \geq 0$ y $p \geq 1$,

$$\mathbb{E}[Z^p] = \int_0^\infty \mathbb{P}\{Z^p \geq \mu\} d\mu = p \int_0^\infty \lambda^{p-1} \mathbb{P}\{Z \geq \lambda\} d\lambda, \quad (5.1)$$

mediante el cambio $\mu = \lambda^p$.

Fijemos $A > 0$ y truncemos: $\tilde{X} = X_N^* \wedge A$, de modo que $\mathbb{E}[\tilde{X}^p] \leq A^p < \infty$. Nótese que $\{\tilde{X} \geq \lambda\} \subset \{X_N^* \geq \lambda\}$ para todo λ , con igualdad si $\lambda \leq A$. Usando (5.1) y el teorema 5.12,

$$\begin{aligned} \mathbb{E}[\tilde{X}^p] &= p \int_0^\infty \lambda^{p-1} \mathbb{P}\{\tilde{X} \geq \lambda\} d\lambda \leq p \int_0^\infty \lambda^{p-1} \frac{1}{\lambda} \int_{\{\tilde{X} \geq \lambda\}} X_N d\mathbb{P} d\lambda \\ &= p \int_0^\infty \int_\Omega \lambda^{p-2} X_N \mathbf{1}_{\{\lambda \leq \tilde{X}\}} d\mathbb{P} d\lambda \stackrel{\text{Fubini}}{=} p \int_\Omega X_N \int_0^{\tilde{X}} \lambda^{p-2} d\lambda d\mathbb{P} \\ &= \frac{p}{p-1} \int_\Omega X_N \tilde{X}^{p-1} d\mathbb{P} = q \mathbb{E}[X_N \tilde{X}^{p-1}]. \end{aligned}$$

Aplicando ahora la desigualdad de Hölder con exponentes p y q ,

$$\mathbb{E}[\tilde{X}^p] \leq q (\mathbb{E}[X_N^p])^{1/p} (\mathbb{E}[\tilde{X}^{q(p-1)}])^{1/q} = q (\mathbb{E}[X_N^p])^{1/p} (\mathbb{E}[\tilde{X}^p])^{1/q},$$

donde se usó $q(p-1) = p$. Como $\mathbb{E}[\tilde{X}^p] < \infty$, podemos dividir por $(\mathbb{E}[\tilde{X}^p])^{1/q}$ y, usando $1 - 1/q = 1/p$,

$$(\mathbb{E}[(X_N^* \wedge A)^p])^{1/p} \leq q (\mathbb{E}[X_N^p])^{1/p}.$$

Haciendo $A \rightarrow \infty$ y aplicando convergencia monótona se concluye. ■

Por qué se trunca

El truncamiento por A no es un adorno: sin él no sabríamos *a priori* que $\mathbb{E}[(X_N^*)^p] < \infty$, y dividir por esa cantidad sería ilegítimo. Es un patrón

recurrente: probar una desigualdad para una cantidad manifestamente finita y luego pasar al límite.

Teorema 5.16. *Si (M_n) es una martingala con $M_n \in L^p$, $p > 1$, entonces*

$$\left\| \max_{0 \leq n \leq N} |M_n| \right\|_p \leq \frac{p}{p-1} \|M_N\|_p.$$

Demostración. $(|M_n|)$ es submartingala positiva; aplicar el [teorema 5.15](#) con $q = p/(p-1)$. ■

Corolario 5.17 (Convergencia en L^p). *Sea (M_n) una martingala tal que $\sup_n \mathbb{E}[|M_n|^p] \leq C < \infty$ para algún $p > 1$. Entonces existe $Z \in L^p$ con*

$$M_n \xrightarrow{\text{c. s.}} Z \quad \text{y} \quad M_n \xrightarrow{L^p} Z.$$

Demostración. Por el [lema 4.5](#) la martingala es uniformemente integrable; en particular está acotada en L^1 y el [corolario 3.4](#) da $M_n \xrightarrow{\text{c. s.}} Z$ con Z integrable, y por Vitali $M_n \xrightarrow{L^1} Z$.

Para la convergencia en L^p , observemos que $|M_n - Z|^p \rightarrow 0$ c. s. y que, por convexidad de $x \mapsto x^p$,

$$|M_n - Z|^p \leq 2^{p-1} (|M_n|^p + |Z|^p) \leq 2^{p-1} \left(\sup_k |M_k|^p + |Z|^p \right).$$

Ahora, $\sup_k |M_k|^p$ es integrable: por el [teorema 5.16](#) y convergencia monótona,

$$\mathbb{E} \left[\sup_k |M_k|^p \right] = \lim_N \mathbb{E} \left[\max_{k \leq N} |M_k|^p \right] \leq \left(\frac{p}{p-1} \right)^p \sup_N \mathbb{E}[|M_N|^p] \leq \left(\frac{p}{p-1} \right)^p C < \infty,$$

y $|Z|^p$ es integrable por Fatou. El teorema de convergencia dominada concluye. ■

Parte II

Aplicaciones en finanzas

CAPÍTULO 6

Mercados financieros a tiempo discreto

Esta parte traduce, casi palabra por palabra, la teoría de martingalas al lenguaje de los mercados financieros. La traducción es literal: una estrategia de inversión es un proceso predecible, el valor de la cartera es una transformada de martingalas, la ausencia de arbitraje es la existencia de una medida bajo la cual los precios son martingala, y el precio de un derivado es una esperanza condicional. Trabajaremos siempre con Ω *finito*, lo que elimina toda dificultad técnica y deja a la vista la estructura.

6.1 El mercado

Definición 6.1 (Mercado). Un *mercado* (a tiempo discreto, con horizonte N) es una terna

$$[\Omega \text{ finito}, (\mathcal{F}_n)_{n=0}^N, (M_n)_{n=0}^N],$$

donde

- (I) $\Omega = \{\omega_1, \dots, \omega_k\}$ es el conjunto de todos los estados del mundo (las trayectorias posibles de los activos);
- (II) $(\mathcal{F}_n)_{n=0}^N$ es una filtración, con $\mathcal{F}_0 = \{\emptyset, \Omega\}$ y $\mathcal{F}_N = 2^\Omega$;
- (III) $M_n : \Omega \rightarrow \mathbb{R}^{\ell+1}$ es el vector de precios,

$$M_n(\omega) = (1, S_n^1(\omega), \dots, S_n^\ell(\omega)),$$

con M_n $\mathcal{F}_n$ -medible.

Salvo mención en contrario tomaremos $\ell = 1$ y escribiremos $M_n = (1, S_n)$.

Por qué la primera coordenada es un 1

La primera coordenada representa el *numerario*: un activo de referencia no riesgoso. En general sería $M_n = (B_n, S_n^1, \dots, S_n^\ell)$ con B_n determinista (por ejemplo $B_n = (1+r)^n$, el dinero en el banco). Tomar $B_n \equiv 1$ no pierde generalidad: equivale a medir todos los precios *en unidades de B*, es decir a trabajar con los precios descontados $\tilde{S}_n = S_n/B_n$. Todo lo que sigue debe leerse en unidades del numerario; en el [capítulo 18](#) el descuento reaparecerá explícitamente como el factor e^{-rt} .

Definición 6.2 (Estrategia). Una *estrategia* o *portafolio* es un proceso

$$\Gamma_n : \Omega \rightarrow \mathbb{R}^{\ell+1}, \quad n = 1, \dots, N,$$

tal que Γ_n es $\mathcal{F}_{n-1}$ -medible para todo n . Es decir, (Γ_n) es *predecible*.

Escribiendo $\Gamma = (\beta, \alpha_1, \dots, \alpha_\ell)$, la convención de signos es:

- $\beta > 0$: el inversor tiene $|\beta|$ unidades del numerario (dinero en el banco);
 $\beta < 0$: las debe;
- $\alpha_i > 0$: el inversor posee $|\alpha_i|$ unidades del activo i ; $\alpha_i < 0$: está *en corto* (las debe).

Definición 6.3 (Valor de una estrategia). El *valor* de la estrategia Γ en cada instante es

$$V_0^\Gamma = \underbrace{\Gamma_1 \cdot M_0}_{\text{lo que cuesta iniciar la estrategia}}, \quad V_n^\Gamma = \underbrace{\Gamma_n \cdot M_n}_{\text{lo que vale la cartera adquirida en } n-1}, \quad n \geq 1.$$

Definición 6.4 (Autofinanciamiento). La estrategia Γ es *autofinanciada* si

$$\Gamma_n \cdot M_n = \Gamma_{n+1} \cdot M_n \quad \text{para todo } n = 1, \dots, N-1.$$

La condición dice: al reajustar la cartera en el instante n , pasando de la composición Γ_n a la composición Γ_{n+1} , no se inyecta ni se retira dinero; el rebalanceo se hace *a los precios vigentes* M_n . Formalmente, descomponiendo la variación de valor,

$$V_{n+1}^\Gamma - V_n^\Gamma = \underbrace{\Gamma_{n+1} \cdot (M_{n+1} - M_n)}_{\text{ganancia por variación de precios}} + \underbrace{M_n \cdot (\Gamma_{n+1} - \Gamma_n)}_{\text{dinero inyectado en } n}, \quad (6.1)$$

y el autofinanciamiento es exactamente la anulación del segundo término.

Proposición 6.5 (El valor como transformada de martingalas). *Si Γ es autofinanciada, entonces*

$$V_n^\Gamma = V_0^\Gamma + \sum_{j=1}^n \Gamma_j \cdot (M_j - M_{j-1}) = V_0^\Gamma + (\Gamma \cdot M)_n.$$

Demostración. Sumar (6.1) desde $j = 1$ hasta n , con el segundo término nulo por autofinanciamiento (y usando $V_0^\Gamma = \Gamma_1 \cdot M_0$ para el primer paso). ■

La identidad central de toda la Parte II

$$V_n^\Gamma = V_0^\Gamma + (\Gamma \cdot M)_n$$

El valor de una cartera autofinanciada es su capital inicial más una *transformada de martingalas* del proceso de precios, con integrando predecible. Todo el resto de la Parte II consiste en explotar la [proposición 2.17](#). En tiempo continuo, esta identidad se convertirá en $V_t = V_0 + \int_0^t \Gamma_s dM_s$ con una integral de Itô ([capítulo 18](#)).

Observación 6.6 (Linealidad). El valor es lineal en la estrategia:

$$V_n^{\Gamma_1 + \Gamma_2} = V_n^{\Gamma_1} + V_n^{\Gamma_2}, \quad V_n^{\lambda \Gamma} = \lambda V_n^\Gamma,$$

y la suma (resp. el múltiplo) de estrategias autofinanciadas es autofinanciada.

Proposición 6.7 (Libertad en la primera coordenada). *Sea $\ell = 1$, $M_n = (1, S_n)$. Para todo proceso predecible $\beta_n : \Omega \rightarrow \mathbb{R}$ (la posición en el activo riesgoso) y toda constante $c \in \mathbb{R}$, existe un único proceso predecible $\alpha_n : \Omega \rightarrow \mathbb{R}$ tal que*

$$\Gamma_n = (\alpha_n, \beta_n)$$

es autofinanciada con $V_0^\Gamma = c$.

Demostración. La condición $V_0^\Gamma = c$ da $\alpha_1 + \beta_1 S_0 = c$, lo que determina α_1 . La condición de autofinanciamiento en el instante n ,

$$\alpha_{n+1} + \beta_{n+1} S_n = \alpha_n + \beta_n S_n,$$

determina $\alpha_{n+1} = \alpha_n + (\beta_n - \beta_{n+1}) S_n$, que es $\mathcal{F}_n$ -medible como se requiere. La recursión define α de manera única. ■

Lectura de la proposición

La posición en el activo riesgoso puede elegirse *libremente*; la posición en el numerario queda entonces forzada por el presupuesto. Dicho de otro modo, una estrategia autofinanciada queda parametrizada por el par (c, β) : capital inicial más posición riesgosa. Esto se usará decisivamente en la demostración del segundo teorema fundamental ([capítulo 7](#)).

6.2 Arbitraje

Definición 6.8 (Oportunidad de arbitraje). Una estrategia autofinanciada Γ es una *oportunidad de arbitraje* en el mercado (M_n) si

- (I) $V_0^\Gamma \leq 0$ (no hace falta invertir nada),
- (II) $V_N^\Gamma \geq 0$ (no hay riesgo de pérdida),
- (III) $V_N^\Gamma(\omega) > 0$ para algún $\omega \in \Omega$ (hay posibilidad de ganancia).

El mercado es *libre de arbitraje* (l.d.a.) si no existe ninguna estrategia autofinanciada que sea oportunidad de arbitraje.

Ejemplo 6.9 (El modelo binomial (s, a, b)). Fijemos $s > 0$ y $-1 < a < b$. Tomamos

$$\Omega = \{a, b\}^N, \quad \rho_j : \Omega \rightarrow \{a, b\} \text{ la } j\text{-ésima coordenada,}$$

y definimos el precio del activo riesgoso por

$$S_0 = s, \quad S_n = S_{n-1}(1 + \rho_n), \quad n = 1, \dots, N,$$

con la filtración canónica $\mathcal{F}_n = \sigma(\rho_1, \dots, \rho_n)$ y $M_n = (1, S_n)$. Así, en cada período el precio sube un factor $1 + b$ o baja un factor $1 + a$.

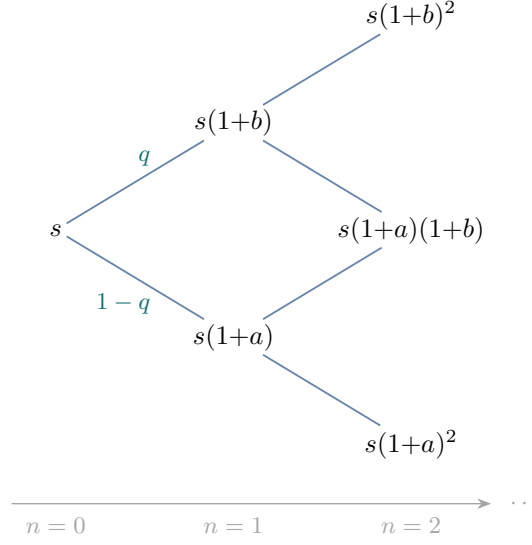


Figura 6.1: Dos períodos del modelo binomial (s, a, b) . El árbol *recombina*: llegar por “sube-baja” o por “baja-suba” produce el mismo precio. Bajo la única PME los rendimientos ρ_n son independientes con $q = \mathbb{Q}\{\rho_n = b\} = \frac{-a}{b-a}$.

Proposición 6.10. *El modelo binomial (s, a, b) es libre de arbitraje si y solo si $a < 0 < b$.*

Demostración. ($\Rightarrow$, por contrarrecíproco.) Si $0 \leq a < b$, el activo riesgoso nunca pierde valor y a veces gana: la estrategia “pedir prestado s unidades del numerario y comprar una acción”, $\Gamma_n = (-s, 1)$, es autofinanciada, tiene $V_0^\Gamma = -s + s = 0$ y $V_N^\Gamma = S_N - s \geq 0$, con desigualdad estricta en el estado “siempre sube”. Es un arbitraje. Si $a < b \leq 0$, la estrategia opuesta $\Gamma_n = (s, -1)$ lo es.

($\Leftarrow$.) Si $a < 0 < b$ construiremos en el [ejemplo 6.14](#) una probabilidad de riesgo neutro, y el [teorema 6.13](#) mostrará que su existencia impide el arbitraje. ■

6.3 Probabilidades martingala equivalentes

Definición 6.11 (Probabilidad martingala equivalente). Dado un mercado $[\Omega, (\mathcal{F}_n), (M_n)]$, una probabilidad $\mathbb{Q}$ sobre Ω es una *probabilidad neutra al riesgo* o *probabilidad martingala equivalente* (PME) si

- (I) $(M_n, \mathcal{F}_n)_{n=0}^N$ es una $\mathbb{Q}$ -martingala (coordenada a coordenada);

(II) $\mathbb{Q}(\{\omega\}) > 0$ para todo $\omega \in \Omega$.

La condición (ii) es la que justifica el adjetivo *equivalente*: $\mathbb{Q}$ y $\mathbb{P}$ tienen los mismos conjuntos de probabilidad nula, es decir coinciden en qué es posible y qué es imposible, aunque discrepen en cuán probable es cada cosa.

Observación 6.12. Si $\mathbb{Q}$ es PME, entonces $(V_n^\Gamma, \mathcal{F}_n)$ es una $\mathbb{Q}$ -martingala para toda estrategia autofinanciada Γ . En efecto, por la [proposición 6.5](#), V^Γ es la transformada de la $\mathbb{Q}$ -martingala M por el proceso predecible Γ —acotado, pues Ω es finito— y la [proposición 2.17](#) concluye.

Teorema 6.13. *Si existe una PME $\mathbb{Q}$, entonces el mercado es libre de arbitraje.*

Demostración. Supongamos que Γ es autofinanciada con $V_0^\Gamma \leq 0$ y $V_N^\Gamma \geq 0$. Por la [observación 6.12](#), (V_n^Γ) es $\mathbb{Q}$ -martingala, luego

$$\mathbb{E}_{\mathbb{Q}}[V_N^\Gamma] = \mathbb{E}_{\mathbb{Q}}[V_0^\Gamma] \leq 0.$$

Pero $V_N^\Gamma \geq 0$, de modo que una variable no negativa tiene esperanza ≤ 0 : necesariamente $V_N^\Gamma = 0$ $\mathbb{Q}$ -c.s.. Como $\mathbb{Q}(\{\omega\}) > 0$ para *todo* $\omega \in \Omega$, esto significa

$$V_N^\Gamma(\omega) = 0 \quad \text{para todo } \omega \in \Omega,$$

y la condición (iii) de la [definición 6.8](#) no puede cumplirse. ■

Dónde se usa cada hipótesis

La propiedad de martingala da $\mathbb{E}_{\mathbb{Q}}[V_N^\Gamma] = \mathbb{E}_{\mathbb{Q}}[V_0^\Gamma] \leq 0$; la *equivalencia* (todos los átomos con masa positiva) convierte “ $V_N^\Gamma = 0$ $\mathbb{Q}$ -c.s.” en “ $V_N^\Gamma = 0$ en todo punto”. Sin (ii) el teorema es falso: bastaría tomar $\mathbb{Q}$ concentrada en un solo estado.

Ejemplo 6.14 (La PME del modelo binomial). En el modelo binomial con $a < 0 < b$ existe una única PME, y bajo ella los rendimientos $\rho_1, \dots, \rho_N$ son independientes, idénticamente distribuidos y de media cero:

$$\mathbb{Q}\{\rho_n = b\} = \frac{-a}{b-a} =: q > 0, \quad \mathbb{Q}\{\rho_n = a\} = \frac{b}{b-a} = 1 - q > 0. \quad (6.2)$$

Demostración. La condición de martingala para $M_n = (1, S_n)$ es $\mathbb{E}_{\mathbb{Q}}[S_n | \mathcal{F}_{n-1}] = S_{n-1}$, es decir $S_{n-1} \mathbb{E}_{\mathbb{Q}}[1 + \rho_n | \mathcal{F}_{n-1}] = S_{n-1}$ y, como $S_{n-1} > 0$,

$$\mathbb{E}_{\mathbb{Q}}[\rho_n | \mathcal{F}_{n-1}] = 0. \quad (6.3)$$

Si $q_n(\omega) := \mathbb{Q}\{\rho_n = b \mid \mathcal{F}_{n-1}\}(\omega)$, (6.3) dice $bq_n + a(1 - q_n) = 0$, o sea $q_n = -a/(b - a)$, constante e independiente de ω y de n . Que q_n sea determinista para todo n equivale a que los ρ_n sean independientes bajo $\mathbb{Q}$ con la ley (6.2). Como $a < 0 < b$, $q \in (0, 1)$ y por lo tanto $\mathbb{Q}(\{\omega\}) = q^k(1 - q)^{N-k} > 0$ para todo ω (donde k es el número de subidas). La unicidad es clara: (6.3) determina q sin ambigüedad. ■

6.4 El primer teorema fundamental

El recíproco del teorema 6.13 también es cierto. Es el resultado central de esta parte.

Teorema 6.15 (Primer teorema fundamental de las finanzas). *Sea $[\Omega \text{ finito}, (\mathcal{F}_n)_{n=0}^N, (M_n)_{n=0}^N]$ un mercado. Entonces*

$$(M_n) \text{ es libre de arbitraje} \iff \text{existe una PME } \mathbb{Q}.$$

Demostración. La implicación ($\Leftarrow$) es el teorema 6.13. Para ($\Rightarrow$), identifiquemos las funciones $\Omega \rightarrow \mathbb{R}$ con vectores de $\mathbb{R}^k$, $k = |\Omega|$, mediante el producto interno

$$\eta \cdot \xi = \sum_{\omega \in \Omega} \eta(\omega) \xi(\omega). \quad (6.4)$$

Una probabilidad $\mathbb{Q}$ sobre Ω se identifica con el vector $(\mathbb{Q}(\{\omega\}))_{\omega}$, y entonces $\mathbb{E}_{\mathbb{Q}}[\eta] = \mathbb{Q} \cdot \eta$.

Consideremos el subespacio de las *ganancias alcanzables sin capital inicial*,

$$\mathcal{A} = \left\{ (\Gamma \cdot M)_N : \Gamma \text{ predecible} \right\} \subset \mathbb{R}^k, \quad (6.5)$$

que es un subespacio vectorial de $\mathbb{R}^k$ (por la observación 6.6), y el cono

$$\mathcal{K} = \left\{ \xi \in \mathbb{R}^k : \xi \geq 0, \sum_{\omega} \xi(\omega) = 1 \right\},$$

que es un conjunto convexo compacto (el simplex). La ausencia de arbitraje dice exactamente que

$$\mathcal{A} \cap \mathcal{K} = \emptyset :$$

en efecto, un elemento de $\mathcal{A} \cap \mathcal{K}$ sería la ganancia terminal de una estrategia con $V_0^{\Gamma} = 0$, $V_N^{\Gamma} \geq 0$ y $V_N^{\Gamma} \not\equiv 0$.

Por el teorema de separación de Hahn–Banach en dimensión finita (separación estricta de un compacto convexo y un subespacio cerrado disjuntos), existe $\mathbb{Q} \in \mathbb{R}^k$ tal que

$$\mathbb{Q} \cdot \xi > 0 \text{ para todo } \xi \in \mathcal{K}, \quad \mathbb{Q} \cdot \eta = 0 \text{ para todo } \eta \in \mathcal{A}.$$

(La segunda igualdad se obtiene de que el funcional está acotado sobre el subespacio $\mathcal{A}$, lo que obliga a que se anule en él.) Tomando $\xi = \delta_\omega$ en la primera condición resulta $\mathbb{Q}(\{\omega\}) > 0$ para todo ω ; normalizando podemos suponer $\sum_\omega \mathbb{Q}(\{\omega\}) = 1$, de modo que $\mathbb{Q}$ es una probabilidad equivalente a $\mathbb{P}$.

Falta ver que (M_n) es $\mathbb{Q}$ -martingala. Fijemos n y $A \in \mathcal{F}_{n-1}$, y tomemos la estrategia predecible $\Gamma_j = \mathbf{1}_A \mathbf{1}_{\{j=n\}} e_i$ (mantener una unidad del activo i solo durante el período n , y solo sobre el suceso A). Su ganancia terminal es $\eta = \mathbf{1}_A (M_n^i - M_{n-1}^i) \in \mathcal{A}$, luego

$$0 = \mathbb{Q} \cdot \eta = \mathbb{E}_{\mathbb{Q}}[(M_n^i - M_{n-1}^i) \mathbf{1}_A] \quad \text{para todo } A \in \mathcal{F}_{n-1},$$

que por la [observación 2.7](#) es exactamente $\mathbb{E}_{\mathbb{Q}}[M_n^i | \mathcal{F}_{n-1}] = M_{n-1}^i$. ■

Observación 6.16 (Caracterización geométrica). La demostración anterior contiene una equivalencia que usaremos varias veces: con el producto interno (6.4) y $\mathcal{A}$ como en (6.5),

$$(M_n, \mathcal{F}_n) \text{ es } \mathbb{Q}\text{-martingala} \iff \mathbb{Q} \perp \mathcal{A}.$$

Ejercicio 6.17. Supongamos que (M_n) es libre de arbitraje. Pruebe que no existe ninguna estrategia autofinanciada Γ ni ningún tiempo de parada $\tau : \Omega \rightarrow \{0, 1, \dots, N\}$ tales que

$$V_0^\Gamma \leq 0, \quad V_\tau^\Gamma \geq 0, \quad V_\tau^\Gamma(\omega) > 0 \text{ para algún } \omega \in \Omega.$$

Indicación: dada tal pareja, construya la estrategia que coincide con Γ hasta τ y que después mantiene todo el capital en el numerario,

$$\tilde{\Gamma}_n = \Gamma_n \mathbf{1}_{\{n \leq \tau\}} + (V_\tau^\Gamma, 0, \dots, 0) \mathbf{1}_{\{n > \tau\}}.$$

Verifique que es predecible (use $\{n \leq \tau\} = \{\tau \leq n-1\}^c \in \mathcal{F}_{n-1}$ y que $V_\tau^\Gamma \mathbf{1}_{\{\tau \leq n-1\}}$ es $\mathcal{F}_{n-1}$ -medible), que es autofinanciada, y que $V_N^{\tilde{\Gamma}} = V_\tau^\Gamma$.

Valoración de derivados y completitud

Admitimos de ahora en adelante que el mercado es libre de arbitraje. El problema es el siguiente: dado un contrato derivado, ¿qué precio hay que ponerle para que, al incorporarlo al mercado, este siga sin ofrecer oportunidades de arbitraje? La respuesta es una esperanza condicional bajo la PME.

7.1 Contratos

Definición 7.1 (Contrato europeo). Dado un mercado $[\Omega, (\mathcal{F}_n), (M_n)]$, un *contrato* (europeo) es una función

$$F : \Omega \rightarrow \mathbb{R} \quad \mathcal{F}_N\text{-medible},$$

interpretada como el pago que el *writer* (vendedor) entrega al *holder* (comprador) en el instante N .

Definición 7.2 (Opción de compra europea). Una *opción de compra europea* (*call*) sobre el subyacente S está determinada por

- (I) el activo subyacente S ;
- (II) el vencimiento N ;
- (III) el precio de ejercicio o *strike* K .

Da al holder el *derecho, pero no la obligación*, de comprar al writer una unidad del subyacente al precio K en el instante N .

El holder no está obligado a ejercer. Si en el instante N el precio S_N supera al strike K , ejercer le reporta $S_N - K$; si $S_N < K$ no ejerce y recibe 0. El contrato corresponde entonces a la variable aleatoria

$$F_c(\omega) = (S_N(\omega) - K)^+.$$

En una *opción de venta* (*put*) el holder tiene el derecho de vender al precio K , de modo que

$$F_v(\omega) = (K - S_N(\omega))^+.$$



Figura 7.1: Pagos al vencimiento de una call y de una put europeas de strike K . Ambos son funciones convexas de S_N y dependen únicamente del precio terminal, no de la trayectoria.

Definición 7.3 (Opción *lookback*).

$$F = S_N - \min_{0 \leq n \leq N} S_n.$$

A diferencia de la call y la put, este contrato *sí* depende de toda la trayectoria: es *path dependent*.

7.2 Precio justo

Definición 7.4 (Precio justo). Diremos que $p \in \mathbb{R}$ *no es un precio justo* para el contrato F si se da alguna de las dos situaciones siguientes.

- **El writer se aprovecha de p :** existe Γ autofinanciada con

$$V_0^\Gamma - p \leq 0, \quad V_N^\Gamma - F \geq 0, \quad V_N^\Gamma(\omega) - F(\omega) > 0 \text{ para algún } \omega.$$

(Vende el contrato en p , monta la cartera Γ sin costo neto, y al vencimiento le sobra dinero después de pagar F .)

- **El holder se aprovecha de p :** existe Γ autofinanciada con

$$V_0^\Gamma + p \leq 0, \quad V_N^\Gamma + F \geq 0, \quad V_N^\Gamma(\omega) + F(\omega) > 0 \text{ para algún } \omega.$$

En caso contrario, p es un *precio justo* para F .

La definición es simétrica y dice exactamente lo que debe decir: p es justo si al incorporar el contrato al mercado, a ese precio, ninguna de las dos partes obtiene una oportunidad de arbitraje.

Proposición 7.5. *Si $\mathbb{Q}$ es una PME para (M_n) , entonces*

$$p^* = \mathbb{E}_{\mathbb{Q}}[F]$$

es un precio justo para F .

Demostración. Supongamos que el writer se aprovecha de p^* : existe Γ autofinanciada con $V_0^\Gamma \leq p^*$, $V_N^\Gamma \geq F$ y $V_N^\Gamma(\omega_0) > F(\omega_0)$ para algún ω_0 . Como (V_n^Γ) es $\mathbb{Q}$ -martingala (*observación 6.12*),

$$\mathbb{E}_{\mathbb{Q}}[V_N^\Gamma - F] = V_0^\Gamma - \mathbb{E}_{\mathbb{Q}}[F] = V_0^\Gamma - p^* \leq 0.$$

Pero $V_N^\Gamma - F \geq 0$, de modo que $V_N^\Gamma - F = 0$ $\mathbb{Q}$ -c. s. y, por ser $\mathbb{Q}$ equivalente, $V_N^\Gamma(\omega) = F(\omega)$ para *todo* ω : esto contradice $V_N^\Gamma(\omega_0) > F(\omega_0)$.

El caso del holder es análogo: de $V_0^\Gamma + p^* \leq 0$ se obtiene $\mathbb{E}_{\mathbb{Q}}[V_N^\Gamma + F] \leq 0$ con $V_N^\Gamma + F \geq 0$, luego $V_N^\Gamma + F \equiv 0$, contradiciendo la tercera condición. ■

Tenemos entonces la inclusión

$$\{\mathbb{E}_{\mathbb{Q}}[F] : \mathbb{Q} \text{ PME}\} \subset \{\text{precios justos para } F\}.$$

De hecho, hay igualdad.

Teorema 7.6 (Caracterización de los precios justos). *En un mercado libre de arbitraje, para todo contrato F ,*

$$\{\mathbb{E}_{\mathbb{Q}}[F] : \mathbb{Q} \text{ PME}\} = \{\text{precios justos para } F\}.$$

Idea de la demostración. La inclusión $\subset$ es la *proposición 7.5*. Para la recíproca, si p es un precio justo, considérese el mercado ampliado $\widehat{M}_n = (M_n, \widehat{S}_n)$ con un activo adicional que vale p en $n = 0$ y F en $n = N$ (interpolado de cualquier manera sin arbitraje). Que p sea justo significa que el mercado ampliado es libre de arbitraje; por el *teorema 6.15* existe una PME $\widehat{\mathbb{Q}}$ para él, y su restricción $\mathbb{Q}$ al mercado original es una PME con $\mathbb{E}_{\mathbb{Q}}[F] = p$. ■

Ejemplo 7.7 (Valoración en el binomial). Tomemos el modelo binomial con $s = 100$, $a = -20\%$, $b = +10\%$, $N = 2$, y una call de strike $K = 91$, es decir $F = (S_2 - 91)^+$. Por (6.2), $q = \frac{-a}{b-a} = \frac{0.2}{0.3} = \frac{2}{3}$. Los precios terminales son

$$S_2 \in \{121, 88, 64\} \quad \text{con} \quad \mathbb{Q}\{\cdot\} = \{q^2, 2q(1-q), (1-q)^2\} = \left\{\frac{4}{9}, \frac{4}{9}, \frac{1}{9}\right\},$$

y los pagos $F \in \{30, 0, 0\}$. Luego

$$p^* = \mathbb{E}_{\mathbb{Q}}[F] = 30 \cdot \frac{4}{9} = \frac{40}{3} \approx 13,33.$$

Proposición 7.8 (Estructura markoviana en el binomial). *En el modelo binomial, si $F = h(S_N)$ depende solo del precio terminal, entonces para todo n*

$$\mathbb{E}_{\mathbb{Q}}[F | \mathcal{F}_n] = \varphi_n(S_n)$$

para una función φ_n determinista. En particular, el precio en cada instante depende solo del precio actual del subyacente, no de cómo se llegó a él.

Demostración. Bajo $\mathbb{Q}$ los rendimientos $\rho_{n+1}, \dots, \rho_N$ son independientes de $\mathcal{F}_n$, y $S_N = S_n \prod_{j=n+1}^N (1 + \rho_j)$. Por lo tanto

$$\mathbb{E}_{\mathbb{Q}}[h(S_N) | \mathcal{F}_n] = \varphi_n(S_n), \quad \varphi_n(x) = \mathbb{E}_{\mathbb{Q}}\left[h\left(x \prod_{j=n+1}^N (1 + \rho_j)\right)\right],$$

por la propiedad de sustitución para la esperanza condicional con independencia. ■

7.3 Cobertura y completitud

Definición 7.9 (Cobertura). Una *cobertura* (hedge) del contrato F es una estrategia autofinanciada Γ tal que

$$V_N^\Gamma = F \quad (\text{en todo punto de } \Omega).$$

Proposición 7.10. *Si F admite una cobertura Γ , entonces el conjunto de precios justos de F se reduce a un punto:*

$$\{\text{precios justos de } F\} = \{V_0^\Gamma\}.$$

Demostración. Para toda PME $\mathbb{Q}$, (V_n^Γ) es $\mathbb{Q}$ -martingala, luego $\mathbb{E}_{\mathbb{Q}}[F] = \mathbb{E}_{\mathbb{Q}}[V_N^\Gamma] = V_0^\Gamma$: todos los $\mathbb{E}_{\mathbb{Q}}[F]$ coinciden. El [teorema 7.6](#) concluye. (Directamente: si $p < V_0^\Gamma$, el holder compra el contrato y vende la cartera; si $p > V_0^\Gamma$, el writer hace lo inverso. En ambos casos hay arbitraje.) ■

Definición 7.11 (Mercado completo). El mercado (M_n) es *completo* si todo contrato F admite una cobertura.

Observación 7.12. Si (M_n) es completo y $\mathbb{Q}, \tilde{\mathbb{Q}}$ son dos PME, entonces $\mathbb{Q} = \tilde{\mathbb{Q}}$ sobre $\mathcal{F}_N$. En efecto, dado $A \in \mathcal{F}_N$, el contrato $F = \mathbf{1}_A$ tiene cobertura Γ y por lo tanto

$$\mathbb{Q}(A) = \mathbb{E}_{\mathbb{Q}}[\mathbf{1}_A] = V_0^\Gamma = \mathbb{E}_{\tilde{\mathbb{Q}}}[\mathbf{1}_A] = \tilde{\mathbb{Q}}(A).$$

El recíproco es el segundo teorema fundamental. Para demostrarlo trabajaremos con los dos subespacios de $\mathbb{R}^{|\Omega|}$ ya introducidos. Recordemos el producto interno $\eta \cdot \xi = \sum_{\omega} \eta(\omega) \xi(\omega)$ y

$$\mathcal{A} = \left\{ (\Gamma \cdot M)_N : \Gamma \text{ predecible} \right\} = \left\{ H : \exists \Gamma \text{ autofinanciada con } V_0^\Gamma = 0, V_N^\Gamma = H \right\},$$

donde la segunda igualdad es consecuencia de la [proposición 6.7](#). Por la [proposición 6.5](#) se tiene entonces:

Observación 7.13. F admite cobertura $\iff$ existe $c \in \mathbb{R}$ con $F - c \in \mathcal{A}$.

Para una probabilidad $\mathbb{Q}$ sobre Ω escribimos

$$B_{\mathbb{Q}} = \{ H : \Omega \rightarrow \mathbb{R} \text{ contrato} : \mathbb{E}_{\mathbb{Q}}[H] = 0 \} = \mathbb{Q}^\perp,$$

un hiperplano de $\mathbb{R}^{|\Omega|}$. Por la [observación 6.16](#), si $\mathbb{Q}$ es PME entonces $\mathcal{A} \subset B_{\mathbb{Q}}$.

Lema 7.14. *Sea $\mathbb{Q}$ una PME. Si la inclusión $\mathcal{A} \subset B_{\mathbb{Q}}$ es estricta, entonces existe otra PME $\tilde{\mathbb{Q}}$ con*

$$\tilde{\mathbb{Q}}|_{\mathcal{F}_N} \neq \mathbb{Q}|_{\mathcal{F}_N}.$$

Demostración. Si $\mathcal{A} \subsetneq B_{\mathbb{Q}}$, existe $g \in B_{\mathbb{Q}}$ con $g \notin \mathcal{A}$. Descomponiendo g en su proyección sobre $\mathcal{A}$ y su componente ortogonal, y observando que esta última sigue estando en $B_{\mathbb{Q}}$ (pues $\mathbb{Q} \perp \mathcal{A}$), podemos suponer además

$$g \perp \mathcal{A}, \quad g \neq 0, \quad \mathbb{Q} \cdot g = 0.$$

Definamos, para $\delta > 0$,

$$\tilde{\mathbb{Q}} = \theta(\mathbb{Q} + \delta g), \quad \theta = \left(1 + \delta \sum_{\omega} g(\omega) \right)^{-1},$$

de modo que $\tilde{\mathbb{Q}}$ suma 1. Como Ω es finito y $\mathbb{Q}(\{\omega\}) > 0$ para todo ω , eligiendo δ suficientemente pequeño se tiene simultáneamente $\theta > 0$ y $\mathbb{Q}(\{\omega\}) + \delta g(\omega) > 0$ para todo ω ; luego $\tilde{\mathbb{Q}}$ es una probabilidad equivalente. Por construcción $\tilde{\mathbb{Q}} \cdot \eta = \theta(\mathbb{Q} \cdot \eta + \delta g \cdot \eta) = 0$ para todo $\eta \in \mathcal{A}$, es decir $\tilde{\mathbb{Q}} \perp \mathcal{A}$, y por la [observación 6.16](#) se trata de una PME.

Finalmente, $\tilde{\mathbb{Q}} \neq \mathbb{Q}$: si fuesen iguales sobre $\mathcal{F}_N = 2^\Omega$, entonces

$$0 = \mathbb{E}_{\mathbb{Q}}[g] = \mathbb{E}_{\tilde{\mathbb{Q}}}[g] = \tilde{\mathbb{Q}} \cdot g = \theta(\mathbb{Q} \cdot g + \delta g \cdot g) = \theta \delta \|g\|^2 > 0,$$

absurdo. ■

Corolario 7.15. Si la PME $\mathbb{Q}$ es única, entonces $\mathcal{A} = B_{\mathbb{Q}}$.

Teorema 7.16 (Segundo teorema fundamental de las finanzas). Sea (M_n) un mercado libre de arbitraje. Son equivalentes:

- (I) (M_n) es completo;
- (II) la PME es única.

Demostración. (i) $\Rightarrow$ (ii) es la observación 7.12.

(ii) $\Rightarrow$ (i): sea $\mathbb{Q}$ la única PME y F un contrato cualquiera. Entonces

$$F - \mathbb{E}_{\mathbb{Q}}[F] \in B_{\mathbb{Q}} \stackrel{\text{corolario 7.15}}{=} \mathcal{A},$$

y por la observación 7.13, F admite cobertura (con $c = \mathbb{E}_{\mathbb{Q}}[F]$, que por la proposición 7.10 es su precio). ■

Corolario 7.17 (Fórmula de valoración). En un mercado completo y libre de arbitraje, con PME $\mathbb{Q}$, todo contrato F tiene un único precio justo, dado en cada instante por

$$\boxed{\pi_n(F) = \mathbb{E}_{\mathbb{Q}}[F | \mathcal{F}_n]}, \quad n = 0, 1, \dots, N,$$

y la cobertura Γ satisface $V_n^{\Gamma} = \pi_n(F)$.

Ejemplo 7.18. El modelo binomial con $a < 0 < b$ es completo, pues por el ejemplo 6.14 su PME es única.

Proposición 7.19 (Toda $\mathbb{Q}$ -martingala es el valor de una cartera). Sea (M_n) completo con PME $\mathbb{Q}$. Entonces, para todo proceso $(X_n)_{n=0}^N$ que sea $\mathbb{Q}$ -martingala, existe una estrategia autofinanciada Γ con

$$V_n^{\Gamma} = X_n \quad \text{para todo } n = 0, 1, \dots, N.$$

Demostración. X_N es un contrato, luego admite una cobertura Γ : $V_N^{\Gamma} = X_N$. Ahora bien, tanto (V_n^{Γ}) como (X_n) son $\mathbb{Q}$ -martingalas, de modo que

$$V_n^{\Gamma} = \mathbb{E}_{\mathbb{Q}}[V_N^{\Gamma} | \mathcal{F}_n] = \mathbb{E}_{\mathbb{Q}}[X_N | \mathcal{F}_n] = X_n. \quad \blacksquare$$

El puente hacia la Parte IV

La [proposición 7.19](#) es el *teorema de representación de martingalas* en versión discreta: en un mercado completo, toda $\mathbb{Q}$ -martingala se escribe como $X_n = X_0 + (\Gamma \cdot M)_n$ para algún proceso predecible Γ . Su análogo continuo —toda martingala respecto de la filtración browniana se escribe como $X_t = X_0 + \int_0^t H_s dB_s$ — es el [teorema 18.3](#), y es exactamente lo que permite cubrir opciones en el modelo de Black–Scholes.

Fórmula de Cox–Ross–Rubinstein

Cerramos el capítulo con la fórmula explícita en el modelo binomial, que reencontraremos como límite en el [capítulo 18](#).

Proposición 7.20 (Fórmula CRR). *En el modelo binomial (s, a, b) con $a < 0 < b$ y $q = \frac{-a}{b-a}$, el precio en $n = 0$ de una call europea de strike K y vencimiento N es*

$$\pi_0 = \sum_{k=0}^N \binom{N}{k} q^k (1-q)^{N-k} \left(s(1+b)^k (1+a)^{N-k} - K \right)^+.$$

Demostración. Bajo $\mathbb{Q}$ el número k de subidas es $\text{Bin}(N, q)$ y $S_N = s(1+b)^k (1+a)^{N-k}$; aplicar el [corolario 7.17](#) con $n = 0$. ■

Opciones americanas y la envolvente de Snell

Una opción americana puede ejercerse en cualquier instante hasta el vencimiento. Valorarla ya no es calcular una esperanza: hay que resolver un *problema de parada óptima*. La herramienta es la envolvente de Snell, la menor supermartingala que domina al pago, y el tiempo óptimo de ejercicio es el primer instante en que la envolvente toca al pago.

Trabajamos en un mercado completo y libre de arbitraje, con PME única $\mathbb{Q}$, sobre Ω finito, con $J = \{0, 1, \dots, N\}$.

8.1 Contratos americanos

Definición 8.1 (Contrato americano). Un *contrato del tipo americano* con conjunto de tiempos de ejercicio $J = \{0, 1, \dots, N\}$ es una familia $(f_n)_{n \in J}$ con $f_n : \Omega \rightarrow \mathbb{R}$ $\mathcal{F}_n$ -medible para cada n .

Ejemplo 8.2. La *call americana* sobre S de strike K es $f_n = (S_n - K)^+$; la *put americana*, $f_n = (K - S_n)^+$.

Definición 8.3 (Tiempo de ejercicio). Dado un contrato americano $(f_n)_{n \in J}$, un *tiempo de ejercicio* es un tiempo de parada $\tau : \Omega \rightarrow J$ respecto de $(\mathcal{F}_n)$.

Que el tiempo de ejercicio sea un *tiempo de parada* es esencial: el holder decide si ejercer hoy usando únicamente la información de hoy, sin conocer el futuro.

Definición 8.4 (Sobrecobertura). Dado $(f_n)_{n \in J}$, una *sobrecobertura* es un par (Γ^*, τ^*) formado por una estrategia autofinanciada Γ^* y un tiempo de ejercicio τ^* tales que

- (I) $V_n^{\Gamma^*} \geq f_n$ para todo $n \in J$;
- (II) $V_{\tau^*}^{\Gamma^*} = f_{\tau^*}$.

La condición (i) dice que el writer, con la cartera Γ^* , puede atender el pago *en cualquier instante* en que el holder decida ejercer; la (ii), que la cartera no es excesiva: existe una forma de ejercer que la agota exactamente.

Proposición 8.5. *Si $(f_n)_{n \in J}$ posee una sobrecobertura (Γ^*, τ^*) , entonces existe un único precio justo para el contrato americano, y vale*

$$p(f) = V_0^{\Gamma^*}.$$

Demostración. Si el precio p fuese mayor que $V_0^{\Gamma^*}$, el writer vendería el contrato en p , montaría la cartera Γ^* por $V_0^{\Gamma^*} < p$, y por (i) podría pagar cualquier ejercicio del holder conservando la diferencia: arbitraje. Si fuese menor, el holder compraría el contrato, vendería en corto la cartera Γ^* y ejercería en τ^* , donde por (ii) el pago iguala exactamente el valor de la cartera: arbitraje de nuevo. ■

8.2 La envolvente de Snell

Definición 8.6 (Envolvente de Snell). Dado un contrato americano $(f_n)_{n \in J}$, su *envolvente de Snell* respecto de $\mathbb{Q}$ es el proceso $(\mathcal{E}_n)_{n \in J}$ definido por *recursión hacia atrás*:

$$\mathcal{E}_N = f_N, \quad \mathcal{E}_n = \max \left\{ f_n, \mathbb{E}_{\mathbb{Q}}[\mathcal{E}_{n+1} | \mathcal{F}_n] \right\}, \quad n = N-1, N-2, \dots, 0. \quad (8.1)$$

La lectura es transparente: en el instante n , el holder compara lo que obtendría ejerciendo *ahora* (f_n) con el valor de *esperar* ($\mathbb{E}_{\mathbb{Q}}[\mathcal{E}_{n+1} | \mathcal{F}_n]$), y se queda con lo mejor. Por construcción,

$$\mathcal{E}_n \geq f_n \quad \text{para todo } n \in J.$$

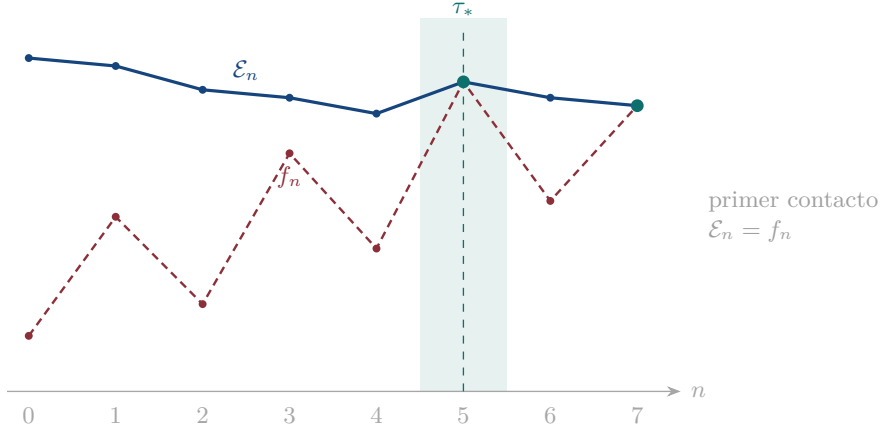


Figura 8.1: La envolvente de Snell $\mathcal{E}$ domina al pago f y es supermartingala. El tiempo óptimo τ_* es el primer instante en que ambas se tocan; hasta ahí, $\mathcal{E}$ es martingala.

Proposición 8.7. $(\mathcal{E}_n)_{n \in J}$ es una $\mathbb{Q}$ -supermartingala.

Demostración. Es adaptado e integrable por construcción (Ω es finito), y de (8.1) se sigue de inmediato $\mathcal{E}_n \geq \mathbb{E}_{\mathbb{Q}}[\mathcal{E}_{n+1} | \mathcal{F}_n]$. ■

Proposición 8.8 ($\mathcal{E}$ es la *menor* supermartingala que domina a f). Si $(Z_n)_{n \in J}$ es una $\mathbb{Q}$ -supermartingala con $Z_n \geq f_n$ para todo $n \in J$, entonces $Z_n \geq \mathcal{E}_n$ para todo $n \in J$.

Demostración. Por inducción hacia atrás. En $n = N$: $Z_N \geq f_N = \mathcal{E}_N$. Si $Z_{n+1} \geq \mathcal{E}_{n+1}$, entonces por monotonía de la esperanza condicional y la propiedad de supermartingala,

$$Z_n \geq \mathbb{E}_{\mathbb{Q}}[Z_{n+1} | \mathcal{F}_n] \geq \mathbb{E}_{\mathbb{Q}}[\mathcal{E}_{n+1} | \mathcal{F}_n],$$

y como también $Z_n \geq f_n$, se tiene $Z_n \geq \max\{f_n, \mathbb{E}_{\mathbb{Q}}[\mathcal{E}_{n+1} | \mathcal{F}_n]\} = \mathcal{E}_n$. ■

Observación. Si $(f_n)_{n \in J}$ es ya una $\mathbb{Q}$ -martingala, entonces $\mathcal{E}_n = f_n$ para todo n : no hay valor alguno en la opcionalidad. Es el caso, por ejemplo, de la call americana sin dividendos (proposición 8.13).

8.3 El tiempo óptimo de ejercicio

Definición 8.9. Definimos

$$\tau_* = \min\{n \in J : \mathcal{E}_n = f_n\}.$$

El mínimo existe porque $\mathcal{E}_N = f_N$, de modo que $\tau_* \leq N$; y es un tiempo de parada porque $\{\tau_* = n\}$ se expresa en términos de $\mathcal{E}_0, f_0, \dots, \mathcal{E}_n, f_n$, todas $\mathcal{F}_n$ -medibles.

Proposición 8.10. *El proceso detenido $(\mathcal{E}_{n \wedge \tau_*})_{n \in J}$ es una $\mathbb{Q}$ -martingala.*

Demostración. Escribimos, como en la [proposición 5.5](#),

$$\mathcal{E}_{(n+1) \wedge \tau_*} - \mathcal{E}_{n \wedge \tau_*} = \mathbf{1}_{\{\tau_* \geq n+1\}} (\mathcal{E}_{n+1} - \mathcal{E}_n),$$

y como $\{\tau_* \geq n+1\} = \{\tau_* \leq n\}^c \in \mathcal{F}_n$,

$$\mathbb{E}_{\mathbb{Q}}[\mathcal{E}_{(n+1) \wedge \tau_*} - \mathcal{E}_{n \wedge \tau_*} \mid \mathcal{F}_n] = \mathbf{1}_{\{\tau_* \geq n+1\}} (\mathbb{E}_{\mathbb{Q}}[\mathcal{E}_{n+1} \mid \mathcal{F}_n] - \mathcal{E}_n).$$

Ahora bien, sobre el suceso $\{\tau_* \geq n+1\}$ se tiene, por definición de τ_* , que $\mathcal{E}_n > f_n$, y entonces la recursión (8.1) fuerza $\mathcal{E}_n = \mathbb{E}_{\mathbb{Q}}[\mathcal{E}_{n+1} \mid \mathcal{F}_n]$. Por lo tanto el paréntesis se anula sobre ese suceso y la esperanza condicional del incremento es 0. ■

Teorema 8.11 (Parada óptima). *Con τ_* como en la [definición 8.9](#),*

$$\mathcal{E}_0 = \mathbb{E}_{\mathbb{Q}}[f_{\tau_*}] = \max \{ \mathbb{E}_{\mathbb{Q}}[f_{\tau}] : \tau \text{ tiempo de parada con valores en } J \}.$$

Es decir, τ_ resuelve el problema de parada óptima y $\mathcal{E}_0$ es su valor.*

Demostración. Por la [proposición 8.10](#),

$$\mathcal{E}_0 = \mathcal{E}_{0 \wedge \tau_*} = \mathbb{E}_{\mathbb{Q}}[\mathcal{E}_{N \wedge \tau_*}] = \mathbb{E}_{\mathbb{Q}}[\mathcal{E}_{\tau_*}] = \mathbb{E}_{\mathbb{Q}}[f_{\tau_*}],$$

usando en la última igualdad la definición de τ_* . Por otro lado, para cualquier tiempo de parada τ con valores en J , la [proposición 8.7](#) y el muestreo opcional con horizonte finito ([lema 5.6](#), versión supermartingala) dan

$$\mathbb{E}_{\mathbb{Q}}[f_{\tau}] \leq \mathbb{E}_{\mathbb{Q}}[\mathcal{E}_{\tau}] \leq \mathcal{E}_0,$$

de modo que $\mathcal{E}_0$ es cota superior y se alcanza en τ_* . ■

8.4 Existencia de sobrecobertura

Teorema 8.12. *Si el mercado (M_n) es completo, entonces todo contrato americano $(f_n)_{n \in J}$ posee una sobrecobertura, y su precio es*

$$p(f) = \mathcal{E}_0 = \max_{\tau} \mathbb{E}_{\mathbb{Q}}[f_{\tau}].$$

Demostración. Por la descomposición de Doob (teorema 2.18, versión supermartingala) aplicada a $\mathcal{E}$ bajo $\mathbb{Q}$,

$$\mathcal{E}_n = X_n - A_n, \quad (8.2)$$

con (X_n) una $\mathbb{Q}$ -martingala, $X_0 = \mathcal{E}_0$, y (A_n) predecible, creciente, $A_0 = 0$.

Como el mercado es completo, la proposición 7.19 da una estrategia autofinanciada Γ^* con

$$V_n^{\Gamma^*} = X_n \quad \text{para todo } n \in J.$$

Verifiquemos que (Γ^*, τ_*) es una sobrecobertura:

- (i) $V_n^{\Gamma^*} = X_n = \mathcal{E}_n + A_n \geq \mathcal{E}_n \geq f_n$, usando $A_n \geq 0$.
- (ii) Por la proposición 8.10, $\mathcal{E}_{n \wedge \tau_*}$ es martingala; comparando con (8.2) detenida en τ_* , el proceso $A_{n \wedge \tau_*}$ es a la vez predecible, creciente y martingala con valor inicial 0, luego $A_{n \wedge \tau_*} \equiv 0$ (unicidad en la descomposición de Doob). En particular $A_{\tau_*} = 0$ y

$$V_{\tau_*}^{\Gamma^*} = X_{\tau_*} = \mathcal{E}_{\tau_*} + A_{\tau_*} = \mathcal{E}_{\tau_*} = f_{\tau_*}.$$

Por la proposición 8.5 el precio es $V_0^{\Gamma^*} = X_0 = \mathcal{E}_0$, y el teorema 8.11 da la representación como máximo. ■

Qué significa el proceso creciente A

A_n es el *consumo acumulado* del writer: la parte del valor de la cartera que puede retirar sin poner en riesgo su capacidad de pagar. Es estrictamente creciente exactamente en los instantes en que ya es óptimo ejercer y el holder no lo hace. Por eso A se anula hasta τ_* : un holder racional no le regala nada al writer.

8.5 La call americana no se ejerce antes

Cerramos con un resultado clásico que ilustra el poder de la maquinaria.

Proposición 8.13. *En un mercado libre de arbitraje con numerario constante $B_n \equiv 1$ (tasa de interés nula) y $S \geq 0$, el precio de la call americana $f_n = (S_n - K)^+$ coincide con el de la call europea $F = (S_N - K)^+$, y $\tau_* = N$ es óptimo: nunca conviene ejercer antes del vencimiento.*

Demostración. Bajo $\mathbb{Q}$, (S_n) es martingala y $x \mapsto (x - K)^+$ es convexa, de modo que por la [proposición 2.9](#) el proceso $f_n = (S_n - K)^+$ es una $\mathbb{Q}$ -submartingala. Para todo tiempo de parada $\tau \leq N$, el muestreo opcional da

$$\mathbb{E}_{\mathbb{Q}}[f_{\tau}] \leq \mathbb{E}_{\mathbb{Q}}[f_N] = \mathbb{E}_{\mathbb{Q}}[(S_N - K)^+].$$

Luego el máximo sobre τ se alcanza en $\tau \equiv N$ y el precio americano coincide con el europeo. ■

Observación. El resultado *falla* para la put americana: $x \mapsto (K - x)^+$ es convexa, pero está acotada por K , y el argumento de la submartingala da la desigualdad en la dirección inútil. En efecto, para la put el ejercicio anticipado sí puede ser óptimo (si S_n cae mucho, conviene cobrar $K - S_n$ de inmediato), y la región de ejercicio $\{\mathcal{E}_n = f_n\}$ es no trivial. También falla para la call si hay dividendos o tasa de interés positiva sobre el numerario.

Resumen de la Parte II

Finanzas	Martingalas
Estrategia Γ	proceso predecible
Valor V_n^{Γ} de una cartera autofinanciada	$V_0^{\Gamma} + (\Gamma \cdot M)_n$: transformada de martingalas
Ausencia de arbitraje	existe $\mathbb{Q} \sim \mathbb{P}$ que hace martingala a M
Precio de un derivado europeo	$\mathbb{E}_{\mathbb{Q}}[F \mathcal{F}_n]$
Mercado completo	la PME es única; toda $\mathbb{Q}$ -martingala es el valor de una cartera
Precio de un derivado americano	$\mathcal{E}_0 = \max_{\tau} \mathbb{E}_{\mathbb{Q}}[f_{\tau}]$: envolvente de Snell
Cobertura del americano	parte martingala de la descomposición de Doob de $\mathcal{E}$

Parte III

Procesos a tiempo continuo

CAPÍTULO 9

Trayectorias càdlàg y tiempos de parada continuos

Al pasar al tiempo continuo aparecen dificultades que no existían en el caso discreto: hay *no numerables* instantes, de modo que sucesos como $\{\sup_{s \leq t} X_s > a\}$ no son automáticamente medibles, y hay que decidir en qué espacio viven las trayectorias. La respuesta estándar es trabajar con trayectorias *càdlàg* y hacer del espacio de trayectorias el propio espacio de probabilidad. Este capítulo monta ese andamiaje; se usará tanto para el proceso de Poisson (capítulo 10) como para el movimiento Browniano (capítulo 13).

9.1 El espacio $\mathbb{D}_S$

Definición 9.1 (Función càdlàg). Sea (S, d) un espacio métrico. Una función $\alpha : [0, \infty) \rightarrow S$ es *càdlàg*¹ si

$$\begin{aligned} \text{para todo } t > 0 : \quad & \lim_{s \rightarrow t^-} \alpha(s) \text{ existe;} \\ \text{para todo } t \geq 0 : \quad & \lim_{s \rightarrow t^+} \alpha(s) = \alpha(t). \end{aligned}$$

Denotamos

$$\mathbb{D}_S = \{\alpha : [0, \infty) \rightarrow S : \alpha \text{ es càdlàg}\}, \quad \Delta\alpha(t) = \alpha(t) - \alpha(t^-), \quad t > 0,$$

y $\mathcal{C}_S \subset \mathbb{D}_S$ el subconjunto de las funciones continuas.

¹Del francés *continue à droite, limite à gauche*.

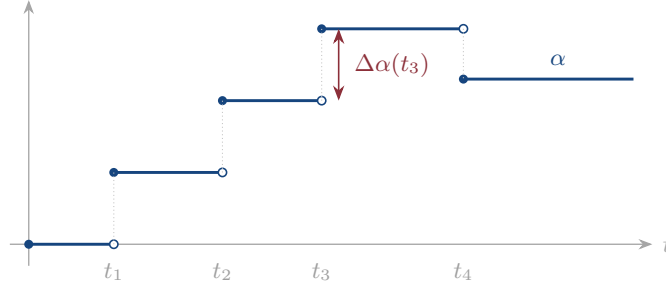


Figura 9.1: Una trayectoria càdlàg. En cada instante de salto el valor es el límite por la derecha (punto lleno); el límite por la izquierda existe pero puede diferir (punto hueco). Las trayectorias de un proceso de Poisson son de este tipo, con saltos de tamaño $+1$.

Proposición 9.2 (Estructura de los saltos). *Sea $\alpha \in \mathbb{D}_S$ y $T > 0$. Entonces:*

- (a) *para todo $\epsilon > 0$, el conjunto $\{0 \leq t \leq T : d(\alpha(t), \alpha(t^-)) > \epsilon\}$ es finito;*
- (b) *el conjunto $\{0 \leq t \leq T : \Delta\alpha(t) \neq 0\}$ es numerable;*
- (c) *$\{\alpha(t) : 0 \leq t \leq T\}$ es acotado en (S, d) .*

Demostración. (a) Fijemos $\epsilon > 0$ y construyamos recursivamente

$$s_0 = 0,$$

$$s_n = \inf \left\{ t > s_{n-1} : d(\alpha(t), \alpha(s_{n-1})) > \frac{\epsilon}{2} \right\},$$

con la convención $s_n = \infty$ si el ínfimo se toma sobre el vacío. Por la desigualdad triangular, el *módulo de oscilación* sobre cada tramo está controlado:

$$w_\alpha[s_{n-1}, s_n] = \sup_{t, t' \in [s_{n-1}, s_n]} d(\alpha(t), \alpha(t')) \leq \epsilon, \quad d(\alpha(s_n), \alpha(s_{n-1})) \geq \frac{\epsilon}{2}.$$

Los s_n no pueden acumularse en un punto finito: si $s_n \uparrow s < \infty$, la existencia del límite por la izquierda $\alpha(s^-)$ obligaría a que $\alpha(s_n)$ fuese de Cauchy, contradiciendo $d(\alpha(s_n), \alpha(s_{n-1})) \geq \epsilon/2$. Por lo tanto solo finitos s_n caen en $[0, T]$, y todo salto de tamaño $> \epsilon$ debe ocurrir en alguno de ellos.

- (b) $\{t \leq T : \Delta\alpha(t) \neq 0\} = \bigcup_{n \in \mathbb{N}} \underbrace{\{t \leq T : d(\alpha(t), \alpha(t^-)) > 1/n\}}_{\text{finito por (a)}}$, unión

numerable de conjuntos finitos.

- (c) Con la misma partición $0 = s_0 < s_1 < \dots < s_m$ de $[0, T]$ (finita, para $\epsilon = 1$), $\alpha([0, T])$ está contenido en la unión finita de bolas de radio 1 centradas en los $\alpha(s_j)$, más los finitos valores $\alpha(s_j)$. ■

9.2 Medibilidad y la filtración canónica

Sobre $\mathbb{D}_S$ definimos el *proceso canónico* de coordenadas

$$X_t : \mathbb{D}_S \rightarrow S, \quad X_t(\alpha) = \alpha(t), \quad t \geq 0,$$

y las σ -álgebras

$$\mathcal{F} = \sigma(X_t : t \geq 0), \quad \mathcal{F}_t = \sigma(X_s : s \leq t), \quad t \geq 0.$$

Queda así construido un *espacio medible filtrado*

$$[(\mathbb{D}_S, \mathcal{F}), (\mathcal{F}_t)_{t \geq 0}].$$

Observación 9.3 (Cilindros). $\mathcal{F}$ está generada por el π -sistema de los *cilindros*:

$$\mathcal{F} = \sigma \left(\underbrace{\bigcup_{\{t_1, \dots, t_k\} \subset [0, \infty)} \sigma(X_{t_1}, \dots, X_{t_k})}_{\pi\text{-sistema}} \right).$$

En consecuencia (teorema π - λ , [teorema A.2](#)), si $\mathbb{P}$ y $\tilde{\mathbb{P}}$ son probabilidades sobre $(\mathbb{D}_S, \mathcal{F})$ tales que

$$\mathbb{P} \circ (X_{t_1}, \dots, X_{t_k})^{-1} = \tilde{\mathbb{P}} \circ (X_{t_1}, \dots, X_{t_k})^{-1}$$

para toda elección finita $t_1 < \dots < t_k$, entonces $\mathbb{P} = \tilde{\mathbb{P}}$. *Las distribuciones finito-dimensionales determinan la ley.*

Observación 9.4. Gracias a la continuidad por la derecha, $X_t = \lim_{q \downarrow t, q \in \mathbb{Q}} X_q$, de modo que

$$\mathcal{F} = \sigma(X_t : t \in [0, \infty) \cap \mathbb{Q}).$$

Esta es la observación que rescata la medibilidad de cantidades como $\sup_{s \leq t} X_s$: basta tomar el supremo sobre los racionales.

El diagrama siguiente resume cómo se transporta un proceso definido sobre un espacio de probabilidad abstracto al espacio de trayectorias.

$$\begin{array}{ccc}
 (\Omega, \mathcal{F}^\Omega) & \xrightarrow{N} & (\mathbb{D}_S, \sigma(X)) \\
 & \searrow (N_{t_1}, \dots, N_{t_k}) & \downarrow (X_{t_1}, \dots, X_{t_k}) \\
 & & (S^k, \mathcal{B}_S^{\otimes k})
 \end{array}$$

Figura 9.2: El diagrama conmuta: $N_t = X_t \circ N$. En consecuencia, $N : \Omega \rightarrow \mathbb{D}_S$ es medible si y solo si $(N_{t_1}, \dots, N_{t_k})$ es medible para toda elección finita de instantes, y la ley $\mathbb{P} \circ N^{-1}$ sobre $\mathbb{D}_S$ queda determinada por las leyes finito-dimensionales.

Definición 9.5 (Operadores de traslación). Para $t \geq 0$ definimos $\theta_t : \mathbb{D}_S \rightarrow \mathbb{D}_S$ por

$$(\theta_t \alpha)(s) = \alpha(t + s), \quad s \geq 0.$$

Ejercicio 9.6. θ_t es medible para todo $t \geq 0$.

Solución. Como $X_s \circ \theta_t = X_{t+s}$ y los cilindros generan $\mathcal{F}$, basta ver que la preimagen de un cilindro es medible:

$$\theta_t^{-1}(\{X_{t_1} \in A_1, \dots, X_{t_n} \in A_n\}) = \{X_{t_1+t} \in A_1, \dots, X_{t_n+t} \in A_n\} \in \mathcal{F}.$$

Es el mismo principio que en la [figura 9.2](#): una aplicación con valores en $\mathbb{D}_S$ es medible si y solo si lo son sus coordenadas. ■

Ejercicio 9.7. Si $A \in \mathcal{F}_s$, entonces $\theta_t^{-1}(A) \in \mathcal{F}_{t+s}$.

Solución. Basta tomar A un cilindro, pues estos generan $\mathcal{F}_s$: $A = \{X_{t_1} \in A_1, \dots, X_{t_n} \in A_n\}$ con $t_i \leq s$. Entonces

$$\theta_t^{-1}(A) = \{X_{t_1+t} \in A_1, \dots, X_{t_n+t} \in A_n\},$$

y $t_i \leq s$ implica $t + t_i \leq t + s$, de modo que el conjunto está en $\mathcal{F}_{t+s}$. ■

Para $\tau : \mathbb{D}_S \rightarrow [0, \infty]$ y $F : \mathbb{D}_S \rightarrow \mathbb{R}$ definimos además

$$F \circ \theta_\tau(\alpha) = \begin{cases} F(\theta_{\tau(\alpha)} \alpha), & \text{si } \tau(\alpha) < \infty, \\ 0, & \text{si } \tau(\alpha) = \infty. \end{cases} \quad (9.1)$$

9.3 Tiempos de parada y tiempos opcionales

Fijemos un espacio medible filtrado $[(\Omega, \mathcal{F}), (\mathcal{F}_t)_{t \geq 0}]$.

Definición 9.8. Sea $\tau : \Omega \rightarrow [0, \infty]$.

- τ es un *tiempo de parada* si $\{\tau \leq t\} \in \mathcal{F}_t$ para todo $t \geq 0$; en tal caso

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t \quad \forall t \geq 0\}.$$

- τ es un *tiempo opcional* si $\{\tau < t\} \in \mathcal{F}_t$ para todo $t \geq 0$; en tal caso

$$\mathcal{F}_{\tau+} = \{A \in \mathcal{F} : A \cap \{\tau < t\} \in \mathcal{F}_t \quad \forall t > 0\}.$$

En tiempo discreto los dos conceptos coinciden

Si el tiempo es discreto, $\{\tau < n\} = \{\tau \leq n - 1\}$ y no hay distinción. En tiempo continuo, $\{\tau < t\}$ y $\{\tau \leq t\}$ son genuinamente distintos, y la diferencia se mide con las σ -álgebras $\mathcal{F}_{t+}$. Toda esta sección consiste en administrar esa diferencia.

Definición 9.9. Dada una filtración $(\mathcal{F}_t)_{t \geq 0}$ ponemos

$$\mathcal{F}_{t-} = \sigma\left(\bigcup_{s < t} \mathcal{F}_s\right), \quad \mathcal{F}_{t+} = \bigcap_{s > t} \mathcal{F}_s = \bigcap_{\epsilon > 0} \mathcal{F}_{t+\epsilon}.$$

La filtración es *continua por la derecha* si $\mathcal{F}_t = \mathcal{F}_{t+}$ para todo t , y *continua por la izquierda* si $\mathcal{F}_t = \mathcal{F}_{t-}$.

Observación 9.10. Todo tiempo de parada es opcional:

$$\{\tau < t\} = \bigcup_{n \geq 1} \underbrace{\{\tau \leq t - \frac{1}{n}\}}_{\in \mathcal{F}_{t-1/n}} \in \mathcal{F}_t.$$

Recíprocamente, si τ es opcional entonces

$$\{\tau \leq t\} = \bigcap_{n \geq 1} \underbrace{\{\tau < t + \frac{1}{n}\}}_{\in \mathcal{F}_{t+1/n}} \in \mathcal{F}_{t+},$$

y de hecho: τ es opcional $\iff \{\tau \leq t\} \in \mathcal{F}_{t+}$ para todo t .

Proposición 9.11. Si todo tiempo opcional es tiempo de parada, entonces la filtración $(\mathcal{F}_t)$ es continua por la derecha.

Demostración. Sea $A \in \mathcal{F}_{t+}$ y definamos $\tau = t \mathbf{1}_A + \infty \mathbf{1}_{A^c}$. Entonces $\{\tau < s\} = \emptyset$ si $s \leq t$ y $= A$ si $s > t$; como $A \in \mathcal{F}_{t+} \subset \mathcal{F}_s$ para $s > t$, resulta que τ es opcional. Por hipótesis τ es de parada, luego $A = \{\tau \leq t\} \in \mathcal{F}_t$. Por lo tanto $\mathcal{F}_{t+} \subset \mathcal{F}_t$, y la otra inclusión es trivial. ■

Proposición 9.12. *Si τ es un tiempo opcional y $\theta > 0$ es una constante, entonces $\tau + \theta$ es un tiempo de parada.*

Demostración. Si $0 \leq t < \theta$, entonces $\{\tau + \theta \leq t\} = \emptyset \in \mathcal{F}_t$. Si $t \geq \theta$,

$$\{\tau + \theta \leq t\} = \{\tau \leq t - \theta\} \in \mathcal{F}_{(t-\theta)+} \subset \mathcal{F}_t,$$

usando la observación 9.10 y $(t - \theta)^+ < t$ cuando $\theta > 0$. ■

Proposición 9.13. *τ es $(\mathcal{F}_t)$ -opcional si y solo si τ es un tiempo de parada para la filtración $\mathcal{H}_t := \mathcal{F}_{t+}$. Además, en tal caso $\mathcal{H}_\tau = \mathcal{F}_{\tau+}$.*

Demostración. La primera equivalencia es la observación 9.10. Para la segunda, sea $A \in \mathcal{H}_\tau$, es decir $A \cap \{\tau \leq t\} \in \mathcal{H}_t$ para todo t . Entonces

$$A \cap \{\tau < t\} = \bigcup_{n \geq 1} \left(A \cap \left\{ \tau \leq t - \frac{1}{n} \right\} \right) \in \mathcal{F}_t,$$

pues $A \cap \{\tau \leq t - 1/n\} \in \mathcal{H}_{t-1/n} = \mathcal{F}_{(t-1/n)+} \subset \mathcal{F}_t$. Luego $A \in \mathcal{F}_{\tau+}$.

Recíprocamente, supongamos $A \cap \{\tau < t\} \in \mathcal{F}_t$ para todo $t > 0$. Entonces, para todo $s > t$,

$$A \cap \{\tau \leq t\} = \bigcap_{n \geq 1} \left(A \cap \left\{ \tau < t + \frac{1}{n} \right\} \right) \in \mathcal{F}_s,$$

de donde $A \cap \{\tau \leq t\} \in \mathcal{F}_{t+} = \mathcal{H}_t$, esto es $A \in \mathcal{H}_\tau$. ■

Ejercicio 9.14. Sea $[(\Omega, \mathcal{F}), (\mathcal{F}_t)]$ un espacio medible filtrado y (τ_n) una sucesión de tiempos de parada con $\tau_n \downarrow \tau$. Pruebe que

- (a) τ es un tiempo opcional;
- (b) $\mathcal{F}_{\tau+} = \bigcap_n \mathcal{F}_{\tau_n+}$;
- (c) si además $\tau < \tau_n$ para todo n , entonces $\mathcal{F}_{\tau+} = \bigcap_n \mathcal{F}_{\tau_n}$.

Solución. (a) Como $\tau_n \downarrow \tau$, se tiene $\{\tau < t\} = \bigcup_n \{\tau_n < t\} \in \mathcal{F}_t$.

(b) De $\tau \leq \tau_n$ se sigue $\mathcal{F}_{\tau+} \subset \mathcal{F}_{\tau_n+}$ para todo n . Recíprocamente, si $A \in \bigcap_n \mathcal{F}_{\tau_n+}$, entonces

$$A \cap \{\tau < t\} = \bigcup_n \underbrace{(A \cap \{\tau_n < t\})}_{\in \mathcal{F}_t} \in \mathcal{F}_t,$$

luego $A \in \mathcal{F}_{\tau+}$.

(c) La inclusión $\bigcap_n \mathcal{F}_{\tau_n} \subset \bigcap_n \mathcal{F}_{\tau_n+} = \mathcal{F}_{\tau+}$ es (b). Para la recíproca, sea $A \in \mathcal{F}_{\tau+}$, de modo que $A \cap \{\tau < t\} \in \mathcal{F}_t$. Como $\tau < \tau_n$, se tiene $\{\tau_n \leq t\} \subset \{\tau < t\}$ y por lo tanto

$$A \cap \{\tau_n \leq t\} = \underbrace{(A \cap \{\tau < t\})}_{\in \mathcal{F}_t} \cap \{\tau_n \leq t\} \in \mathcal{F}_t,$$

es decir $A \in \mathcal{F}_{\tau_n}$ para todo n . ■

Observación 9.15 (Aproximación diádica). Dado un tiempo opcional $\tau : \Omega \rightarrow [0, \infty]$, definamos

$$\tau_n = \frac{k}{2^n} \text{ sobre } \left\{ \frac{k-1}{2^n} \leq \tau < \frac{k}{2^n} \right\}, \quad k = 1, 2, \dots, \quad \tau_n = \infty \text{ sobre } \{\tau = \infty\}.$$

Entonces cada τ_n es un tiempo de parada con imagen numerable, $\tau_n \downarrow \tau$, $\tau_n > \tau$ salvo cuando τ es diádico, y por el [ejercicio 9.14](#), $\mathcal{F}_{\tau+} = \bigcap_n \mathcal{F}_{\tau_n}$. Esta aproximación es la herramienta estándar para pasar de tiempos de parada discretos a tiempos generales, y se usará en la demostración de la propiedad de Markov fuerte ([capítulo 11](#)).

Verificación. Basta ver que $\{\tau_n = i/2^n\} \in \mathcal{F}_{i/2^n}$. Ahora,

$$\left\{ \tau < \frac{i}{2^n} \right\} \in \mathcal{F}_{i/2^n} \quad (\text{por ser } \tau \text{ opcional}),$$

y entonces

$$\left\{ \frac{i-1}{2^n} \leq \tau < \frac{i}{2^n} \right\} = \left\{ \tau < \frac{i}{2^n} \right\} \cap \underbrace{\left\{ \tau < \frac{i-1}{2^n} \right\}^c}_{\in \mathcal{F}_{(i-1)/2^n} \subset \mathcal{F}_{i/2^n}} \in \mathcal{F}_{i/2^n},$$

$$\text{y } \{\tau_n \leq i/2^n\} = \bigcup_{j \leq i} \{\tau_n = j/2^n\} \in \mathcal{F}_{i/2^n}. \quad \blacksquare$$

9.4 Procesos medibles, adaptados y progresivos

Definición 9.16. Sea $X = (X_t)_{t \geq 0}$ un proceso con valores en un espacio medible $(S, \mathcal{B}_S)$.

- X es *medible* si la aplicación $(\omega, t) \mapsto X_t(\omega)$, definida sobre $\Omega \times \mathbb{R}_+$, es $\mathcal{F} \otimes \mathcal{B}_{\mathbb{R}_+}$ -medible.
- X es *adaptado* si X_t es $\mathcal{F}_t$ -medible para todo $t \geq 0$.
- X es *progresivo* (o *progresivamente medible*) si para todo $t \geq 0$ la aplicación $(\omega, s) \mapsto X_s(\omega)$, definida sobre $\Omega \times [0, t]$, es $\mathcal{F}_t \otimes \mathcal{B}_{[0, t]}$ -medible.

Progresivo $\Rightarrow$ medible y adaptado; el recíproco es falso en general, pero la continuidad por la derecha lo salva.

Proposición 9.17. Sea $X = (X_t)_{t \geq 0}$ un proceso adaptado con valores en un espacio métrico (S, d) cuyas trayectorias $t \mapsto X_t(\omega)$ son continuas por la derecha. Entonces X es progresivo.

Demostración. Fijemos $t > 0$. Para $n \geq 1$ y $s \in [0, t]$ definimos la discretización

$$X_s^n = X_{kt/n} \text{ si } s \in \left[\frac{(k-1)t}{n}, \frac{kt}{n} \right), \quad k \in \{1, \dots, n\}, \quad X_t^n = X_t.$$

La continuidad por la derecha asegura que $X_s^n(\omega) \rightarrow X_s(\omega)$ para todo (ω, s) . Ahora, para $A \in \mathcal{B}_S$,

$$\{(\omega, s) \in \Omega \times [0, t] : X_s^n(\omega) \in A\} = (\{X_t \in A\} \times \{t\}) \cup \bigcup_{k=1}^n \left(\underbrace{\{X_{kt/n} \in A\}}_{\in \mathcal{F}_t} \times \underbrace{\left[\frac{(k-1)t}{n}, \frac{kt}{n} \right)}_{\in \mathcal{B}_{[0, t]}} \right),$$

que claramente pertenece a $\mathcal{F}_t \otimes \mathcal{B}_{[0, t]}$. Como el límite puntual de funciones medibles es medible, X restringido a $\Omega \times [0, t]$ es $\mathcal{F}_t \otimes \mathcal{B}_{[0, t]}$ -medible. ■

Proposición 9.18. Sobre $\mathbb{D}_S$, la aplicación

$$X : (\mathbb{D}_S \times [0, \infty), \mathcal{F} \otimes \mathcal{B}_{[0, \infty)}) \rightarrow (S, \mathcal{B}_S), \quad (\alpha, s) \mapsto \alpha(s),$$

es medible.

Demostración. Para $n \in \mathbb{N}$ definimos $X^{(n)}(\alpha, s) = X(\alpha, k/2^n)$ siempre que $\frac{k-1}{2^n} < s \leq \frac{k}{2^n}$, y $X^{(n)}(\alpha, 0) = X(\alpha, 0)$. Entonces, para $A \in \mathcal{B}_S$,

$$\{(\alpha, s) : X^{(n)}(\alpha, s) \in A\} = \underbrace{\left(X_0^{-1}(A) \times \{0\} \right)}_{\in \mathcal{F}} \cup \bigcup_{k=1}^{\infty} \underbrace{X_{k/2^n}^{-1}(A)}_{\in \mathcal{F}} \times \underbrace{\left(\frac{k-1}{2^n}, \frac{k}{2^n} \right]}_{\in \mathcal{B}_{[0, \infty)}} ,$$

que está en $\mathcal{F} \otimes \mathcal{B}_{\mathbb{R}_+}$. Como α es càdlàg, $X^{(n)} \rightarrow X$ puntualmente, y el límite de medibles es medible. ■

Corolario 9.19. Si $\tau : \mathbb{D}_S \rightarrow [0, \infty]$ es un tiempo de parada, entonces

$$X_\tau : \{\tau < \infty\} \rightarrow S, \quad \alpha \mapsto X_{\tau(\alpha)}(\alpha),$$

es $\mathcal{F}_\tau$ -medible; es decir, $\{\tau < \infty, X_\tau \in A\} \in \mathcal{F}_\tau$ para todo $A \in \mathcal{B}_S$.

Demostración. La aplicación $\alpha \mapsto X_\tau(\alpha)$ es la composición

$$\{\tau \leq t\} \xrightarrow{(\text{Id}, \tau)} \mathbb{D}_S \times [0, t] \xrightarrow{X} S,$$

donde la primera es $\mathcal{F}_t/\mathcal{F}_t \otimes \mathcal{B}_{[0, t]}$ -medible (por serlo cada coordenada, usando que τ restringido a $\{\tau \leq t\}$ es $\mathcal{F}_t$ -medible) y la segunda es medible por la [proposición 9.17](#) aplicada al proceso canónico, que es càdlàg y adaptado. Luego la restricción de X_τ a $\{\tau \leq t\}$ es $\mathcal{F}_t$ -medible, o sea

$$\{X_\tau \in A\} \cap \{\tau \leq t\} \in \mathcal{F}_t \quad \text{para todo } t,$$

que es exactamente la definición de $\mathcal{F}_\tau$ -medibilidad. ■

Proposición 9.20 (Criterio práctico). Sea τ un tiempo de parada y $Z : \{\tau < \infty\} \rightarrow (E, \Sigma)$. Entonces

$$\{\tau < \infty, Z \in A\} \in \mathcal{F}_\tau \quad \forall A \in \Sigma \iff \{\tau \leq t, Z \in A\} \in \mathcal{F}_t \quad \forall A \in \Sigma, \forall t \geq 0.$$

Ejercicio 9.21. Sea (S, d) métrico y $[(\mathbb{D}_S, \mathcal{F}), (\mathcal{F}_t)]$ el espacio filtrado canónico. Fijemos un tiempo de parada τ y definamos $\theta_\tau : \{\tau < \infty\} \rightarrow \mathbb{D}_S$ por $\theta_\tau(\alpha)(t) = \alpha(\tau(\alpha) + t)$.

- (a) Para todo $s \geq 0$ y $E \in \mathcal{F}_s$, $\{\tau < \infty, \theta_\tau \in E\} \in \mathcal{F}_{\tau+s}$.
- (b) Si ρ es otro tiempo de parada, entonces $\tau + \rho \circ \theta_\tau$ (definido como $+\infty$ sobre $\{\tau = \infty\}$) es un tiempo opcional.

Solución. (a) Como $\mathcal{F}_s = \sigma(X_r : r \leq s)$, basta probarlo para $E = \{X_r \in D\}$ con $r \leq s$ y $D \in \mathcal{B}_S$. Entonces

$$\{\tau < \infty, \theta_\tau \in \{X_r \in D\}\} = \{\tau < \infty, X_{\tau+r} \in D\} \in \mathcal{F}_{\tau+r} \subset \mathcal{F}_{\tau+s},$$

donde la pertenencia es el [corolario 9.19](#) aplicado al tiempo de parada $\tau + r$, y la inclusión es el [ejercicio 5.3\(c\)](#) con $\tau + r \leq \tau + s$.

(b) Queremos ver que $\{\tau + \rho \circ \theta_\tau < s\} \in \mathcal{F}_s$. Ahora bien,

$$\{\tau + \rho \circ \theta_\tau < s\} = \bigcup_{q \in \mathbb{Q}, 0 \leq q < s} \left\{ \underbrace{\tau < q}_{\in \mathcal{F}_q \subset \mathcal{F}_s}, \underbrace{\rho \circ \theta_\tau < s - q}_{= \{\theta_\tau \in \{\rho < s - q\}\}} \right\},$$

usando que $\theta_\tau^{-1}(\{\rho < s - q\}) = (\rho \circ \theta_\tau)^{-1}([0, s - q])$. Como $\{\rho < s - q\} \in \mathcal{F}_{s-q}$, por (a) el segundo suceso está en $\mathcal{F}_{\tau+s-q}$, y sobre $\{\tau < q\}$ se tiene $\tau + s - q < s$: el suceso completo está en $\mathcal{F}_s$. ■

El proceso de Poisson

El proceso de Poisson es el proceso de saltos más simple posible: cuenta ocurrencias que suceden “al azar y sin memoria”. Lo construiremos de dos maneras independientes —como una medida aleatoria de Poisson sobre la recta y como una sucesión de tiempos de espera exponenciales— y comprobaremos que ambas producen la misma ley sobre $\mathbb{D}_{\mathbb{Z}_+}$.

10.1 La distribución de Poisson

Definición 10.1. La *distribución de Poisson* de parámetro $\lambda > 0$ es la probabilidad sobre $\mathbb{Z}_+$ dada por

$$\text{Poiss}(\lambda)(\{k\}) = \frac{e^{-\lambda}\lambda^k}{k!}, \quad k \in \mathbb{Z}_+.$$

Convenimos $\text{Poiss}(0) = \delta_0$ y $\text{Poiss}(\infty) = \delta_\infty$. Una *variable aleatoria de Poisson* es $X : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{Z}_+$ con $\mathbb{P} \circ X^{-1} = \text{Poiss}(\lambda)$.

Observación 10.2 (Transformada de Laplace). Si $X \sim \text{Poiss}(\lambda)$ entonces, para $s \in \mathbb{R}$,

$$\mathbb{E}[e^{sX}] = \sum_{k=0}^{\infty} e^{sk} \frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^s)^k}{k!} = e^{\lambda(e^s-1)}. \quad (10.1)$$

En particular $\mathbb{E}[X] = \text{Var}(X) = \lambda$.

Proposición 10.3 (Estabilidad por sumas). Sean $X_1, X_2, \dots$ independientes con $X_n \sim \text{Poiss}(\lambda_n)$. Entonces:

- (a) $X_1 + \dots + X_n \sim \text{Poiss}(\lambda_1 + \dots + \lambda_n)$;
- (b) si $\hat{\lambda} = \sum_n \lambda_n < \infty$, entonces $Z = \sum_n X_n \sim \text{Poiss}(\hat{\lambda})$;
- (c) si $\hat{\lambda} = \sum_n \lambda_n = \infty$, entonces $\sum_n X_n = \infty$ c. s..

Demostración. (a) Basta el caso $n = 2$. Por convolución y el binomio de Newton,

$$\begin{aligned}\mathbb{P}\{X_1 + X_2 = k\} &= \sum_{j=0}^k \mathbb{P}\{X_1 = j\} \mathbb{P}\{X_2 = k - j\} = \sum_{j=0}^k \frac{e^{-\lambda_1} \lambda_1^j}{j!} \frac{e^{-\lambda_2} \lambda_2^{k-j}}{(k-j)!} \\ &= \frac{e^{-(\lambda_1 + \lambda_2)}}{k!} \sum_{j=0}^k \binom{k}{j} \lambda_1^j \lambda_2^{k-j} = \frac{e^{-(\lambda_1 + \lambda_2)}}{k!} (\lambda_1 + \lambda_2)^k.\end{aligned}$$

(b) Sea $Z_n = X_1 + \dots + X_n$. Por convergencia monótona, $\mathbb{E}[Z] = \sum_n \mathbb{E}[X_n] = \hat{\lambda} < \infty$, de modo que $Z < \infty$ c.s.. Como $Z_n \uparrow Z$ y las variables son enteras, $\mathbf{1}_{\{Z_n = k\}} \rightarrow \mathbf{1}_{\{Z = k\}}$ c.s. y, por convergencia dominada,

$$\mathbb{P}\{Z = k\} = \lim_n \mathbb{P}\{Z_n = k\} = \lim_n \frac{e^{-\sum_{j \leq n} \lambda_j} (\sum_{j \leq n} \lambda_j)^k}{k!} = \frac{e^{-\hat{\lambda}} \hat{\lambda}^k}{k!}.$$

(c) Definamos

$$Y = \begin{cases} e^{-\sum_n X_n}, & \text{si } \sum_n X_n < \infty, \\ 0, & \text{en caso contrario.} \end{cases}$$

Sobre $\{\sum_n X_n < \infty\}$, por convergencia monótona y la independencia,

$$\begin{aligned}\mathbb{E}[Y] &= \int_{\{\sum_n X_n < \infty\}} e^{-\sum_n X_n} = \lim_k \int_{\{\sum_n X_n < \infty\}} \prod_{j=1}^k e^{-X_j} \leq \lim_k \prod_{j=1}^k \mathbb{E}[e^{-X_j}] \\ &\stackrel{(10.1)}{=} \lim_k \exp \left\{ (e^{-1} - 1) \sum_{j=1}^k \lambda_j \right\} = 0,\end{aligned}$$

pues $e^{-1} - 1 < 0$ y $\sum_j \lambda_j = \infty$. Luego $Y = 0$ c.s., lo que significa $\mathbb{P}\{\sum_n X_n < \infty\} = 0$. ■

10.2 Medidas aleatorias de Poisson

Fijemos un espacio métrico S y el espacio medible $(S, \mathcal{B}_S)$. Denotemos por

$$\mathcal{M} = \mathcal{M}_{S, \mathcal{B}_S} = \{\mu \text{ medida sobre } (S, \mathcal{B}_S) \text{ con valores en } \overline{\mathbb{Z}}_+ \text{ y } \sigma\text{-finita sobre compactos}\}$$

el espacio de las *medidas de conteo*. Todo $\mu \in \mathcal{M}$ se escribe como $\mu = \sum_i \delta_{x_i}$ para una familia numerable de puntos (x_i) (con repeticiones si hay átomos múltiples). Definimos la familia de aplicaciones

$$\#_A : \mathcal{M} \rightarrow \overline{\mathbb{Z}}_+, \quad \#_A(\mu) = \mu(A), \quad A \in \mathcal{B}_S,$$

y dotamos a $\mathcal{M}$ de la σ -álgebra

$$\sigma(\#) = \sigma(\#_A : A \in \mathcal{B}_S) = \sigma\left(\bigcup_{A_1, \dots, A_k} \{\#_{A_1} = i_1, \dots, \#_{A_k} = i_k\}\right).$$

Definición 10.4 (Medida aleatoria de Poisson). Sea ν una medida σ -finita sobre $(S, \mathcal{B}_S)$, llamada *intensidad*. Una probabilidad $\mathbb{P}$ sobre $(\mathcal{M}, \sigma(\#))$ es la *distribución de una medida aleatoria de Poisson* (*Poisson random measure*, PRM) de intensidad ν si

- (I) $\#_A \sim \text{Poiss}(\nu(A))$ bajo $\mathbb{P}$, para todo $A \in \mathcal{B}_S$;
- (II) si $A_1, \dots, A_k \in \mathcal{B}_S$ son disjuntos dos a dos, entonces $\#_{A_1}, \dots, \#_{A_k}$ son $\mathbb{P}$ -independientes.

Una PRM es una aplicación medible $M : (\Omega, \mathcal{F}, \mathbb{P}') \rightarrow (\mathcal{M}, \sigma(\#))$ cuya ley es tal distribución.

Teorema 10.5 (Existencia). Para toda medida σ -finita ν sobre $(S, \mathcal{B}_S)$ existe una PRM de intensidad ν .

Demostración. Por σ -finitud podemos escribir $S = \bigsqcup_{n \geq 1} U_n$ con $\nu(U_n) < \infty$. Sobre un espacio de probabilidad auxiliar $(\Omega, \mathcal{F}, \mathbb{P}')$ construimos, *todas independientes*:

- variables $N_n \sim \text{Poiss}(\nu(U_n))$, $n \geq 1$;
- variables $\xi_j^{(n)} : \Omega \rightarrow U_n$, $j \geq 1$, con ley $\nu(\cdot \cap U_n)/\nu(U_n)$, es decir

$$\mathbb{P}'\{\xi_j^{(n)} \in A\} = \frac{1}{\nu(U_n)} \int_A \mathbf{1}_{U_n} d\nu.$$

Definimos entonces

$$\bar{\mu} = \sum_{n=1}^{\infty} \sum_{j=1}^{N_n} \delta_{\xi_j^{(n)}} : \Omega \longrightarrow \mathcal{M}. \quad (10.2)$$

La aplicación es medible porque para cada A , $\#_A \circ \bar{\mu} = \sum_n \sum_{j \leq N_n} \mathbf{1}_A(\xi_j^{(n)})$ es medible. Veamos que tiene la ley correcta.

Primero, si $A \subset U_n$, condicionando a $N_n = m$ el número de puntos que caen en A es $\text{Bin}(m, \nu(A)/\nu(U_n))$, y por el adelgazamiento de una Poisson (proposición 10.6 más abajo, cuya demostración es independiente) resulta $\bar{\mu}(A) \sim \text{Poiss}(\nu(A))$. Para A general, $\bar{\mu}(A) = \sum_n \bar{\mu}(A \cap U_n)$ es una suma de variables de Poisson independientes de parámetros $\nu(A \cap U_n)$, y la proposición 10.3 da $\bar{\mu}(A) \sim \text{Poiss}(\nu(A))$.

Segundo, si $A_1, \dots, A_k$ son disjuntos, el mismo adelgazamiento (versión multinomial) muestra que los conteos son independientes dentro de cada U_n , y la independencia entre distintos n es por construcción. ■

Proposición 10.6 (Adelgazamiento y superposición). (a) (Adelgazamiento.)

Sea $N \sim \text{Poiss}(\lambda)$ y, dado $N = m$, sea $(Y_1, \dots, Y_k)$ multinomial de parámetros m y $(p_1, \dots, p_k)$ con $\sum_i p_i = 1$. Entonces $Y_1, \dots, Y_k$ son independientes con $Y_i \sim \text{Poiss}(\lambda p_i)$.

(b) (Superposición.) Si M_1, M_2 son PRM independientes sobre $(S, \mathcal{B}_S)$ con intensidades ν_1, ν_2 , entonces $M_1 + M_2$ es una PRM de intensidad $\nu_1 + \nu_2$.

Demostración. (a) Para $m_1, \dots, m_k \in \mathbb{Z}_+$ con $m = \sum_i m_i$,

$$\mathbb{P}\{Y_1 = m_1, \dots, Y_k = m_k\} = \frac{e^{-\lambda} \lambda^m}{m!} \cdot \frac{m!}{m_1! \dots m_k!} p_1^{m_1} \dots p_k^{m_k} = \prod_{i=1}^k \frac{e^{-\lambda p_i} (\lambda p_i)^{m_i}}{m_i!},$$

usando $\sum_i p_i = 1$ para repartir $e^{-\lambda} = \prod_i e^{-\lambda p_i}$.

(b) Para $A \in \mathcal{B}_S$, $M_1(A) \sim \text{Poiss}(\nu_1(A))$ y $M_2(A) \sim \text{Poiss}(\nu_2(A))$ son independientes, luego por la proposición 10.3(a)

$$(M_1 + M_2)(A) = M_1(A) + M_2(A) \sim \text{Poiss}((\nu_1 + \nu_2)(A)).$$

Falta la independencia de $(M_1 + M_2)(A_1), \dots, (M_1 + M_2)(A_k)$ para A_j disjuntos dos a dos; esto es inmediato porque el vector completo

$$(M_1(A_1), \dots, M_1(A_k), M_2(A_1), \dots, M_2(A_k))$$

tiene $2k$ coordenadas independientes, y sumas de subfamilias disjuntas de variables independientes son independientes. ■

10.3 Primera construcción del proceso de Poisson

Definición 10.7 (Proceso de Poisson). Sea $\lambda > 0$. Decimos que $[(\Omega, \mathcal{F}, \mathbb{P}), (N_t)_{t \geq 0}]$, con $N_t : \Omega \rightarrow \mathbb{Z}_+$, es un *proceso de Poisson de intensidad λ* si

- (I) $N_0 = 0$ y $t \mapsto N_t(\omega)$ es càdlàg para todo ω ;
- (II) (incrementos independientes) para todos $0 \leq t_0 < t_1 < \dots < t_k$, las variables $N_{t_1} - N_{t_0}, \dots, N_{t_k} - N_{t_{k-1}}$ son independientes;
- (III) (incrementos estacionarios de Poisson) para $0 \leq s < t$,

$$N_t - N_s \sim \text{Poiss}(\lambda(t-s)), \quad \text{es decir} \quad \mathbb{P}\{N_t - N_s = k\} = \frac{e^{-\lambda(t-s)} (\lambda(t-s))^k}{k!}.$$

Ejemplo 10.8 (Construcción vía PRM). Tomemos $S = \mathbb{R}$ con la intensidad $\nu = \lambda \cdot m$, donde m es la medida de Lebesgue, y sea $M : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathcal{M}_{\mathbb{R}, \mathcal{B}_{\mathbb{R}}}$ una PRM de intensidad ν (existe por el [teorema 10.5](#)). Definamos

$$N_0 = 0, \quad N_t = \#_{(0,t]} \circ M, \quad t > 0.$$

Entonces $(N_t)_{t \geq 0}$ es un proceso de Poisson de intensidad λ :

- $t \mapsto N_t$ es càdlàg (cuenta puntos en $(0, t]$, un intervalo cerrado por la derecha);
- $N_t - N_s = \#_{(s,t]} \circ M \sim \text{Poiss}(\lambda(t-s))$;
- los intervalos $(t_{j-1}, t_j]$ son disjuntos, de modo que los incrementos son independientes.

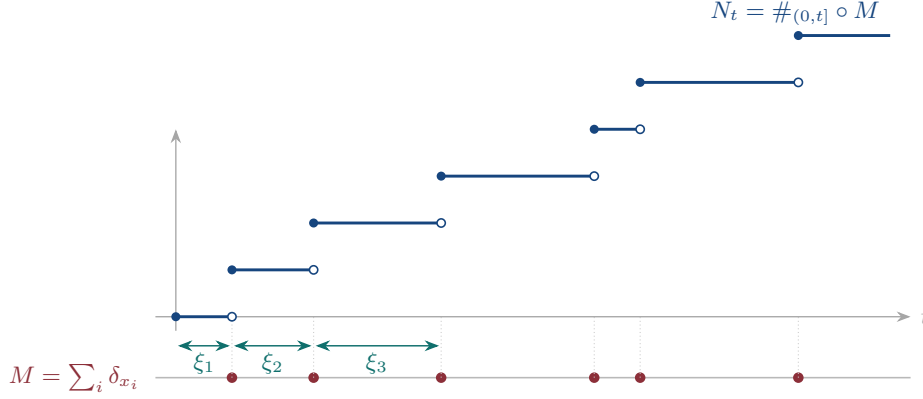


Figura 10.1: Las dos caras del proceso de Poisson. Abajo: una medida aleatoria de Poisson sobre $\mathbb{R}$, una nube de puntos. Arriba: el proceso de conteo N_t , càdlàg, con saltos de tamaño $+1$. Los tiempos entre saltos ξ_n son independientes $\text{Exp}(\lambda)$ (teorema 10.20).

Teorema 10.9 (Unicidad de la ley). *Sea $\lambda > 0$. Existe una única probabilidad $\mathbb{P}_{\text{Pois}}^\lambda$ sobre $(\mathbb{D}_{\mathbb{Z}_+}, \sigma(X))$ tal que, bajo ella, el proceso canónico $(X_t)_{t \geq 0}$ cumple (i)–(iii) de la definición 10.7.*

Demostración. Existencia. Sea M la PRM del ejemplo 10.8 y definamos la aplicación

$$\begin{aligned} \mathcal{N} : \mathcal{M} &\longrightarrow \mathbb{D}_{\mathbb{Z}_+}, \\ \mu &\longmapsto (t \mapsto \#_{(0,t]}(\mu) = \mu((0, t])), \end{aligned}$$

que es medible respecto de $\sigma(\#)$ y $\sigma(X)$, pues $\mathcal{N}^{-1}\{X_t \in A\} = \{\#_{(0,t]} \in A\} \in \sigma(\#)$ y los cilindros generan $\sigma(X)$. Tomamos $\mathbb{P}_{\text{Pois}}^\lambda = \mathbb{P}' \circ \mathcal{N}^{-1}$, donde $\mathbb{P}'$ es la ley de la PRM. Verifiquemos las tres propiedades.

- $\mathbb{P}_{\text{Pois}}^\lambda\{X_0 = 0\} = \mathbb{P}'\{\mu : \mu((0, 0]) = 0\} = 1$, pues $(0, 0] = \emptyset$.
- Para $0 \leq s < t$ y $A \subset \mathbb{Z}_+$,

$$\begin{aligned} \mathbb{P}_{\text{Pois}}^\lambda\{X_t - X_s \in A\} &= \mathbb{P}'\{\mu : \mathcal{N}(\mu)(t) - \mathcal{N}(\mu)(s) \in A\} = \mathbb{P}'\{\mu : \mu((s, t]) \in A\} \\ &= \text{Poiss}(\nu((s, t]))(A) = \text{Poiss}(\lambda(t - s))(A). \end{aligned}$$
- Para $t_0 < t_1 < \dots < t_k$, los incrementos corresponden a $\#_{(t_{i-1}, t_i]}(\mu)$, conteos sobre *conjuntos disjuntos*, luego independientes por la definición 10.4(ii).

Unicidad. Sean $\mathbb{P}$ y $\tilde{\mathbb{P}}$ dos probabilidades sobre $(\mathbb{D}_{\mathbb{Z}_+}, \sigma(X))$ con las propiedades (i)–(iii). Por la [observación 9.3](#) basta ver que coinciden en las leyes finito-dimensionales. Ahora, para $t_1 < \dots < t_k$ y $a_1, \dots, a_k \in \mathbb{Z}_+$, escribiendo el vector en términos de incrementos,

$$\begin{aligned} \mathbb{P}\{X_{t_1} = a_1, \dots, X_{t_k} = a_k\} &= \mathbb{P}\{X_{t_1} - X_0 = a_1, X_{t_2} - X_{t_1} = a_2 - a_1, \dots\} \\ &= \prod_{i=1}^k \text{Pois}(\lambda(t_i - t_{i-1}))(\{a_i - a_{i-1}\}), \end{aligned}$$

con $t_0 = 0$, $a_0 = 0$, y lo mismo para $\tilde{\mathbb{P}}$: las dos expresiones coinciden. $\blacksquare$

Notación. De ahora en adelante trabajamos sobre el espacio canónico

$$\left[(\mathbb{D}_{\mathbb{Z}_+}, \mathcal{F}), (\mathcal{F}_t)_{t \geq 0}, (X_t)_{t \geq 0}, \mathbb{P}_{\text{Pois}}^\lambda \right],$$

que es la realización canónica del proceso de Poisson.

Ejercicio 10.10. Bajo $\mathbb{P}_{\text{Pois}}^\lambda$, para $0 \leq s < t$ el incremento $X_t - X_s$ es independiente de $\mathcal{F}_s = \sigma(X_u : u \leq s)$.

Solución. La clase

$$\mathcal{L} = \{A \in \mathcal{F}_s : \mathbb{P}\{A \cap \{X_t - X_s = k\}\} = \mathbb{P}(A) \mathbb{P}\{X_t - X_s = k\} \quad \forall k\}$$

es un λ -sistema. Por el teorema π - λ ([teorema A.2](#)) basta verificar la propiedad sobre un π -sistema que genere $\mathcal{F}_s$, por ejemplo los cilindros $A = \{X_{t_1} = a_1, \dots, X_{t_n} = a_n\}$ con $t_i \leq s$. Escribiendo A en términos de incrementos sobre intervalos disjuntos de $(t_n, t]$ y usando la independencia de los incrementos, la factorización es inmediata. $\blacksquare$

Ejercicio 10.11 (Martingala compensada). Pruebe que

$$\tilde{N}_t := N_t - \lambda t, \quad t \geq 0,$$

es una martingala respecto de $\mathcal{F}_t = \sigma(N_u : u \leq t)$.

Solución. Para $s < t$, usando el [ejercicio 10.10](#),

$$\begin{aligned} \mathbb{E}[\tilde{N}_t - \tilde{N}_s \mid \mathcal{F}_s] &= \mathbb{E}[N_t - N_s \mid \mathcal{F}_s] - \lambda(t - s) = \mathbb{E}[N_t - N_s] - \lambda(t - s) \\ &= \lambda(t - s) - \lambda(t - s) = 0. \end{aligned} \quad \blacksquare$$

Ejercicio 10.12. Pruebe que $\tilde{N}_t^2 - \lambda t$ también es martingala, es decir que $\langle \tilde{N} \rangle_t = \lambda t$.

10.4 Propiedades de las trayectorias

Las trayectorias de un proceso de Poisson son escaleras que suben de uno en uno. El resultado siguiente lo precisa; es la clave para reconocer un proceso de Poisson “desde afuera”.

Ejercicio 10.13. Consideremos $(\mathbb{D}_{\mathbb{Z}_+}, \mathcal{F}, \mathbb{P}_{\text{Pois}}^\lambda)$ y, para $t > 0$,

$$\Delta_t : \mathbb{D}_{\mathbb{Z}_+} \rightarrow \mathbb{Z}_+, \quad \alpha \mapsto \alpha(t) - \alpha(t^-).$$

Pruebe que

- (a) Δ_t es medible;
- (b) $\mathbb{P}_{\text{Pois}}^\lambda\{\Delta_t \neq 0\} = 0$ para cada $t > 0$ fijo;
- (c) para todo $T > 0$, $\alpha \mapsto \max_{0 \leq t \leq T} |\Delta_t \alpha|$ es medible;
- (d) $\mathbb{P}_{\text{Pois}}^\lambda\{\max_{0 \leq t \leq T} |\Delta_t \alpha| \leq 1\} = 1$.

Solución. (a) Por ser α càdlàg, $\alpha(t^-) = \lim_n \alpha(t - 1/n)$ y como los valores son enteros el límite se alcanza para n grande:

$$\{\Delta_t = k\} = \bigcup_{m \geq 1} \bigcap_{n \geq m} \underbrace{\{\alpha(t) - \alpha(t - \frac{1}{n}) = k\}}_{= \{X_t - X_{t-1/n} = k\} \in \sigma(X)}.$$

(b) Basta ver que $\mathbb{P}\{\Delta_t = k\} = 0$ para $k \geq 1$. Por (a),

$$\mathbb{P}\{\Delta_t = k\} \leq \mathbb{P}\left\{\bigcap_{n \geq m} \{X_t - X_{t-1/n} = k\}\right\} \leq \lim_n \mathbb{P}\{X_t - X_{t-1/n} = k\},$$

y

$$\mathbb{P}\{X_t - X_{t-1/n} = k\} = e^{-\lambda/n} \frac{(\lambda/n)^k}{k!} \xrightarrow{n \rightarrow \infty} 0 \quad (k \geq 1).$$

(c) Como los saltos de tamaño $> \epsilon$ son finitos ([proposición 9.2](#)), el máximo de los saltos se puede leer sobre los racionales:

$$\left\{\max_{0 \leq t \leq T} |\Delta_t| \leq k\right\} = \bigcup_{n \geq 1} \bigcap_{\substack{q_1, q_2 \in \mathbb{Q} \cap [0, T] \\ |q_1 - q_2| < 1/n}} \{|\alpha(q_2) - \alpha(q_1)| \leq k\} \in \sigma(X).$$

(d) Equivale a probar $\mathbb{P}\{\max_{0 \leq t \leq T} |\Delta_t| \geq 2\} = 0$. Si algún salto es ≥ 2 , entonces para todo n alguno de los n intervalos $(\frac{(k-1)T}{n}, \frac{kT}{n}]$ contiene un incremento ≥ 2 :

$$\left\{\max_{0 \leq t \leq T} |\Delta_t| \geq 2\right\} \subset \bigcup_{k \leq n} \left\{X_{kT/n} - X_{(k-1)T/n} \geq 2\right\}.$$

Por subaditividad, con $x = \lambda T/n$,

$$\mathbb{P}\left\{\max_{0 \leq t \leq T} |\Delta_t| \geq 2\right\} \leq n(1 - e^{-x} - xe^{-x}) = n \cdot O(x^2) = O\left(\frac{(\lambda T)^2}{n}\right) \xrightarrow{n \rightarrow \infty} 0. \quad \blacksquare$$

Observación 10.14. De (b) y (d) se sigue que, para cualquier proceso $N : (\Omega, \mathcal{F}, \mathbb{P}') \rightarrow \mathbb{D}_{\mathbb{Z}_+}$ con ley $\mathbb{P}'_{\text{Pois}}$,

$$\begin{aligned} \mathbb{P}'\{\Delta_t N \neq 0\} &= 0 \quad \text{para todo } t > 0, \\ \mathbb{P}'\left\{\max_{0 \leq t \leq T} |\Delta_t N| \leq 1\right\} &= 1 \quad \text{para todo } T > 0. \end{aligned}$$

En palabras: los saltos ocurren en instantes aleatorios —ningún instante *determinista* es un salto— y son todos de tamaño exactamente 1.

Definición 10.15 (Procesos de Poisson independientes). Decimos que $(N_t^1), \dots, (N_t^k)$ son k procesos de Poisson independientes si las σ -álgebras $\sigma(N^j) = \sigma(N_t^j : t \geq 0)$, $j = 1, \dots, k$, son independientes.

Observación 10.16 (Construcción conjunta). Se pueden construir procesos de Poisson independientes de intensidades $\lambda_1, \dots, \lambda_k$ con una sola PRM. Consideremos

$$\left(\{1, \dots, k\} \times \mathbb{R}, \quad 2^{\{1, \dots, k\}} \otimes \mathcal{B}_{\mathbb{R}}\right)$$

y la intensidad

$$\Lambda(A) = \sum_{j=1}^k \lambda_j m(A \cap (\{j\} \times \mathbb{R})), \quad A \in \mathcal{B}_S.$$

Como Λ es σ -finita, existe una PRM $M : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathcal{M}_{S, \mathcal{B}_S}$ de intensidad Λ , y poniendo

$$N_t^j = \#_{\{j\} \times (0, t]} \circ M$$

se obtienen k procesos de Poisson de intensidades λ_j , que son independientes porque los conjuntos $\{j\} \times \mathbb{R}$ son disjuntos. Por la [proposición 10.3](#),

$$N_t = N_t^1 + \dots + N_t^k \sim \text{Pois}((\lambda_1 + \dots + \lambda_k)t),$$

y de hecho (N_t) es un proceso de Poisson de intensidad $\lambda_1 + \dots + \lambda_k$.

Ejercicio 10.17. Sean $(N_t^1), \dots, (N_t^k)$ procesos de Poisson independientes. Pruebe que para $i \neq j$,

$$\mathbb{P}\{\Delta_t N^i \cdot \Delta_t N^j = 0 \text{ para todo } t \geq 0\} = 1,$$

es decir, dos procesos independientes nunca saltan simultáneamente.

Solución. Tomando complementos, hay que ver que $\mathbb{P}\{\exists t \leq T : \Delta_t N^i \cdot \Delta_t N^j > 0\} = 0$ para todo T . Si hay un salto simultáneo en $[0, T]$, para todo n alguno de los n intervalos de longitud T/n contiene un incremento positivo de ambos procesos:

$$\{\exists t \leq T : \Delta_t N^i \Delta_t N^j > 0\} \subset \bigcap_{n \in \mathbb{N}} \bigcup_{k \leq n} \{N_{kT/n}^i - N_{(k-1)T/n}^i > 0\} \cap \{N_{kT/n}^j - N_{(k-1)T/n}^j > 0\}.$$

Por independencia y subaditividad, la probabilidad del conjunto de la derecha está acotada por

$$n(1 - e^{-\lambda_i T/n})(1 - e^{-\lambda_j T/n}) = n \cdot O(1/n^2) \xrightarrow{n \rightarrow \infty} 0. \quad \blacksquare$$

10.5 Segunda construcción: tiempos de espera exponenciales

La segunda construcción es la que usa un simulador: esperar un tiempo $\text{Exp}(\lambda)$, saltar, repetir.

Sea $\xi_1, \xi_2, \dots$ una sucesión i.i.d. $\text{Exp}(\lambda)$ sobre un espacio $(\Omega, \mathcal{F}, \mathbb{P}')$, y consideremos la aplicación

$$\Phi : [0, \infty]^\mathbb{N} \longrightarrow \mathbb{D}_{\mathbb{Z}_+}$$

definida así: si $\sum_k u_k < \infty$ ponemos $\Phi(u_1, u_2, \dots) \equiv 0$ (convención); si $\sum_k u_k = \infty$, definimos $\alpha = \Phi(u_1, u_2, \dots)$ por

$$t_0 = 0, \quad t_n = \sum_{k=1}^n u_k \quad (n \geq 1), \quad \alpha(t) = n \quad \text{si } t_n \leq t < t_{n+1}. \quad (10.3)$$

Observación. Φ es medible, pues $\{X_t \circ \Phi = n\} = \{t_n \leq t < t_{n+1}\}$ es un suceso que depende medibly de $(u_1, \dots, u_{n+1})$, y los cilindros generan $\sigma(X)$.

Teorema 10.18 (Segunda construcción). *En el diagrama*

$$(\Omega, \mathcal{F}, \mathbb{P}') \xrightarrow{(\xi_1, \xi_2, \dots)} [0, \infty]^\mathbb{N} \xrightarrow{\Phi} \mathbb{D}_{\mathbb{Z}_+},$$

la medida imagen

$$\mathbb{P} = [\Phi \circ (\xi_1, \xi_2, \dots)]_* \mathbb{P}' = \mathbb{P}' \circ [\Phi \circ (\xi_1, \xi_2, \dots)]^{-1}$$

es exactamente $\mathbb{P}_{\text{Pois}}^\lambda$.

Recíprocamente, sobre el espacio canónico podemos *leer* los tiempos de espera a partir de la trayectoria.

Definición 10.19 (Tiempos de salto y de espera). Sobre $\mathbb{D}_{\mathbb{Z}_+}$ definimos

$$J_1 = \inf\{t > 0 : X_t \neq X_0\},$$

$$J_n = \begin{cases} \inf\{t > J_{n-1} : X_t \neq X_{J_{n-1}}\}, & \text{si } J_{n-1} < \infty, \\ \infty, & \text{si } J_{n-1} = \infty, \end{cases}$$

y los *tiempos entre saltos* $\xi_n = J_n - J_{n-1}$ (con $J_0 = 0$).

Observación. En el diagrama

$$\mathbb{D}_{\mathbb{Z}_+} \xrightarrow{(\xi_1, \xi_2, \dots)} [0, \infty]^{\mathbb{N}} \xrightarrow{\Phi} \mathbb{D}_{\mathbb{Z}_+},$$

la composición $\Phi \circ (\xi_1, \xi_2, \dots)$ es la identidad $\mathbb{P}_{\text{Pois}}^\lambda$ -c. s.; esto es consecuencia de la [observación 10.14](#): las trayectorias son escaleras de paso 1, y por lo tanto quedan determinadas por sus tiempos de salto.

Teorema 10.20 (Tiempos de espera exponenciales). *Bajo $\mathbb{P}_{\text{Pois}}^\lambda$, los tiempos entre saltos $\xi_1, \xi_2, \dots$ son independientes e idénticamente distribuidos $\text{Exp}(\lambda)$. Equivalentemente,*

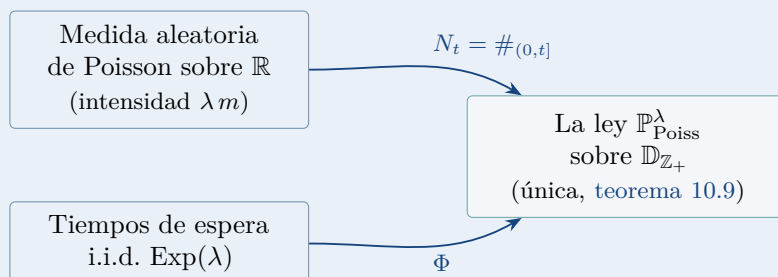
$$\mathbb{P}_{\text{Pois}}^\lambda \circ (\xi_1, \xi_2, \dots)^{-1} = \text{Exp}(\lambda)^{\otimes \mathbb{N}}.$$

La demostración de $\xi_1 \sim \text{Exp}(\lambda)$ es inmediata:

$$\mathbb{P}\{\xi_1 > a\} = \mathbb{P}\{X_a - X_0 = 0\} = e^{-\lambda a}, \quad \text{luego} \quad \mathbb{P}\{\xi_1 \leq a\} = 1 - e^{-\lambda a}. \quad (10.4)$$

Para los demás hace falta la propiedad de Markov fuerte, que es el tema del [capítulo 11](#); completaremos allí la demostración ([sección 11.4](#)). Aceptando el [teorema 10.20](#), el [teorema 10.18](#) se sigue: las dos construcciones inducen la misma ley sobre $\mathbb{D}_{\mathbb{Z}_+}$ porque ambas quedan determinadas por la ley de $(\xi_1, \xi_2, \dots)$, y Φ es una biyección c. s. entre ambas descripciones.

Dos construcciones, una sola ley



La PRM se generaliza a espacios arbitrarios (y da los procesos de Lévy con saltos); los tiempos exponenciales se generalizan a cadenas de Markov a tiempo continuo. Que ambas den lo mismo es lo que hace tan flexible al proceso de Poisson.

CAPÍTULO 11

Las propiedades de Markov

“El futuro depende del pasado solo a través del presente”. Formalizar esa frase exige tres ingredientes: una familia de leyes $(\mathbb{P}_x)$ indexada por el punto de partida, los operadores de traslación θ_t , y la observación de que “el futuro visto desde t ” es $F \circ \theta_t$. La versión *fuerte* reemplaza el instante determinista t por un tiempo de parada, y es la que tiene verdadero poder.

Desarrollamos la teoría para el proceso de Poisson; en el [capítulo 13](#) los mismos enunciados —con las mismas demostraciones— valdrán para el movimiento Browniano, cambiando $\mathbb{D}_{\mathbb{Z}_+}$ por $\mathcal{C}_{\mathbb{R}}$.

11.1 La familia $(\mathbb{P}_x)_{x \in \mathbb{Z}_+}$

Sobre el espacio canónico $[(\mathbb{D}_{\mathbb{Z}_+}, \mathcal{F}), (\mathcal{F}_t)_{t \geq 0}, (X_t)_{t \geq 0}]$ tenemos la ley $\mathbb{P}_{\text{Pois}}^\lambda$ del [teorema 10.9](#). Para “arrancar en x ” definimos, para cada $x \in \mathbb{Z}_+$, la traslación

$$T_x : \mathbb{D}_{\mathbb{Z}_+} \rightarrow \mathbb{D}_{\mathbb{Z}_+}, \quad (T_x \alpha)(t) = x + \alpha(t),$$

y ponemos

$$\mathbb{P}_0 = \mathbb{P}_{\text{Pois}}^\lambda, \quad \mathbb{P}_x = \mathbb{P}_0 \circ T_x^{-1}. \quad (11.1)$$

Bajo $\mathbb{P}_x$ el proceso canónico arranca en $X_0 = x$ y tiene los mismos incrementos. Escribimos $\mathbb{E}_x$ para la esperanza correspondiente. Queda así el objeto

$$\left[(\mathbb{D}_{\mathbb{Z}_+}, \mathcal{F}), (\mathcal{F}_t)_{t \geq 0}, (X_t)_{t \geq 0}, \{\mathbb{P}_x\}_{x \in \mathbb{Z}_+} \right].$$

Observación 11.1 (Continuidad de Feller). Para $F : \mathbb{D}_{\mathbb{Z}_+} \rightarrow \mathbb{R}$ medible acotada (o $F \geq 0$), la aplicación

$$\mathbb{Z}_+ \rightarrow \mathbb{R}, \quad x \mapsto \mathbb{E}_x[F]$$

es medible; y como $\mathbb{Z}_+$ lleva la topología discreta, es automáticamente continua. Esta “continuidad de Feller” se usará al pasar al límite en $\mathbb{E}_{X_{\tau_n}}[F] \rightarrow \mathbb{E}_{X_\tau}[F]$.

Ejercicio 11.2. Bajo cada $\mathbb{P}_x$ y para todos $0 \leq s < t < \infty$,

$$(X_t - X_s) \perp \mathcal{F}_s.$$

Solución. La clase $\{A \in \mathcal{F}_s : \mathbb{P}_x\{A \cap \{X_t - X_s = k\}\} = \mathbb{P}_x(A) \mathbb{P}_x\{X_t - X_s = k\} \forall k\}$ es un λ -sistema; por el teorema π - λ basta verificarlo en el π -sistema de los cilindros $A = \{X_{t_1} = a_1, \dots, X_{t_n} = a_n\}$, $t_i \leq s$. Escribiendo todo en términos de incrementos y usando que $\mathbb{P}_x$ es la imagen de $\mathbb{P}_0$ por T_x (que no altera incrementos),

$$\begin{aligned} \mathbb{P}_x\{A, X_t - X_s = k\} &= \mathbb{P}_0\{X_{t_1} = a_1 - x, X_{t_2} - X_{t_1} = a_2 - a_1, \dots, X_t - X_s = k\} \\ &= \mathbb{P}_0\{X_{t_1} = a_1 - x, \dots\} \cdot \mathbb{P}_0\{X_t - X_s = k\} \\ &= \mathbb{P}_x(A) \cdot \mathbb{P}_x\{X_t - X_s = k\}, \end{aligned}$$

usando la independencia de los incrementos sobre intervalos disjuntos. ■

Proposición 11.3. Bajo cada $\mathbb{P}_x$ y para $0 \leq s < t$, el incremento $X_t - X_s$ es independiente de $\mathcal{F}_{s+}$.

Demostración. Sea $A \in \mathcal{F}_{s+}$ y $k \in \mathbb{Z}_+$. Para $0 < h < t - s$ se tiene $\mathcal{F}_{s+} \subset \mathcal{F}_{s+h}$, de modo que, por el [ejercicio 11.2](#) aplicado en el instante $s + h$,

$$\mathbb{P}_x(\{X_t - X_{s+h} = k\} \cap A) = \mathbb{P}_x\{X_t - X_{s+h} = k\} \cdot \mathbb{P}_x(A).$$

Ahora hacemos $h \downarrow 0$. Por continuidad por la derecha de las trayectorias, $X_{s+h} \rightarrow X_s$ puntualmente, y como los valores son enteros,

$$\mathbf{1}_A \mathbf{1}_{\{k\}}(X_t - X_{s+h}) \xrightarrow{h \downarrow 0} \mathbf{1}_A \mathbf{1}_{\{k\}}(X_t - X_s).$$

Por convergencia dominada (las indicadoras están acotadas por 1),

$$\mathbb{E}_x[\mathbf{1}_A \mathbf{1}_{\{k\}}(X_t - X_{s+h})] \longrightarrow \mathbb{E}_x[\mathbf{1}_A \mathbf{1}_{\{k\}}(X_t - X_s)],$$

y análogamente $\mathbb{P}_x\{X_t - X_{s+h} = k\} \rightarrow \mathbb{P}_x\{X_t - X_s = k\}$. Pasando al límite en la igualdad se concluye. ■

11.2 La propiedad de Markov débil

Ejercicio 11.4. Para todo $x_0 \in \mathbb{Z}_+$, todo $k \in \mathbb{Z}_+$ y todos $0 \leq s < t$,

$$\mathbb{P}\{X_t = k \mid \mathcal{F}_{s+}\} = \varphi(X_s) \quad \mathbb{P}_{x_0}\text{-c. s.},$$

donde $\varphi(x) = \mathbb{P}_x\{X_{t-s} = k\}$.

Solución. Descomponiendo según el valor de X_s y usando la [proposición 11.3](#),

$$\begin{aligned}
 \mathbb{E}[\mathbf{1}_{\{X_t=k\}} \mid \mathcal{F}_{s+}] &= \sum_{i=0}^{\infty} \mathbb{E}[\mathbf{1}_{\{X_s=i\}} \mathbf{1}_{\{X_t-X_s=k-i\}} \mid \mathcal{F}_{s+}] \\
 &= \sum_{i=0}^{\infty} \mathbf{1}_{\{X_s=i\}} \mathbb{E}[\mathbf{1}_{\{X_t-X_s=k-i\}} \mid \mathcal{F}_{s+}] && (\mathbf{1}_{\{X_s=i\}} \text{ es } \mathcal{F}_{s+}\text{-medible}) \\
 &= \sum_{i=0}^{\infty} \mathbf{1}_{\{X_s=i\}} \mathbb{P}_{x_0}\{X_t - X_s = k - i\} && (\text{independencia}) \\
 &= \sum_{i=0}^{\infty} \mathbf{1}_{\{X_s=i\}} \mathbb{P}_i\{X_{t-s} = k\} = \mathbb{P}_{X_s}\{X_{t-s} = k\},
 \end{aligned}$$

donde en el penúltimo paso se usó que, bajo $\mathbb{P}_i$, la variable $X_{t-s} - X_0 = X_{t-s} - i$ tiene la ley de $X_t - X_s$ bajo $\mathbb{P}_{x_0}$. $\blacksquare$

Teorema 11.5 (Propiedad de Markov débil). *Sea $F : \mathbb{D}_{\mathbb{Z}_+} \rightarrow \mathbb{R}$ medible acotada (o $F \geq 0$). Para todo $s > 0$ y todo $x_0 \in \mathbb{Z}_+$,*

$$\mathbb{E}_{x_0}[F \circ \theta_s \mid \mathcal{F}_{s+}] = \varphi(X_s) \quad \mathbb{P}_{x_0}\text{-c. s.}, \quad \text{donde } \varphi(x) = \mathbb{E}_x[F].$$

Demostración. Paso 1: el caso $\mathcal{F}_s$ y F indicadora de un cilindro. Sea $F = \mathbf{1}_A$ con

$$A = \{X_{t_1} = a_1, \dots, X_{t_n} = a_n\}.$$

La familia de conjuntos A para los que la identidad vale es un λ -sistema y los cilindros son un π -sistema generador, de modo que por el teorema π - λ basta con ellos; luego se pasa a F simple por linealidad y a F general por convergencia monótona.

Queremos ver que, para todo B en un π -sistema generador de $\mathcal{F}_s$,

$$\int_B \mathbb{E}_x[F \circ \theta_s \mid \mathcal{F}_s] d\mathbb{P}_x = \int_B \mathbb{E}_{X_s}[F] d\mathbb{P}_x.$$

Tomamos $B = \{X_{s_1} = b_1, \dots, X_{s_m} = b_m, X_s = b\}$ con $s_j < s$. Por un lado,

$$\begin{aligned}
 \int_B \mathbb{E}_x[F \circ \theta_s \mid \mathcal{F}_s] d\mathbb{P}_x &= \int \mathbf{1}_B \mathbf{1}_{\{\theta_s \in A\}} d\mathbb{P}_x = \mathbb{P}_x\{B, \theta_s \in A\} \\
 &= \mathbb{P}_x\{B, X_{t_1+s} = a_1, \dots, X_{t_n+s} = a_n\}.
 \end{aligned}$$

Por otro lado, usando que sobre B se tiene $X_s = b$,

$$\begin{aligned}
 \int_B \mathbb{E}_{X_s}[\mathbf{1}_A] d\mathbb{P}_x &= \mathbb{P}_b\{A\} \cdot \mathbb{P}_x(B) \\
 &= \mathbb{P}_b\{X_{t_1} = a_1, \dots, X_{t_n} = a_n\} \cdot \mathbb{P}_x(B) \\
 &= \mathbb{P}_0\{X_{t_1} = a_1 - b, \dots, X_{t_n} = a_n - b\} \cdot \mathbb{P}_x(B) \quad (\text{por (11.1)}) \\
 &= \mathbb{P}_0\{X_{t_1+s} - X_s = a_1 - b, \dots\} \cdot \mathbb{P}_x(B) \quad (\text{estacionariedad}) \\
 &= \mathbb{P}_x\{X_s = b, X_{t_1+s} = a_1, \dots, X_{t_n+s} = a_n, B\},
 \end{aligned}$$

donde el último paso usa la independencia entre $B \in \mathcal{F}_s$ y los incrementos posteriores a s (ejercicio 11.2). Ambas expresiones coinciden.

Paso 2: de $\mathcal{F}_s$ a $\mathcal{F}_{s+}$. Por el Paso 1, para todo $\epsilon > 0$,

$$\mathbb{E}_x[F \circ \theta_{s+\epsilon} \mid \mathcal{F}_{s+\epsilon}] = \mathbb{E}_{X_{s+\epsilon}}[F].$$

Tomemos $\epsilon = 1/n$ y $B \in \mathcal{F}_{s+} \subset \mathcal{F}_{s+1/n}$, y consideremos $F = \mathbf{1}_{\{X_{t_1}=i_1, \dots, X_{t_m}=i_m\}}$. Integrando sobre B ,

$$\int_B \mathbf{1}_{\{X_{t_1+s+1/n}=i_1, \dots, X_{t_m+s+1/n}=i_m\}} d\mathbb{P}_x = \int_B \mathbb{E}_{X_{s+1/n}}[F] d\mathbb{P}_x.$$

Ahora hacemos $n \rightarrow \infty$. En el miembro izquierdo, la continuidad por la derecha de las trayectorias y el hecho de que los valores son enteros dan $X_{t_j+s+1/n} \rightarrow X_{t_j+s}$ puntualmente, y por convergencia dominada el límite es $\int_B \mathbf{1}_{\{X_{t_1+s}=i_1, \dots\}} d\mathbb{P}_x = \int_B F \circ \theta_s d\mathbb{P}_x$. En el miembro derecho, $X_{s+1/n} \rightarrow X_s$ y la continuidad de Feller (observación 11.1) dan $\mathbb{E}_{X_{s+1/n}}[F] \rightarrow \mathbb{E}_{X_s}[F]$, de nuevo con convergencia dominada. Se concluye

$$\int_B F \circ \theta_s d\mathbb{P}_x = \int_B \mathbb{E}_{X_s}[F] d\mathbb{P}_x \quad \text{para todo } B \in \mathcal{F}_{s+},$$

que es la afirmación. ■

11.3 La propiedad de Markov fuerte

Recordemos que $\mathcal{F}_\tau$, $\mathcal{F}_{\tau+}$ y $F \circ \theta_\tau$ se definieron en el capítulo 9 (definición 9.8 y (9.1)).

Proposición 11.6. *Si τ es un tiempo opcional, entonces para todo $k \in \mathbb{Z}_+$*

$$\{\tau < \infty, X_\tau = k\} \in \mathcal{F}_{\tau+}.$$

Demostración. Sea (τ_n) la aproximación diádica de la observación 9.15; cada τ_n es un tiempo de parada con $\tau_n \downarrow \tau$, y $\mathcal{F}_{\tau+} = \bigcap_n \mathcal{F}_{\tau_n}$. Por el corolario 9.19 aplicado al tiempo de parada τ_m ,

$$\{\tau < \infty, X_{\tau_m} = k\} \in \mathcal{F}_{\tau_m} \subset \mathcal{F}_{\tau_n} \quad \text{para } m \geq n,$$

de modo que $\mathbf{1}_{\{\tau < \infty\}} \mathbf{1}_{\{k\}}(X_{\tau_m})$ es $\mathcal{F}_{\tau_n}$ -medible para todo $m \geq n$. Haciendo $m \uparrow \infty$ y usando la continuidad por la derecha ($X_{\tau_m} \rightarrow X_\tau$ sobre $\{\tau < \infty\}$), obtenemos que $\mathbf{1}_{\{\tau < \infty, X_\tau = k\}}$ es $\mathcal{F}_{\tau_n}$ -medible para todo n , es decir el suceso está en $\bigcap_n \mathcal{F}_{\tau_n} = \mathcal{F}_{\tau+}$. ■

Teorema 11.7 (Propiedad de Markov fuerte). *Sea τ un tiempo $(\mathcal{F}_t)$ -opcional. Para toda $F : \mathbb{D}_{\mathbb{Z}_+} \rightarrow \mathbb{R}$ medible acotada (o $F \geq 0$) y todo $x_0 \in \mathbb{Z}_+$,*

$$\mathbb{E}_{x_0}[F \circ \theta_\tau \mid \mathcal{F}_{\tau+}] = \varphi(X_\tau) \quad \mathbb{P}_{x_0}\text{-c. s. sobre } \{\tau < \infty\},$$

donde $\varphi(x) = \mathbb{E}_x[F]$.

Demostración. Paso 1: τ con imagen numerable. Supongamos que $\tau : \mathbb{D}_{\mathbb{Z}_+} \rightarrow \{s_0, s_1, s_2, \dots\} \cup \{\infty\}$ es un tiempo de parada. Fijemos $A \in \mathcal{F}_\tau$; entonces $A \cap \{\tau = s_k\} \in \mathcal{F}_{s_k}$. Descomponiendo según el valor de τ y aplicando la propiedad de Markov débil en cada instante determinista s_k ,

$$\begin{aligned} \int_{A \cap \{\tau < \infty\}} F \circ \theta_\tau d\mathbb{P}_{x_0} &= \sum_{k=0}^{\infty} \int_{A \cap \{\tau = s_k\}} F \circ \theta_{s_k} d\mathbb{P}_{x_0} \\ &= \sum_{k=0}^{\infty} \int_{A \cap \{\tau = s_k\}} \mathbb{E}_{X_{s_k}}[F] d\mathbb{P}_{x_0} && \text{(teorema 11.5)} \\ &= \sum_{k=0}^{\infty} \int_{A \cap \{\tau = s_k\}} \mathbb{E}_{X_\tau}[F] d\mathbb{P}_{x_0} = \int_{A \cap \{\tau < \infty\}} \mathbb{E}_{X_\tau}[F] d\mathbb{P}_{x_0}. \end{aligned}$$

Paso 2: aproximación. Sea τ opcional y (τ_n) su aproximación diádica (observación 9.15): tiempos de parada con imagen numerable, $\tau_n \downarrow \tau$ y $\mathcal{F}_{\tau+} = \bigcap_n \mathcal{F}_{\tau_n}$. Basta probar la identidad para

$$F = \mathbf{1}_{\{k_1\}}(X_{t_1}) \cdots \mathbf{1}_{\{k_\ell\}}(X_{t_\ell}),$$

pues estas funciones generan (teorema π - λ y clase monótona). Fijemos $A \in \mathcal{F}_{\tau+} \subset \mathcal{F}_{\tau_n}$. Por el Paso 1,

$$\underbrace{\int_{A \cap \{\tau < \infty\}} F \circ \theta_{\tau_n} d\mathbb{P}_{x_0}}_{(*)} = \int_{A \cap \{\tau < \infty\}} \varphi(X_{\tau_n}) d\mathbb{P}_{x_0}. \quad (11.2)$$

En el miembro izquierdo,

$$(*) = \int_{A \cap \{\tau < \infty\}} \mathbf{1}_{\{k_1\}}(X_{\tau_n+t_1}) \cdots \mathbf{1}_{\{k_\ell\}}(X_{\tau_n+t_\ell}) d\mathbb{P}_{x_0},$$

y como $\tau_n \downarrow \tau$ y las trayectorias son continuas por la derecha, $X_{\tau_n+t_j} \rightarrow X_{\tau+t_j}$ puntualmente sobre $\{\tau < \infty\}$; por convergencia dominada,

$$(*) \longrightarrow \int_{A \cap \{\tau < \infty\}} F \circ \theta_\tau d\mathbb{P}_{x_0}.$$

En el miembro derecho, $X_{\tau_n} \rightarrow X_\tau$ y la continuidad de Feller ([observación 11.1](#)) dan $\varphi(X_{\tau_n}) \rightarrow \varphi(X_\tau)$; otra vez por convergencia dominada el límite es $\int_{A \cap \{\tau < \infty\}} \varphi(X_\tau) d\mathbb{P}_{x_0}$. Pasando al límite en ([11.2](#)),

$$\int_{A \cap \{\tau < \infty\}} F \circ \theta_\tau d\mathbb{P}_{x_0} = \int_A \varphi(X_\tau) d\mathbb{P}_{x_0} \quad \text{para todo } A \in \mathcal{F}_{\tau+},$$

que es exactamente la afirmación, pues $\varphi(X_\tau)$ es $\mathcal{F}_{\tau+}$ -medible por la [proposición 11.6](#). ■

El esquema de la demostración, para recordarlo

1. Probar la propiedad para $\tau = s$ determinista, usando indicadoras de cilindros y el teorema π - λ .
2. Extender a τ con imagen numerable, partiendo $\{\tau < \infty\} = \biguplus_k \{\tau = s_k\}$.
3. Aproximar un τ general por arriba con $\tau_n \downarrow \tau$ diádicos.
4. Pasar al límite usando (i) que las trayectorias son continuas por la derecha y (ii) la continuidad de Feller de $x \mapsto \mathbb{E}_x[F]$.

Este es exactamente el esquema que reaparecerá en el [teorema 13.25](#).

11.4 Los tiempos de espera son exponenciales i.i.d.

Cerramos la deuda contraída en el [capítulo 10](#).

Demostración del teorema 10.20. *Ley de cada ξ_n .* Ya vimos en ([10.4](#)) que $\xi_1 \sim \text{Exp}(\lambda)$. Procedemos por inducción. El punto clave es la identidad

$$\xi_{n+1} = \xi_1 \circ \theta_{J_n},$$

válida porque el primer salto de la trayectoria trasladada $\theta_{J_n}\alpha$ ocurre exactamente ξ_{n+1} unidades después de J_n . Aplicando el [teorema 11.7](#) al tiempo de parada J_n y a $F = \mathbf{1}_{\{\xi_1 \leq \ell\}}$,

$$\begin{aligned}\mathbb{P}_0\{\xi_{n+1} \leq \ell\} &= \mathbb{E}_0[\mathbf{1}_{\{\xi_1 \leq \ell\}} \circ \theta_{J_n}] = \mathbb{E}_0[\mathbb{E}_0[\mathbf{1}_{\{\xi_1 \leq \ell\}} \circ \theta_{J_n} \mid \mathcal{F}_{J_n^+}]] \\ &= \mathbb{E}_0[\varphi(X_{J_n})], \quad \varphi(x) = \mathbb{P}_x\{\xi_1 \leq \ell\}.\end{aligned}$$

Pero $\varphi(x) = \mathbb{P}_x\{\xi_1 \leq \ell\} = \mathbb{P}_0\{\xi_1 \leq \ell\} = 1 - e^{-\lambda\ell}$ para todo x , pues bajo $\mathbb{P}_x$ el proceso es el de $\mathbb{P}_0$ desplazado por la constante x y los tiempos de salto son los mismos. Por lo tanto

$$\mathbb{P}_0\{\xi_{n+1} \leq \ell\} = 1 - e^{-\lambda\ell},$$

esto es $\xi_{n+1} \sim \text{Exp}(\lambda)$.

Independencia. Veamos que ξ_{n+1} es independiente de $(\xi_1, \dots, \xi_n)$. Nótese primero que $\{\xi_1 \leq t_1, \dots, \xi_n \leq t_n\} \in \mathcal{F}_{J_n} \subset \mathcal{F}_{J_n^+}$. Por la propiedad de Markov fuerte,

$$\begin{aligned}\mathbb{P}_0\{\xi_1 \leq t_1, \dots, \xi_n \leq t_n, \xi_{n+1} \leq t\} &= \int_{\{\xi_1 \leq t_1, \dots, \xi_n \leq t_n\}} \mathbb{E}_0[\mathbf{1}_{\{\xi_1 \leq t\}} \circ \theta_{J_n} \mid \mathcal{F}_{J_n^+}] d\mathbb{P}_0 \\ &= \int_{\{\xi_1 \leq t_1, \dots, \xi_n \leq t_n\}} \underbrace{\varphi(X_{J_n})}_{=\mathbb{P}_0\{\xi_1 \leq t\}} d\mathbb{P}_0 \\ &= \mathbb{P}_0\{\xi_1 \leq t_1, \dots, \xi_n \leq t_n\} \cdot \mathbb{P}_0\{\xi_{n+1} \leq t\}.\end{aligned}$$

Por inducción sobre n se obtiene que $\xi_1, \xi_2, \dots$ son independientes, todas $\text{Exp}(\lambda)$. ■

Observación. El argumento tiene un sabor inconfundible: *el proceso de Poisson no tiene memoria*. Cada vez que salta, “se olvida” de todo y vuelve a empezar. Eso es exactamente lo que dice la propiedad de Markov fuerte aplicada a los tiempos de salto, y es también la razón por la que la distribución exponencial —la única sin memoria en tiempo continuo— es la que aparece.

Corolario 11.8 (*Ley de los tiempos de salto*). $J_n = \xi_1 + \dots + \xi_n$ sigue una ley $\Gamma(n, \lambda)$, de densidad

$$f_{J_n}(t) = \lambda^n \frac{t^{n-1}}{(n-1)!} e^{-\lambda t} \mathbf{1}_{\{t>0\}},$$

y $\{N_t \geq n\} = \{J_n \leq t\}$, de donde se reobtiene $\mathbb{P}\{N_t = n\} = e^{-\lambda t} (\lambda t)^n / n!$.

Vectores y procesos Gaussianos

El movimiento Browniano es, antes que nada, un proceso Gaussiano: sus distribuciones finito-dimensionales son normales multivariadas y su ley está completamente determinada por una función de covarianza, $K(s, t) = s \wedge t$. Este capítulo reúne el álgebra lineal probabilística que hace falta — espacios Gaussianos, proyecciones, medidas Gaussianas — y que convierte la construcción del Browniano en un ejercicio de espacios de Hilbert.

12.1 Variables Gaussianas

Fijemos $(\Omega, \mathcal{F}, \mathbb{P})$.

Definición 12.1. Una variable aleatoria X es *Gaussiana centrada reducida* si su ley $P_X(B) = \mathbb{P}\{X \in B\}$ tiene densidad

$$p_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad x \in \mathbb{R}.$$

Escribimos $X \sim \mathcal{N}(0, 1)$.

Observación 12.2. Si $X \sim \mathcal{N}(0, 1)$ entonces, para $z \in \mathbb{C}$ y $\xi \in \mathbb{R}$,

$$\underbrace{\mathbb{E}[e^{zX}]}_{\text{transformada de Laplace}} = e^{z^2/2}, \quad \underbrace{\mathbb{E}[e^{i\xi X}]}_{\text{función característica}} = e^{-\xi^2/2},$$

y

$$\mathbb{E}[X] = 0, \quad \mathbb{E}[X^2] = 1, \quad \mathbb{E}[X^{2n}] = \frac{(2n)!}{2^n n!}, \quad \mathbb{E}[X^{2n+1}] = 0.$$

Proposición 12.3. Sean $\sigma > 0$, $m \in \mathbb{R}$. Para una variable Y son equivalentes:

- (i) $Y = \sigma X + m$ con $X \sim \mathcal{N}(0, 1)$;

- (II) Y tiene densidad $p_Y(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y-m)^2}{2\sigma^2}\right)$;
- (III) $\mathbb{E}\left[e^{i\xi Y}\right] = \exp\left(im\xi - \frac{\sigma^2\xi^2}{2}\right)$ para todo $\xi \in \mathbb{R}$.

En tal caso escribimos $Y \sim \mathcal{N}(m, \sigma^2)$, y $\mathbb{E}[Y] = m$, $\text{Var}(Y) = \sigma^2$. Convenimos además $\mathcal{N}(m, 0) = \delta_m$.

Proposición 12.4 (Suma de Gaussianas). Sean $Y \sim \mathcal{N}(m, \sigma^2)$ e $\tilde{Y} \sim \mathcal{N}(\tilde{m}, \tilde{\sigma}^2)$ independientes. Entonces

$$Y + \tilde{Y} \sim \mathcal{N}(m + \tilde{m}, \sigma^2 + \tilde{\sigma}^2).$$

La independencia es indispensable

Sin independencia el enunciado es falso, y de manera espectacular: si $Y \sim \mathcal{N}(0, 1)$ y $\tilde{Y} = -Y$, entonces $Y + \tilde{Y} = 0$, que no es $\mathcal{N}(0, 2)$. Lo que sí es cierto en general es que la suma de las componentes de un *vector* Gaussiano es Gaussiana (definición 12.6); el caso independiente es un caso particular de esto.

Demostración. Por independencia, la función característica de la suma es el producto: $\mathbb{E}\left[e^{i\xi(Y+\tilde{Y})}\right] = e^{im\xi - \sigma^2\xi^2/2} e^{i\tilde{m}\xi - \tilde{\sigma}^2\xi^2/2} = e^{i(m+\tilde{m})\xi - (\sigma^2 + \tilde{\sigma}^2)\xi^2/2}$, y se aplica la proposición 12.3(iii). ■

Proposición 12.5 (Estabilidad de la clase Gaussiana). Sea (X_n) con $X_n \sim \mathcal{N}(m_n, \sigma_n^2)$ y $X_n \xrightarrow{\text{ley}} X$. Entonces los límites $m = \lim_n m_n$ y $\sigma^2 = \lim_n \sigma_n^2$ existen y $X \sim \mathcal{N}(m, \sigma^2)$. Más aún, si $X_n \xrightarrow{\mathbb{P}} X$ entonces $X_n \xrightarrow{L^p} X$ para todo $p \geq 1$.

Indicación. La convergencia en ley equivale a la convergencia puntual de las funciones características, y $e^{im_n\xi - \sigma_n^2\xi^2/2}$ converge a una función continua en 0 si y solo si convergen m_n y σ_n^2 . Para la segunda afirmación, si $X_n \xrightarrow{\mathbb{P}} X$ entonces $X_n \xrightarrow{\text{ley}} X$, luego $\sup_n \mathbb{E}[|X_n|^{2p}] < \infty$ (los momentos de una $\mathcal{N}(m_n, \sigma_n^2)$ son funciones continuas de (m_n, σ_n^2) , que convergen); por el lema 4.5, $(|X_n - X|^p)$ es uniformemente integrable, y Vitali (teorema 4.7) concluye. ■

12.2 Vectores Gaussianos

Definición 12.6 (Vector Gaussiano). Sea E un espacio euclídeo de dimensión d con producto interno $\langle \cdot, \cdot \rangle$, y $X : \Omega \rightarrow E$. Decimos que X es un *vector Gaussiano* si $\langle u, X \rangle$ es una variable Gaussiana para todo $u \in E$.

Ejemplo 12.7. Si $X_1, \dots, X_d$ son Gaussianas independientes, $X = (X_1, \dots, X_d)$ es un vector Gaussiano: $\langle \alpha, X \rangle = \sum_i \alpha_i X_i$ es Gaussiana por la [proposición 12.4](#).

Observación 12.8 (Media y forma cuadrática). Si X es un vector Gaussiano, existen $m_X \in E$ y una forma cuadrática positiva q_X sobre E tales que

$$\mathbb{E}[\langle u, X \rangle] = \langle u, m_X \rangle, \quad \text{Var}(\langle u, X \rangle) = q_X(u), \quad u \in E.$$

En una base ortonormal $\{e_j\}$,

$$m_X = \sum_{j=1}^d \mathbb{E}[X_j] e_j, \quad q_X(u) = \sum_{1 \leq j, k \leq d} u_j u_k \text{Cov}(X_j, X_k).$$

A q_X se le asocia un endomorfismo simétrico positivo γ_X —la *matriz de covarianzas*— mediante $q_X(u) = \langle u, \gamma_X u \rangle$.

Proposición 12.9 (Independencia $\iff$ covarianzas nulas). Sea $X = (X_1, \dots, X_d)$ un vector Gaussiano centrado. Son equivalentes:

- (I) $X_1, \dots, X_d$ son independientes;
- (II) $\text{Cov}(X_j, X_k) = 0$ para todos $j \neq k$, es decir, la matriz de covarianzas γ_X es diagonal en la base $\{e_1, \dots, e_d\}$.

Diagonal, no la identidad

La condición es que γ_X sea diagonal, no que sea $\mathbb{I}_d$: las varianzas pueden ser distintas entre sí (y hasta nulas). Pedir $\gamma_X = \mathbb{I}_d$ sería exigir además que todas las componentes sean $\mathcal{N}(0, 1)$.

Demostración. (i) $\Rightarrow$ (ii) es inmediato. Para (ii) $\Rightarrow$ (i), calculamos la función característica conjunta. Para $t = (t_1, \dots, t_d) \in \mathbb{R}^d$, por definición de vector Gaussiano

$$\sum_{i=1}^d t_i X_i = \langle t, X \rangle \sim \mathcal{N}(0, \langle t, \gamma_X t \rangle),$$

de modo que, evaluando su función característica en 1,

$$\varphi_X(t) = \mathbb{E}\left[e^{i\langle t, X \rangle}\right] = e^{-\frac{1}{2}\langle t, \gamma_X t \rangle}.$$

Si γ_X es diagonal, $\langle t, \gamma_X t \rangle = \sum_j t_j^2 \mathbb{E}[X_j^2]$ y entonces

$$\varphi_X(t) = \prod_{j=1}^d e^{-\frac{1}{2}t_j^2 \mathbb{E}[X_j^2]} = \varphi_{X_1}(t_1) \cdots \varphi_{X_d}(t_d),$$

lo que equivale a la independencia. ■

Teorema 12.10 (Estructura de los vectores Gaussianos). (a) *Para todo endomorfismo simétrico positivo γ de E existe un vector Gaussiano centrado X con $\gamma_X = \gamma$.*

(b) *Sea X un vector Gaussiano centrado y $\{e_1, \dots, e_d\}$ una base ortonormal de vectores propios de γ_X , $\gamma_X e_j = \lambda_j e_j$, con*

$$\lambda_1 \geq \cdots \geq \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_d.$$

Entonces

$$X = \sum_{j=1}^r Y_j e_j, \quad Y_j = \langle e_j, X \rangle \sim \mathcal{N}(0, \lambda_j) \text{ independientes.}$$

En particular X vive c. s. en el subespacio de dimensión r generado por $e_1, \dots, e_r$.

(c) *Si $r = d$ (es decir, γ_X es invertible), X tiene densidad*

$$p_X(x) = \frac{1}{(2\pi)^{d/2} \sqrt{\det \gamma_X}} \exp\left(-\frac{1}{2} \langle x, \gamma_X^{-1} x \rangle\right).$$

12.3 Espacios y procesos Gaussianos

Definición 12.11 (Espacio Gaussiano). Un *subespacio Gaussiano centrado* de $L^2(\Omega, \mathcal{F}, \mathbb{P})$ es un subespacio vectorial *cerrado* formado íntegramente por variables Gaussianas centradas.

Ejemplo 12.12. Si $X = (X_1, \dots, X_d)$ es un vector Gaussiano centrado en $\mathbb{R}^d$, entonces $\text{gen}\{X_1, \dots, X_d\}$ es un espacio Gaussiano.

Definición 12.13 (Proceso Gaussiano). Sea T un conjunto de índices. Un proceso $(X_t)_{t \in T}$ con valores reales es un *proceso Gaussiano centrado* si toda combinación lineal finita $\sum_i c_i X_{t_i}$ es una variable Gaussiana centrada. El subespacio cerrado de L^2 generado por $(X_t)_{t \in T}$ es entonces un espacio Gaussiano, llamado el *espacio Gaussiano engendrado* por el proceso.

Teorema 12.14 (Ortogonalidad $\iff$ independencia). Sea $\mathcal{H}$ un espacio Gaussiano y $(\mathcal{H}_i)_{i \in I}$ subespacios de $\mathcal{H}$. Entonces

$$(\mathcal{H}_i)_{i \in I} \text{ ortogonales dos a dos en } L^2 \iff (\sigma(\mathcal{H}_i))_{i \in I} \text{ independientes dos a dos.}$$

El enunciado exige un espacio Gaussiano común

Es esencial que los $\mathcal{H}_i$ sean subespacios del *mismo* espacio Gaussiano $\mathcal{H}$. Contraejemplo: sea $X \sim \mathcal{N}(0, 1)$ y ϵ independiente con $\mathbb{P}\{\epsilon = 1\} = \mathbb{P}\{\epsilon = -1\} = \frac{1}{2}$. Poniendo $X_1 = X$ y $X_2 = \epsilon X$, ambas son $\mathcal{N}(0, 1)$ y

$$\mathbb{E}[X_1 X_2] = \mathbb{E}[\epsilon] \mathbb{E}[X^2] = 0,$$

pero $|X_1| = |X_2|$: no hay independencia. El punto es que (X_1, X_2) *no* es un vector Gaussiano —por ejemplo, $X_1 + X_2$ vale 0 o $2X$ con probabilidad $1/2$ cada uno, y no es Gaussiana.

Corolario 12.15 (Esperanza condicional = proyección). Sea $\mathcal{H}$ un espacio Gaussiano, $K \subset \mathcal{H}$ un subespacio cerrado, y p_K la proyección ortogonal sobre K . Para $X \in \mathcal{H}$:

- (a) $\mathbb{E}[X | \sigma(K)] = p_K(X)$;
- (b) con $\sigma^2 = \mathbb{E}[(X - p_K(X))^2]$, la ley condicional de X dado $\sigma(K)$ es $\mathcal{N}(p_K(X), \sigma^2)$; es decir, para todo $A \in \mathcal{B}_{\mathbb{R}}$,

$$\mathbb{P}\{X \in A | \sigma(K)\} = \frac{1}{\sigma\sqrt{2\pi}} \int_A \exp\left(-\frac{(y - p_K(X))^2}{2\sigma^2}\right) dy.$$

Demostración. (a) Escribimos $X = p_K(X) + (X - p_K(X))$. El segundo sumando es ortogonal a K y pertenece a $\mathcal{H}$, de modo que por el [teorema 12.14](#) es independiente de $\sigma(K)$; luego $\mathbb{E}[X - p_K(X) | \sigma(K)] = \mathbb{E}[X - p_K(X)] = 0$. Como $p_K(X)$ es $\sigma(K)$ -medible, se concluye.

(b) Con la misma descomposición, $X = p_K(X) + Z$ con $Z \sim \mathcal{N}(0, \sigma^2)$ independiente de $\sigma(K)$, de donde la ley condicional se obtiene por sustitución. ■

Por qué los Gaussianos son fáciles

En un espacio Gaussiano, *la geometría de L^2 agota la probabilidad*: la ortogonalidad es independencia, la proyección ortogonal es esperanza condicional, y toda la ley está codificada en el producto interno $\Gamma(s, t) = \mathbb{E}[X_s X_t]$. Es la razón por la que el movimiento Browniano se puede construir con puro análisis funcional (sección 12.4).

Dado un proceso Gaussiano centrado $(X_t)_{t \in T}$, su *función de covarianza* es

$$\Gamma : T \times T \rightarrow \mathbb{R}, \quad \Gamma(s, t) = \mathbb{E}[X_s X_t].$$

Teorema 12.16 (Existencia dada la covarianza). *Sea $\Gamma : T \times T \rightarrow \mathbb{R}$ simétrica y de tipo positivo, esto es,*

$$\sum_{s, t \in T} c(s)c(t) \Gamma(s, t) \geq 0$$

para toda función $c : T \rightarrow \mathbb{R}$ de soporte finito. Entonces existe un proceso Gaussiano centrado con función de covarianza Γ .

Observación. La condición es necesaria, pues $\sum_{s, t} c(s)c(t) \Gamma(s, t) = \mathbb{E}[(\sum_s c(s) X_s)^2] \geq 0$. La construcción usa el teorema de extensión de Kolmogorov, o bien —y esto es lo que haremos para el Browniano— una isometría explícita desde un espacio L^2 .

12.4 Medidas Gaussianas (ruido blanco)

Recordemos primero el marco de Hilbert.

Teorema 12.17 (Bases ortonormales). *Sea $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ un espacio de Hilbert y $\{u_\alpha : \alpha \in \Lambda\}$ un sistema ortonormal con Λ numerable. Son equivalentes:*

- (I) *el sistema es completo: si $\langle x, u_\alpha \rangle = 0$ para todo α , entonces $x = 0$;*
- (II) *(identidad de Parseval) $\|x\|^2 = \sum_{\alpha \in \Lambda} |\langle x, u_\alpha \rangle|^2$ para todo $x \in \mathcal{H}$;*
- (III) *$x = \sum_{\alpha \in \Lambda} \langle x, u_\alpha \rangle u_\alpha$ para todo $x \in \mathcal{H}$ (con convergencia en norma).*

Si se cumplen, $\{u_\alpha\}$ es una base ortonormal.

Proposición 12.18. *$(\mathcal{H}, \langle \cdot, \cdot \rangle)$ es separable si y solo si admite una base ortonormal numerable. En tal caso, dos espacios de Hilbert separables de igual dimensión son isométricamente isomorfos.*

Ejemplo 12.19. $(L^2(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, m), \langle f, g \rangle = \int fg \, dm)$ es separable, pues

$$\left\{ \sum_{i=1}^k q_i \mathbf{1}_{(a_i, b_i)} : k \geq 1, q_i, a_i, b_i \in \mathbb{Q}, a_i < b_i \right\}$$

es numerable y denso. Lo mismo vale para $L^2(\mathbb{R}_+, \mathcal{B}_{\mathbb{R}_+}, m)$.

Ejemplo 12.20 (Un espacio Gaussiano concreto). Sea $(\zeta_\alpha)_{\alpha \in \Lambda}$, Λ numerable, una familia i.i.d. $\mathcal{N}(0, 1)$ sobre $(\Omega, \mathcal{F}, \mathbb{P})$. Entonces

$$\mathcal{H} = \overline{\text{gen}}\{\zeta_\alpha : \alpha \in \Lambda\} \subset L^2(\Omega, \mathcal{F}, \mathbb{P})$$

es un espacio Gaussiano: si $X = \sum_\alpha c_\alpha \zeta_\alpha$ con $\sum_\alpha c_\alpha^2 < \infty$, la [proposición 12.5](#) da $X \sim \mathcal{N}(0, \sum_\alpha c_\alpha^2) = \mathcal{N}(0, \|X\|_{L^2}^2)$. Además $\{\zeta_\alpha\}$ es una base ortonormal de $\mathcal{H}$.

Definición 12.21 (Medida Gaussiana). Sea $(E, \mathcal{E})$ un espacio medible y μ una medida σ -finita sobre él. Una *medida Gaussiana de intensidad* μ es una isometría lineal

$$G : L^2(E, \mathcal{E}, \mu) \longrightarrow \mathcal{H}$$

sobre un espacio Gaussiano $\mathcal{H}$. En particular, para $f \in L^2(E, \mathcal{E}, \mu)$, $G(f)$ es una variable Gaussiana centrada con

$$\mathbb{E}[G(f)^2] = \|G(f)\|_{L^2(\Omega)}^2 = \|f\|_{L^2(\mu)}^2, \quad \mathbb{E}[G(f)G(g)] = \int_E fg \, d\mu. \quad (12.1)$$

Teorema 12.22 (Existencia). Para toda medida σ -finita μ sobre $(E, \mathcal{E})$ existe una medida Gaussiana de intensidad μ .

Demostración. $L^2(E, \mathcal{E}, \mu)$ es separable, luego admite una base ortonormal numerable $(f_i)_{i \in I}$. Sea $(\zeta_i)_{i \in I}$ i.i.d. $\mathcal{N}(0, 1)$ y $\mathcal{H}$ como en el [ejemplo 12.20](#). Para $f = \sum_i \alpha_i f_i$ con $\sum_i \alpha_i^2 < \infty$ definimos

$$G(f) = \sum_{i \in I} \alpha_i \zeta_i,$$

serie que converge en $L^2(\Omega)$ y satisface $\mathbb{E}[G(f)^2] = \sum_i \alpha_i^2 = \|f\|_{L^2(\mu)}^2$. Es manifiestamente lineal, y por polarización preserva el producto interno. ■

Observación 12.23 (Ruido blanco). Escribimos $G(A) := G(\mathbf{1}_A)$ para $A \in \mathcal{E}$ con $\mu(A) < \infty$; entonces $G(A) \sim \mathcal{N}(0, \mu(A))$. Si $A_1, \dots, A_n$ son *disjuntos*

con medida finita, el vector $(G(A_1), \dots, G(A_n))$ es Gaussiano con matriz de covarianzas diagonal, pues

$$\mathbb{E}[G(A_i)G(A_j)] = \langle \mathbf{1}_{A_i}, \mathbf{1}_{A_j} \rangle = \mu(A_i \cap A_j) = 0 \quad (i \neq j),$$

de modo que por la [proposición 12.9](#) son *independientes*. Además, si $A = \biguplus_{j \geq 1} A_j$ con $\mu(A) < \infty$, la convergencia $\mathbf{1}_A = \sum_j \mathbf{1}_{A_j}$ en $L^2(\mu)$ y la isometría dan

$$G(A) = \sum_{j=1}^{\infty} G(A_j) \quad \text{en } L^2(\Omega).$$

G se comporta, pues, como una “medida con signo aleatoria”, salvo que *no* es una medida verdadera: sus trayectorias no tienen variación acotada. Es lo que se llama *ruido blanco*.

La construcción del Browniano en dos líneas

Tomemos $E = \mathbb{R}_+$, $\mu = m$ (Lebesgue) y G una medida Gaussiana de intensidad m . Definiendo

$$B_t = G(\mathbf{1}_{[0,t]}), \quad t \geq 0,$$

la isometría (12.1) da de inmediato

$$\mathbb{E}[B_s B_t] = \int_0^{\infty} \mathbf{1}_{[0,s]} \mathbf{1}_{[0,t]} dr = s \wedge t,$$

que es exactamente la covarianza del movimiento Browniano. El [capítulo 13](#) desarrolla esta idea.

Proposición 12.24 (Variación cuadrática del ruido blanco). *Sea G una medida Gaussiana sobre $(E, \mathcal{E})$ de intensidad μ , y $A \in \mathcal{E}$ con $\mu(A) < \infty$. Supongamos que existe una sucesión de particiones medibles*

$$A = A_1^n \uplus \dots \uplus A_{k_n}^n \quad \text{con} \quad \lim_n \left(\max_{1 \leq j \leq k_n} \mu(A_j^n) \right) = 0.$$

Entonces

$$\lim_{n \rightarrow \infty} \sum_{j=1}^{k_n} G(A_j^n)^2 = \mu(A) \quad \text{en } L^2(\Omega).$$

Demostración. Sea $S_n = \sum_j G(A_j^n)^2$. Las variables $G(A_j^n)$, $j = 1, \dots, k_n$, son independientes $\mathcal{N}(0, \mu(A_j^n))$, de modo que $\mathbb{E}[S_n] = \sum_j \mu(A_j^n) = \mu(A)$ y,

usando $\text{Var}(Z^2) = 2\sigma^4$ para $Z \sim \mathcal{N}(0, \sigma^2)$,

$$\text{Var}(S_n) = \sum_{j=1}^{k_n} 2\mu(A_j^n)^2 \leq 2\left(\max_j \mu(A_j^n)\right) \sum_j \mu(A_j^n) = 2\mu(A) \max_j \mu(A_j^n) \rightarrow 0.$$

Luego $\mathbb{E}[(S_n - \mu(A))^2] = \text{Var}(S_n) \rightarrow 0$. ■

12.5 Ejercicios resueltos

Ejercicio 12.25. Sea $(X_t)_{t \in [0,1]}$ un proceso Gaussiano centrado tal que $(t, \omega) \mapsto X_t(\omega)$ es medible de $[0, 1] \times \Omega$ en $\mathbb{R}$. Sea K su función de covarianza.

- (a) Pruebe que $t \mapsto X_t$ es continua de $[0, 1]$ en $L^2(\Omega)$ si y solo si K es continua sobre $[0, 1]^2$. (*Mantendremos este supuesto en lo que sigue.*)
- (b) Sea $h : [0, 1] \rightarrow \mathbb{R}$ medible con $\int_0^1 |h(t)| \sqrt{K(t, t)} dt < \infty$. Pruebe que, para casi todo ω , la integral $\int_0^1 h(t) X_t(\omega) dt$ es absolutamente convergente. Denotamos $Z = \int_0^1 h(t) X_t dt$.
- (c) Suponiendo además $\int_0^1 |h(t)| dt < \infty$, pruebe que Z es el límite en L^2 de

$$Z_n = \sum_{i=1}^n X_{i/n} \int_{(i-1)/n}^{i/n} h(t) dt,$$

y deduzca que Z es Gaussiana.

Solución. (a) Supongamos K continua. Entonces

$$\|X_{t+h} - X_t\|_{L^2(\Omega)}^2 = \mathbb{E}[|X_{t+h} - X_t|^2] = K(t+h, t+h) - 2K(t+h, t) + K(t, t),$$

que tiende a 0 cuando $h \rightarrow 0$ por continuidad de K ; luego $t \mapsto X_t$ es continua en L^2 .

Recíprocamente, supongamos $t \mapsto X_t$ continua en L^2 . Sumando y restando y aplicando Cauchy-Schwarz,

$$\begin{aligned} |K(u+t, v+s) - K(u, v)| &\leq |K(u+t, v+s) - K(u, v+s)| + |K(u, v+s) - K(u, v)| \\ &\leq \mathbb{E}[|(X_{u+t} - X_u)X_{v+s}|] + \mathbb{E}[|(X_{v+s} - X_v)X_u|] \\ &\leq \|X_{u+t} - X_u\|_{L^2} \|X_{v+s}\|_{L^2} + \|X_{v+s} - X_v\|_{L^2} \|X_u\|_{L^2}, \end{aligned}$$

y el miembro derecho tiende a 0 cuando $(t, s) \rightarrow (0, 0)$, pues $\|X_r\|_{L^2} = \sqrt{K(r, r)}$ está acotada sobre el compacto $[0, 1]$ por continuidad en L^2 .

(b) Basta ver que la integral doble es finita; por Tonelli y Cauchy–Schwarz,

$$\begin{aligned} \int_{\Omega} \int_0^1 |X_t(\omega)| |h(t)| dt \mathbb{P}(d\omega) &= \int_0^1 |h(t)| \|X_t\|_{L^1} dt \\ &\leq \int_0^1 |h(t)| \|X_t\|_{L^2} dt = \int_0^1 |h(t)| \sqrt{K(t, t)} dt < \infty. \end{aligned}$$

Luego $\int_0^1 |h(t)X_t(\omega)| dt < \infty$ para casi todo ω .

(c) Cada Z_n es una combinación lineal finita de Gaussianas centradas del proceso, luego es Gaussiana centrada; por la [proposición 12.5](#), si $Z_n \rightarrow Z$ en L^2 entonces Z es Gaussiana. Escribamos

$$Z_n(\omega) = \int_0^1 h(t) \left(\sum_{i=1}^n X_{i/n}(\omega) \mathbf{1}_{[\frac{i-1}{n}, \frac{i}{n})}(t) \right) dt.$$

Por la desigualdad integral de Minkowski,

$$\|Z - Z_n\|_{L^2(\Omega)} \leq \int_0^1 |h(t)| \left\| X_t - \sum_{i=1}^n X_{i/n} \mathbf{1}_{[\frac{i-1}{n}, \frac{i}{n})}(t) \right\|_{L^2(\Omega)} dt.$$

Para $t \in [\frac{k-1}{n}, \frac{k}{n})$ el integrando vale $|h(t)| \|X_t - X_{k/n}\|_{L^2}$, que por (a) tiende a 0 para cada t , y está dominado por

$$|h(t)| (\|X_t\|_{L^2} + \|X_{k/n}\|_{L^2}) \leq 2 |h(t)| \sup_{s \in [0, 1]} \sqrt{K(s, s)} =: C |h(t)|,$$

que es integrable por hipótesis. El teorema de convergencia dominada da $\|Z - Z_n\|_{L^2} \rightarrow 0$. ■

Ejercicio 12.26. Sea $\mathcal{H}$ un espacio Gaussiano centrado, $\mathcal{H}_1, \mathcal{H}_2 \subset \mathcal{H}$ subespacios lineales y $K \subset \mathcal{H}$ un subespacio cerrado, con proyección ortogonal p_K . Demuestre que

$$\mathbb{E}[X_1 X_2] = \mathbb{E}[p_K(X_1) p_K(X_2)] \quad \text{para todos } X_1 \in \mathcal{H}_1, X_2 \in \mathcal{H}_2 \quad (12.2)$$

implica que $\sigma(\mathcal{H}_1)$ y $\sigma(\mathcal{H}_2)$ son *condicionalmente independientes dado* $\sigma(K)$: para toda $Y_1 \geq 0$ $\sigma(\mathcal{H}_1)$ -medible y toda $Y_2 \geq 0$ $\sigma(\mathcal{H}_2)$ -medible,

$$\mathbb{E}[Y_1 Y_2 | \sigma(K)] = \mathbb{E}[Y_1 | \sigma(K)] \mathbb{E}[Y_2 | \sigma(K)].$$

Solución. Por el teorema de la clase monótona ([teorema A.5](#)) basta probar la identidad para $Y_s = \prod_i \mathbf{1}_{\Gamma_i^s}(X_i^s)$ con $X_i^s \in \mathcal{H}_s$ y $\Gamma_i^s \in \mathcal{B}_{\mathbb{R}}$, $s = 1, 2$.

Fijemos $M_s = \text{gen}\{X_1^s, \dots, X_{n_s}^s\} \subset \mathcal{H}_s$ y elijamos un sistema ortonormal $\{Z_i^s : i = 1, \dots, m_s\}$ de M_s . Como $\{Z_i^s\}_i$ es ortonormal dentro de un espacio Gaussiano, por el [teorema 12.14](#) las $Z_1^s, \dots, Z_{m_s}^s$ son Gaussianas independientes $\mathcal{N}(0, 1)$. Basta probar la factorización para las Z , pues $\sigma(M_s) = \sigma(Z_1^s, \dots, Z_{m_s}^s)$.

Pongamos

$$Y_i^s = Z_i^s - p_K(Z_i^s), \quad (\sigma_i^s)^2 = \mathbb{E}[(Y_i^s)^2].$$

Dentro de cada bloque, la independencia de los Z_i^s da $\mathbb{E}[Y_i^s Y_j^s] = 0$ para $i \neq j$. Entre bloques distintos, usando (12.2) y el [corolario 12.15\(a\)](#),

$$\begin{aligned} \mathbb{E}[Y_i^1 Y_j^2] &= \mathbb{E}[Z_i^1 Z_j^2] + \mathbb{E}[p_K(Z_i^1) p_K(Z_j^2)] - \mathbb{E}[Z_i^1 p_K(Z_j^2)] - \mathbb{E}[p_K(Z_i^1) Z_j^2] \\ &= 2 \mathbb{E}[p_K(Z_i^1) p_K(Z_j^2)] - \mathbb{E}[\mathbb{E}[Z_i^1 | \sigma(K)] p_K(Z_j^2)] - \mathbb{E}[p_K(Z_i^1) \mathbb{E}[Z_j^2 | \sigma(K)]] \\ &= 2 \mathbb{E}[p_K(Z_i^1) p_K(Z_j^2)] - 2 \mathbb{E}[p_K(Z_i^1) p_K(Z_j^2)] = 0. \end{aligned}$$

Por lo tanto la familia completa $\{Y_i^s\}$ es ortogonal dentro de $\mathcal{H}$ y, además, cada Y_i^s es ortogonal a K ; por el [teorema 12.14](#) las Y_i^s son independientes entre sí e independientes de $\sigma(K)$.

Definiendo

$$F(z_1^1, \dots, z_{m_2}^2) = \prod_{s=1,2} \prod_{i=1}^{m_s} \mathbf{1}_{\Gamma_i^s}(z_i^s)$$

y escribiendo $Z_i^s = Y_i^s + p_K(Z_i^s)$ con $p_K(Z_i^s)$ $\sigma(K)$ -medible, la propiedad de sustitución da

$$\begin{aligned} \mathbb{E}[F(Z_1^1, \dots, Z_{m_2}^2) | \sigma(K)] &= \prod_{s=1,2} \prod_{i=1}^{m_s} \int \mathbf{1}_{\Gamma_i^s}(y + p_K(Z_i^s)) \mathbb{P}_{Y_i^s}(dy) \\ &= \prod_{s=1,2} \prod_{i=1}^{m_s} \mathbb{P}\{Z_i^s \in \Gamma_i^s | \sigma(K)\}, \end{aligned}$$

donde cada factor es, por el [corolario 12.15\(b\)](#),

$$\mathbb{P}\{Z_i^s \in \Gamma_i^s | \sigma(K)\} = \frac{1}{\sigma_i^s \sqrt{2\pi}} \int_{\Gamma_i^s} \exp\left(-\frac{(y - p_K(Z_i^s))^2}{2(\sigma_i^s)^2}\right) dy.$$

Agrupando los factores con $s = 1$ y los factores con $s = 2$ se obtiene exactamente $\mathbb{E}[Y_1 | \sigma(K)] \mathbb{E}[Y_2 | \sigma(K)]$. ■

El movimiento Browniano

El movimiento Browniano es el proceso de Lévy continuo por excelencia y el objeto central del resto de estas notas. Lo construiremos como imagen de una medida Gaussiana, probaremos que admite una versión con trayectorias continuas (Kolmogorov–Čentsov), estudiaremos la ley de Wiener sobre $\mathcal{C}_{\mathbb{R}}$, y estableceremos sus propiedades finas: la ley 0–1 de Blumenthal, la propiedad de Markov fuerte, el principio de reflexión y la variación cuadrática.

13.1 Motivación: procesos de Lévy

Definición 13.1 (Proceso de Lévy). Una probabilidad $\mathbb{P}$ sobre $(\mathbb{D}_{\mathbb{R}}, \mathcal{F})$ es la ley de un proceso de Lévy (en $(\mathbb{R}, +)$) si, bajo ella, el proceso canónico verifica

- (I) $X_0 = 0$ c. s.;
- (II) para $0 = t_0 < t_1 < \dots < t_k$, los incrementos $X_{t_1} - X_{t_0}, \dots, X_{t_k} - X_{t_{k-1}}$ son independientes;
- (III) $X_t - X_s \sim X_{t-s}$ para $0 \leq s < t$: la ley del incremento depende solo de la longitud del intervalo.

Un proceso $\eta : (\Omega, \mathcal{F}', \mathbb{P}') \rightarrow (\mathbb{D}_{\mathbb{R}}, \mathcal{F})$ cuya ley cumple (i)–(iii) se llama proceso de Lévy.

Ejemplo 13.2. Un proceso de Poisson (N_t) es un proceso de Lévy, con $N_t - N_s \sim \text{Poiss}(\lambda(t-s))$ y valores en $\mathbb{Z}_+$.

Ejercicio 13.3 (Proceso de Poisson compuesto). Sea (N_t) un proceso de Poisson de intensidad λ y $(Y_j)_{j \geq 1}$ una sucesión i.i.d. de ley μ , independiente de N . Pruebe que

$$\eta_t = \sum_{j=1}^{N_t} Y_j, \quad t \geq 0,$$

es un proceso de Lévy, llamado *Poissonian* (λ, μ) .

Solución. La vía más limpia es la función característica. Condicionando a $N_t = n$ y usando la independencia,

$$\mathbb{E}\left[e^{i\xi\eta_t}\right] = \sum_{n \geq 0} \mathbb{P}\{N_t = n\} \hat{\mu}(\xi)^n = e^{-\lambda t} \sum_{n \geq 0} \frac{(\lambda t \hat{\mu}(\xi))^n}{n!} = \exp\left\{\lambda t(\hat{\mu}(\xi) - 1)\right\},$$

donde $\hat{\mu}(\xi) = \int e^{i\xi y} \mu(dy)$. El mismo cálculo, condicionando sobre los incrementos de N —que son independientes— y usando que los Y_j son i.i.d., da para $t_0 < \dots < t_k$

$$\mathbb{E}\left[\exp\left(i \sum_{j=1}^k \xi_j(\eta_{t_j} - \eta_{t_{j-1}})\right)\right] = \prod_{j=1}^k \exp\left\{\lambda(t_j - t_{j-1})(\hat{\mu}(\xi_j) - 1)\right\},$$

lo que establece simultáneamente (ii) y (iii). Finalmente $\eta_0 = 0$ pues $N_0 = 0$, y las trayectorias son càdlàg por serlo las de N . ■

Por qué el Browniano es inevitable

En un proceso de Lévy podemos escribir, para todo n ,

$$X_t = \sum_{k=1}^n \underbrace{(X_{kt/n} - X_{(k-1)t/n})}_{\text{i.i.d.}},$$

de modo que la ley de X_t es *infinitamente divisible*. Si además exigimos que $t \mapsto X_t$ sea *continua*, la única posibilidad (teorema de Lévy–Khintchine, sin saltos) es

$$X_t \sim \mathcal{N}(at, \sigma^2 t), \quad \text{es decir} \quad X_t = at + \sigma B_t,$$

con B un proceso de Lévy continuo estándar. La ley de B es la *medida de Wiener* y B es el *movimiento Browniano*.

13.2 El pre-movimiento Browniano

Definición 13.4 (Pre-movimiento Browniano). Sea G una medida Gaussiana (definición 12.21) sobre $E = \mathbb{R}_+$ con intensidad la medida de Lebesgue m . El proceso

$$B_t = G(\mathbf{1}_{[0,t]}), \quad t \geq 0,$$

se llama *pre-movimiento Browniano*.

Proposición 13.5. $(B_t)_{t \geq 0}$ es un proceso Gaussiano centrado con función de covarianza

$$K(s, t) = \mathbb{E}[B_s B_t] = \min\{s, t\} = s \wedge t.$$

Demostración. Como G toma valores en un espacio Gaussiano, toda combinación lineal finita de los B_t es Gaussiana centrada. Por la isometría (12.1),

$$\mathbb{E}[B_s B_t] = \mathbb{E}[G(\mathbf{1}_{[0,s]})G(\mathbf{1}_{[0,t]})] = \int_0^\infty \mathbf{1}_{[0,s]}(r)\mathbf{1}_{[0,t]}(r) dr = s \wedge t. \quad \blacksquare$$

Proposición 13.6 (Caracterizaciones equivalentes). Sea $(X_t)_{t \geq 0}$ un proceso con valores reales. Son equivalentes:

- (I) (X_t) es un pre-movimiento Browniano;
- (II) (X_t) es un proceso Gaussiano centrado con covarianza $K(s, t) = s \wedge t$;
- (III) $X_0 = 0$ c. s. y, para $0 \leq s < t$, el incremento $X_t - X_s$ es independiente de $\sigma(X_r : r \leq s)$ y $X_t - X_s \sim \mathcal{N}(0, t - s)$;
- (IV) $X_0 = 0$ c. s. y, para todos $0 = t_0 < t_1 < \dots < t_p$, los incrementos $X_{t_i} - X_{t_{i-1}}$ son independientes con $X_{t_i} - X_{t_{i-1}} \sim \mathcal{N}(0, t_i - t_{i-1})$.

Demostración. (i) $\Rightarrow$ (ii) es la [proposición 13.5](#).

(ii) $\Rightarrow$ (iii). Sea $\mathcal{H}$ el espacio Gaussiano generado por (X_t) . De $K(0, 0) = 0$ se sigue $X_0 \sim \mathcal{N}(0, 0)$, es decir $X_0 = 0$ c. s.. Fijemos $s > 0$ y sean

$$\mathcal{H}_s = \overline{\text{gen}}\{X_r : r \leq s\}, \quad \tilde{\mathcal{H}}_s = \overline{\text{gen}}\{X_{s+u} - X_s : u \geq 0\}.$$

Son ortogonales: para $r \leq s$ y $u \geq 0$,

$$\mathbb{E}[X_r(X_{s+u} - X_s)] = \mathbb{E}[X_r X_{s+u}] - \mathbb{E}[X_r X_s] = r \wedge (s+u) - r \wedge s = r - r = 0.$$

Por el [teorema 12.14](#), $\sigma(\mathcal{H}_s)$ y $\sigma(\tilde{\mathcal{H}}_s)$ son independientes; en particular, para $t > s$, $X_t - X_s$ es independiente de $\sigma(X_r : r \leq s)$. Su varianza es

$$\mathbb{E}[(X_t - X_s)^2] = \mathbb{E}[X_t^2] - 2\mathbb{E}[X_t X_s] + \mathbb{E}[X_s^2] = t - 2(s \wedge t) + s = t - 2s + s = t - s.$$

(iii) $\Rightarrow$ (iv) es inmediato por inducción.

(iv) $\Rightarrow$ (i). Las combinaciones lineales finitas de los X_t son combinaciones lineales de incrementos independientes Gaussianos, luego Gaussianas: (X_t)

es un proceso Gaussiano centrado. Definamos, para una función escalonada $f = \sum_{i=1}^n \lambda_i \mathbf{1}_{(t_{i-1}, t_i]}$,

$$G(f) = \sum_{i=1}^n \lambda_i (X_{t_i} - X_{t_{i-1}}).$$

Por la independencia de los incrementos,

$$\mathbb{E}[G(f)G(g)] = \int_{\mathbb{R}_+} f(t)g(t) dt,$$

de modo que G es una isometría sobre el conjunto (denso en $L^2(\mathbb{R}_+)$) de las funciones escalonadas; se extiende por densidad a una isometría $G : L^2(\mathbb{R}_+) \rightarrow \mathcal{H}$, que es una medida Gaussiana de intensidad m . Finalmente, por construcción $G(\mathbf{1}_{[0,t]}) = X_t - X_0 = X_t$ c. s.. ■

Corolario 13.7 (Distribución conjunta). *Sea (B_t) un pre-movimiento Browniano. Para $0 = t_0 < t_1 < \dots < t_n$, el vector $(B_{t_1}, \dots, B_{t_n})$ tiene densidad*

$$p(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2} \sqrt{t_1(t_2 - t_1) \dots (t_n - t_{n-1})}} \exp\left(-\sum_{i=1}^n \frac{(x_i - x_{i-1})^2}{2(t_i - t_{i-1})}\right), \quad (13.1)$$

con la convención $x_0 = 0$.

Demostración. Las variables $B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}$ son independientes con leyes $\mathcal{N}(0, t_1), \dots, \mathcal{N}(0, t_n - t_{n-1})$, de modo que el vector de incrementos tiene densidad

$$q(y_1, \dots, y_n) = \frac{1}{(2\pi)^{n/2} \sqrt{t_1(t_2 - t_1) \dots (t_n - t_{n-1})}} \exp\left(-\sum_{i=1}^n \frac{y_i^2}{2(t_i - t_{i-1})}\right).$$

El cambio de variable $x_i = y_1 + \dots + y_i$ tiene jacobiano 1 y transforma q en p . ■

Proposición 13.8 (Invariancias). *Si (B_t) es un pre-movimiento Browniano, también lo son:*

- (a) (simetría) $(-B_t)_{t \geq 0}$;
- (b) (invariancia de escala) $B_t^\lambda = \lambda^{-1} B_{\lambda^2 t}$, para cada $\lambda > 0$;
- (c) (Markov simple) $B_t^{(s)} = B_{s+t} - B_s$, para cada $s \geq 0$; y $(B_t^{(s)})_{t \geq 0}$ es independiente de $\sigma(B_r : r \leq s)$.

Demostración. (a) y (b) se comprueban con la caracterización (ii) de la [proposición 13.6](#): la covarianza de $-B$ es la misma, y

$$\mathbb{E}[B_s^\lambda B_t^\lambda] = \lambda^{-2} \mathbb{E}[B_{\lambda^2 s} B_{\lambda^2 t}] = \lambda^{-2} (\lambda^2 s \wedge \lambda^2 t) = s \wedge t.$$

(c) La σ -álgebra generada por $B^{(s)}$ es $\sigma(\tilde{\mathcal{H}}_s)$, independiente de $\sigma(\mathcal{H}_s) = \sigma(B_r : r \leq s)$ por la demostración de la [proposición 13.6](#). Que $B^{(s)}$ sea pre-Browniano se verifica con (iv): $B_{t_i}^{(s)} - B_{t_{i-1}}^{(s)} = B_{t_i+s} - B_{t_{i-1}+s} \sim \mathcal{N}(0, t_i - t_{i-1})$, independientes. ■

13.2.1 La integral de Wiener

Si B es un pre-movimiento Browniano con medida Gaussiana asociada G , escribimos para $f \in L^2(\mathbb{R}_+, \mathcal{B}_{\mathbb{R}_+}, dt)$

$$\int_0^\infty f(s) dB_s := G(f), \quad \int_s^t f(r) dB_r := G(f \mathbf{1}_{(s,t]}). \quad (13.2)$$

La notación se justifica porque $\int_u^v dB_s = G(\mathbf{1}_{(u,v]}) = B_v - B_u$. La aplicación $f \mapsto \int_0^\infty f dB_s$ se llama *integral de Wiener*, y por la isometría (12.1),

$$\int_0^\infty f(s) dB_s \sim \mathcal{N}\left(0, \int_0^\infty f(s)^2 ds\right). \quad (13.3)$$

Una integral que no es una integral

Pese a la notación, dB_s no define una medida: veremos en la [sección 13.7](#) que las trayectorias del Browniano no tienen variación acotada, de modo que no se puede integrar trayectoria a trayectoria con la teoría de Lebesgue–Stieltjes. La integral de Wiener existe por razones puramente L^2 , y su extensión a integrandos *aleatorios* —que es lo que se necesita en las aplicaciones— es la integral de Itô del [capítulo 15](#).

13.3 La continuidad de las trayectorias

Definición 13.9. Sean $(X_t)_{t \in T}$ y $(\tilde{X}_t)_{t \in T}$ dos procesos indexados por el mismo conjunto y con valores en el mismo espacio.

- $\tilde{X}$ es una *modificación* (o versión) de X si $\mathbb{P}\{\tilde{X}_t = X_t\} = 1$ para cada $t \in T$.

- X y $\tilde{X}$ son *indistinguibles* si existe $N \subset \Omega$ con $\mathbb{P}(N) = 0$ tal que $X_t(\omega) = \tilde{X}_t(\omega)$ para todo $\omega \notin N$ y todo $t \in T$.

La diferencia importa

Indistinguibilidad es estrictamente más fuerte: en una modificación, el conjunto excepcional N_t depende de t , y $\bigcup_{t \in T} N_t$ puede tener probabilidad positiva cuando T es no numerable. Ejemplo: sobre $\Omega = [0, 1]$ con la medida de Lebesgue, $X_t \equiv 0$ y $\tilde{X}_t(\omega) = \mathbf{1}_{\{t=\omega\}}$ son modificaciones una de la otra, pero *ninguna* trayectoria de $\tilde{X}$ es idénticamente nula.

Observación. Hasta aquí no tenemos siquiera medibilidad de las trayectorias $t \mapsto B_t(\omega)$: el pre-movimiento Browniano se definió variable a variable. El teorema siguiente arregla esto.

Teorema 13.10 (Criterio de continuidad de Kolmogorov–Čentsov). *Sea $X = (X_t)_{t \in I}$ un proceso indexado por un intervalo acotado $I \subset \mathbb{R}$, con valores en un espacio métrico completo (E, d) . Supongamos que existen $q, \epsilon, C > 0$ tales que*

$$\mathbb{E}[d(X_s, X_t)^q] \leq C |t - s|^{1+\epsilon} \quad \text{para todos } s, t \in I.$$

Entonces existe una modificación $\tilde{X}$ de X tal que, para todo $\alpha \in (0, \epsilon/q)$, existe una variable aleatoria $C_\alpha(\omega)$ con

$$d(\tilde{X}_s(\omega), \tilde{X}_t(\omega)) \leq C_\alpha(\omega) |t - s|^\alpha \quad \text{para todos } s, t \in I.$$

En particular $\tilde{X}$ tiene trayectorias continuas, y de hecho α -hölderianas.

Demostración. Véase Le Gall [7], Teorema 2.9, o Revuz–Yor [8], Teorema I.2.1. La idea es controlar los incrementos diádicos: por Markov, $\mathbb{P}\{d(X_{(k+1)2^{-n}}, X_{k2^{-n}}) > 2^{-\alpha n}\} \leq C 2^{-n(1+\epsilon-\alpha q)}$, y al sumar en $k \leq 2^n$ y en n se obtiene una serie convergente cuando $\alpha < \epsilon/q$; Borel–Cantelli da entonces la continuidad hölderiana sobre los diádicos, y la extensión por densidad define la modificación. ■

Corolario 13.11. *Sea $(B_t)_{t \geq 0}$ un pre-movimiento Browniano. Entonces B admite una modificación con trayectorias continuas; más aún, localmente hölderianas de exponente $\frac{1}{2} - \delta$ para todo $\delta \in (0, \frac{1}{2})$.*

Demostración. Para $s < t$, $B_t - B_s \sim \mathcal{N}(0, t - s)$ tiene la misma ley que $\sqrt{t - s} N$ con $N \sim \mathcal{N}(0, 1)$. Por lo tanto, para todo $q > 0$,

$$\mathbb{E}[|B_t - B_s|^q] = (t - s)^{q/2} \mathbb{E}[|N|^q] = C_q (t - s)^{q/2}, \quad C_q = \mathbb{E}[|N|^q] < \infty.$$

Para $q > 2$ podemos aplicar el [teorema 13.10](#) con $1 + \epsilon = q/2$, es decir $\epsilon = \frac{q}{2} - 1$. El exponente de Hölder admisible es entonces cualquier

$$\alpha < \frac{\epsilon}{q} = \frac{q-2}{2q} = \frac{1}{2} - \frac{1}{q},$$

y haciendo $q \rightarrow \infty$ se alcanza cualquier $\alpha < \frac{1}{2}$. ■

Definición 13.12 (Movimiento Browniano). Un proceso $(B_t)_{t \geq 0}$ es un *movimiento Browniano* (estándar) si

- (I) (B_t) es un pre-movimiento Browniano;
- (II) $t \mapsto B_t(\omega)$ es continua para todo $\omega \in \Omega$.

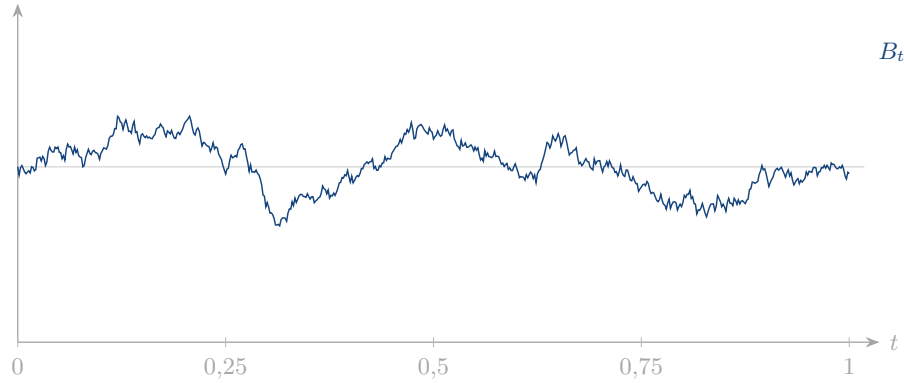


Figura 13.1: Una trayectoria del movimiento Browniano sobre $[0, 1]$ (simulación con 600 pasos). La trayectoria es continua pero, como veremos, en ningún punto diferenciable, y su variación cuadrática sobre $[0, t]$ vale exactamente t .

13.4 La medida de Wiener

Definición 13.13 (Medida de Wiener). Sea $\mathcal{C}_{\mathbb{R}} = C(\mathbb{R}_+, \mathbb{R})$ el espacio de funciones continuas de $\mathbb{R}_+$ en $\mathbb{R}$, con

$$\mathcal{C} = \sigma(w \mapsto w(t) : t \geq 0),$$

la menor σ -álgebra que hace medibles a las evaluaciones. Dado un movimiento Browniano B , la aplicación

$$\Omega \longrightarrow \mathcal{C}_{\mathbb{R}}, \quad \omega \longmapsto (t \mapsto B_t(\omega))$$

es medible, y su medida imagen $\mathbb{W} = \mathbb{P} \circ B^{-1}$ se llama *medida de Wiener*, o ley del movimiento Browniano.

Observación. La analogía con el capítulo 10 es exacta: allí el espacio de trayectorias era $\mathbb{D}_{\mathbb{Z}_+}$ y la ley $\mathbb{P}_{\text{Poiiss}}^\lambda$; aquí es $\mathcal{C}_{\mathbb{R}}$ y la ley $\mathbb{W}$. En ambos casos $X_t(w) = w(t)$ es el proceso canónico y $\mathcal{C} = \sigma(X)$.

Del corolario 13.7 se obtiene la expresión explícita: para $0 = t_0 < t_1 < \dots < t_n$ y $A_0, \dots, A_n \in \mathcal{B}_{\mathbb{R}}$,

$$\begin{aligned} & \mathbb{W}\{w : w(t_0) \in A_0, \dots, w(t_n) \in A_n\} \\ &= \mathbf{1}_{A_0}(0) \int_{A_1 \times \dots \times A_n} \frac{\exp\left\{-\sum_{i=1}^n \frac{(x_i - x_{i-1})^2}{2(t_i - t_{i-1})}\right\}}{(2\pi)^{n/2} \sqrt{t_1(t_2 - t_1) \dots (t_n - t_{n-1})}} dx_1 \dots dx_n, \quad (13.4) \end{aligned}$$

con $x_0 = 0$.

Proposición 13.14. *Los valores (13.4) caracterizan a $\mathbb{W}$. En consecuencia, la ley del movimiento Browniano es única: si B y B' son dos movimientos Brownianos (posiblemente sobre espacios distintos), entonces*

$$\mathbb{P}\{(B_t)_{t \geq 0} \in A\} = \mathbb{W}(A) = \mathbb{P}'\{(B'_t)_{t \geq 0} \in A\} \quad \text{para todo } A \in \mathcal{C}.$$

Demostración. Los conjuntos $\{w : w(t_0) \in A_0, \dots, w(t_n) \in A_n\}$ forman un π -sistema que genera $\mathcal{C}$; el teorema π - λ (teorema A.2) concluye. ■

Observación 13.15 (El proceso canónico). Tomando

$$\Omega = \mathcal{C}_{\mathbb{R}}, \quad \mathcal{F} = \mathcal{C}, \quad \mathbb{P} = \mathbb{W},$$

el proceso canónico $X_t(w) = w(t)$ es un movimiento Browniano. Igual que en el caso de Poisson, definimos para $x \in \mathbb{R}$ la traslación $T_x w = x + w$ y la familia

$$\mathbb{P}_x = \mathbb{W} \circ T_x^{-1} = \text{Ley}(x + B),$$

obteniendo $[(\mathcal{C}_{\mathbb{R}}, \mathcal{C}), (\mathcal{C}_t), (X_t), \{\mathbb{P}_x\}_{x \in \mathbb{R}}]$.

Ejercicio 13.16. Sea B un movimiento Browniano y $\mathcal{F}_t = \sigma(B_s : 0 \leq s \leq t)$. Pruebe que $(B_t, \mathcal{F}_t)$ es una martingala y que, de hecho, $B_t - B_s$ es independiente de $\mathcal{F}_{s+}$.

Solución. Para $s \leq t$, usando la independencia del incremento respecto de $\mathcal{F}_s$ (proposición 13.8(c)),

$$\mathbb{E}[B_t | \mathcal{F}_s] = \mathbb{E}[B_t - B_s | \mathcal{F}_s] + B_s = \mathbb{E}\left[\underbrace{B_t - B_s}_{\sim \mathcal{N}(0, t-s)}\right] + B_s = B_s.$$

Para la segunda afirmación, sabemos que $B_t - B_{s+\varepsilon}$ es independiente de $\mathcal{F}_{s+\varepsilon} \supset \mathcal{F}_{s+}$ para $0 < \varepsilon < t - s$. Sea $D \in \mathcal{F}_{s+}$ y $A \in \mathcal{B}_{\mathbb{R}}$ con $\mathbb{P}\{B_t - B_s \in \partial A\} = 0$; por continuidad de las trayectorias, $B_{s+\varepsilon} \rightarrow B_s$ c. s., y por convergencia dominada

$$\begin{aligned} \mathbb{P}\{B_t - B_s \in A, D\} &= \lim_{\varepsilon \downarrow 0} \mathbb{P}\{B_t - B_{s+\varepsilon} \in A, D\} \\ &= \lim_{\varepsilon \downarrow 0} \mathbb{P}\{B_t - B_{s+\varepsilon} \in A\} \mathbb{P}(D) = \mathbb{P}\{B_t - B_s \in A\} \mathbb{P}(D) \blacksquare \end{aligned}$$

Ejercicio 13.17. Pruebe que también son martingalas respecto de $(\mathcal{F}_t)$:

$$B_t^2 - t, \quad \mathcal{E}_t^\sigma := \exp\left(\sigma B_t - \frac{\sigma^2 t}{2}\right) \quad (\sigma \in \mathbb{R} \text{ fijo}).$$

Indicación: use $\mathbb{E}[e^{\sigma Z}] = e^{\sigma^2 v/2}$ para $Z \sim \mathcal{N}(0, v)$.

13.5 La propiedad de Markov y la ley 0–1 de Blumenthal

Teorema 13.18 (Markov débil). Para toda $F : \mathcal{C}_{\mathbb{R}} \rightarrow \mathbb{R}$ medible acotada (o $F \geq 0$), todo $s \geq 0$ y todo $x_0 \in \mathbb{R}$,

$$\mathbb{E}_{x_0}[F \circ \theta_s | \mathcal{C}_{s+}] \stackrel{\text{c.s.}}{=} \varphi(X_s), \quad \varphi(x) = \mathbb{E}_x[F].$$

Esquema. Idéntico al de la [teorema 11.5](#). Se toma $F = \mathbf{1}_A$ con $A = \{X_{t_1} \leq a_1, \dots, X_{t_n} \leq a_n\}$; se prueba la identidad primero para $\mathcal{C}_s$, integrando sobre conjuntos $B = \{X_{s_1} \leq b_1, \dots, X_{s_m} \leq b_m, X_s \leq b\}$ con $s_j < s$, que generan $\mathcal{C}_s$. El cálculo se reduce a manipular la densidad (13.1): poniendo

$$\widehat{B} = (-\infty, b_1 - x_0] \times \dots \times (-\infty, b_m - x_0] \times (-\infty, b - x_0],$$

y usando que bajo $\mathbb{P}_{x_0}$ el proceso es $x_0 + X$ bajo $\mathbb{P}_0$, se obtiene

$$\int_B \mathbb{P}_{X_s}\{A\} d\mathbb{P}_{x_0} = \mathbb{P}_0\{X_{s_1} \leq b_1 - x_0, \dots, X_s \leq b - x_0, X_{t_1+s} \leq a_1 - x_0, \dots\},$$

que coincide con $\mathbb{P}_{x_0}\{B, X_{t_1+s} \leq a_1, \dots\}$. La extensión de $\mathcal{C}_s$ a $\mathcal{C}_{s+}$ se hace por convergencia, usando la continuidad de las trayectorias y la continuidad de Feller del [ejercicio 13.19](#). $\blacksquare$

Ejercicio 13.19. Para toda $F : \mathcal{C}_{\mathbb{R}} \rightarrow \mathbb{R}$ medible acotada, la aplicación $x \mapsto \mathbb{E}_x[F]$ es continua.

Solución. Basta considerar $F = \mathbf{1}_A$ con $A = \{X_{t_1} \leq a_1, \dots, X_{t_n} \leq a_n\}$, pues tales A forman un π -sistema generador y el conjunto de F para los que vale la conclusión es estable por límites acotados. Entonces

$$\mathbb{E}_x[F] = \mathbb{P}_0\{X_{t_1} \leq a_1 - x, \dots, X_{t_n} \leq a_n - x\} = \int_{\prod_i (-\infty, a_i - x]} p(y_1, \dots, y_n) dy,$$

con p la densidad (13.1), que es continua y acotada; la continuidad en x se sigue por convergencia dominada. ■

Teorema 13.20 (Ley 0-1 de Blumenthal). Sea B un movimiento Browniano, $\mathcal{F}_t = \sigma(B_s : s \leq t)$ y

$$\mathcal{F}_{0+} = \bigcap_{s>0} \mathcal{F}_s.$$

Entonces, para todo $A \in \mathcal{F}_{0+}$, $\mathbb{P}(A) \in \{0, 1\}$.

Demostración. Sean $0 < t_1 < \dots < t_k$ y $g : \mathbb{R}^k \rightarrow \mathbb{R}$ continua y acotada, y sea $A \in \mathcal{F}_{0+}$. Por continuidad de las trayectorias y convergencia dominada,

$$\mathbb{E}[\mathbf{1}_A g(B_{t_1}, \dots, B_{t_k})] = \lim_{\epsilon \downarrow 0} \mathbb{E}[\mathbf{1}_A g(B_{t_1} - B_{\epsilon}, \dots, B_{t_k} - B_{\epsilon})].$$

Ahora bien, en cuanto $\epsilon < t_1$, el vector $(B_{t_1} - B_{\epsilon}, \dots, B_{t_k} - B_{\epsilon})$ es independiente de $\mathcal{F}_{\epsilon}$, y por lo tanto de $\mathcal{F}_{0+} \subset \mathcal{F}_{\epsilon}$. Luego cada término del límite se factoriza y

$$\mathbb{E}[\mathbf{1}_A g(B_{t_1}, \dots, B_{t_k})] = \mathbb{P}(A) \mathbb{E}[g(B_{t_1}, \dots, B_{t_k})].$$

Es decir, $\mathcal{F}_{0+}$ es independiente de $\sigma(B_{t_1}, \dots, B_{t_k})$ para toda elección finita $\{t_1, \dots, t_k\} \subset (0, \infty)$; por el teorema π - λ , $\mathcal{F}_{0+}$ es independiente de $\sigma(B_t : t > 0)$. Finalmente $\sigma(B_t : t > 0) = \sigma(B_t : t \geq 0)$, pues $B_0 = \lim_{t \downarrow 0} B_t$ por continuidad. Como $\mathcal{F}_{0+} \subset \sigma(B_t : t \geq 0)$, la σ -álgebra $\mathcal{F}_{0+}$ es independiente de sí misma: $\mathbb{P}(A) = \mathbb{P}(A \cap A) = \mathbb{P}(A)^2$, de donde $\mathbb{P}(A) \in \{0, 1\}$. ■

Corolario 13.21 (Oscilación inmediata y recurrencia). Casi seguramente,

$$\sup_{0 \leq s \leq \epsilon} B_s > 0 \quad y \quad \inf_{0 \leq s \leq \epsilon} B_s < 0 \quad \text{para todo } \epsilon > 0.$$

En consecuencia, para $a \in \mathbb{R}$ el tiempo de llegada

$$\tau_a = \inf\{t \geq 0 : B_t = a\} \quad (\inf \emptyset = \infty)$$

cumple $\mathbb{P}\{\tau_a < \infty\} = 1$ para todo a , y

$$\limsup_{t \rightarrow \infty} B_t = +\infty, \quad \liminf_{t \rightarrow \infty} B_t = -\infty.$$

Observación. A priori no es obvio que $\sup_{0 \leq s \leq \epsilon} B_s$ sea medible —es un supremo sobre un conjunto no numerable— pero la continuidad de las trayectorias permite restringirlo a los racionales de $[0, \epsilon]$ (observación 9.4).

Demostración. Sea $\varepsilon_p \downarrow 0$ y

$$A = \bigcap_{p \geq 1} \left\{ \sup_{0 \leq s \leq \varepsilon_p} B_s > 0 \right\} \in \mathcal{F}_{0+}.$$

Por continuidad decreciente, $\mathbb{P}(A) = \lim_p \downarrow \mathbb{P}\{\sup_{0 \leq s \leq \varepsilon_p} B_s > 0\}$, y para cada p ,

$$\mathbb{P}\left\{ \sup_{0 \leq s \leq \varepsilon_p} B_s > 0 \right\} \geq \mathbb{P}\{B_{\varepsilon_p} > 0\} = \frac{1}{2}.$$

Luego $\mathbb{P}(A) \geq \frac{1}{2}$ y, por Blumenthal, $\mathbb{P}(A) = 1$. El caso del ínfimo se obtiene aplicando lo anterior a $-B$, que también es un movimiento Browniano.

Para la segunda afirmación, por lo anterior $\mathbb{P}\{\sup_{s \geq 0} B_s > 0\} = 1$ y, por continuidad monótona,

$$1 = \mathbb{P}\left\{ \sup_{0 \leq s \leq 1} B_s > 0 \right\} = \lim_{\delta \downarrow 0} \uparrow \mathbb{P}\left\{ \sup_{0 \leq s \leq 1} B_s > \delta \right\}.$$

Usando la invariancia de escala (proposición 13.8(b)) con el Browniano $B_u^\delta = \delta^{-1} B_{\delta^2 u}$,

$$\mathbb{P}\left\{ \sup_{0 \leq s \leq 1} B_s > \delta \right\} = \mathbb{P}\left\{ \sup_{0 \leq u \leq 1/\delta^2} B_u^\delta > 1 \right\} = \mathbb{P}\left\{ \sup_{0 \leq u \leq 1/\delta^2} B_u > 1 \right\},$$

donde la última igualdad usa que todo movimiento Browniano induce la misma medida de Wiener (proposición 13.14). Haciendo $\delta \downarrow 0$,

$$\mathbb{P}\left\{ \sup_{u \geq 0} B_u > 1 \right\} = 1,$$

y un nuevo cambio de escala da $\mathbb{P}\{\sup_{u \geq 0} B_u > M\} = 1$ para todo $M > 0$; pasando a $-B$, $\mathbb{P}\{\inf_{u \geq 0} B_u < -M\} = 1$. Por continuidad de las trayectorias y el teorema del valor intermedio, B alcanza todo nivel a , es decir $\tau_a < \infty$ c. s. ■

13.6 La propiedad de Markov fuerte y el principio de reflexión

Definición 13.22. Un tiempo de parada es $\tau : \Omega \rightarrow [0, \infty]$ con $\{\tau \leq t\} \in \mathcal{F}_t$ para todo $t \geq 0$, y

$$\mathcal{F}_\tau = \{A \in \mathcal{F}_\infty : A \cap \{\tau \leq t\} \in \mathcal{F}_t \quad \forall t \geq 0\}.$$

Ejemplo 13.23. Son tiempos de parada: $\tau \equiv t$ constante; $\tau + \theta$ con $\theta \geq 0$ constante; y $\tau_a = \inf\{t \geq 0 : B_t = a\}$, pues por continuidad

$$\{\tau_a \leq t\} = \left\{ \inf_{s \in \mathbb{Q} \cap [0, t]} |B_s - a| = 0 \right\} \in \mathcal{F}_t.$$

Observación 13.24 (B_τ es $\mathcal{F}_\tau$ -medible). Definiendo

$$\mathbf{1}_{\{\tau < \infty\}} B_\tau(\omega) = \begin{cases} B_{\tau(\omega)}(\omega), & \tau(\omega) < \infty, \\ 0, & \tau(\omega) = \infty, \end{cases}$$

la variable resultante es $\mathcal{F}_\tau$ -medible. En efecto, por continuidad por la derecha,

$$\mathbf{1}_{\{\tau < \infty\}} B_\tau = \lim_{n \rightarrow \infty} \sum_{i=0}^{\infty} \mathbf{1}_{\{i2^{-n} \leq \tau\}} \mathbf{1}_{\{\tau < (i+1)2^{-n}\}} B_{i2^{-n}},$$

y basta ver que cada $B_s \mathbf{1}_{\{s \leq \tau\}}$ es $\mathcal{F}_\tau$ -medible. Para $A \in \mathcal{B}_\mathbb{R}$ con $0 \notin A$,

$$\{B_s \mathbf{1}_{\{s \leq \tau\}} \in A\} \cap \{\tau \leq t\} = \begin{cases} \emptyset, & \text{si } t < s, \\ \underbrace{\{B_s \in A\} \cap \{s \leq \tau \leq t\}}_{\in \mathcal{F}_s \subset \mathcal{F}_t}, & \text{si } t \geq s, \end{cases}$$

y $\{s \leq \tau \leq t\} = \{\tau \leq t\} \cap \{\tau < s\}^c \in \mathcal{F}_t$.

Teorema 13.25 (Markov fuerte). Sea τ un tiempo de parada con $\mathbb{P}\{\tau < \infty\} > 0$, y definamos para $t \geq 0$

$$B_t^\tau = \mathbf{1}_{\{\tau < \infty\}} (B_{\tau+t} - B_\tau).$$

Entonces, bajo la probabilidad condicional $\mathbb{P}(\cdot \mid \tau < \infty)$, el proceso $(B_t^\tau)_{t \geq 0}$ es un movimiento Browniano independiente de $\mathcal{F}_\tau$.

Demostración. Supongamos primero $\tau < \infty$ c.s.. Mostraremos que, si $A \in \mathcal{F}_\tau$, $0 \leq t_1 < \dots < t_p$ y $F : \mathbb{R}^p \rightarrow \mathbb{R}_+$ es continua y acotada, entonces

$$\mathbb{E}[\mathbf{1}_A F(B_{t_1}^\tau, \dots, B_{t_p}^\tau)] = \mathbb{P}(A) \mathbb{E}[F(B_{t_1}, \dots, B_{t_p})]. \quad (13.5)$$

Esto prueba a la vez la independencia (el factor $\mathbb{P}(A)$ se separa) y que B^τ es un movimiento Browniano (tiene las mismas distribuciones finito-dimensionales que B , y trayectorias continuas).

Para $n \geq 1$ denotemos por $[\tau]_n$ el menor número de la forma $k2^{-n}$ mayor o igual que τ (con $[\tau]_n = \infty$ si $\tau = \infty$). Por continuidad de las trayectorias,

$$F(B_{t_1}^\tau, \dots, B_{t_p}^\tau) = \lim_{n \rightarrow \infty} F(B_{t_1}^{[\tau]_n}, \dots, B_{t_p}^{[\tau]_n}),$$

de modo que, por convergencia dominada,

$$\begin{aligned} & \mathbb{E} \left[\mathbf{1}_A F(B_{t_1}^\tau, \dots, B_{t_p}^\tau) \right] \\ &= \lim_n \sum_{k=0}^{\infty} \mathbb{E} \left[\mathbf{1}_{A \cap \{(k-1)2^{-n} < \tau \leq k2^{-n}\}} F(B_{k2^{-n}+t_1} - B_{k2^{-n}}, \dots, B_{k2^{-n}+t_p} - B_{k2^{-n}}) \right], \end{aligned}$$

donde simplemente se desagregó según el valor de $[\tau]_n$. Ahora, para $A \in \mathcal{F}_\tau$,

$$A \cap \{(k-1)2^{-n} < \tau \leq k2^{-n}\} = (A \cap \{\tau \leq k2^{-n}\}) \cap \{\tau \leq (k-1)2^{-n}\}^c \in \mathcal{F}_{k2^{-n}},$$

y por la propiedad de Markov simple ([proposición 13.8\(c\)](#)) el vector de incrementos posteriores a $k2^{-n}$ es independiente de $\mathcal{F}_{k2^{-n}}$, con la ley de $(B_{t_1}, \dots, B_{t_p})$. Por lo tanto cada sumando vale

$$\mathbb{P}[A \cap \{(k-1)2^{-n} < \tau \leq k2^{-n}\}] \cdot \mathbb{E}[F(B_{t_1}, \dots, B_{t_p})],$$

y sumando en k se obtiene (13.5).

Si $\mathbb{P}\{\tau = \infty\} > 0$, el mismo argumento con $A \cap \{\tau < \infty\}$ da

$$\mathbb{E} \left[\mathbf{1}_{A \cap \{\tau < \infty\}} F(B_{t_1}^\tau, \dots, B_{t_p}^\tau) \right] = \mathbb{P}(A \cap \{\tau < \infty\}) \mathbb{E}[F(B_{t_1}, \dots, B_{t_p})],$$

que es el enunciado bajo $\mathbb{P}(\cdot \mid \tau < \infty)$. ■

Teorema 13.26 (Principio de reflexión). Sea $S_t = \sup_{0 \leq s \leq t} B_s$. Para $a \geq 0$ y $b \leq a$,

$$\mathbb{P}\{S_t \geq a, B_t \leq b\} = \mathbb{P}\{B_t \geq 2a - b\}. \quad (13.6)$$

En particular, S_t tiene la misma ley que $|B_t|$, y

$$\mathbb{P}\{S_t \geq a\} = 2\mathbb{P}\{B_t \geq a\} = \frac{2}{\sqrt{2\pi t}} \int_a^\infty e^{-x^2/(2t)} dx.$$

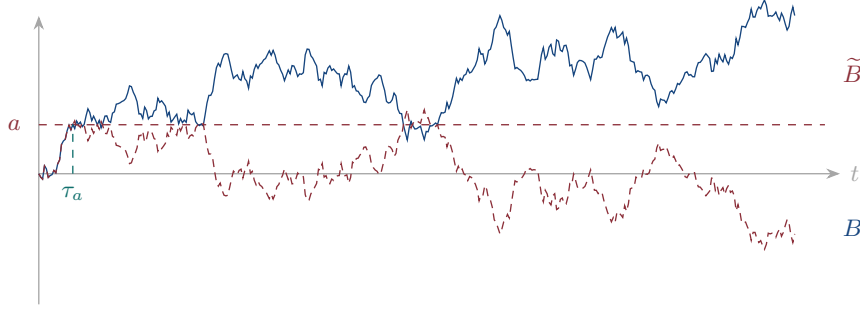


Figura 13.2: El principio de reflexión. A partir de τ_a se refleja la trayectoria respecto del nivel a : por la propiedad de Markov fuerte, la trayectoria reflejada $\tilde{B}$ tiene la misma ley que B . Las trayectorias con $S_t \geq a$ y $B_t \leq b$ se corresponden biyectivamente con las que terminan por encima de $2a - b$.

Demostración. Consideremos $\tau_a = \inf\{t \geq 0 : B_t = a\}$, que por el [corolario 13.21](#) cumple $\tau_a < \infty$ c.s.. Como B es continuo, $\{S_t \geq a\} = \{\tau_a \leq t\}$, de modo que

$$\mathbb{P}\{S_t \geq a, B_t \leq b\} = \mathbb{P}\{\tau_a \leq t, B_t \leq b\} = \mathbb{P}\{\tau_a \leq t, B_{t-\tau_a}^{\tau_a} \leq b - a\},$$

usando $B_{t-\tau_a}^{\tau_a} = B_t - B_{\tau_a} = B_t - a$.

Por Markov fuerte ([teorema 13.25](#)), B^{τ_a} es un movimiento Browniano independiente de $\mathcal{F}_{\tau_a}$; como $-B^{\tau_a}$ tiene la misma ley que B^{τ_a} y τ_a es $\mathcal{F}_{\tau_a}$ -medible, los pares (τ_a, B^{τ_a}) y $(\tau_a, -B^{\tau_a})$ tienen la misma ley. Aplicando esto al conjunto medible

$$H = \{(s, w) \in \mathbb{R}_+ \times \mathcal{C}_{\mathbb{R}} : s \leq t, w(t-s) \leq b - a\},$$

obtenemos

$$\begin{aligned} \mathbb{P}\{(\tau_a, B^{\tau_a}) \in H\} &= \mathbb{P}\{(\tau_a, -B^{\tau_a}) \in H\} = \mathbb{P}\{\tau_a \leq t, -B_{t-\tau_a}^{\tau_a} \leq b - a\} \\ &= \mathbb{P}\{\tau_a \leq t, B_t \geq 2a - b\} = \mathbb{P}\{B_t \geq 2a - b\}, \end{aligned}$$

donde la última igualdad usa que $\{B_t \geq 2a - b\} \subset \{\tau_a \leq t\}$ c.s., pues $2a - b \geq a$.

Finalmente, tomando $b = a$ en (13.6),

$$\begin{aligned} \mathbb{P}\{S_t \geq a\} &= \mathbb{P}\{S_t \geq a, B_t \geq a\} + \mathbb{P}\{S_t \geq a, B_t \leq a\} \\ &= \mathbb{P}\{B_t \geq a\} + \mathbb{P}\{B_t \geq a\} = \mathbb{P}\{|B_t| \geq a\}, \end{aligned}$$

usando $\{B_t \geq a\} \subset \{S_t \geq a\}$ y la simetría de B_t . ■

13.7 Variación cuadrática y no diferenciabilidad

Proposición 13.27 (Variación cuadrática). *Sea $t > 0$ y $0 = t_0^n < t_1^n < \dots < t_{p_n}^n = t$ una sucesión de particiones de $[0, t]$ con paso que tiende a cero, es decir $\max_{1 \leq i \leq p_n} (t_i^n - t_{i-1}^n) \rightarrow 0$. Entonces*

$$\sum_{i=1}^{p_n} (B_{t_i^n} - B_{t_{i-1}^n})^2 \xrightarrow{n \rightarrow \infty} t \quad \text{en } L^2(\Omega).$$

Demostración. Aplicar la [proposición 12.24](#) con $E = \mathbb{R}_+$, $\mu = m$, $A = [0, t]$ y las particiones $A_i^n = (t_{i-1}^n, t_i^n]$, teniendo en cuenta que

$$B_{t_i^n} - B_{t_{i-1}^n} = G(\mathbf{1}_{(t_{i-1}^n, t_i^n]}) \quad \text{y} \quad \max_i m(A_i^n) = \max_i (t_i^n - t_{i-1}^n) \rightarrow 0. \quad \blacksquare$$

Corolario 13.28 (Variación infinita). *Casi seguramente, las trayectorias del movimiento Browniano no tienen variación acotada sobre ningún intervalo $[0, t]$ con $t > 0$.*

Demostración. Supongamos que sobre un conjunto de probabilidad positiva la variación total $V_t(\omega) = \sup \sum_i |B_{t_i} - B_{t_{i-1}}|$ fuese finita. Entonces, sobre ese conjunto,

$$\sum_i (B_{t_i^n} - B_{t_{i-1}^n})^2 \leq \left(\max_i |B_{t_i^n} - B_{t_{i-1}^n}| \right) \cdot V_t(\omega) \xrightarrow{n \rightarrow \infty} 0,$$

pues el máximo tiende a 0 por continuidad uniforme de B sobre $[0, t]$. Esto contradice la [proposición 13.27](#) (pasando a una subsucesión que converja c. s.). $\blacksquare$

El corazón del cálculo de Itô

La variación cuadrática del Browniano sobre $[0, t]$ vale t , no 0. Para una función diferenciable f la variación cuadrática es nula, porque $\sum (f(t_i) - f(t_{i-1}))^2 \approx \sum f'(\xi_i)^2 (t_i - t_{i-1})^2 \rightarrow 0$. La diferencia se resume en la regla mnemotécnica

$$(dB_t)^2 = dt,$$

que es precisamente el término extra de la fórmula de Itô ([capítulo 16](#)) y la razón de que el cálculo estocástico no sea el cálculo ordinario.

13.8 Ejercicios resueltos

Ejercicio 13.29 (Inversión temporal). Sea B un movimiento Browniano y definamos $W_0 = 0$, $W_t = t B_{1/t}$ para $t > 0$. Pruebe que $(W_t)_{t \geq 0}$ es indistinguible de un movimiento Browniano. Deduzca que $\lim_{t \rightarrow \infty} B_t/t = 0$ c. s..

Solución. W es pre-Browniano. Como B es Gaussiano centrado, toda combinación lineal finita de los $W_t = t B_{1/t}$ lo es. Para $s, t > 0$,

$$\mathbb{E}[W_s W_t] = st \mathbb{E}[B_{1/s} B_{1/t}] = st \left(\frac{1}{s} \wedge \frac{1}{t} \right) = s \wedge t,$$

y $\mathbb{E}[W_t W_0] = 0$. Por la [proposición 13.6\(ii\)](#), W es un pre-movimiento Browniano.

Continuidad en 0. W es continuo sobre $(0, \infty)$ por serlo B ; falta ver que $W_t \rightarrow 0$ cuando $t \downarrow 0$, lo que equivale a

$$\lim_{u \rightarrow \infty} \frac{B_u}{u} = 0 \quad \text{c. s.}$$

Para probarlo veamos que $\mathbb{P}\{\limsup_u |B_u|/u > \epsilon\} = 0$ para todo $\epsilon > 0$. Primero, una cota exponencial: como e^{aB_s} es una submartingala positiva (por Jensen, con $\varphi(x) = e^{ax}$ convexa), la desigualdad maximal de Doob ([teorema 5.12](#), versión continua) da para $a > 0$

$$\mathbb{P}\left\{\sup_{s \leq t} B_s \geq \lambda\right\} = \mathbb{P}\left\{\sup_{s \leq t} e^{aB_s} \geq e^{a\lambda}\right\} \leq \frac{\mathbb{E}[e^{aB_t}]}{e^{a\lambda}} = e^{a^2 t/2 - a\lambda},$$

y optimizando con $a = \lambda/t$,

$$\mathbb{P}\left\{\sup_{s \leq t} |B_s| \geq \lambda\right\} \leq 2e^{-\lambda^2/(2t)}. \quad (13.7)$$

Ahora, si $\sup_{n \leq s \leq n+1} |B_s|/s > \epsilon$, entonces existe $s \in [n, n+1]$ con $|B_s| > \epsilon s \geq \epsilon n \geq \epsilon(n+1)/2$ para $n \geq 1$, de modo que $\sup_{s \leq n+1} |B_s| \geq \epsilon(n+1)/2$. Por (13.7),

$$\begin{aligned} \mathbb{P}\left\{\sup_{t \geq N} \frac{|B_t|}{t} > \epsilon\right\} &\leq \sum_{n=N}^{\infty} \mathbb{P}\left\{\sup_{n \leq t \leq n+1} \frac{|B_t|}{t} > \epsilon\right\} \leq \sum_{n=N}^{\infty} \mathbb{P}\left\{\sup_{s \leq n+1} |B_s| \geq \frac{\epsilon(n+1)}{2}\right\} \\ &\leq 2 \sum_{n=N+1}^{\infty} e^{-\epsilon^2 n/8} \xrightarrow[N \rightarrow \infty]{} 0. \end{aligned}$$

Por lo tanto $\mathbb{P}\{\limsup_t |B_t|/t > \epsilon\} = \lim_N \mathbb{P}\{\sup_{t \geq N} |B_t|/t > \epsilon\} = 0$.

Se sigue que W tiene trayectorias continuas c.s., luego es indistinguible de un movimiento Browniano. ■

Ejercicio 13.30 (Los tiempos de llegada forman un subordinator). Para $a \geq 0$ sea $\tau_a = \inf\{t \geq 0 : B_t = a\}$. Pruebe que $(\tau_a)_{a \geq 0}$ tiene incrementos independientes y estacionarios: para $0 \leq a \leq b$, la variable $\tau_b - \tau_a$ es independiente de $\sigma(\tau_c : 0 \leq c \leq a)$ y tiene la misma ley que τ_{b-a} .

Solución. Estacionariedad. Definamos $\tilde{B}_t = \mathbf{1}_{\{\tau_a < \infty\}}(B_{\tau_a+t} - B_{\tau_a})$. Como $\tau_a < \infty$ c.s., el teorema 13.25 dice que $\tilde{B}$ es un movimiento Browniano independiente de $\mathcal{F}_{\tau_a}$. Sea $\tilde{\tau}_c = \inf\{t \geq 0 : \tilde{B}_t = c\}$; por igualdad de leyes, $\tilde{\tau}_{b-a}$ tiene la misma distribución que τ_{b-a} . Y, casi seguramente,

$$\begin{aligned} \tilde{\tau}_{b-a} &= \inf\{t \geq 0 : B_{\tau_a+t} - a = b - a\} = \inf\{s \geq \tau_a : B_s = b\} - \tau_a \\ &= \inf\{s \geq 0 : B_s = b\} - \tau_a = \tau_b - \tau_a, \end{aligned}$$

donde la penúltima igualdad usa que, por continuidad y $a \leq b$, el Browniano debe pasar por a antes de alcanzar b , es decir $\tau_a \leq \tau_b$.

Independencia. Basta ver que $\sigma(\tau_c : 0 \leq c \leq a) \subset \mathcal{F}_{\tau_a}$, pues $\tilde{B}$ —y por tanto $\tilde{\tau}_{b-a} = \tau_b - \tau_a$ — es independiente de $\mathcal{F}_{\tau_a}$. Ahora bien, para $c \leq a$ se tiene $\tau_c \leq \tau_a$ y, por el ejercicio 5.3(c), $\mathcal{F}_{\tau_c} \subset \mathcal{F}_{\tau_a}$; como τ_c es $\mathcal{F}_{\tau_c}$ -medible, los conjuntos $\{\tau_{c_1} \leq t_1, \dots, \tau_{c_n} \leq t_n\}$ con $c_i \leq a$ están en $\mathcal{F}_{\tau_a}$, y forman un π -sistema que genera $\sigma(\tau_c : 0 \leq c \leq a)$. ■

Ejercicio 13.31 (Los máximos locales son distintos). Pruebe que, casi seguramente, para todos los racionales $p < q < r < s$,

$$\sup_{p \leq t \leq q} B_t \neq \sup_{r \leq t \leq s} B_t.$$

En consecuencia, los máximos locales del movimiento Browniano son todos distintos.

Solución. Fijemos $p < q < r < s$ racionales y definamos

$$U = \sup_{p \leq t \leq q} B_t - B_r, \quad V = \sup_{r \leq t \leq s} B_t - B_r.$$

Escribiendo $B_t - B_r = (B_t - B_q) - (B_r - B_q)$ para $t \in [p, q]$, se ve que U es medible respecto de $\sigma(B_u - B_v : u, v \leq r)$, mientras que

$$V = \sup_{r \leq t \leq s} (B_t - B_r)$$

lo es respecto de $\sigma(B_u - B_r : u \geq r)$. Por la independencia de los incrementos del movimiento Browniano, estas dos σ -álgebras son independientes, y por lo tanto U y V también lo son. Además V tiene una ley sin átomos: por el [teorema 13.26](#) aplicado al movimiento Browniano $(B_{r+u} - B_r)_{u \geq 0}$ se tiene $V \stackrel{d}{=} |B_{s-r}|$, que es absolutamente continua. Por lo tanto

$$\mathbb{P}\left\{\sup_{p \leq t \leq q} B_t = \sup_{r \leq t \leq s} B_t\right\} = \mathbb{P}\{U = V\} = \int_{\mathbb{R}} \underbrace{\mathbb{P}\{V = u\}}_{=0} \mathbb{P}_U(du) = 0,$$

usando la independencia y el teorema de Fubini. Como hay una cantidad numerable de cuádruplas racionales, la unión de los sucesos excepcionales sigue teniendo probabilidad 0. ■

Ejercicio 13.32 (No diferenciabilidad). Pruebe que

$$\limsup_{t \downarrow 0} \frac{B_t}{\sqrt{t}} = +\infty, \quad \liminf_{t \downarrow 0} \frac{B_t}{\sqrt{t}} = -\infty \quad \text{c. s.}$$

y deduzca que, para todo $s \geq 0$ fijo, la función $t \mapsto B_t$ no es diferenciable por la derecha en s .

Solución. Fijemos $M > 0$. Como

$$\limsup_{t \downarrow 0} \frac{B_t}{\sqrt{t}} = \lim_{c \downarrow 0} \downarrow \sup_{0 < t \leq c} \frac{B_t}{\sqrt{t}}$$

es medible respecto de $\mathcal{F}_{0+}$, el suceso $\{\limsup_{t \downarrow 0} B_t/\sqrt{t} \geq M\}$ está en $\mathcal{F}_{0+}$. Por Fatou inverso (las indicadoras están acotadas),

$$\begin{aligned} \mathbb{P}\left\{\limsup_{t \downarrow 0} \frac{B_t}{\sqrt{t}} \geq M\right\} &\geq \mathbb{P}\left\{\limsup_{n \rightarrow \infty} \frac{B_{1/n}}{\sqrt{1/n}} \geq M\right\} \geq \limsup_n \mathbb{P}\left\{\frac{B_{1/n}}{\sqrt{1/n}} \geq M\right\} \\ &= \int_M^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx > 0, \end{aligned}$$

pues $B_{1/n}/\sqrt{1/n} \sim \mathcal{N}(0, 1)$ para todo n . Como $\mathcal{F}_{0+}$ es degenerada ([teorema 13.20](#)), esa probabilidad vale 1. Siendo M arbitrario,

$$\mathbb{P}\left\{\limsup_{t \downarrow 0} \frac{B_t}{\sqrt{t}} = +\infty\right\} = 1,$$

y pasando a $-B$ se obtiene la afirmación sobre el $\liminf$.

Para la no diferenciabilidad en s , aplicamos lo anterior al movimiento Browniano $(B_{s+u} - B_s)_{u \geq 0}$:

$$\limsup_{t \downarrow s} \frac{B_t - B_s}{t - s} = \limsup_{t \downarrow s} \frac{B_t - B_s}{\sqrt{t - s}} \cdot \frac{1}{\sqrt{t - s}} = +\infty \quad \text{c. s.,}$$

pues el primer factor tiene $\limsup = +\infty$ y el segundo tiende a $+\infty$. Análogamente el $\liminf$ vale $-\infty$, de modo que la derivada por la derecha no existe. ■

Un matiz importante

El ejercicio anterior fija s *de antemano*: para cada s , casi toda trayectoria no es diferenciable en s . El enunciado mucho más fuerte —casi toda trayectoria no es diferenciable en *ningún* punto— es el teorema de Paley–Wiener–Zygmund, y requiere un argumento distinto, precisamente porque el conjunto de instantes es no numerable (compárese con la [definición 13.9](#)).

Ejercicio 13.33 (Ceros del Browniano). Sea $H = \{t \in [0, 1] : B_t = 0\}$. Pruebe que, casi seguramente, H es un compacto de $[0, 1]$, de medida de Lebesgue nula y sin puntos aislados.

Solución. Compacto. $H = B^{-1}(\{0\}) \cap [0, 1]$ es cerrado por continuidad de B , y acotado.

Medida nula. Por Tonelli,

$$\mathbb{E}[m(H)] = \int_{\Omega} \int_0^1 \mathbf{1}_{\{B_t(\omega)=0\}} dt \mathbb{P}(d\omega) = \int_0^1 \mathbb{P}\{B_t = 0\} dt = 0,$$

pues $B_t \sim \mathcal{N}(0, t)$ no tiene átomos para $t > 0$. Luego $m(H) = 0$ c. s..

Sin puntos aislados. Para $q \in [0, 1) \cap \mathbb{Q}$ definamos

$$\varrho_q = \inf\{t \geq q : B_t = 0\},$$

un tiempo de parada, finito c. s. y ≤ 1 con probabilidad positiva. Por la propiedad de Markov fuerte ([teorema 13.25](#)) aplicada a ϱ_q y por el [corolario 13.21](#) aplicado al Browniano B^{ϱ_q} ,

$$\mathbb{P}\left\{\sup_{0 \leq u \leq \epsilon} B_{\varrho_q+u} > 0 \text{ y } \inf_{0 \leq u \leq \epsilon} B_{\varrho_q+u} < 0, \forall \epsilon \in (0, 1 - q) \cap \mathbb{Q}\right\} = 1.$$

Por el teorema del valor intermedio, en cada intervalo $(\varrho_q, \varrho_q + \epsilon)$ hay un cero de B : ningún ϱ_q es punto aislado de H . Sea

$$E = \bigcap_{q \in [0,1) \cap \mathbb{Q}} \bigcap_{\epsilon \in (0,1-q) \cap \mathbb{Q}} \{\exists t \in (\varrho_q, \varrho_q + \epsilon) : B_t = 0\},$$

que tiene probabilidad 1. Fijemos $\omega \in E$ y sea $t \in H \setminus \{\varrho_q : q \in \mathbb{Q} \cap [0,1)\}$; elijamos racionales $q_n \uparrow t$. Como $q_n < t$ y $B_t = 0$, se tiene $q_n \leq \varrho_{q_n} \leq t$ y por tanto $\varrho_{q_n} \uparrow t$ con $\varrho_{q_n} \in H$: tampoco t es aislado. ■

Ejercicio 13.34 (Reversión temporal en $[0,1]$). Sea $B'_t = B_1 - B_{1-t}$ para $t \in [0,1]$. Pruebe que $(B_t)_{t \in [0,1]}$ y $(B'_t)_{t \in [0,1]}$ tienen la misma ley.

Solución. Ambos son procesos Gaussianos centrados y continuos, de modo que basta comparar covarianzas. Para $s, t \in [0,1]$,

$$\begin{aligned} \mathbb{E}[B'_s B'_t] &= \mathbb{E}[(B_1 - B_{1-s})(B_1 - B_{1-t})] \\ &= \mathbb{E}[B_1^2] - \mathbb{E}[B_1 B_{1-t}] - \mathbb{E}[B_1 B_{1-s}] + \mathbb{E}[B_{1-s} B_{1-t}] \\ &= 1 - (1-t) - (1-s) + ((1-s) \wedge (1-t)) \\ &= s + t - 1 + 1 - (s \vee t) = s \wedge t, \end{aligned}$$

que es la covarianza de B . Como ambos son procesos Gaussianos centrados con la misma covarianza, tienen las mismas distribuciones finito-dimensionales y por lo tanto la misma ley sobre $C([0,1], \mathbb{R})$ (proposición 13.14). ■

Parte IV

Cálculo estocástico

Martingalas a tiempo continuo y variación cuadrática

Sobre esta parte

Los cuatro capítulos que siguen desarrollan el cálculo estocástico: la integral de Itô, su fórmula de cambio de variable, las ecuaciones diferenciales estocásticas y el teorema de Girsanov. El destino es la fórmula de Black–Scholes, que es la Parte II llevada al límite continuo.

El punto de partida es el hecho ya establecido en el [corolario 13.28](#): las trayectorias del movimiento Browniano *no tienen variación acotada*. En consecuencia, $\int_0^t H_s dB_s$ no puede definirse trayectoria a trayectoria con la integral de Lebesgue–Stieltjes. Hay que construirla de otra manera, y el precio a pagar es que el cálculo resultante *no* obedece la regla de la cadena ordinaria.

14.1 Martingalas a tiempo continuo

Fijemos un espacio de probabilidad filtrado $((\Omega, \mathcal{F}, \mathbb{P}), (\mathcal{F}_t)_{t \geq 0})$. Supondremos en adelante que la filtración satisface las *condiciones habituales*: es continua por la derecha ($\mathcal{F}_t = \mathcal{F}_{t+}$) y $\mathcal{F}_0$ contiene todos los conjuntos $\mathbb{P}$ -nulos de $\mathcal{F}$.

Definición 14.1 (Martingala a tiempo continuo). Un proceso $(M_t)_{t \geq 0}$ es una *martingala* si

- (I) (M_t) es adaptado;
- (II) $\mathbb{E}[|M_t|] < \infty$ para todo $t \geq 0$;
- (III) $\mathbb{E}[M_t | \mathcal{F}_s] = M_s$ c. s. para todos $0 \leq s \leq t$.

Sub- y supermartingalas se definen reemplazando = por $\geq$ y $\leq$.

Ejemplo 14.2. Con B un movimiento Browniano y N un proceso de Poisson de intensidad λ , respecto de sus filtraciones naturales:

$$B_t, \quad B_t^2 - t, \quad \exp\left(\sigma B_t - \frac{\sigma^2 t}{2}\right), \quad N_t - \lambda t, \quad (N_t - \lambda t)^2 - \lambda t$$

son todas martingalas (ejercicio 13.16, ejercicio 13.17, ejercicio 10.11).

La teoría discreta se transfiere por aproximación a lo largo de mallas diádicas. Enunciamos los resultados que usaremos.

Teorema 14.3 (Regularización). *Sea $(M_t)_{t \geq 0}$ una martingala respecto de una filtración que satisface las condiciones habituales. Entonces M admite una modificación con trayectorias càdlàg. En particular, toda martingala puede suponerse càdlàg, y toda martingala continua en probabilidad admite una modificación continua.*

Teorema 14.4 (Desigualdades de Doob, versión continua). *Sea $(M_t)_{t \geq 0}$ una martingala càdlàg y $T > 0$.*

(a) *Para todo $\lambda > 0$,*

$$\mathbb{P}\left\{\sup_{0 \leq t \leq T} |M_t| \geq \lambda\right\} \leq \frac{\mathbb{E}[|M_T|]}{\lambda}.$$

(b) *Para todo $p > 1$,*

$$\left\|\sup_{0 \leq t \leq T} |M_t|\right\|_p \leq \frac{p}{p-1} \|M_T\|_p.$$

$$\text{En particular, para } p = 2, \mathbb{E}\left[\sup_{t \leq T} M_t^2\right] \leq 4 \mathbb{E}[M_T^2].$$

Demostración. Sea $D_n = \{kT2^{-n} : 0 \leq k \leq 2^n\}$ la malla diádica de orden n y $D = \bigcup_n D_n$, denso en $[0, T]$. Sobre D_n el proceso es una martingala discreta, de modo que los teoremas 5.13 y 5.16 dan

$$\mathbb{P}\left\{\max_{t \in D_n} |M_t| \geq \lambda\right\} \leq \frac{\mathbb{E}[|M_T|]}{\lambda}, \quad \left\|\max_{t \in D_n} |M_t|\right\|_p \leq \frac{p}{p-1} \|M_T\|_p.$$

Como $\max_{t \in D_n} |M_t| \uparrow \sup_{t \in D} |M_t|$, la convergencia monótona da las desigualdades con el supremo sobre D ; y por continuidad por la derecha de las trayectorias, $\sup_{t \in D} |M_t| = \sup_{0 \leq t \leq T} |M_t|$. ■

Teorema 14.5 (Muestreo opcional). Sea (M_t) una martingala càdlàg y $\tau \leq \rho$ tiempos de parada acotados (digamos $\rho \leq T$). Entonces $M_\tau, M_\rho \in L^1$ y

$$\mathbb{E}[M_\rho | \mathcal{F}_\tau] = M_\tau \quad \text{c. s.}$$

Demostración. Sea $\tau_n = \lceil 2^n \tau \rceil 2^{-n} \wedge T$ y ρ_n análogamente: son tiempos de parada con valores en la malla diádica, $\tau_n \downarrow \tau$, $\rho_n \downarrow \rho$ y $\tau_n \leq \rho_n$. Sobre la malla, el teorema 5.10 da $\mathbb{E}[M_{\rho_n} | \mathcal{F}_{\tau_n}] = M_{\tau_n}$, es decir $\int_A M_{\rho_n} = \int_A M_{\tau_n}$ para $A \in \mathcal{F}_{\tau_n}$. Como $\mathcal{F}_\tau \subset \mathcal{F}_{\tau_n}$, la identidad vale para $A \in \mathcal{F}_\tau$. Las familias (M_{τ_n}) y (M_{ρ_n}) son uniformemente integrables —son esperanzas condicionales de M_T , por proposición 4.6— y convergen c. s. a M_τ y M_ρ por continuidad por la derecha; el teorema 4.7 permite pasar al límite. ■

Definición 14.6 (Martingala local). Un proceso adaptado y continuo $(M_t)_{t \geq 0}$ con $M_0 = 0$ es una *martingala local* si existe una sucesión creciente de tiempos de parada $\tau_n \uparrow \infty$ c. s. tal que, para cada n , el proceso detenido $M^{\tau_n} = (M_{t \wedge \tau_n})_{t \geq 0}$ es una martingala uniformemente integrable. A tal sucesión se le llama *sucesión localizante*.

Para qué sirven las martingalas locales

Son la clase natural en la que vive la integral de Itô. La razón es que para que $\int_0^t H_s dB_s$ sea una verdadera martingala hace falta $\mathbb{E}\left[\int_0^t H_s^2 ds\right] < \infty$, una condición de integrabilidad que muchas veces no se puede verificar *a priori*; en cambio siempre se puede localizar con $\tau_n = \inf\{t : \int_0^t H_s^2 ds \geq n\}$. Una martingala local es una martingala “salvo integrabilidad”.

Proposición 14.7. Una martingala local acotada es una martingala. Más generalmente, si M es una martingala local con $\mathbb{E}\left[\sup_{s \leq t} |M_s|\right] < \infty$ para todo t , entonces M es una martingala.

Demostración. Sea (τ_n) localizante. Para $s \leq t$ y $A \in \mathcal{F}_s$, $\int_A M_{t \wedge \tau_n} = \int_A M_{s \wedge \tau_n}$. Como $|M_{t \wedge \tau_n}| \leq \sup_{u \leq t} |M_u| \in L^1$ y $M_{t \wedge \tau_n} \rightarrow M_t$ c. s., el teorema de convergencia dominada permite pasar al límite. ■

Proposición 14.8 (Una martingala local positiva es supermartingala). Si M es una martingala local con $M_t \geq 0$ para todo t y $\mathbb{E}[M_0] < \infty$, entonces M es una supermartingala.

Demostración. Por Fatou condicional, para $s \leq t$,

$$\mathbb{E}[M_t | \mathcal{F}_s] = \mathbb{E}[\liminf_n M_{t \wedge \tau_n} | \mathcal{F}_s] \leq \liminf_n \mathbb{E}[M_{t \wedge \tau_n} | \mathcal{F}_s] = \liminf_n M_{s \wedge \tau_n} = M_s. \quad \blacksquare$$

14.2 Variación cuadrática

Teorema 14.9 (Existencia de la variación cuadrática). *Sea $(M_t)_{t \geq 0}$ una martingala local continua con $M_0 = 0$. Existe un único proceso $(\langle M \rangle_t)_{t \geq 0}$, adaptado, continuo, creciente, con $\langle M \rangle_0 = 0$, tal que*

$$M_t^2 - \langle M \rangle_t \quad \text{es una martingala local continua.}$$

Además, para toda sucesión de particiones $0 = t_0^n < \dots < t_{p_n}^n = t$ de paso tendiendo a cero,

$$\sum_{i=1}^{p_n} (M_{t_i^n} - M_{t_{i-1}^n})^2 \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \langle M \rangle_t. \quad (14.1)$$

Al proceso $\langle M \rangle$ se le llama variación cuadrática de M .

El análogo discreto

Es la versión continua del [ejemplo 2.19](#): allí vimos que, para una martingala discreta, $M_n^2 - \langle M \rangle_n$ es martingala con $\langle M \rangle_n = \sum_{k \leq n} \mathbb{E}[(M_k - M_{k-1})^2 | \mathcal{F}_{k-1}]$. El [teorema 14.9](#) es la descomposición de Doob–Meyer de la submartingala M^2 , y (14.1) dice que la parte creciente se puede leer directamente de las trayectorias.

Ejemplo 14.10 (El caso Browniano). Si B es un movimiento Browniano, $\langle B \rangle_t = t$. En efecto, $B_t^2 - t$ es una martingala ([ejercicio 13.17](#)), el proceso t es continuo creciente y nulo en 0, y la unicidad concluye. La convergencia (14.1) es exactamente la [proposición 13.27](#).

Definición 14.11 (Covariación). Para M, N martingalas locales continuas nulas en 0 definimos, por polarización,

$$\langle M, N \rangle_t = \frac{1}{2} (\langle M + N \rangle_t - \langle M \rangle_t - \langle N \rangle_t).$$

Es el único proceso continuo, adaptado, de variación acotada, nulo en 0, tal que $M_t N_t - \langle M, N \rangle_t$ es una martingala local. Además

$$\sum_{i=1}^{p_n} (M_{t_i^n} - M_{t_{i-1}^n})(N_{t_i^n} - N_{t_{i-1}^n}) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \langle M, N \rangle_t.$$

Proposición 14.12 (Propiedades). Sean M, N, L martingalas locales continuas nulas en 0.

- (a) $\langle \cdot, \cdot \rangle$ es bilineal y simétrica, y $\langle M, M \rangle = \langle M \rangle \geq 0$.
- (b) (Kunita–Watanabe) $|\langle M, N \rangle_t| \leq \sqrt{\langle M \rangle_t} \sqrt{\langle N \rangle_t}$.
- (c) $\langle M \rangle \equiv 0$ si y solo si $M \equiv 0$.
- (d) Si M es además de variación acotada, entonces $M \equiv 0$.
- (e) $\langle M^\tau \rangle_t = \langle M \rangle_{t \wedge \tau}$ para todo tiempo de parada τ .

Indicación de (c) y (d). (c) Si $\langle M \rangle \equiv 0$ entonces M^2 es martingala local con $M_0 = 0$ y $M^2 \geq 0$; por la [proposición 14.8](#) es una supermartingala, luego $\mathbb{E}[M_t^2] \leq \mathbb{E}[M_0^2] = 0$.

(d) Si M tiene variación acotada V_t sobre $[0, t]$, entonces como en el [corolario 13.28](#),

$$\sum_i (M_{t_i^n} - M_{t_{i-1}^n})^2 \leq \left(\max_i |M_{t_i^n} - M_{t_{i-1}^n}| \right) V_t \rightarrow 0$$

por continuidad uniforme, de modo que $\langle M \rangle_t = 0$ y se aplica (c). ■

El enunciado (d) es la clave de todo

*Una martingala local continua no trivial nunca tiene variación acotada. Por eso el movimiento Browniano no la tiene, por eso $\int H dB$ no se puede definir con Lebesgue–Stieltjes, y por eso hay que construir la integral de Itô. Recíprocamente, esto hace que la descomposición de un proceso de Itô en “parte martingala + parte de variación acotada” sea *única*: las dos clases se intersecan solo en el proceso nulo.*

14.3 Semimartingalas continuas

Definición 14.13 (Semimartingala continua). Un proceso $(X_t)_{t \geq 0}$ adaptado y continuo es una *semimartingala continua* si admite una descomposición

$$X_t = X_0 + M_t + A_t, \quad (14.2)$$

donde M es una martingala local continua con $M_0 = 0$ y A es un proceso adaptado, continuo, de variación acotada sobre compactos, con $A_0 = 0$.

Proposición 14.14 (Unicidad de la descomposición). *La descomposición (14.2) es única (salvo indistinguibilidad).*

Demostración. Si $X_0 + M + A = X_0 + \widetilde{M} + \widetilde{A}$, entonces $M - \widetilde{M} = \widetilde{A} - A$ es a la vez martingala local continua nula en 0 y de variación acotada; por la proposición 14.12(d) es idénticamente nulo. ■

Definición 14.15. Para una semimartingala continua $X = X_0 + M + A$ definimos $\langle X \rangle := \langle M \rangle$ y, para otra semimartingala $Y = Y_0 + N + C$, $\langle X, Y \rangle := \langle M, N \rangle$.

Un diccionario para recordar

	Variación acotada	Martingala local
Ejemplo típico	$A_t = \int_0^t a_s ds$	$M_t = \int_0^t h_s dB_s$
Variación cuadrática	0	$\int_0^t h_s^2 ds$
“Tamaño” de d	dt	$\sqrt{dt}$
Contribución a Itô	$f'(X) dA$	$f'(X) dM$ $+ \frac{1}{2} f''(X) d\langle M \rangle$

La última fila es la fórmula de Itô: la parte martingala, por tener incrementos de tamaño $\sqrt{dt}$, contribuye también en segundo orden.

La integral de Itô

Queremos dar sentido a

$$\int_0^t H_s dB_s,$$

donde H es un proceso *aleatorio*. La estrategia es la misma que en la construcción de la integral de Lebesgue: definirla primero para integrandos elementales, donde no hay nada que discutir, y extenderla por continuidad. Lo que hace especial al caso estocástico es que la norma respecto de la que se extiende no es la del supremo, sino una isometría L^2 —la *isometría de Itô*— y que el integrando debe ser *predecible*, exactamente como en el [capítulo 2](#).

15.1 Procesos elementales

Fijemos un espacio filtrado que satisface las condiciones habituales y un movimiento Browniano $(B_t)_{t \geq 0}$ adaptado, tal que $B_t - B_s$ es independiente de $\mathcal{F}_s$ para $s < t$.

Definición 15.1 (Proceso elemental). Un proceso $H = (H_t)_{t \geq 0}$ es *elemental* si es de la forma

$$H_t(\omega) = \sum_{i=0}^{p-1} h_i(\omega) \mathbf{1}_{(t_i, t_{i+1}]}(t), \quad (15.1)$$

con $0 = t_0 < t_1 < \dots < t_p$ y h_i una variable aleatoria $\mathcal{F}_{t_i}$ -medible y acotada.

$\mathcal{F}_{t_i}$ -medible, no $\mathcal{F}_{t_{i+1}}$

El valor h_i , que se usa durante el intervalo $(t_i, t_{i+1}]$, debe conocerse *al comienzo* del intervalo. Es exactamente la predecibilidad del [definición 2.16](#), y es indispensable: si se permitiera h_i $\mathcal{F}_{t_{i+1}}$ -medible, tomando $h_i = \text{sgn}(B_{t_{i+1}} - B_{t_i})$ obtendríamos una “integral” con esperanza positiva a partir de una martingala.

Definición 15.2. Para H elemental como en (15.1) definimos

$$(H \cdot B)_t = \int_0^t H_s dB_s := \sum_{i=0}^{p-1} h_i (B_{t_{i+1} \wedge t} - B_{t_i \wedge t}). \quad (15.2)$$

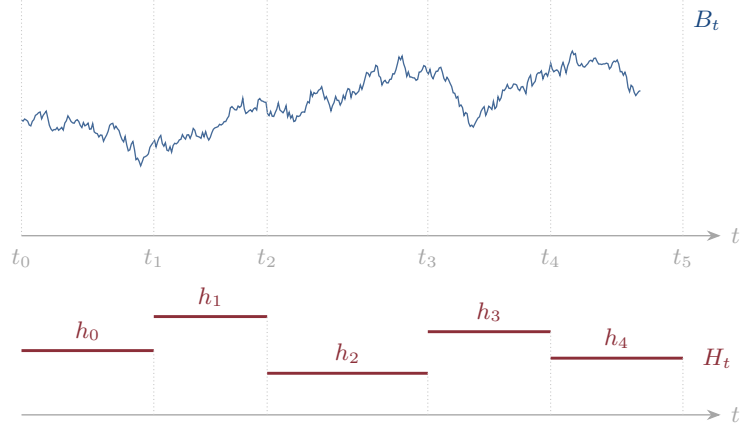


Figura 15.1: Un proceso elemental H (abajo) y el Browniano que integra (arriba). La integral $\int_0^t H dB$ acumula $h_i (B_{t_{i+1}} - B_{t_i})$: cada “posición” h_i se fija al comienzo de su intervalo y multiplica el incremento que viene después.

Proposición 15.3 (Propiedades para integrandos elementales). *Sea H elemental. Entonces:*

- (a) $((H \cdot B)_t)_{t \geq 0}$ es una martingala continua nula en 0;
- (b) (isometría de Itô)

$$\mathbb{E}[(H \cdot B)_t^2] = \mathbb{E}\left[\int_0^t H_s^2 ds\right]; \quad (15.3)$$

$$(c) \quad \langle H \cdot B \rangle_t = \int_0^t H_s^2 ds.$$

Demostración. (a) Continuidad y adaptación son claras. Para la propiedad de martingala, basta verla en un intervalo: para $t_i \leq s < t \leq t_{i+1}$, como h_i es $\mathcal{F}_{t_i} \subset \mathcal{F}_s$ -medible y acotada,

$$\mathbb{E}[h_i(B_t - B_s) \mid \mathcal{F}_s] = h_i \mathbb{E}[B_t - B_s \mid \mathcal{F}_s] = 0.$$

- (b) Escribiendo $\Delta_i B = B_{t_{i+1} \wedge t} - B_{t_i \wedge t}$,

$$\mathbb{E}[(H \cdot B)_t^2] = \sum_i \mathbb{E}[h_i^2 (\Delta_i B)^2] + 2 \sum_{i < j} \mathbb{E}[h_i h_j \Delta_i B \Delta_j B].$$

Los términos cruzados se anulan: para $i < j$, el factor $h_i h_j \Delta_i B$ es $\mathcal{F}_{t_j}$ -medible y $\Delta_j B$ es independiente de $\mathcal{F}_{t_j}$ con media 0, luego

$$\mathbb{E}[h_i h_j \Delta_i B \Delta_j B] = \mathbb{E}[h_i h_j \Delta_i B \mathbb{E}[\Delta_j B \mid \mathcal{F}_{t_j}]] = 0.$$

Para los términos diagonales, h_i^2 es $\mathcal{F}_{t_i}$ -medible y $(\Delta_i B)^2$ independiente de $\mathcal{F}_{t_i}$ con esperanza $t_{i+1} \wedge t - t_i \wedge t$, de modo que

$$\mathbb{E}[(H \cdot B)_t^2] = \sum_i \mathbb{E}[h_i^2] (t_{i+1} \wedge t - t_i \wedge t) = \mathbb{E}\left[\int_0^t H_s^2 ds\right].$$

(c) El mismo cálculo, aplicado condicionalmente, muestra que $(H \cdot B)_t^2 - \int_0^t H_s^2 ds$ es una martingala; la unicidad en el [teorema 14.9](#) concluye. ■

15.2 Extensión por isometría

Definición 15.4 (La clase $\mathcal{H}^2$). Denotamos por $\mathcal{H}^2$ el espacio de los procesos H progresivamente medibles ([definición 9.16](#)) tales que

$$\|H\|_{\mathcal{H}^2(T)}^2 := \mathbb{E}\left[\int_0^T H_s^2 ds\right] < \infty \quad \text{para todo } T > 0.$$

Identificamos procesos que coinciden $dt \otimes d\mathbb{P}$ -c.s..

Lema 15.5 (Densidad de los elementales). *Los procesos elementales son densos en $\mathcal{H}^2$ para la norma $\|\cdot\|_{\mathcal{H}^2(T)}$, para cada $T > 0$.*

Esquema. Se procede en tres etapas.

1. Por truncamiento ($H^{(n)} = H \mathbf{1}_{\{|H| \leq n\}}$) y convergencia dominada, basta aproximar procesos *acotados* de $\mathcal{H}^2$.
2. Para H progresivo y acotado, los procesos regularizados

$$H_t^{(n)} = n \int_{(t-1/n)^+}^t H_s ds$$

son adaptados, con trayectorias continuas, acotados por la misma constante, y $H^{(n)} \rightarrow H$ en $\mathcal{H}^2(T)$ (por el teorema de diferenciación de Lebesgue aplicado trayectoria a trayectoria, más convergencia dominada).

3. Un proceso continuo y acotado se aproxima por su discretización

$$\tilde{H}_t^{(n)} = \sum_i H_{t_i} \mathbf{1}_{(t_i, t_{i+1}]}(t), \quad t_i = \frac{iT}{n},$$

que es elemental y converge por continuidad uniforme sobre $[0, T]$. ■

Teorema 15.6 (La integral de Itô sobre $\mathcal{H}^2$). *Existe una única aplicación lineal*

$$H \mapsto \left(\int_0^t H_s dB_s \right)_{t \geq 0}$$

definida sobre $\mathcal{H}^2$, con valores en el espacio de las martingalas continuas nulas en 0, que extiende a (15.2) y satisface la isometría de Itô (15.3). Además, para todo $H \in \mathcal{H}^2$:

(a) $\left(\int_0^t H dB \right)_{t \geq 0}$ es una martingala continua de cuadrado integrable, nula en 0;

(b) $\left\langle \int_0^\cdot H dB \right\rangle_t = \int_0^t H_s^2 ds$;

(c) más generalmente, para $H, K \in \mathcal{H}^2$,

$$\left\langle \int_0^\cdot H dB, \int_0^\cdot K dB \right\rangle_t = \int_0^t H_s K_s ds, \quad \mathbb{E} \left[\int_0^t H dB \int_0^t K dB \right] = \mathbb{E} \left[\int_0^t H_s K_s ds \right];$$

(d) (localización) para todo tiempo de parada τ ,

$$\int_0^{t \wedge \tau} H_s dB_s = \int_0^t H_s \mathbf{1}_{\{s \leq \tau\}} dB_s.$$

Demostración. Fijemos $T > 0$. Por el lema 15.5 existen procesos elementales $H^{(n)}$ con $\|H^{(n)} - H\|_{\mathcal{H}^2(T)} \rightarrow 0$. Por la isometría (15.3) aplicada a la diferencia (que es elemental),

$$\mathbb{E} \left[\left(\int_0^T H^{(n)} dB - \int_0^T H^{(m)} dB \right)^2 \right] = \|H^{(n)} - H^{(m)}\|_{\mathcal{H}^2(T)}^2 \xrightarrow{n, m \rightarrow \infty} 0,$$

de modo que $\left(\int_0^T H^{(n)} dB \right)_n$ es de Cauchy en $L^2(\Omega)$ y converge. Para obtener convergencia de los procesos usamos la desigualdad maximal de Doob (teorema 14.4(b) con $p = 2$) aplicada a la martingala $\int_0^\cdot (H^{(n)} - H^{(m)}) dB$:

$$\mathbb{E} \left[\sup_{t \leq T} \left(\int_0^t H^{(n)} dB - \int_0^t H^{(m)} dB \right)^2 \right] \leq 4 \|H^{(n)} - H^{(m)}\|_{\mathcal{H}^2(T)}^2.$$

Por lo tanto, pasando a una subsucesión, la convergencia es uniforme sobre $[0, T]$ c. s., y el límite —que llamamos $\int_0^t H dB$ — tiene trayectorias continuas. La propiedad de martingala pasa al límite porque la convergencia es en L^2 , y lo mismo la isometría. La unicidad es clara, pues los elementales son densos y la isometría fuerza el valor del límite.

(c) se obtiene por polarización de (b), y (b) pasando al límite en [proposición 15.3\(c\)](#). Para (d), basta observar que si H es elemental y τ toma finitos valores en la partición, la identidad es evidente término a término; el caso general se sigue por aproximación. ■

La isometría de Itô, en una frase

$$\mathbb{E}\left[\left(\int_0^t H_s dB_s\right)^2\right] = \mathbb{E}\left[\int_0^t H_s^2 ds\right]$$

La integral estocástica convierte la norma $L^2(dt \otimes d\mathbb{P})$ del integrando en la norma $L^2(d\mathbb{P})$ del resultado. Toda la construcción consiste en explotar esta igualdad: es lo que permite definir el objeto por continuidad, y es la herramienta de cálculo principal cuando hay que estimar momentos.

15.3 Localización

La condición $\mathbb{E}\left[\int_0^T H_s^2 ds\right] < \infty$ es a veces incómoda de verificar. Se puede relajar al precio de obtener una martingala *local*.

Definición 15.7. Sea $L_{\text{loc}}^2(B)$ el espacio de los procesos progresivamente medibles H tales que

$$\int_0^T H_s^2 ds < \infty \quad \text{c. s.} \quad \text{para todo } T > 0.$$

Teorema 15.8. La integral de Itô se extiende de manera única a $L_{\text{loc}}^2(B)$, con valores en el espacio de las martingalas locales continuas nulas en 0, conservando las propiedades (b), (c) y (d) del [teorema 15.6](#).

Demostración. Dado $H \in L_{\text{loc}}^2(B)$, definamos

$$\tau_n = \inf \left\{ t \geq 0 : \int_0^t H_s^2 ds \geq n \right\},$$

que son tiempos de parada (por continuidad de $t \mapsto \int_0^t H^2$) con $\tau_n \uparrow \infty$ c.s.. El proceso $H\mathbf{1}_{\{s \leq \tau_n\}}$ pertenece a $\mathcal{H}^2$, pues $\int_0^T H_s^2 \mathbf{1}_{\{s \leq \tau_n\}} ds \leq n$. Definimos

$$\int_0^t H dB := \int_0^t H \mathbf{1}_{\{s \leq \tau_n\}} dB_s \quad \text{sobre } \{t \leq \tau_n\},$$

definición consistente por la propiedad (d) de localización. Por construcción (τ_n) es una sucesión localizante. $\blacksquare$

Martingala local, no martingala

Sobre $L_{\text{loc}}^2(B)$ la integral es solo una martingala *local*: en general $\mathbb{E}\left[\int_0^t H dB\right] \neq 0$. Un contraejemplo clásico: si $\tau_a = \inf\{t : B_t = a\}$ con $a > 0$, el proceso B^{τ_a} es una martingala local acotada superiormente por a , pero

$$\mathbb{E}[B_{\tau_a}] = a \neq 0 = \mathbb{E}[B_0].$$

Es el mismo fenómeno que el jugador terco del [ejemplo 5.8](#). Para asegurar que la integral es una martingala verdadera basta $\mathbb{E}\left[\int_0^t H_s^2 ds\right] < \infty$, o bien la condición de Novikov del [capítulo 18](#).

15.4 Integración respecto de semimartingalas

Definición 15.9. Sea $X = X_0 + M + A$ una semimartingala continua y H progresivo tal que

$$\int_0^T H_s^2 d\langle M \rangle_s < \infty \quad \text{y} \quad \int_0^T |H_s| |dA_s| < \infty \quad \text{c. s., } \forall T > 0.$$

Definimos

$$\int_0^t H_s dX_s = \underbrace{\int_0^t H_s dM_s}_{\text{integral de Itô}} + \underbrace{\int_0^t H_s dA_s}_{\text{Lebesgue-Stieltjes, trayectoria a trayectoria}}.$$

La descomposición está bien definida gracias a la [proposición 14.14](#). Vale

$$\left\langle \int_0^\cdot H dX \right\rangle_t = \int_0^t H_s^2 d\langle X \rangle_s,$$

y la *asociatividad*: si $Y_t = \int_0^t H dX$, entonces

$$\int_0^t K_s dY_s = \int_0^t K_s H_s dX_s.$$

15.5 Procesos de Itô

Definición 15.10 (Proceso de Itô). Un *proceso de Itô* (unidimensional) es un proceso de la forma

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dB_s, \quad (15.4)$$

con X_0 $\mathcal{F}_0$ -medible, b y σ progresivos y tales que

$$\int_0^T |b_s| ds < \infty \quad \text{y} \quad \int_0^T \sigma_s^2 ds < \infty \quad \text{c. s., } \forall T > 0.$$

Escribimos en forma diferencial

$$dX_t = b_t dt + \sigma_t dB_t,$$

que es *notación* para (15.4), no una igualdad entre diferenciales. A b se le llama *deriva* (*drift*) y a σ *volatilidad* o coeficiente de difusión.

Un proceso de Itô es en particular una semimartingala continua, con

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

y la unicidad de la descomposición dice que b y σ^2 están determinados $dt \otimes d\mathbb{P}$ -c. s..

Ejemplo 15.11. Calculemos $\int_0^t B_s dB_s$ directamente desde la definición, para ver aparecer el término de Itô. Tomemos $t_i = it/n$ y escribamos, usando la identidad $a(b-a) = \frac{1}{2}(b^2 - a^2) - \frac{1}{2}(b-a)^2$,

$$\begin{aligned} \sum_{i=0}^{n-1} B_{t_i} (B_{t_{i+1}} - B_{t_i}) &= \frac{1}{2} \sum_{i=0}^{n-1} (B_{t_{i+1}}^2 - B_{t_i}^2) - \frac{1}{2} \sum_{i=0}^{n-1} (B_{t_{i+1}} - B_{t_i})^2 \\ &= \frac{B_t^2}{2} - \frac{1}{2} \sum_{i=0}^{n-1} (B_{t_{i+1}} - B_{t_i})^2. \end{aligned}$$

El primer término no depende de n ; el segundo converge en L^2 a $\frac{1}{2}t$ por la [proposición 13.27](#). Por lo tanto

$$\boxed{\int_0^t B_s dB_s = \frac{B_t^2}{2} - \frac{t}{2}} \quad (15.5)$$

Aquí está toda la diferencia

Si f fuese una función diferenciable, el cálculo ordinario daría $\int_0^t f df = \frac{1}{2}f(t)^2 - \frac{1}{2}f(0)^2$, sin el término $-t/2$. Ese término extra es exactamente $-\frac{1}{2}\langle B \rangle_t$, y aparece porque la variación cuadrática del Browniano no es nula.

Nótese también *dónde* se evalúa el integrando: en el extremo *izquierdo* B_{t_i} de cada subintervalo. Si se evaluara en el punto medio se obtendría $\frac{1}{2}B_t^2$ sin corrección —esa es la integral de Stratonovich— pero se perdería la propiedad de martingala, que es lo que hace útil a la integral de Itô en finanzas: el integrando izquierdo es la posición que el inversor *ya tenía* cuando ocurrió el movimiento del precio.

Ejemplo 15.12 (Integral de Wiener como caso particular). Si $f \in L^2(\mathbb{R}_+)$ es determinista, entonces $f \in \mathcal{H}^2$ y $\int_0^t f(s) dB_s$ coincide con la integral de Wiener (13.2); en particular, por (13.3), es Gaussiana $\mathcal{N}(0, \int_0^t f(s)^2 ds)$. Para integrandos aleatorios esto *deja de ser cierto*: $\int_0^t B_s dB_s$ no es Gaussiana, como muestra (15.5).

La fórmula de Itô

La fórmula de Itô es la regla de la cadena del cálculo estocástico. Dice que, al componer una función suave con una semimartingala, aparece un término de segundo orden que no existe en el cálculo ordinario, y que ese término está gobernado por la variación cuadrática. Es, sin exageración, la herramienta más útil de toda la teoría.

16.1 El enunciado

Teorema 16.1 (Fórmula de Itô, caso unidimensional). *Sea X una semimartingala continua y $f \in C^2(\mathbb{R})$. Entonces $(f(X_t))_{t \geq 0}$ es una semimartingala continua y*

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) dX_s + \frac{1}{2} \int_0^t f''(X_s) d\langle X \rangle_s \quad (16.1)$$

para todo $t \geq 0$, casi seguramente.

Corolario 16.2 (Caso de un proceso de Itô). *Si $dX_t = b_t dt + \sigma_t dB_t$ y $f \in C^2(\mathbb{R})$, entonces*

$$df(X_t) = \left(f'(X_t) b_t + \frac{1}{2} f''(X_t) \sigma_t^2 \right) dt + f'(X_t) \sigma_t dB_t. \quad (16.2)$$

En particular, con $X = B$ ($b \equiv 0$, $\sigma \equiv 1$),

$$f(B_t) = f(0) + \int_0^t f'(B_s) dB_s + \frac{1}{2} \int_0^t f''(B_s) ds. \quad (16.3)$$

Corolario 16.3 (Versión con dependencia temporal). *Si $f \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R})$ y $dX_t = b_t dt + \sigma_t dB_t$, entonces*

$$df(t, X_t) = \left(\partial_t f + b_t \partial_x f + \frac{1}{2} \sigma_t^2 \partial_{xx} f \right)(t, X_t) dt + (\sigma_t \partial_x f)(t, X_t) dB_t. \quad (16.4)$$

La regla mnemotécnica

Todo se obtiene desarrollando f en serie de Taylor hasta segundo orden y aplicando la tabla de multiplicación

$\times$	dt	dB_t
dt	0	0
dB_t	0	dt

En efecto,

$$df(X_t) \approx f'(X_t) dX_t + \frac{1}{2} f''(X_t) (dX_t)^2, \quad (dX_t)^2 = (b_t dt + \sigma_t dB_t)^2 = \sigma_t^2 dt,$$

que es exactamente (16.2). La entrada $(dB_t)^2 = dt$ es la [proposición 13.27](#); las otras tres son la afirmación de que t y B tienen covariación nula y que t tiene variación cuadrática nula.

16.2 Demostración

Demostración del teorema 16.1. Paso 0: reducciones. Por localización podemos suponer que X es acotado, que $\langle X \rangle$ y la variación total de A son acotadas, y que f, f', f'' son acotadas sobre el rango de X (basta detener en $\tau_n = \inf\{t : |X_t| + \langle X \rangle_t + |A|_t \geq n\}$ y hacer $n \rightarrow \infty$).

Paso 1: desarrollo de Taylor sobre una partición. Fijemos $t > 0$ y la partición $t_i^n = it/n$, $i = 0, \dots, n$. Escribimos la suma telescópica

$$f(X_t) - f(X_0) = \sum_{i=0}^{n-1} (f(X_{t_{i+1}^n}) - f(X_{t_i^n})),$$

y aplicamos Taylor con resto de Lagrange a cada sumando: existe ξ_i^n entre $X_{t_i^n}$ y $X_{t_{i+1}^n}$ con

$$f(X_{t_{i+1}^n}) - f(X_{t_i^n}) = f'(X_{t_i^n}) \Delta_i^n X + \frac{1}{2} f''(X_{t_i^n}) (\Delta_i^n X)^2 + \frac{1}{2} (f''(\xi_i^n) - f''(X_{t_i^n})) (\Delta_i^n X)^2, \quad (16.5)$$

donde $\Delta_i^n X = X_{t_{i+1}^n} - X_{t_i^n}$.

Paso 2: el primer término. La suma $\sum_i f'(X_{t_i^n}) \Delta_i^n X$ es exactamente la integral del proceso elemental $H_s^n = \sum_i f'(X_{t_i^n}) \mathbf{1}_{(t_i^n, t_{i+1}^n]}(s)$ respecto de X . Como $f'(X)$ es continuo y acotado, $H^n \rightarrow f'(X)$ puntualmente y de forma

acotada, de modo que

$$\sum_i f'(X_{t_i^n}) \Delta_i^n X \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \int_0^t f'(X_s) dX_s,$$

por convergencia dominada (para la parte de variación acotada) y por la isometría de Itô más convergencia dominada (para la parte martingala).

Paso 3: el segundo término. Afirmamos que

$$\sum_i f''(X_{t_i^n}) (\Delta_i^n X)^2 \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \int_0^t f''(X_s) d\langle X \rangle_s. \quad (16.6)$$

Esto es la versión “con peso” de la convergencia (14.1): la demostración estándar es probarlo primero para f'' de la forma $\mathbf{1}_{(u,v]}$ (donde se reduce a (14.1)), extenderlo por linealidad a funciones escalonadas y pasar al límite uniforme usando la continuidad de $f'' \circ X$ sobre el compacto $[0, t]$.

Paso 4: el resto. Como X es continua sobre $[0, t]$, es uniformemente continua, de modo que $\varepsilon_n := \max_i |f''(\xi_i^n) - f''(X_{t_i^n})| \rightarrow 0$ c. s. (por continuidad uniforme de f'' sobre el rango compacto de X). Entonces

$$\left| \sum_i (f''(\xi_i^n) - f''(X_{t_i^n})) (\Delta_i^n X)^2 \right| \leq \varepsilon_n \sum_i (\Delta_i^n X)^2,$$

y el segundo factor converge en probabilidad a $\langle X \rangle_t < \infty$; el producto tiende a 0 en probabilidad.

Combinando los pasos 1–4 en (16.5) se obtiene (16.1). ■

Observación. La demostración deja ver por qué hace falta $f \in C^2$: el desarrollo de Taylor debe llegar hasta el segundo orden porque $\sum_i (\Delta_i^n X)^2$ no tiende a cero. Con una semimartingala de variación acotada bastaría C^1 , y se recuperaría la regla de la cadena ordinaria.

16.3 Versiones multidimensionales

Teorema 16.4 (Fórmula de Itô multidimensional). *Sean $X^1, \dots, X^d$ semimartingalas continuas y $f \in C^2(\mathbb{R}^d)$. Entonces*

$$f(X_t) = f(X_0) + \sum_{i=1}^d \int_0^t \partial_i f(X_s) dX_s^i + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \partial_{ij} f(X_s) d\langle X^i, X^j \rangle_s. \quad (16.7)$$

Corolario 16.5 (Regla del producto / integración por partes). Si X, Y son semimartingalas continuas,

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle X, Y \rangle_t. \quad (16.8)$$

En forma diferencial, $d(XY) = X dY + Y dX + d\langle X, Y \rangle$.

Demostración. Aplicar (16.7) a $f(x, y) = xy$, cuyas derivadas son $\partial_1 f = y$, $\partial_2 f = x$, $\partial_{12} f = \partial_{21} f = 1$, $\partial_{11} f = \partial_{22} f = 0$. ■

El término de corrección

La regla del producto tiene un término extra $\langle X, Y \rangle$ respecto de la fórmula clásica $d(fg) = f dg + g df$. Si X o Y es de variación acotada, ese término se anula y se recupera la fórmula usual.

Ejemplo 16.6 (Browniano multidimensional). Si $B = (B^1, \dots, B^d)$ es un movimiento Browniano d -dimensional (coordenadas independientes), entonces $\langle B^i, B^j \rangle_t = \delta_{ij} t$ y (16.7) se escribe

$$df(B_t) = \nabla f(B_t) \cdot dB_t + \frac{1}{2} \Delta f(B_t) dt,$$

con Δ el laplaciano. La aparición de $\frac{1}{2} \Delta$ es el origen de la conexión entre el movimiento Browniano y la ecuación del calor.

16.4 Ejemplos de cálculo

Ejemplo 16.7 (Reobtención de (15.5)). Con $f(x) = x^2$ y $X = B$: $f'(x) = 2x$, $f''(x) = 2$, $\langle B \rangle_t = t$, de modo que

$$B_t^2 = 0 + \int_0^t 2B_s dB_s + \frac{1}{2} \int_0^t 2 ds = 2 \int_0^t B_s dB_s + t,$$

es decir $\int_0^t B_s dB_s = \frac{1}{2}(B_t^2 - t)$, exactamente (15.5). Nótese de paso que esto prueba que $B_t^2 - t$ es una martingala: es una integral de Itô.

Ejemplo 16.8 (La exponencial estocástica). Sea M una martingala local continua con $M_0 = 0$. La *exponencial estocástica* de M es

$$\mathcal{E}(M)_t := \exp \left(M_t - \frac{1}{2} \langle M \rangle_t \right). \quad (16.9)$$

Aplicando Itô a $f(x, v) = e^{x-v/2}$ con $X = (M, \langle M \rangle)$ —o directamente a $f(x) = e^x$ tras un cambio—, y usando que $\langle M \rangle$ es de variación acotada,

$$d\mathcal{E}(M)_t = \mathcal{E}(M)_t \left(dM_t - \frac{1}{2} d\langle M \rangle_t \right) + \frac{1}{2} \mathcal{E}(M)_t d\langle M \rangle_t = \mathcal{E}(M)_t dM_t.$$

Es decir, $\mathcal{E}(M)$ es la única solución de $dZ_t = Z_t dM_t$, $Z_0 = 1$, y en particular es una *martingala local*. Con $M_t = \sigma B_t$ recuperamos el [ejercicio 13.17](#): $\exp(\sigma B_t - \sigma^2 t/2)$.

Local, no necesariamente verdadera

$\mathcal{E}(M)$ es una martingala local positiva, luego una supermartingala ([proposición 14.8](#)) y por tanto $\mathbb{E}[\mathcal{E}(M)_t] \leq 1$. La igualdad —es decir, que sea una martingala verdadera— *no* es automática. La condición suficiente estándar es la de Novikov ([teorema 18.2](#)), y es exactamente la hipótesis que hace funcionar el teorema de Girsanov.

Ejemplo 16.9 (Movimiento Browniano geométrico). Sean $\mu \in \mathbb{R}$, $\sigma > 0$, $S_0 > 0$ y definamos

$$S_t = S_0 \exp \left(\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma B_t \right). \quad (16.10)$$

Aplicando Itô con $f(t, x) = S_0 e^{(\mu - \sigma^2/2)t + \sigma x}$ y $X = B$: $\partial_t f = (\mu - \sigma^2/2)f$, $\partial_x f = \sigma f$, $\partial_{xx} f = \sigma^2 f$, de modo que por [\(16.4\)](#)

$$dS_t = \left(\mu - \frac{\sigma^2}{2} + \frac{1}{2} \sigma^2 \right) S_t dt + \sigma S_t dB_t = \mu S_t dt + \sigma S_t dB_t.$$

Así, [\(16.10\)](#) es la solución de

$$\frac{dS_t}{S_t} = \mu dt + \sigma dB_t, \quad (16.11)$$

que es el modelo de Black–Scholes: el *rendimiento relativo* del activo tiene una tendencia determinista μ más un ruido de intensidad σ . Nótese la corrección $-\sigma^2/2$ en el exponente: sin ella, $\mathbb{E}[S_t]$ no sería $S_0 e^{\mu t}$.

Ejemplo 16.10 (Proceso de Ornstein–Uhlenbeck). Sea $\theta, \sigma > 0$ y

$$dX_t = -\theta X_t dt + \sigma dB_t, \quad X_0 = x.$$

Aplicando la regla del producto a $e^{\theta t} X_t$ —el factor $e^{\theta t}$ es determinista, luego de variación acotada y con covariación nula con X —,

$$d(e^{\theta t} X_t) = \theta e^{\theta t} X_t dt + e^{\theta t} dX_t = e^{\theta t} \sigma dB_t,$$

y por lo tanto

$$X_t = x e^{-\theta t} + \sigma \int_0^t e^{-\theta(t-s)} dB_s. \quad (16.12)$$

El integrando es determinista, de modo que X_t es Gaussiana con

$$\mathbb{E}[X_t] = x e^{-\theta t}, \quad \text{Var}(X_t) = \sigma^2 \int_0^t e^{-2\theta(t-s)} ds = \frac{\sigma^2}{2\theta} (1 - e^{-2\theta t}).$$

Cuando $t \rightarrow \infty$, $X_t \xrightarrow{\text{ley}} \mathcal{N}(0, \sigma^2/(2\theta))$: el proceso es *recurrente* y tiene una distribución estacionaria, a diferencia del Browniano.

16.5 La caracterización de Lévy

Cerramos el capítulo con un teorema notable: el movimiento Browniano queda caracterizado por ser una martingala continua con $\langle B \rangle_t = t$. Es la llave que hace funcionar el teorema de Girsanov.

Teorema 16.11 (Caracterización de Lévy). *Sea $(M_t)_{t \geq 0}$ una martingala local continua con $M_0 = 0$ y $\langle M \rangle_t = t$ para todo $t \geq 0$. Entonces M es un movimiento Browniano.*

Demostración. Fijemos $\xi \in \mathbb{R}$ y apliquemos la fórmula de Itô a $f(t, x) = \exp(i\xi x + \frac{\xi^2}{2}t)$ —separando parte real e imaginaria si se quiere evitar funciones complejas— con $X = M$:

$$\partial_t f = \frac{\xi^2}{2} f, \quad \partial_x f = i\xi f, \quad \partial_{xx} f = -\xi^2 f,$$

de modo que, usando $d\langle M \rangle_t = dt$,

$$d\left(e^{i\xi M_t + \frac{\xi^2}{2}t}\right) = \left(\frac{\xi^2}{2} - \frac{\xi^2}{2}\right) e^{i\xi M_t + \frac{\xi^2}{2}t} dt + i\xi e^{i\xi M_t + \frac{\xi^2}{2}t} dM_t = i\xi e^{i\xi M_t + \frac{\xi^2}{2}t} dM_t.$$

El proceso $Z_t^\xi = e^{i\xi M_t + \xi^2 t/2}$ es entonces una martingala local; como está acotado sobre cada $[0, T]$ en módulo por $e^{\xi^2 T/2}$, la [proposición 14.7](#) dice que es una martingala verdadera. Por lo tanto, para $s < t$ y $A \in \mathcal{F}_s$,

$$\mathbb{E}[\mathbf{1}_A e^{i\xi M_t + \frac{\xi^2}{2}t}] = \mathbb{E}[\mathbf{1}_A e^{i\xi M_s + \frac{\xi^2}{2}s}],$$

es decir

$$\mathbb{E}[\mathbf{1}_A e^{i\xi(M_t - M_s)}] = \mathbb{P}(A) e^{-\frac{\xi^2}{2}(t-s)}.$$

Tomando $A = \Omega$ vemos que $M_t - M_s \sim \mathcal{N}(0, t - s)$; y la identidad general dice que $M_t - M_s$ es independiente de $\mathcal{F}_s$ (su función característica condicional es determinista). Por la [proposición 13.6\(iii\)](#), M es un pre-movimiento Browniano; siendo continuo, es un movimiento Browniano. ■

Por qué es sorprendente

No se pide nada sobre las distribuciones de M : solo que sea una martingala continua cuya variación cuadrática sea exactamente t . De ahí sale la normalidad de los incrementos, su independencia y su estacionariedad. En el [capítulo 18](#) usaremos esto así: tras cambiar de medida, el proceso $B_t - \int_0^t \theta_s ds$ sigue siendo una martingala continua con la misma variación cuadrática, luego es un movimiento Browniano bajo la nueva medida.

Ecuaciones diferenciales estocásticas

Una ecuación diferencial estocástica (EDE) describe la evolución de un sistema sometido a una fuerza determinista y a un ruido. Su solución es un proceso de Markov —una *difusión*— cuyo generador es un operador diferencial de segundo orden; esa correspondencia entre procesos y ecuaciones en derivadas parciales es una de las razones por las que el cálculo estocástico es tan útil fuera de la probabilidad.

17.1 Existencia y unicidad

Definición 17.1 (Solución fuerte). Sean $b, \sigma : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$ medibles. Dado un movimiento Browniano (B_t) sobre un espacio filtrado y una condición inicial η $\mathcal{F}_0$ -medible, una *solución fuerte* de la EDE

$$dX_t = b(t, X_t) dt + \sigma(t, X_t) dB_t, \quad X_0 = \eta, \quad (17.1)$$

es un proceso continuo y adaptado (X_t) tal que, para todo $t \geq 0$,

$$X_t = \eta + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s \quad \text{c. s.},$$

con ambas integrales bien definidas.

Teorema 17.2 (Existencia y unicidad bajo condiciones de Lipschitz). Supongamos que existe $K > 0$ tal que, para todos $t \geq 0$ y $x, y \in \mathbb{R}$,

$$|b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq K |x - y|, \quad (17.2)$$

$$|b(t, x)| + |\sigma(t, x)| \leq K (1 + |x|). \quad (17.3)$$

Sea η $\mathcal{F}_0$ -medible con $\mathbb{E}[\eta^2] < \infty$. Entonces (17.1) tiene una única solución fuerte (salvo indistinguibilidad), y para todo $T > 0$

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} X_t^2 \right] < \infty. \quad (17.4)$$

Demostración. La demostración es una iteración de Picard, con la isometría de Itô reemplazando a la estimación trivial de la integral ordinaria. Fijemos $T > 0$ y trabajemos en el espacio

$$\mathcal{S}_T = \left\{ X \text{ continuo, adaptado : } \|X\|_{\mathcal{S}_T}^2 := \mathbb{E} \left[\sup_{t \leq T} X_t^2 \right] < \infty \right\},$$

que es completo. Definamos $X_t^0 = \eta$ y

$$X_t^{n+1} = \eta + \int_0^t b(s, X_s^n) ds + \int_0^t \sigma(s, X_s^n) dB_s.$$

Estimación básica. Para dos procesos $X, Y \in \mathcal{S}_T$ y $t \leq T$, usando $(u + v)^2 \leq 2u^2 + 2v^2$,

$$\begin{aligned} \mathbb{E} \left[\sup_{r \leq t} (X_r^{(1)} - Y_r^{(1)})^2 \right] &\leq 2 \mathbb{E} \left[\sup_{r \leq t} \left(\int_0^r (b(s, X_s) - b(s, Y_s)) ds \right)^2 \right] \\ &\quad + 2 \mathbb{E} \left[\sup_{r \leq t} \left(\int_0^r (\sigma(s, X_s) - \sigma(s, Y_s)) dB_s \right)^2 \right], \end{aligned}$$

donde $X^{(1)}, Y^{(1)}$ denotan los resultados de una iteración. Para el primer término, Cauchy–Schwarz y (17.2) dan

$$\left(\int_0^r (b(s, X_s) - b(s, Y_s)) ds \right)^2 \leq t \int_0^r K^2 (X_s - Y_s)^2 ds.$$

Para el segundo, la desigualdad maximal de Doob (teorema 14.4(b) con $p = 2$) y la isometría de Itô (15.3) dan

$$\mathbb{E} \left[\sup_{r \leq t} \left(\int_0^r (\sigma(s, X_s) - \sigma(s, Y_s)) dB_s \right)^2 \right] \leq 4 \mathbb{E} \left[\int_0^t K^2 (X_s - Y_s)^2 ds \right].$$

Combinando, con $C = 2K^2(T + 4)$,

$$\mathbb{E} \left[\sup_{r \leq t} (X_r^{(1)} - Y_r^{(1)})^2 \right] \leq C \int_0^t \mathbb{E} \left[\sup_{u \leq s} (X_u - Y_u)^2 \right] ds. \quad (17.5)$$

Convergencia. Aplicando (17.5) iteradamente a $g_n(t) = \mathbb{E} \left[\sup_{r \leq t} (X_r^{n+1} - X_r^n)^2 \right]$ se obtiene

$$g_n(t) \leq g_0(T) \frac{(Ct)^n}{n!},$$

de modo que $\sum_n \|X^{n+1} - X^n\|_{\mathcal{S}_T} < \infty$ (la serie $\sum_n \sqrt{(CT)^n/n!}$ converge). Por lo tanto (X^n) es de Cauchy en $\mathcal{S}_T$ y converge a un límite X , continuo y

adaptado; pasando al límite en la recursión —lícito por (17.5)— resulta que X resuelve (17.1). La condición de crecimiento (17.3) garantiza $g_0(T) < \infty$ y, por el mismo tipo de estimación más el lema de Grönwall, (17.4).

Unicidad. Si X e Y son dos soluciones, (17.5) con $X^{(1)} = X$, $Y^{(1)} = Y$ da

$$h(t) := \mathbb{E} \left[\sup_{r \leq t} (X_r - Y_r)^2 \right] \leq C \int_0^t h(s) ds,$$

y como h es acotada sobre $[0, T]$, el lema de Grönwall da $h \equiv 0$. ■

Dónde interviene lo estocástico

La demostración es la de Picard–Lindelöf, con una sola sustitución: la estimación $|\int_0^t \sigma dB| \leq \int_0^t |\sigma|$ —que sería falsa— se reemplaza por la isometría de Itô combinada con Doob. La condición de Lipschitz aparece exactamente en el mismo lugar que en el caso determinista.

17.2 Ejemplos

Ejemplo 17.3 (Movimiento Browniano geométrico). Con $b(t, x) = \mu x$ y $\sigma(t, x) = \sigma x$, ambos Lipschitz en x , la solución de $dS_t = \mu S_t dt + \sigma S_t dB_t$, $S_0 = s > 0$, es

$$S_t = s \exp \left(\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma B_t \right),$$

como se verificó en el ejemplo 16.9. En particular $S_t > 0$ para todo t : el precio nunca se vuelve negativo, que es la razón principal para usar este modelo.

Ejemplo 17.4 (Ornstein–Uhlenbeck / Vasicek). $dX_t = \theta(\bar{\mu} - X_t) dt + \sigma dB_t$ tiene solución explícita

$$X_t = \bar{\mu} + (x - \bar{\mu})e^{-\theta t} + \sigma \int_0^t e^{-\theta(t-s)} dB_s,$$

obtenida como en el ejemplo 16.10 multiplicando por $e^{\theta t}$. Es un proceso Gaussiano con *reversión a la media* $\bar{\mu}$, usado para modelar tasas de interés.

Ejemplo 17.5 (Puente Browniano). Fijemos $x_0, x_1 \in \mathbb{R}$. La EDE

$$dX_t = \frac{x_1 - X_t}{1 - t} dt + dB_t, \quad t \in [0, 1), \quad X_0 = x_0,$$

tiene por solución

$$X_t = x_0(1-t) + x_1t + (1-t) \int_0^t \frac{dB_s}{1-s},$$

y $X_t \rightarrow x_1$ c. s. cuando $t \uparrow 1$: es el *punte Browniano* de x_0 a x_1 . Nótese que el coeficiente de deriva explota en $t = 1$, de modo que el [teorema 17.2](#) solo garantiza existencia y unicidad sobre $[0, T]$ con $T < 1$.

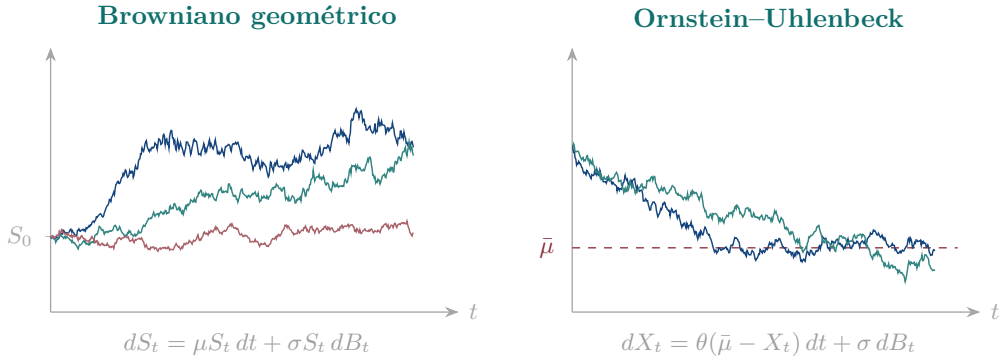


Figura 17.1: Dos difusiones simuladas. A la izquierda, tres trayectorias de un movimiento Browniano geométrico: permanecen positivas y crecen en media exponencialmente. A la derecha, dos trayectorias de un proceso de Ornstein–Uhlenbeck que arrancan lejos del nivel $\bar{\mu}$ y revierten rápidamente a él, oscilando después alrededor de su distribución estacionaria $\mathcal{N}(\bar{\mu}, \sigma^2/(2\theta))$.

17.3 Difusiones, generador y Feynman–Kac

Supongamos que b y σ no dependen de t (caso *homogéneo*) y satisfacen las hipótesis del [teorema 17.2](#). Denotamos por X^x la solución con $X_0 = x$, y por $\mathbb{P}_x$ su ley.

Teorema 17.6 (Propiedad de Markov). *La familia $(X^x)_{x \in \mathbb{R}}$ es un proceso de Markov: para toda $F : \mathcal{C}_{\mathbb{R}} \rightarrow \mathbb{R}$ medible acotada y todo $s \geq 0$,*

$$\mathbb{E}_x[F \circ \theta_s \mid \mathcal{F}_s] = \varphi(X_s), \quad \varphi(y) = \mathbb{E}_y[F].$$

Idea. La unicidad de la solución implica que, a partir del instante s , el proceso $(X_{s+u})_{u \geq 0}$ resuelve la misma EDE con condición inicial X_s y con el Browniano $B_u^{(s)} = B_{s+u} - B_s$, que es independiente de $\mathcal{F}_s$. La propiedad de sustitución para la esperanza condicional concluye. ■

Definición 17.7 (Generador infinitesimal). El *generador* de la difusión es el operador diferencial

$$(\mathcal{L}f)(x) = b(x) f'(x) + \frac{1}{2} \sigma(x)^2 f''(x), \quad f \in C^2(\mathbb{R}). \quad (17.6)$$

Con esta notación la fórmula de Itô (16.2) se lee

$$f(X_t) = f(X_0) + \int_0^t (\mathcal{L}f)(X_s) ds + \underbrace{\int_0^t f'(X_s) \sigma(X_s) dB_s}_{\text{martingala local}}, \quad (17.7)$$

es decir: $f(X_t) - \int_0^t \mathcal{L}f(X_s) ds$ es una *martingala local*. Esta es la formulación del *problema de martingalas* de Stroock–Varadhan, y es la manera moderna de definir una difusión sin pasar por la EDE.

Teorema 17.8 (Fórmula de Dynkin). Sea $f \in C^2$ con f y $\mathcal{L}f$ acotadas, y τ un tiempo de parada con $\mathbb{E}[\tau] < \infty$. Entonces

$$\mathbb{E}_x[f(X_\tau)] = f(x) + \mathbb{E}_x\left[\int_0^\tau (\mathcal{L}f)(X_s) ds\right].$$

Demostración. Tomar esperanzas en (17.7) evaluada en $t \wedge \tau$ (la integral estocástica detenida es una martingala verdadera, pues su integrando es acotado) y pasar al límite $t \rightarrow \infty$ por convergencia dominada. ■

Teorema 17.9 (Fórmula de Feynman–Kac). Sean b, σ como arriba, $r \geq 0$ constante, y $g : \mathbb{R} \rightarrow \mathbb{R}$ continua con crecimiento a lo sumo polinomial. Supongamos que $u \in C^{1,2}([0, T] \times \mathbb{R})$ resuelve el problema

$$\begin{cases} \partial_t u(t, x) + (\mathcal{L}u)(t, x) - r u(t, x) = 0, & (t, x) \in [0, T] \times \mathbb{R}, \\ u(T, x) = g(x), & x \in \mathbb{R}, \end{cases} \quad (17.8)$$

con $\partial_x u$ de crecimiento polinomial. Entonces

$$u(t, x) = \mathbb{E}\left[e^{-r(T-t)} g(X_T) \mid X_t = x\right] \quad (17.9)$$

Demostración. Fijemos (t, x) y consideremos la difusión X con $X_t = x$. Apliquemos la fórmula de Itô (16.4) al proceso

$$Y_s = e^{-r(s-t)} u(s, X_s), \quad s \in [t, T].$$

La regla del producto (corolario 16.5, con el factor determinista $e^{-r(s-t)}$, de covariación nula) da

$$\begin{aligned} dY_s &= e^{-r(s-t)} \left(-ru + \partial_s u + b \partial_x u + \frac{1}{2} \sigma^2 \partial_{xx} u \right) (s, X_s) ds + e^{-r(s-t)} (\sigma \partial_x u)(s, X_s) dB_s \\ &= e^{-r(s-t)} \underbrace{\left(\partial_s u + \mathcal{L}u - ru \right)}_{=0 \text{ por (17.8)}} (s, X_s) ds + e^{-r(s-t)} (\sigma \partial_x u)(s, X_s) dB_s. \end{aligned}$$

Por lo tanto Y es una martingala local sobre $[t, T]$; las hipótesis de crecimiento polinomial y la estimación (17.4) garantizan que el integrando está en $\mathcal{H}^2$, de modo que Y es una martingala verdadera. Igualando esperanzas en $s = t$ y $s = T$,

$$u(t, x) = Y_t = \mathbb{E}[Y_T] = \mathbb{E}[e^{-r(T-t)} u(T, X_T)] = \mathbb{E}[e^{-r(T-t)} g(X_T)]. \quad \blacksquare$$

El puente entre EDPs y probabilidad

Feynman–Kac dice que resolver la EDP parabólica (17.8) equivale a calcular una esperanza a lo largo de las trayectorias de una difusión. Se lee en ambas direcciones:

- *Probabilidad $\rightarrow$ EDP*: se puede calcular $\mathbb{E}[g(X_T)]$ resolviendo numéricamente una EDP;
- *EDP $\rightarrow$ probabilidad*: se puede resolver una EDP simulando trayectorias (método de Monte Carlo).

En el capítulo 18 veremos que el precio de una opción europea es exactamente una expresión de la forma (17.9), y que la EDP (17.8) correspondiente es la ecuación de Black–Scholes.

Ejemplo 17.10 (Caso $b = 0$, $\sigma = 1$, $r = 0$). El generador es $\mathcal{L} = \frac{1}{2} \partial_{xx}$ y (17.8) se convierte en la ecuación del calor retrógrada $\partial_t u + \frac{1}{2} \partial_{xx} u = 0$, $u(T, \cdot) = g$. La fórmula (17.9) da

$$u(t, x) = \mathbb{E}[g(x + B_{T-t})] = \int_{\mathbb{R}} g(y) \frac{e^{-(y-x)^2/(2(T-t))}}{\sqrt{2\pi(T-t)}} dy,$$

que es exactamente la solución clásica por convolución con el núcleo del calor.

Girsanov, representación de martingalas y Black–Scholes

Cerramos el curso donde empezó la Parte II, pero en tiempo continuo. El teorema de Girsanov construye la probabilidad martingala equivalente; el teorema de representación de martingalas da la cobertura; y juntos producen la fórmula de Black–Scholes. El paralelo con la Parte II es exacto, y lo haremos explícito al final.

18.1 El teorema de Girsanov

Teorema 18.1 (Girsanov). *Sea $(B_t)_{0 \leq t \leq T}$ un movimiento Browniano sobre $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ y $(\theta_t)_{0 \leq t \leq T}$ un proceso progresivo con $\int_0^T \theta_s^2 ds < \infty$ c. s.. Definamos*

$$Z_t = \exp \left(- \int_0^t \theta_s dB_s - \frac{1}{2} \int_0^t \theta_s^2 ds \right) = \mathcal{E} \left(- \int_0^t \theta dB \right)_t. \quad (18.1)$$

Si $(Z_t)_{0 \leq t \leq T}$ es una martingala verdadera —en particular si se cumple la condición de Novikov del [teorema 18.2](#)—, entonces $\mathbb{E}[Z_T] = 1$ y la fórmula

$$\left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_T} = Z_T \quad (18.2)$$

define una probabilidad $\mathbb{Q}$ equivalente a $\mathbb{P}$ sobre $\mathcal{F}_T$, bajo la cual el proceso

$$\tilde{B}_t := B_t + \int_0^t \theta_s ds, \quad 0 \leq t \leq T, \quad (18.3)$$

es un movimiento Browniano.

Demostración. Que $\mathbb{Q}$ es una probabilidad equivalente es claro: $Z_T > 0$ y $\mathbb{E}[Z_T] = 1$ por ser Z martingala con $Z_0 = 1$.

Para ver que $\tilde{B}$ es un $\mathbb{Q}$ -movimiento Browniano usamos la caracterización de Lévy ([teorema 16.11](#)): basta probar que $\tilde{B}$ es una $\mathbb{Q}$ -martingala local

continua con $\langle \tilde{B} \rangle_t = t$. La variación cuadrática es inmediata, pues $\tilde{B} - B = \int_0^\cdot \theta ds$ es de variación acotada y $\langle B \rangle_t = t$.

Para la propiedad de martingala usamos el criterio siguiente, elemental pero central:

$$Y \text{ es } \mathbb{Q}\text{-martingala} \iff YZ \text{ es } \mathbb{P}\text{-martingala.} \quad (18.4)$$

(En efecto, $\mathbb{E}_{\mathbb{Q}}[Y_t | \mathcal{F}_s] = Z_s^{-1} \mathbb{E}[Y_t Z_t | \mathcal{F}_s]$ por la fórmula de Bayes para esperanzas condicionales.) Calculemos entonces $d(\tilde{B}_t Z_t)$ con la regla del producto (corolario 16.5). Por el ejemplo 16.8,

$$dZ_t = -Z_t \theta_t dB_t,$$

y $d\tilde{B}_t = dB_t + \theta_t dt$, de modo que $d\langle \tilde{B}, Z \rangle_t = d\langle B, Z \rangle_t = -Z_t \theta_t dt$. Por lo tanto

$$\begin{aligned} d(\tilde{B}_t Z_t) &= \tilde{B}_t dZ_t + Z_t d\tilde{B}_t + d\langle \tilde{B}, Z \rangle_t \\ &= -\tilde{B}_t Z_t \theta_t dB_t + Z_t dB_t + Z_t \theta_t dt - Z_t \theta_t dt \\ &= Z_t(1 - \tilde{B}_t \theta_t) dB_t, \end{aligned}$$

que no tiene parte de deriva: $\tilde{B}Z$ es una $\mathbb{P}$ -martingala local y, por (18.4), $\tilde{B}$ es una $\mathbb{Q}$ -martingala local. La teorema 16.11 concluye. ■

Teorema 18.2 (Condición de Novikov). *Si*

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

entonces $(Z_t)_{0 \leq t \leq T}$ definido por (18.1) es una martingala uniformemente integrable sobre $[0, T]$.

Qué hace y qué no hace Girsanov

Girsanov *cambia la deriva* sin tocar la volatilidad. Bajo $\mathbb{P}$, el proceso B no tiene deriva; bajo $\mathbb{Q}$, el que no tiene deriva es $\tilde{B} = B + \int \theta ds$, de modo que “bajo $\mathbb{Q}$, el proceso B tiene deriva $-\int \theta ds$ ”. Lo que *no* puede hacer es cambiar la variación cuadrática: dos medidas equivalentes ven la misma $\langle \cdot \rangle$, porque esta se lee de las trayectorias (14.1) y las medidas equivalentes tienen los mismos conjuntos nulos. Por eso, en el modelo de Black–Scholes, la volatilidad σ es observable y el rendimiento μ no aparece en el precio de la opción.

18.2 Representación de martingalas

Teorema 18.3 (Representación de martingalas brownianas). *Sea $(B_t)_{0 \leq t \leq T}$ un movimiento Browniano y $\mathcal{F}_t = \sigma(B_s : s \leq t)$ su filtración natural completada. Entonces toda martingala $(M_t)_{0 \leq t \leq T}$ respecto de $(\mathcal{F}_t)$ con $\mathbb{E}[M_T^2] < \infty$ se escribe de manera única como*

$$M_t = M_0 + \int_0^t H_s dB_s, \quad 0 \leq t \leq T, \quad (18.5)$$

para algún $H \in \mathcal{H}^2$. En particular, toda martingala respecto de la filtración browniana es continua.

Esquema. Sea $\mathcal{I} = \{\int_0^T H dB : H \in \mathcal{H}^2\}$, que por la isometría de Itô es un subespacio cerrado de $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$. Hay que ver que $\mathcal{I} \oplus \mathbb{R} = L^2(\Omega, \mathcal{F}_T, \mathbb{P})$, es decir que el ortogonal de $\mathcal{I}$ dentro de las variables centradas es nulo.

El paso clave es que las exponenciales

$$\mathcal{E}\left(\int_0^T f dB\right)_T = \exp\left(\int_0^T f(s) dB_s - \frac{1}{2} \int_0^T f(s)^2 ds\right), \quad f \in L^2([0, T]) \text{ determinista,}$$

pertenecen a $\mathcal{I} \oplus \mathbb{R}$ —por el [ejemplo 16.8](#), $d\mathcal{E} = \mathcal{E} f dB$ — y generan un subconjunto total de $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$: si Ξ es ortogonal a todas ellas, tomando $f = \sum_j \lambda_j \mathbf{1}_{(t_{j-1}, t_j]}$ se deduce que $\mathbb{E}\left[\Xi e^{\sum_j \lambda_j (B_{t_j} - B_{t_{j-1}})}\right] = 0$ para todos λ_j , y por analiticidad/inversión de Fourier $\mathbb{E}[\Xi | B_{t_1}, \dots, B_{t_k}] = 0$; como los cilindros generan $\mathcal{F}_T$, $\Xi = 0$.

Dada entonces M , se escribe $M_T = M_0 + \int_0^T H dB$ y se toman esperanzas condicionales: $M_t = \mathbb{E}[M_T | \mathcal{F}_t] = M_0 + \int_0^t H dB$, pues la integral de Itô es una martingala. La unicidad se sigue de la isometría: si $\int_0^T (H - K) dB = 0$ entonces $\mathbb{E}\left[\int_0^T (H - K)^2 ds\right] = 0$. ■

El mismo teorema que en la Parte II

Compárese con la [proposición 7.19](#): en un mercado completo a tiempo discreto, toda $\mathbb{Q}$ -martingala se escribe como $X_n = X_0 + (\Gamma \cdot M)_n$. Aquí, toda martingala browniana se escribe como $M_t = M_0 + \int_0^t H dB$. En ambos casos, el resultado es *exactamente* lo que se necesita para cubrir un derivado: el integrando H es la estrategia.

18.3 El modelo de Black–Scholes

Definición 18.4 (El mercado de Black–Scholes). Sobre $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ con B un movimiento Browniano y $\mathcal{F}_t$ su filtración natural, el mercado consta de:

- un *activo sin riesgo* (el numerario) de precio

$$B_t^0 = e^{rt}, \quad \text{es decir} \quad dB_t^0 = r B_t^0 dt,$$

con $r \geq 0$ la tasa de interés instantánea;

- un *activo riesgoso* de precio S dado por

$$dS_t = \mu S_t dt + \sigma S_t dB_t, \quad S_0 = s > 0, \quad (18.6)$$

con $\mu \in \mathbb{R}$ y $\sigma > 0$ constantes.

Una *estrategia* es un par de procesos progresivos $\Gamma_t = (\alpha_t, \beta_t)$ —unidades del numerario y del activo riesgoso— cuyo valor es

$$V_t = \alpha_t B_t^0 + \beta_t S_t,$$

y que es *autofinanciada* si

$$dV_t = \alpha_t dB_t^0 + \beta_t dS_t. \quad (18.7)$$

Esta es la traducción exacta de la [definición 6.4](#): comparando con la [proposición 6.5](#), $V_t = V_0 + \int_0^t \Gamma_s dM_s$ con $M = (B^0, S)$.

Definición 18.5 (Precios descontados). Escribimos $\tilde{S}_t = e^{-rt} S_t$ y $\tilde{V}_t = e^{-rt} V_t$.

Lema 18.6. Γ es autofinanciada si y solo si

$$d\tilde{V}_t = \beta_t d\tilde{S}_t, \quad \text{es decir} \quad \tilde{V}_t = V_0 + \int_0^t \beta_u d\tilde{S}_u. \quad (18.8)$$

Demostración. Por la regla del producto con el factor determinista e^{-rt} —que es de variación acotada, luego con covariación nula con V — se tiene $d\tilde{V}_t = e^{-rt}(dV_t - rV_t dt)$. Sustituyendo (18.7), $dB_t^0 = rB_t^0 dt$ y $V_t = \alpha_t B_t^0 + \beta_t S_t$:

$$\begin{aligned} d\tilde{V}_t &= e^{-rt}(dV_t - rV_t dt) \\ &= e^{-rt}(\alpha_t r B_t^0 dt + \beta_t dS_t - r\alpha_t B_t^0 dt - r\beta_t S_t dt) \\ &= \beta_t e^{-rt}(dS_t - rS_t dt) = \beta_t d\tilde{S}_t, \end{aligned}$$

donde la última igualdad es la regla del producto aplicada a $\tilde{S}_t = e^{-rt} S_t$. El recíproco es el mismo cálculo al revés. ■

La probabilidad de riesgo neutro

De (18.6) y la regla del producto,

$$d\tilde{S}_t = (\mu - r)\tilde{S}_t dt + \sigma\tilde{S}_t dB_t = \sigma\tilde{S}_t \underbrace{\left(dB_t + \frac{\mu - r}{\sigma} dt \right)}_{= d\tilde{B}_t}. \quad (18.9)$$

Aparece naturalmente el *precio de mercado del riesgo*

$$\theta = \frac{\mu - r}{\sigma} \quad (\text{constante}), \quad (18.10)$$

y la condición de Novikov se cumple trivialmente por ser θ constante. Por el teorema 18.1, con

$$Z_T = \exp\left(-\theta B_T - \frac{\theta^2}{2}T\right), \quad \left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_T} = Z_T, \quad (18.11)$$

el proceso $\tilde{B}_t = B_t + \theta t$ es un $\mathbb{Q}$ -movimiento Browniano y (18.9) se convierte en

$$d\tilde{S}_t = \sigma\tilde{S}_t d\tilde{B}_t, \quad \text{es decir} \quad \tilde{S}_t = s \mathcal{E}(\sigma\tilde{B})_t = s \exp\left(\sigma\tilde{B}_t - \frac{\sigma^2}{2}t\right). \quad (18.12)$$

Teorema 18.7 (Ausencia de arbitraje y completitud). *En el modelo de Black–Scholes:*

- (a) $\mathbb{Q}$ definida por (18.11) es la única probabilidad equivalente a $\mathbb{P}$ bajo la cual $\tilde{S}$ es una martingala;
- (b) el mercado es libre de arbitraje y completo: todo contrato $F \in L^2(\Omega, \mathcal{F}_T, \mathbb{Q})$ admite una cobertura;
- (c) el precio del contrato F en el instante t es

$$\boxed{\pi_t(F) = \mathbb{E}_{\mathbb{Q}}[e^{-r(T-t)} F \mid \mathcal{F}_t]} \quad (18.13)$$

Demostración. (a) Por (18.12), $\tilde{S}$ es bajo $\mathbb{Q}$ una exponencial estocástica, luego una martingala local; por la condición de Novikov (o porque $\mathbb{E}[\tilde{S}_t^2] < \infty$) es una martingala verdadera. La unicidad se sigue de que cualquier otra medida equivalente $\mathbb{Q}'$ debe hacer de $\tilde{S}$ una martingala; por (18.9) eso obliga a que la deriva de $\tilde{S}$ bajo $\mathbb{Q}'$ se anule, lo que determina θ y, por el teorema 18.3 aplicado a la densidad $d\mathbb{Q}'/d\mathbb{P}$, determina $\mathbb{Q}' = \mathbb{Q}$.

(b) Sea $F \in L^2(\Omega, \mathcal{F}_T, \mathbb{Q})$ y definamos la $\mathbb{Q}$ -martingala

$$\tilde{V}_t = \mathbb{E}_{\mathbb{Q}}[e^{-rT} F \mid \mathcal{F}_t].$$

Como $\tilde{B}$ es un $\mathbb{Q}$ -movimiento Browniano que genera la misma filtración, el [teorema 18.3](#) da $H \in \mathcal{H}^2$ con

$$\tilde{V}_t = \tilde{V}_0 + \int_0^t H_u d\tilde{B}_u.$$

Poniendo

$$\beta_t = \frac{H_t}{\sigma \tilde{S}_t}, \quad \alpha_t = \tilde{V}_t - \beta_t \tilde{S}_t, \quad (18.14)$$

se tiene $\beta_t d\tilde{S}_t = \beta_t \sigma \tilde{S}_t d\tilde{B}_t = H_t d\tilde{B}_t = d\tilde{V}_t$, de modo que por el [lema 18.6](#) la estrategia (α, β) es autofinanciada, y su valor descontado en T es $\tilde{V}_T = e^{-rT} F$, es decir $V_T = F$: es una cobertura.

(c) Si Γ cubre a F , entonces $\tilde{V}^\Gamma$ es una $\mathbb{Q}$ -martingala por (18.8), luego

$$e^{-rt} V_t^\Gamma = \tilde{V}_t^\Gamma = \mathbb{E}_{\mathbb{Q}}[\tilde{V}_T^\Gamma \mid \mathcal{F}_t] = \mathbb{E}_{\mathbb{Q}}[e^{-rT} F \mid \mathcal{F}_t],$$

que es (18.13). Que ese sea el único precio justo se argumenta igual que en la [proposición 7.10](#). En particular no hay arbitraje, pues el valor descontado de toda estrategia autofinanciada es una $\mathbb{Q}$ -martingala y se repite la demostración del [teorema 6.13](#). ■

18.4 La fórmula de Black–Scholes

Teorema 18.8 (Ecuación y fórmula de Black–Scholes). *Sea $F = g(S_T)$ con g continua de crecimiento polinomial. Entonces el precio en t es $\pi_t = u(t, S_t)$, donde $u \in C^{1,2}([0, T) \times (0, \infty))$ resuelve la ecuación de Black–Scholes*

$$\begin{cases} \partial_t u + r x \partial_x u + \frac{1}{2} \sigma^2 x^2 \partial_{xx} u - r u = 0, & (t, x) \in [0, T) \times (0, \infty), \\ u(T, x) = g(x), & x > 0. \end{cases} \quad (18.15)$$

Para la call europea $g(x) = (x - K)^+$ la solución es

$$u(t, x) = x \Phi(d_+) - K e^{-r(T-t)} \Phi(d_-) \quad (18.16)$$

donde Φ es la función de distribución de $\mathcal{N}(0, 1)$ y

$$d_{\pm} = \frac{\log(x/K) + (r \pm \frac{\sigma^2}{2})(T-t)}{\sigma \sqrt{T-t}}. \quad (18.17)$$

La cobertura correspondiente es $\beta_t = \partial_x u(t, S_t) = \Phi(d_+)$ (el delta de la opción).

Demostración. La EDP. Bajo $\mathbb{Q}$, por (18.12),

$$dS_t = rS_t dt + \sigma S_t d\tilde{B}_t, \quad (18.18)$$

de modo que S es una difusión con generador

$$(\mathcal{L}f)(x) = r x f'(x) + \frac{1}{2} \sigma^2 x^2 f''(x).$$

La fórmula (18.13) dice $\pi_t = \mathbb{E}_{\mathbb{Q}}[e^{-r(T-t)} g(S_T) \mid \mathcal{F}_t]$, que por la propiedad de Markov de S (teorema 17.6) es una función $u(t, S_t)$. El teorema 17.9, aplicado bajo $\mathbb{Q}$ con este generador y esta tasa r , dice que u resuelve (17.8), que es exactamente (18.15).

La fórmula. De (18.18), bajo $\mathbb{Q}$,

$$S_T = S_t \exp\left((r - \frac{\sigma^2}{2})(T - t) + \sigma(\tilde{B}_T - \tilde{B}_t)\right),$$

y $\tilde{B}_T - \tilde{B}_t \sim \mathcal{N}(0, T - t)$ es independiente de $\mathcal{F}_t$. Escribiendo $\tau = T - t$, $x = S_t$ y $G \sim \mathcal{N}(0, 1)$,

$$u(t, x) = e^{-r\tau} \mathbb{E}\left[\left(x e^{(r - \sigma^2/2)\tau + \sigma\sqrt{\tau}G} - K\right)^+\right].$$

El integrando es positivo exactamente cuando

$$x e^{(r - \sigma^2/2)\tau + \sigma\sqrt{\tau}G} > K \iff G > \frac{\log(K/x) - (r - \sigma^2/2)\tau}{\sigma\sqrt{\tau}} = -d_-.$$

Separando los dos sumandos,

$$u(t, x) = e^{-r\tau} x e^{(r - \sigma^2/2)\tau} \int_{-d_-}^{\infty} e^{\sigma\sqrt{\tau}y} \frac{e^{-y^2/2}}{\sqrt{2\pi}} dy - e^{-r\tau} K \int_{-d_-}^{\infty} \frac{e^{-y^2/2}}{\sqrt{2\pi}} dy.$$

El segundo término es $K e^{-r\tau} \Phi(d_-)$ por simetría de la normal. Para el primero, completando el cuadrado, $\sigma\sqrt{\tau}y - \frac{y^2}{2} = \frac{\sigma^2\tau}{2} - \frac{(y - \sigma\sqrt{\tau})^2}{2}$, de modo que

$$e^{-\sigma^2\tau/2} \int_{-d_-}^{\infty} e^{\sigma\sqrt{\tau}y} \frac{e^{-y^2/2}}{\sqrt{2\pi}} dy = \int_{-d_-}^{\infty} \frac{e^{-(y - \sigma\sqrt{\tau})^2/2}}{\sqrt{2\pi}} dy = \Phi(d_- + \sigma\sqrt{\tau}) = \Phi(d_+),$$

usando $d_+ = d_- + \sigma\sqrt{\tau}$, que se comprueba directamente en (18.17). Reuniendo términos se obtiene (18.16).

El delta. Por (18.14) y $\tilde{V}_t = e^{-rt}u(t, S_t)$, aplicando Itô a $e^{-rt}u(t, S_t)$ bajo $\mathbb{Q}$ y comparando el coeficiente de $d\tilde{B}_t$,

$$H_t = e^{-rt} \sigma S_t \partial_x u(t, S_t) \implies \beta_t = \frac{H_t}{\sigma \tilde{S}_t} = \partial_x u(t, S_t).$$

Derivando (18.16) (los términos en Φ' se cancelan) se obtiene $\partial_x u = \Phi(d_+)$. ■

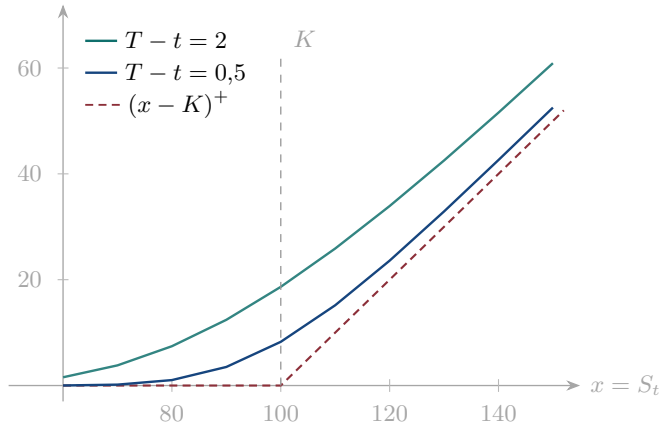


Figura 18.1: La fórmula de Black–Scholes (18.16) para una call con $K = 100$, $r = 5\%$, $\sigma = 25\%$ y dos vencimientos. El precio domina siempre al pago $(x - K)^+$ —la diferencia es el *valor temporal*— y converge a él cuando $T - t \downarrow 0$. La pendiente de cada curva es el delta $\Phi(d_+)$, la cantidad de acciones que hay que mantener en cartera.

Observación 18.9 (μ no aparece). La fórmula (18.16) no depende de μ , el rendimiento esperado del activo. Esto es consecuencia directa de Girsanov: el cambio de medida absorbe μ en el precio de mercado del riesgo θ y lo reemplaza por r . Dos inversores que discrepen completamente sobre el rendimiento futuro del activo —pero coincidan sobre su volatilidad— deben coincidir en el precio de la opción, so pena de arbitraje.

18.5 Del modelo binomial a Black–Scholes

Cerramos con el puente explícito entre la Parte II y la Parte IV.

Teorema 18.10 (Convergencia de Cox–Ross–Rubinstein). *Fijemos $T > 0$, $r \geq 0$, $\sigma > 0$ y, para cada n , consideremos el modelo binomial de n períodos*

de longitud $h = T/n$ con

$$1 + b_n = e^{\sigma\sqrt{h}}, \quad 1 + a_n = e^{-\sigma\sqrt{h}}, \quad \text{tasa por período } r_n = rh,$$

y probabilidad de riesgo neutro $q_n = \frac{e^{r_n} - (1 + a_n)}{(1 + b_n) - (1 + a_n)}$. Entonces el precio CRR ([proposición 7.20](#), ajustado por el descuento) de la call de strike K converge, cuando $n \rightarrow \infty$, al precio de Black–Scholes ([18.16](#)) con $t = 0$.

Idea. Bajo $\mathbb{Q}_n$, $\log(S_T/s) = \sum_{k=1}^n Y_k$ donde los Y_k son i.i.d. con valores $\pm\sigma\sqrt{h}$. Desarrollando,

$$q_n = \frac{e^{rh} - e^{-\sigma\sqrt{h}}}{e^{\sigma\sqrt{h}} - e^{-\sigma\sqrt{h}}} = \frac{1}{2} + \frac{\sqrt{h}}{2\sigma} \left(r - \frac{\sigma^2}{2} \right) + O(h),$$

de donde

$$\mathbb{E}_{\mathbb{Q}}[Y_k] = \left(r - \frac{\sigma^2}{2} \right) h + O(h^{3/2}), \quad \text{Var}_{\mathbb{Q}}(Y_k) = \sigma^2 h + O(h^2).$$

Por el teorema central del límite (versión de Lindeberg para arreglos triangulares),

$$\log \frac{S_T}{s} \xrightarrow{\text{ley}} \mathcal{N}\left(\left(r - \frac{\sigma^2}{2} \right) T, \sigma^2 T \right),$$

que es exactamente la ley de $\log(S_T/s)$ bajo la medida de riesgo neutro de Black–Scholes, por ([18.18](#)). Como la función $x \mapsto (x - K)^+$ es continua y los momentos están uniformemente acotados, las esperanzas convergen. ■

El diccionario completo

	Tiempo discreto (Parte II)	Tiempo continuo (Parte IV)
Ruido	± 1 con probabilidad q	dB_t
Precios	$S_n = S_{n-1}(1 + \rho_n)$	$dS_t = \mu S_t dt + \sigma S_t dB_t$
Estrategia	Γ_n predecible	Γ_t progresivo
Valor	$V_0^\Gamma + (\Gamma \cdot M)_n$	$V_0 + \int_0^t \Gamma_s dM_s$
Cambio de medida	$q = \frac{-a}{b-a}$ (ejemplo 6.14)	Girsanov, $\theta = \frac{\mu-r}{\sigma}$
1.º teo. fundamental	teorema 6.15	teorema 18.7(a)
Representación	proposición 7.19	teorema 18.3
Completitud	teorema 7.16	teorema 18.7(b)
Precio	$\mathbb{E}_Q[F \mathcal{F}_n]$	$\mathbb{E}_Q[e^{-r(T-t)} F \mathcal{F}_t]$
Fórmula cerrada	CRR (proposición 7.20)	Black-Scholes (18.16)

La columna de la derecha no es una teoría nueva: es la de la izquierda, escrita con la integral de Itô en lugar de la transformada de martingalas.

Apéndices

Resultados de teoría de la medida

Reunimos aquí, sin demostración, los resultados de teoría de la medida que se usan a lo largo del texto. Las demostraciones están en cualquier curso estándar; véase por ejemplo Durrett [2], apéndice A, o Billingsley [1].

A.1 Clases de conjuntos y unicidad de medidas

Definición A.1. Sea E un conjunto y $\mathcal{C} \subset \mathcal{P}(E)$.

- $\mathcal{C}$ es un π -sistema si es estable por intersecciones finitas: $A, B \in \mathcal{C} \implies A \cap B \in \mathcal{C}$.
- $\mathcal{M} \subset \mathcal{P}(E)$ es un λ -sistema (o clase monótona) si
 - (I) $E \in \mathcal{M}$;
 - (II) si $A, B \in \mathcal{M}$ y $A \subset B$, entonces $B \setminus A \in \mathcal{M}$;
 - (III) si (A_n) es creciente con $A_n \in \mathcal{M}$, entonces $\bigcup_n A_n \in \mathcal{M}$.

Teorema A.2 (Dynkin, π - λ). Si $\mathcal{C}$ es un π -sistema y $\mathcal{M}$ un λ -sistema con $\mathcal{C} \subset \mathcal{M}$, entonces $\sigma(\mathcal{C}) \subset \mathcal{M}$. En particular, si $\mathcal{M}$ es un λ -sistema que contiene a un π -sistema generador de $\mathcal{F}$, entonces $\mathcal{M} \supset \mathcal{F}$.

Corolario A.3 (Unicidad de medidas). Sean μ, ν dos probabilidades sobre $(E, \mathcal{F})$ y $\mathcal{C} \subset \mathcal{F}$ un π -sistema con $\sigma(\mathcal{C}) = \mathcal{F}$. Si $\mu = \nu$ sobre $\mathcal{C}$, entonces $\mu = \nu$ sobre $\mathcal{F}$.

Corolario A.4 (Criterios de independencia). (a) Sean $\{X_i\}_{i \in I}$ variables aleatorias y $\mathcal{G}$ una σ -álgebra. Para probar que $\sigma(X_i : i \in I)$ y $\mathcal{G}$ son independientes basta ver que $(X_{i_1}, \dots, X_{i_p})$ es independiente de $\mathcal{G}$ para toda elección finita de índices. En efecto, los sucesos que dependen de finitas X_i forman un π -sistema que genera $\sigma(X_i : i \in I)$.

- (b) Sean $\{X_i\}_{i \in I}$ variables aleatorias, Z acotada e $i_0 \in I$. Para ver que $\mathbb{E}[Z | X_i : i \in I] = \mathbb{E}[Z | X_{i_0}]$ basta probar $\mathbb{E}[Z | X_{i_1}, \dots, X_{i_p}] = \mathbb{E}[Z | X_{i_0}]$ para toda elección finita.

Teorema A.5 (Clase monótona funcional). Sea $\mathcal{C}$ un π -sistema sobre E y $\mathcal{H}$ un espacio vectorial de funciones acotadas $E \rightarrow \mathbb{R}$ tal que

- (I) $\mathbf{1}_A \in \mathcal{H}$ para todo $A \in \mathcal{C}$, y $1 \in \mathcal{H}$;
- (II) si $(f_n) \subset \mathcal{H}$ es creciente, no negativa, y $f = \lim_n f_n$ es acotada, entonces $f \in \mathcal{H}$.

Entonces $\mathcal{H}$ contiene a todas las funciones acotadas $\sigma(\mathcal{C})$ -medibles.

A.2 Integración

Teorema A.6 (Convergencia monótona). Si $0 \leq f_n \uparrow f$ puntualmente, entonces $\int f_n d\mu \uparrow \int f d\mu$.

Teorema A.7 (Lema de Fatou). Si $f_n \geq 0$, entonces $\int \liminf_n f_n d\mu \leq \liminf_n \int f_n d\mu$.

Teorema A.8 (Convergencia dominada). Si $f_n \rightarrow f$ c. s. y $|f_n| \leq g$ con $\int g d\mu < \infty$, entonces $\int f_n d\mu \rightarrow \int f d\mu$ y $\int |f_n - f| d\mu \rightarrow 0$.

Teorema A.9 (Tonelli y Fubini). Sean $(E_1, \mathcal{F}_1, \mu_1)$ y $(E_2, \mathcal{F}_2, \mu_2)$ espacios σ -finitos. Si $f \geq 0$ es $\mathcal{F}_1 \otimes \mathcal{F}_2$ -medible (Tonelli), o bien f es $\mu_1 \otimes \mu_2$ -integrable (Fubini), entonces

$$\int f d(\mu_1 \otimes \mu_2) = \int \left(\int f(x, y) \mu_2(dy) \right) \mu_1(dx) = \int \left(\int f(x, y) \mu_1(dx) \right) \mu_2(dy).$$

Teorema A.10 (Desigualdad de Markov). Si $f \geq 0$ y $t > 0$, entonces $\mu\{f \geq t\} \leq \frac{1}{t} \int f d\mu$.

Teorema A.11 (Hölder y Minkowski). Sean $p, q \in (1, \infty)$ con $1/p + 1/q = 1$. Entonces

$$\|fg\|_1 \leq \|f\|_p \|g\|_q, \quad \|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

La desigualdad integral de Minkowski afirma que

$$\left\| \int f(\cdot, y) \mu_2(dy) \right\|_{L^p(\mu_1)} \leq \int \|f(\cdot, y)\|_{L^p(\mu_1)} \mu_2(dy).$$

Teorema A.12 (Continuidad absoluta de la integral). Si $f \in L^1(\mu)$ y $\varepsilon > 0$, existe $\delta > 0$ tal que $\mu(A) < \delta$ implica $\int_A |f| d\mu < \varepsilon$.

A.3 Radon–Nikodym

Definición A.13. Decimos que ν es *absolutamente continua* respecto de μ , $\nu \ll \mu$, si $\mu(A) = 0$ implica $\nu(A) = 0$. Si además $\mu \ll \nu$ decimos que son *equivalentes*, $\mu \sim \nu$.

Teorema A.14 (Radon–Nikodym). Sean μ y ν medidas σ -finitas sobre $(E, \mathcal{F})$ con $\nu \ll \mu$. Entonces existe una función $\mathcal{F}$ -medible $f \geq 0$, única μ -c. s., tal que

$$\nu(A) = \int_A f d\mu \quad \text{para todo } A \in \mathcal{F}.$$

Se escribe $f = \frac{d\nu}{d\mu}$ y se llama derivada de Radon–Nikodym.

Este teorema es el que da existencia a la esperanza condicional (teorema 1.5) y el que permite cambiar de medida en el teorema de Girsanov (teorema 18.1).

A.4 Convergencia y funciones características

Teorema A.15 (Subsucesiones). Si $X_n \xrightarrow{\mathbb{P}} X$, existe una subsucesión (X_{n_k}) con $X_{n_k} \xrightarrow{\text{c. s.}} X$.

Teorema A.16 (Borel–Cantelli). Si $\sum_n \mathbb{P}(A_n) < \infty$, entonces $\mathbb{P}\{\limsup_n A_n\} = 0$. Si además los A_n son independientes y $\sum_n \mathbb{P}(A_n) = \infty$, entonces $\mathbb{P}\{\limsup_n A_n\} = 1$.

Definición A.17. La *función característica* de una variable X con valores en $\mathbb{R}^k$ es

$$\varphi_X(t) = \mathbb{E}\left[e^{i\langle t, X \rangle}\right], \quad t \in \mathbb{R}^k.$$

Teorema A.18 (Unicidad e independencia). φ_X determina la ley de X . Además, las componentes $X_1, \dots, X_k$ son independientes si y solo si

$$\varphi_{(X_1, \dots, X_k)}(t_1, \dots, t_k) = \varphi_{X_1}(t_1) \cdots \varphi_{X_k}(t_k) \quad \text{para todo } t \in \mathbb{R}^k.$$

Teorema A.19 (Lévy). $X_n \xrightarrow{\text{ley}} X$ si y solo si $\varphi_{X_n}(t) \rightarrow \varphi_X(t)$ para todo t . Si $\varphi_{X_n} \rightarrow \psi$ puntualmente con ψ continua en 0, entonces ψ es la función característica de alguna variable X y $X_n \xrightarrow{\text{ley}} X$.

A.5 Lema de Grönwall

Lema A.20 (Grönwall). Sea $h : [0, T] \rightarrow \mathbb{R}_+$ acotada y medible, y supongamos que existe $C \geq 0$ con

$$h(t) \leq a + C \int_0^t h(s) ds \quad \text{para todo } t \in [0, T].$$

Entonces $h(t) \leq a e^{Ct}$ para todo $t \in [0, T]$. En particular, si $a = 0$ entonces $h \equiv 0$.

APÉNDICE B

Síntesis: formulario

Este apéndice recoge, sin demostraciones, los enunciados que conviene tener a mano. Es la versión revisada de la síntesis para el examen parcial.

B.1 Esperanza condicional

Propiedades (proposición 1.9)

Para $X, Y \in L^1$ y $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$:

- $X \in \mathcal{G} \implies \mathbb{E}[X | \mathcal{G}] = X$;
- $X \perp \mathcal{G} \implies \mathbb{E}[X | \mathcal{G}] = \mathbb{E}[X]$;
- linealidad y monotonía;
- torre: $\mathbb{E}[\mathbb{E}[X | \mathcal{G}] | \mathcal{H}] = \mathbb{E}[X | \mathcal{H}]$;
- $\mathbb{E}[\mathbb{E}[X | \mathcal{G}]] = \mathbb{E}[X]$;
- sacar lo conocido: $\mathbb{E}[XY | \mathcal{G}] = X\mathbb{E}[Y | \mathcal{G}]$ si $X \in \mathcal{G}$;
- $|\mathbb{E}[X | \mathcal{G}]| \leq \mathbb{E}[|X| | \mathcal{G}]$;
- Jensen: $\varphi(\mathbb{E}[X | \mathcal{G}]) \leq \mathbb{E}[\varphi(X) | \mathcal{G}]$, φ convexa;
- TCM, Fatou y TCD condicionales;
- en L^2 : proyección ortogonal sobre $L^2(\Omega, \mathcal{G}, \mathbb{P})$.

B.2 Martingalas a tiempo discreto

Convergencia

- **Upcrossings** (proposición 3.2): para (X_n) submartingala y $\alpha < \beta$,
$$(\beta - \alpha)\mathbb{E}[U_n^{\alpha\beta}] \leq \mathbb{E}[(X_n - \alpha)^+] - \mathbb{E}[(X_0 - \alpha)^+] \leq \mathbb{E}[|X_n|] + |\alpha|.$$
- **Convergencia c.s.** (teorema 3.3): si (X_n) es submartingala con $\sup_n \mathbb{E}[X_n^+] < \infty$, existe Z integrable con $X_n \xrightarrow{\text{c.s.}} Z$. Casos particulares: martingala con $\sup_n \mathbb{E}[|X_n|] < \infty$; supermartingala con $\sup_n \mathbb{E}[X_n^-] < \infty$; supermartingala positiva.

- **Vitali** (teorema 4.7): si $X_n \xrightarrow{\mathbb{P}} X$, entonces (X_n) u. i. $\iff X_n \xrightarrow{L^1} X \iff \mathbb{E}[|X_n|] \rightarrow \mathbb{E}[|X|] < \infty$.
- **Último elemento** (teorema 4.12): una martingala tiene último elemento $\iff$ es u. i. $\iff X_n = \mathbb{E}[X_\infty | \mathcal{F}_n]$.
- **Convergencia en L^p** (corolario 5.17): si $\sup_n \mathbb{E}[|M_n|^p] < \infty$ con $p > 1$, entonces $M_n \xrightarrow{\text{c. s.}} y \xrightarrow{L^p}$.

Tiempos de parada y desigualdades

- τ es tiempo de parada si $\{\tau \leq n\} \in \mathcal{F}_n$; y $\mathcal{F}_\tau = \{A : A \cap \{\tau \leq n\} \in \mathcal{F}_n \forall n\}$.
- **Proceso detenido** (proposición 5.5): X^τ conserva el carácter de (sub/super)martingala.
- **Muestreo opcional** (teorema 5.9): con último elemento y $\tau \leq \rho$,

$$\mathbb{E}[X_0] \bowtie \mathbb{E}[X_\tau] \bowtie \mathbb{E}[X_\rho] \bowtie \mathbb{E}[X_\infty], \quad X_\tau \bowtie \mathbb{E}[X_\rho | \mathcal{F}_\tau],$$

con $\bowtie \in \{=, \leq, \geq\}$ según sea martingala, sub o super. Sin hipótesis adicional el resultado es falso (ejemplo 5.8).

- **Doob L^1** (teorema 5.12): para (X_n) submartingala positiva, $\mathbb{P}\{\max_{n \leq N} X_n \geq \lambda\} \leq \mathbb{E}[X_N] / \lambda$.
- **Doob L^p** (teorema 5.15): para $p > 1$ y $q = p/(p-1)$, $\|\max_{n \leq N} X_n\|_p \leq q \|X_N\|_p$.
- **Descomposición de Doob** (teorema 2.18): toda submartingala se escribe de manera única como $X_n = M_n + A_n$ con M martingala y A predecible creciente, $A_0 = 0$.

B.3 Mercados a tiempo discreto

Los dos teoremas fundamentales

Mercado $[\Omega \text{ finito}, (\mathcal{F}_n), (M_n)]$, Γ predecible.

- **Autofinanciamiento**: $\Gamma_n \cdot M_n = \Gamma_{n+1} \cdot M_n$; equivale a $V_n^\Gamma = V_0^\Gamma + (\Gamma \cdot M)_n$.

- **Arbitraje:** $V_0^\Gamma \leq 0$, $V_N^\Gamma \geq 0$, $V_N^\Gamma(\omega) > 0$ para algún ω .
- **PME:** $\mathbb{Q}$ con $\mathbb{Q}(\{\omega\}) > 0$ y (M_n) $\mathbb{Q}$ -martingala.
- **1.º teorema** (teorema 6.15): libre de arbitraje $\iff$ existe PME.
- **Precio justo** (teorema 7.6): $\{\text{precios justos de } F\} = \{\mathbb{E}_{\mathbb{Q}}[F] : \mathbb{Q} \text{ PME}\}$.
- **2.º teorema** (teorema 7.16): completo $\iff$ la PME es única. En tal caso $\pi_n(F) = \mathbb{E}_{\mathbb{Q}}[F | \mathcal{F}_n]$.
- **Modelo binomial** (s, a, b) : libre de arbitraje $\iff a < 0 < b$; en tal caso es completo, con $\mathbb{Q}\{\rho_n = b\} = \frac{-a}{b-a}$ y los ρ_n i.i.d. centrados.
- **Americanas** (teorema 8.12): $\mathcal{E}_N = f_N$, $\mathcal{E}_n = \max\{f_n, \mathbb{E}_{\mathbb{Q}}[\mathcal{E}_{n+1} | \mathcal{F}_n]\}$; precio $= \mathcal{E}_0 = \max_{\tau} \mathbb{E}_{\mathbb{Q}}[f_{\tau}]$, alcanzado en $\tau_* = \min\{n : \mathcal{E}_n = f_n\}$.

B.4 Procesos a tiempo continuo

Poisson y Browniano

	Poisson (λ)	Browniano
Espacio	$\mathbb{D}_{\mathbb{Z}_+}$	$\mathcal{C}_{\mathbb{R}}$
Ley	$\mathbb{P}_{\text{Pois}}^\lambda$	$\mathbb{W}$ (Wiener)
Incrementos	$\text{Poiss}(\lambda(t-s))$	$\mathcal{N}(0, t-s)$
Trayectorias	càdlàg, saltos $+1$	continuas, α -Hölder $\forall \alpha < \frac{1}{2}$
Martingala	$N_t - \lambda t$	$B_t, B_t^2 - t, e^{\sigma B_t - \sigma^2 t/2}$
Variación cuadrática	$\langle N - \lambda \cdot \rangle_t = \lambda t$	$\langle B \rangle_t = t$
Markov fuerte	teorema 11.7	teorema 13.25
Tiempos de espera	ξ_n i.i.d. $\text{Exp}(\lambda)$	—
Ley 0-1	—	Blumenthal (teorema 13.20)

Procesos Gaussianos

- X es vector Gaussiano si $\langle u, X \rangle$ es Gaussiana para todo u .
- Independencia $\iff$ matriz de covarianzas *diagonal* (proposición 12.9). No la identidad.
- La suma de dos Gaussianas es Gaussiana si son *independientes* —o si forman un vector Gaussiano.
- En un espacio Gaussiano: ortogonalidad $\iff$ independencia (teorema 12.14); y $\mathbb{E}[X \mid \sigma(K)] = p_K(X)$ (corolario 12.15).
- El pre-Browniano es el proceso Gaussiano centrado de covarianza $K(s, t) = s \wedge t$; Kolmogorov–Čentsov (teorema 13.10) le da una modificación continua.

B.5 Cálculo estocástico

Fórmulas clave

- **Isometría de Itô:** $\mathbb{E}\left[\left(\int_0^t H dB\right)^2\right] = \mathbb{E}\left[\int_0^t H_s^2 ds\right]$.
- **Variación cuadrática:** $\langle \int_0^\cdot H dB \rangle_t = \int_0^t H_s^2 ds$; $\langle \int H dB, \int K dB \rangle_t = \int_0^t H_s K_s ds$.
- **Fórmula de Itô** (teorema 16.1):

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

Regla mnemotécnica: $(dB_t)^2 = dt$, $dt dB_t = 0$, $(dt)^2 = 0$.

- **Regla del producto:** $d(XY) = X dY + Y dX + d\langle X, Y \rangle$.
- **Exponencial estocástica:** $\mathcal{E}(M)_t = e^{M_t - \langle M \rangle_t / 2}$ resuelve $dZ = Z dM$, $Z_0 = 1$; es martingala local positiva, y martingala verdadera bajo Novikov.
- **Caracterización de Lévy** (teorema 16.11): martingala local continua con $\langle M \rangle_t = t \implies$ es un Browniano.
- **Girsanov** (teorema 18.1): con $d\mathbb{Q}/d\mathbb{P} = \mathcal{E}(-\int \theta dB)_T$, el proceso $\tilde{B}_t = B_t + \int_0^t \theta_s ds$ es $\mathbb{Q}$ -Browniano.

- **Representación** (teorema 18.3): toda martingala browniana en L^2 es $M_0 + \int_0^t H dB$.
- **Feynman–Kac** (teorema 17.9): $\partial_t u + \mathcal{L}u - ru = 0$, $u(T, \cdot) = g \implies u(t, x) = \mathbb{E}\left[e^{-r(T-t)} g(X_T) \mid X_t = x\right]$.
- **Black–Scholes** (teorema 18.8): $u(t, x) = x\Phi(d_+) - Ke^{-r(T-t)}\Phi(d_-)$, con $d_{\pm} = \frac{\log(x/K) + (r \pm \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}$; delta = $\Phi(d_+)$.

Ejercicios adicionales

Los ejercicios que aparecen dentro del texto están resueltos. Los que siguen se dejan propuestos; se indica entre paréntesis el capítulo cuyo material se necesita.

Esperanza condicional y martingalas

1. (*capítulo 1*) Sea $X \in L^1$ y $\mathcal{G} \subset \mathcal{F}$. Pruebe que $\mathbb{E}[X | \mathcal{G}] = X$ c.s. si y solo si X es $\mathcal{G}$ -medible c.s. (es decir, existe $\tilde{X}$ $\mathcal{G}$ -medible con $X = \tilde{X}$ c.s.).
2. (*capítulo 1*) Sea $(\Omega, \mathcal{F}, \mathbb{P}) = ([0, 1], \mathcal{B}_{[0,1]}, \text{Leb})$ y $\mathcal{G}_n$ la filtración diádica. Calcule $\mathbb{E}[X | \mathcal{G}_n]$ para $X(\omega) = \omega$ y describa explícitamente la martingala resultante.
3. (*capítulo 2*) Sea (ξ_n) i.i.d. con $\mathbb{P}\{\xi_n = 1\} = p$, $\mathbb{P}\{\xi_n = -1\} = 1 - p$, $p \neq \frac{1}{2}$, y $S_n = \sum_{j \leq n} \xi_j$. Pruebe que

$$M_n = \left(\frac{1-p}{p} \right)^{S_n}$$

es una martingala. Úsela, con el teorema de muestreo opcional, para calcular la probabilidad de que el paseo alcance $b > 0$ antes que $-a < 0$ (*ruina del jugador*).

4. (*capítulo 2*) Sea (M_n) una martingala con $M_n \in L^2$ e incrementos $d_k = M_k - M_{k-1}$. Pruebe la *ortogonalidad de los incrementos*: $\mathbb{E}[d_j d_k] = 0$ para $j \neq k$, y deduzca $\mathbb{E}[M_n^2] = \mathbb{E}[M_0^2] + \sum_{k \leq n} \mathbb{E}[d_k^2]$.
5. (*capítulo 3*) **Ley fuerte vía martingalas.** Sea (ξ_n) i.i.d. e integrables, $S_n = \sum_{j \leq n} \xi_j$ y $\mathcal{G}_{-n} = \sigma(S_n, S_{n+1}, \dots)$. Pruebe que $\mathbb{E}[\xi_1 | \mathcal{G}_{-n}] = S_n/n$ y deduzca, usando el teorema de convergencia para martingalas inversas, que $S_n/n \rightarrow \mathbb{E}[\xi_1]$ c.s..
6. (*capítulo 4*) Dé un ejemplo de una martingala (M_n) acotada en L^1 que converja c.s. pero no en L^1 , distinto del proceso de ramificación crítico.

Indicación: considere $M_n = \prod_{j \leq n} \xi_j$ con $\xi_j \geq 0$ i.i.d. de media 1 y $\mathbb{P}\{\xi_j = 0\} > 0$.

7. (*capítulo 5*) Sea (Z_n) una supermartingala positiva y $Z_\infty = \lim_n Z_n$. Pruebe que $\mathbb{E}[Z_0] = \mathbb{E}[Z_\infty]$ si y solo si (Z_n) es uniformemente integrable, y que en ese caso (Z_n) es de hecho una martingala cerrada. Exhiba una supermartingala positiva para la que la desigualdad $\mathbb{E}[Z_0] \geq \mathbb{E}[Z_\infty]$ sea estricta.
8. (*capítulo 5*) **Identidad de Wald.** Sea (ξ_n) i.i.d. integrables y τ un tiempo de parada con $\mathbb{E}[\tau] < \infty$. Pruebe que $\mathbb{E}[S_\tau] = \mathbb{E}[\xi_1] \mathbb{E}[\tau]$.

Finanzas

9. (*capítulo 6*) En el modelo binomial con $s = 100$, $a = -10\%$, $b = +20\%$, $N = 3$ y tasa nula, calcule la PME y valore la call europea de strike $K = 110$. Calcule además la cobertura β_n en cada nodo del árbol.
10. (*capítulo 7*) **Paridad put–call.** Pruebe que, en cualquier mercado libre de arbitraje con numerario constante,

$$\pi_0((S_N - K)^+) - \pi_0((K - S_N)^+) = S_0 - K.$$

Deduzca la versión con descuento en el modelo de Black–Scholes: $C_0 - P_0 = S_0 - Ke^{-rT}$.

11. (*capítulo 7*) Valore la opción *lookback* $F = S_N - \min_{n \leq N} S_n$ en el modelo binomial de dos períodos del [ejemplo 7.7](#). Compare con el precio de la call.
12. (*capítulo 8*) Pruebe que el precio de la put americana es, en general, estrictamente mayor que el de la put europea. Exhiba un ejemplo binomial explícito donde el ejercicio anticipado sea óptimo.
13. (*capítulo 8*) Sea (f_n) un contrato americano y $\mathcal{E}$ su envolvente de Snell. Pruebe que

$$\tau^* = \min\{n : A_{n+1} > 0\}$$

—con A el proceso creciente de la descomposición de Doob de $\mathcal{E}$ — es el *mayor* tiempo de ejercicio óptimo, mientras que τ_* es el menor.

Procesos a tiempo continuo

14. (*capítulo 10*) Sea N un proceso de Poisson de intensidad λ . Calcule $\mathbb{E}[N_s N_t]$ para $s \leq t$ y deduzca la función de covarianza de la martingala compensada $\tilde{N}_t = N_t - \lambda t$.
15. (*capítulo 10*) **Adelgazamiento.** Sea N un proceso de Poisson de intensidad λ y marquemos cada salto, independientemente, con probabilidad p . Pruebe que el proceso de saltos marcados y el de saltos no marcados son procesos de Poisson independientes de intensidades λp y $\lambda(1-p)$.
16. (*capítulo 10*) Sea N Poisson de intensidad λ . Condicionado a $\{N_t = n\}$, pruebe que los instantes de salto $(J_1, \dots, J_n)$ tienen la misma ley que los estadísticos de orden de n variables uniformes independientes sobre $[0, t]$.
17. (*capítulo 13*) Pruebe que $\mathbb{E}[\tau_a] = \infty$ para todo $a \neq 0$, donde $\tau_a = \inf\{t : B_t = a\}$, pese a que $\tau_a < \infty$ c. s.. *Indicación:* use el principio de reflexión para obtener la densidad de τ_a , o aplique muestreo opcional a $B_{t \wedge \tau_a}$ con dos barreras.
18. (*capítulo 13*) Sea $S_t = \sup_{s \leq t} B_s$. Usando el [teorema 13.26](#), calcule la densidad conjunta de (S_t, B_t) y deduzca la ley de $S_t - B_t$.
19. (*capítulo 13*) **Ley del arco seno.** Pruebe que $\mathbb{P}\{\text{el último cero de } B \text{ en } [0, 1] \text{ es } \leq u\} = \frac{2}{\pi} \arcsin \sqrt{u}$.

Cálculo estocástico

20. (*capítulo 15*) Calcule directamente $\int_0^t s dB_s$ y verifique con la regla del producto que vale $tB_t - \int_0^t B_s ds$.
21. (*capítulo 16*) Aplique la fórmula de Itô para probar que, para $n \geq 2$,

$$B_t^n = n \int_0^t B_s^{n-1} dB_s + \frac{n(n-1)}{2} \int_0^t B_s^{n-2} ds,$$

y deduzca por recursión que $\mathbb{E}[B_t^{2n}] = \frac{(2n)!}{2^n n!} t^n$.

22. (*capítulo 16*) Sea f armónica en $\mathbb{R}^d$ ($\Delta f = 0$) y B un Browniano d -dimensional. Pruebe que $f(B_t)$ es una martingala local. ¿Qué dice esto sobre la recurrencia del Browniano en dimensión d ?

-
- 23. (capítulo 16) Integral de Stratonovich.** Defina $\int_0^t H \circ dB = \int_0^t H dB + \frac{1}{2}\langle H, B \rangle_t$ y pruebe que, para $f \in C^3$,

$$f(B_t) = f(0) + \int_0^t f'(B_s) \circ dB_s$$

—es decir, la integral de Stratonovich sí obedece la regla de la cadena ordinaria. Explique por qué, pese a ello, se prefiere la de Itô en finanzas.

- 24. (capítulo 17)** Resuelva explícitamente $dX_t = (\alpha - \beta X_t) dt + \sigma \sqrt{X_t} dB_t$ en el caso $\sigma = 0$ y discuta por qué el [teorema 17.2](#) no se aplica directamente cuando $\sigma > 0$ (proceso de Cox–Ingersoll–Ross).
- 25. (capítulo 18)** Use Girsanov para calcular $\mathbb{P}\{\max_{t \leq T} (B_t + \nu t) \geq a\}$ para $\nu \in \mathbb{R}$, $a > 0$ (máximo de un Browniano con deriva). *Indicación:* cambie a la medida bajo la cual $B_t + \nu t$ es un Browniano y aplique el principio de reflexión.
- 26. (capítulo 18)** Deduzca de [\(18.16\)](#) las *griegas* de la call:

$$\Delta = \partial_x u = \Phi(d_+), \quad \Gamma = \partial_{xx} u = \frac{\Phi'(d_+)}{x\sigma\sqrt{T-t}}, \quad \mathcal{V} = \partial_\sigma u = x\sqrt{T-t}\Phi'(d_+),$$

y verifique que satisfacen la ecuación de Black–Scholes [\(18.15\)](#).

- 27. (capítulo 18) Opción binaria.** Valore el contrato $F = \mathbf{1}_{\{S_T \geq K\}}$ en el modelo de Black–Scholes y compruebe que su precio es $e^{-rT}\Phi(d_-)$. Relaciónelo con ∂_K del precio de la call.

Bibliografía

Bibliografía

- [1] P. Billingsley. *Probability and Measure*, 3.^a ed. Wiley, 1995. (Referencia general de teoría de la medida y probabilidad; apéndice A.)
- [2] R. Durrett. *Probability: Theory and Examples*, 5.^a ed. Cambridge University Press, 2019. (Capítulos 4 y 5: martingalas a tiempo discreto. Es la referencia principal de la Parte I.)
- [3] D. Williams. *Probability with Martingales*. Cambridge University Press, 1991. (Tratamiento muy legible de martingalas discretas, upcrossings e integrabilidad uniforme.)
- [4] D. Lamberton, B. Lapeyre. *Introduction to Stochastic Calculus Applied to Finance*, 2.^a ed. Chapman & Hall/CRC, 2007. (Referencia principal de la Parte II: mercados discretos, teoremas fundamentales, envolvente de Snell y opciones americanas.)
- [5] S. E. Shreve. *Stochastic Calculus for Finance I: The Binomial Asset Pricing Model*. Springer, 2004.
- [6] S. E. Shreve. *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer, 2004. (Parte IV, con énfasis en las aplicaciones financieras; capítulos 4–6 para Itô, Girsanov y Black–Scholes.)
- [7] J.-F. Le Gall. *Mouvement brownien, martingales et calcul stochastique*. Springer, 2013. (Traducción inglesa: *Brownian Motion, Martingales, and Stochastic Calculus*, Springer, 2016.) (Referencia principal de las Partes III y IV: procesos Gaussianos, medidas Gaussianas, construcción del movimiento Browniano, Kolmogorov–Čentsov, integral de Itô.)
- [8] D. Revuz, M. Yor. *Continuous Martingales and Brownian Motion*, 3.^a ed. Springer, 1999. (Tratamiento exhaustivo; referencia para tiempos opcionales, procesos progresivos y variación cuadrática.)
- [9] I. Karatzas, S. E. Shreve. *Brownian Motion and Stochastic Calculus*, 2.^a ed. Springer, 1991. (Especialmente los capítulos 1 y 3, para la teoría general de procesos y la integración estocástica.)
- [10] O. Kallenberg. *Foundations of Modern Probability*, 3.^a ed. Springer, 2021. (Medidas aleatorias de Poisson, capítulo 10.)

- [11] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*, 6.^a ed. Springer, 2003. (*Ecuaciones diferenciales estocásticas y Feynman–Kac, capítulo 17, con muchos ejemplos.*)
- [12] P. E. Protter. *Stochastic Integration and Differential Equations*, 2.^a ed. Springer, 2005. (*Integración respecto de semimartingalas generales, con saltos.*)