Lección n° 1: Bases teóricas de los procesos estocásticos AMARUN 2025

El presente curso es una introducción al estudio de las Cadenas de Markov homogéneas definidas en tiempo y espacio discretos. Formalmente, se trata de sucesiones recursivas aleatorias de la forma:

$$X_{n+1} = g(X_n, \varepsilon_{n+1}),$$

donde $(\varepsilon_n)_{n\in\mathbb{N}}$ es una sucesión de variables aleatorias independientes e idénticamente distribuidas (i.i.d.). Este tipo de dependencia secuencial, a diferencia del modelo i.i.d., introduce una estructura probabilística con ciertas propiedades y aplicaciones profundas en modelización.

El concepto de cadena de Markov fue introducido por el matemático ruso Andrei A. Markov alrededor de 1906 [3], a la edad de cincuenta años. Su objetivo era generalizar la ley de los grandes números más allá del caso independiente, en un contexto filosófico donde se debatía el aparente conflicto entre el libre albedrío individual y el determinismo colectivo: ¿cómo es posible que de decisiones individuales aparentemente aleatorias emerja un comportamiento colectivo predecible?

Algunas referencias para profundizar: el libro de [2] tiene un punto de vista muy matemático. Por el contrario, el libro de [4] ofrece un panorama bastante completo, manteniendo un nivel de formalismo relativamente bajo: es ideal para forjarse una intuición sobre el tema. Por último, el libro de [1] es una recopilación de ejercicios bastante útil para repasar (aunque en general sean bastante difíciles). También contiene breves recordatorios teóricos muy útiles (y sobre los cuales se basan muchas de las demostraciones de estas lecciones).

Herramientas

Espacios, conjuntos y funciones medibles

Definición 0.1 (Espacio medible)**.** Un *espacio medible* es un par $(X, \mathcal{M})$ donde:

- X es un conjunto cualquiera.
- $\mathcal{M}$ es una σ-álgebra de subconjuntos de X, es decir, satisface:
  1. $X \in \mathcal{M}$.
  2. $A \in \mathcal{M} \implies X \setminus A \in \mathcal{M}$.
  3. Si $A_1, A_2, \dots \in \mathcal{M}$ entonces $\bigcup_{i=1}^{\infty} A_i \in \mathcal{M}$.

Ejemplo 0.2 (σ-álgebra trivial)**.** Para cualquier conjunto X, la σ-álgebra $\mathcal{M} = \{\emptyset, X\}$ es la más pequeña posible.

Ejemplo 0.3 (σ-álgebra de Borel)**.** En $\mathbb{R}$, la σ-álgebra de Borel $\mathcal{B}(\mathbb{R})$ es generada por todos los intervalos abiertos (a, b) con $a < b$.

Definición 0.4 (Conjunto medible)**.** Dado un espacio medible $(X, \mathcal{M})$, un subconjunto $A \subset X$ es un *conjunto medible* si $A \in \mathcal{M}$. A dichos conjuntos se les puede asignar una medida (o probabilidad) de forma consistente.

Definición 0.5 (Función medible)**.** Sean $(X, \mathcal{M}_X)$ y $(Y, \mathcal{M}_Y)$ dos espacios medibles. Una aplicación

$$f\colon X \longrightarrow Y$$

es $\mathcal{M}_X$–$\mathcal{M}_Y$–*medible* si, para todo conjunto medible $B \in \mathcal{M}_Y$,

$$f^{-1}(B) \;=\; \{\, x \in X : f(x) \in B \,\} \;\in\; \mathcal{M}_X.$$

Ejemplo 0.6 (Función no medible)**.** La función indicadora de un conjunto de Vitali en $\mathbb{R}$ no es Borel-medible respecto a la σ-álgebra de Borel.

Ejemplo 0.7 (Espacio de probabilidad y variable aleatoria)**.** En teoría de la probabilidad tomamos

$$(\Omega, \mathcal{F}, \mathbb{P}),$$

donde $(\Omega, \mathcal{F})$ es un espacio medible y $\mathbb{P}$ una medida de probabilidad. Si además consideramos $(\mathbb{R}, \mathcal{B})$ con $\mathcal{B}$ la σ-álgebra de Borel, una aplicación

$$X \colon \Omega \longrightarrow \mathbb{R}$$

es una *variable aleatoria* exactamente si es $\mathcal{F}$–$\mathcal{B}$-medible. Por ejemplo, el evento $\{\omega : X(\omega) \leq 2\}$ es medible porque $(-\infty, 2] \in \mathcal{B}$ implica $\{\omega : X(\omega) \leq 2\} = X^{-1}((-\infty, 2]) \in \mathcal{F}$, y por tanto tiene probabilidad bien definida.

Ley (distribución) de una variable aleatoria

Definición 0.8 (Ley (o distribución) de una variable aleatoria)**.** Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y

$$X \colon \Omega \longrightarrow E$$

una variable aleatoria con espacio de estados $(E, \mathcal{E})$. La *ley* (o *distribución*) de X es la medida de probabilidad

$$\mu_X \ : \ \mathcal{E} \longrightarrow [0,1], \qquad \mu_X(A) \ = \ \mathbb{P}\big(X \in A\big) \ = \ \mathbb{P}\big(\{\omega : X(\omega) \in A\}\big).$$

Observación 0.9. En muchos textos, a μ_X se le llama *distribución* de X, y sus sinónimos son *ley* o *medida inducida por* X. Todos estos términos designan la misma construcción.

Definición 0.10 (Función de distribución acumulada)**.** Si X toma valores en $\mathbb{R}$ y $\mathcal{E}$ es la σ-álgebra de Borel, la *función de distribución acumulada* (o *FDC*) de X es

$$F_X(x) \ = \ \mathbb{P}\big(X \leq x\big), \quad x \in \mathbb{R}.$$

Esta función caracteriza completamente la ley de X.

Ejemplo 0.11 (Variable discreta)**.** Si X es discreta con valores en $\{x_1, x_2, \dots\}$ y $\mathbb{P}(X = x_i) = p_i$, entonces su ley es

$$\mu_X(\{x_i\}) = p_i, \quad \sum_i p_i = 1,$$

y su función de distribución acumulada es

$$F_X(x) = \sum_{i \colon x_i \leq x} p_i.$$

Ejemplo 0.12 (Variable continua)**.** Si X es absolutamente continua con densidad $f_X(x)$, entonces

$$\mu_X(A) = \int_A f_X(x)\,\mathrm{d}x, \quad F_X(x) = \int_{-\infty}^{x} f_X(t)\,\mathrm{d}t.$$

Ejemplo 0.13 (Contraejemplo: distribución singular)**.** La distribución de Cantor tiene función de distribución F_X continua pero no absolutamente continua (no tiene densidad respecto a la medida de Lebesgue).

1. Procesos estocásticos

En todo el curso, consideraremos un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, un espacio medible $(E, \mathcal{E})$ y un conjunto $\mathcal{T}$.

Definición 1.1 (Proceso estocástico)**.** Se llama proceso estocástico, o proceso aleatorio, a una familia $(X_t)_{t \in \mathcal{T}}$ de variables aleatorias con valores en E. Es decir, para todo $t \in \mathcal{T}$, la aplicación $\omega \mapsto X_t(\omega)$ es medible de $(\Omega, \mathcal{F})$ a $(E, \mathcal{E})$. Se llama E el espacio de estados del proceso.

En la práctica, se suelen utilizar los procesos para la modelización dinámica: X_t es el valor de una variable de interés en la fecha t. El conjunto $\mathcal{T}$ representa entonces el conjunto de fechas posibles.

Definición 1.2. Cuando $\mathcal{T} = \mathbb{N}$ o $\mathcal{T} = \mathbb{Z}$ se dice que $(X_t)_{t\in\mathcal{T}}$ es un proceso en tiempo discreto. Cuando $\mathcal{T} = \mathbb{R}$, o un intervalo de $\mathbb{R}$, se habla de proceso en tiempo continuo.

Ejemplo 1.3. Un ejemplo en tiempo discreto: sea X_t el PIB de Ecuador en el año t (entonces $E = \mathbb{R}_+$). Por el contrario, en finanzas, los precios de los activos, divisas, etc. se actualizan con una frecuencia tan alta que se prefiere utilizar una modelización en tiempo continuo: sea X_t el tipo de cambio EURO/DÓLAR en la fecha t, t medido en horas con $t = 0$ correspondiente al 1 de enero de 1999 a las 0h.

Por lo tanto, se puede ver un proceso como una función de dos variables:

$$\begin{aligned} \Omega \times \mathcal{T} &\to E \\ (\omega, t) &\mapsto X_t(\omega) \end{aligned}$$

que debe verificar la condición de que, para t fijo, $\omega \mapsto X_t(\omega)$ es medible (es decir, X_t cumple la definición de variable aleatoria). Para ω fijo, la función $t \mapsto X_t(\omega)$ se llama trayectoria del proceso.

Definición 1.4. Se llama filtración a una familia $(\mathcal{F}_t)_{t\in\mathcal{T}}$ de σ-álgebras que verifica

$$s \leq t \Rightarrow \mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$$

Cuando se observa un proceso a lo largo del tiempo, en la fecha t, se conocen los valores de X_s para $s \leq t$, pero aún no se conocen los valores de X_s para $s > t$. En términos de condicionamiento, eso significa que a menudo se condicionará por las variables $(X_s)_{s\leq t}$, o de forma equivalente por la σ-álgebra $\mathcal{F}_t := \sigma(X_s, s \leq t)$. Se verifica inmediatamente que $(\mathcal{F}_t)_{t\in\mathcal{T}}$ es una filtración, llamada filtración canónica.

Así, la idea de una filtración $(\mathcal{F}_t)_{t\in\mathcal{T}}$ es representar la información disponible en la fecha t. Puede surgir la pregunta natural: ¿por qué introducir un concepto general de filtración en lugar de utilizar siempre la filtración canónica? Intuitivamente: en la fecha t, se puede tener más información que los valores pasados $(X_s)_{s\leq t}$. En el ejemplo donde X_t es el PIB de Ecuador en el año t, al final del año t, se conoce el PIB de Ecuador en ese año, pero también el de Estaods Unidos, o el de Francia, o de otros países de la zona europea, el precio del petróleo, los diferentes tipos de cambio, etc., que pueden dar información sobre los valores futuros del PIB de Ecuador. Más formalmente, se pueden tener varios procesos definidos en el mismo $(\Omega, \mathcal{F}, \mathbb{P})$, por ejemplo $(X_t)_{t\in\mathcal{T}}$ y $(Y_t)_{t\in\mathcal{T}}$, y considerar la filtración $\mathcal{F}_t := \sigma(X_s, Y_s, s \leq t)$ que NO es la filtración canónica para el proceso (X_t).

Definición 1.5. El proceso $(X_t)_{t\in\mathcal{T}}$ se dice adaptado a la filtración $(\mathcal{F}_t)_{t\in\mathcal{T}}$ si para todo $t \in \mathcal{T}$, X_t es $\mathcal{F}_t$-medible.

Ejemplo 1.6. En este ejemplo $\mathcal{T} = \mathbb{N}$, sea $(\varepsilon_t)_{t\in\mathbb{N}}$ una sucesión de variables aleatorias i.i.d. de ley $\mathcal{N}(0,1)$, $(\alpha, \beta) \in \mathbb{R}^2$, $X_0 = 0$ y

$$X_{t+1} = \alpha X_t + \beta + \varepsilon_t.$$

Este proceso se estudia en el curso de series temporales bajo el nombre de proceso autorregresivo (AR). Se define $\mathcal{F}_t = \sigma(\varepsilon_s, s \leq t)$ la filtración canónica para $(\varepsilon_t)_{t\in\mathbb{N}}$.

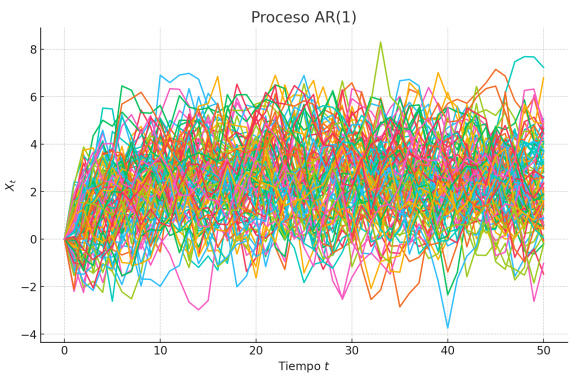

Figura 1: Proceso AR (1)

En la Figura 1 se observa que, tras unos pocos pasos, las trayectorias tienden a oscilar alrededor de un valor medio. En nuestro caso, como definimos

$$X_{t+1} = 0.8\, X_t + 0.5 + \varepsilon_{t+1},$$

la constante $\beta = 0.5$ "lleva" ligeramente el proceso hacia valores positivos, mientras que el coeficiente $\alpha = 0.8$ hace que cada paso dependa fuertemente del anterior. Por otro lado, los "picos" y "valles" de cada trayectoria no crecen indefinidamente: la influencia de $\alpha < 1$ amortigua las oscilaciones, y por eso las curvas permanecen en un intervalo relativamente estrecho. Se observa que la nube de trayectorias forma una banda con anchura aproximadamente constante. Finalmente, tras cierto tiempo (llamado "periodo de establecimiento"), la distribución de X_t se estabiliza: la media y la varianza de las trayectorias dejan de crecer, y todas oscilan alrededor de un valor cercano a

$$\frac{\beta}{1-\alpha} = \frac{0.5}{1-0.8} = 2.5.$$

Se puede verificar de manera trivial que el proceso $(X_t)_{t\in\mathbb{N}}$ es adaptado a la filtración $(\mathcal{F}_t)_{t\in\mathbb{N}}$.

Ejemplo 1.7. Sea $\mathcal{T} = \mathbb{N}$. Consideremos una sucesión $\big(\varepsilon_t\big)_{t\in\mathbb{N}}$ de variables aleatorias i.i.d. con

$$\mathbb{P}(\varepsilon_t = +1) = \mathbb{P}(\varepsilon_t = -1) = \tfrac{1}{2}.$$

Definimos la *trayectoria* del paseo aleatorio (*Random Walk*) por

$$X_0 = 0, \qquad X_{t+1} = X_t + \varepsilon_{t+1}, \quad t \geq 0.$$

Aquí $E = \mathbb{Z}$ y, para cada t, el valor X_t indica la posición tras t pasos. Tomamos la filtración canónica

$$\mathcal{F}_t = \sigma\big(\varepsilon_1, \ldots, \varepsilon_t\big),$$

de modo que es inmediato ver que X_t es $\mathcal{F}_t$-medible y, por tanto, el proceso $\{X_t\}_{t\in\mathbb{N}}$ está adaptado.

Este paseo aleatorio sirve como modelo básico en física estadística (difusión de partículas) y en finanzas (modelos simplificados de precios discretos).

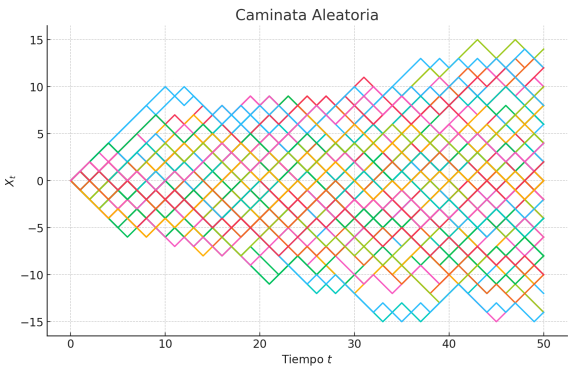

Figura 2: Caminata aleatoria (Random Walk)

A diferencia del AR(1), en la Figura 2 no hay tendencia a volver a ningún nivel. Cada paso es ± 1 con igual probabilidad, de modo que el proceso "no recuerda" la posición previa más allá de sumarle o restarle uno. Se aprecia claramente que conforme t crece, la dispersión entre trayectorias se hace mayor: algunas se alejan mucho hacia arriba, otras hacia abajo. Esto refleja que la varianza de un *random walk* crece linealmente con el tiempo. Finalmente, se puede visualizar que no existe un valor de equilibrio: la media es siempre cero, pero la amplitud de las oscilaciones crece sin límite. Gráficamente, la nube de trayectorias se expande con el tiempo, formando un "abanico" de caminos cada vez más separados.

Ejemplo 1.8 (Caminata aleatoria gaussiana)**.** Consideramos un proceso estocástico $\{X_t\}_{t\in\mathbb{N}}$ definido por:

$$X_0 = 0, \quad X_{t+1} = X_t + \varepsilon_{t+1},$$

donde $\{\varepsilon_t\}_{t\in\mathbb{N}}$ es una sucesión de variables aleatorias independientes e idénticamente distribuidas con ley normal estándar:

$$\varepsilon_t \sim \mathcal{N}(0,1).$$

Este proceso se conoce como *caminata aleatoria gaussiana* (o *Gaussian random walk*).

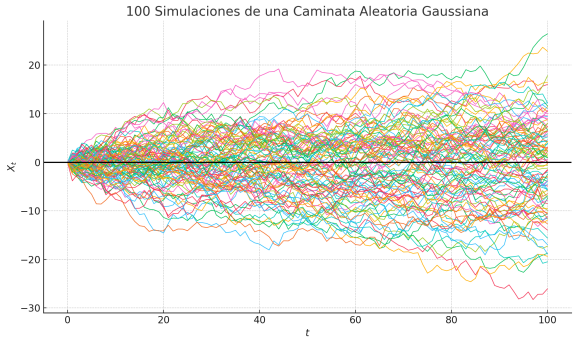

Figura 3: Simulación de una caminata aleatoria gaussiana con 100 trayectorias.

La Figura 3 muestra 100 trayectorias simuladas de la caminata aleatoria gaussiana. Podemos observar que: todas las trayectorias parten de $X_0 = 0$, los valores futuros X_t se obtienen acumulando incrementos gausianos independientes. Por lo tanto, X_t sigue una distribución normal:

$$X_t \sim \mathcal{N}(0,t).$$

La media se mantiene constante en cero, pero la *varianza crece linealmente* con el tiempo. Por eso las trayectorias se dispersan cada vez más; además, este proceso no es estacionario, ya que su distribución depende de t. La caminata aleatoria gaussiana sirve como base para el estudio del *movimiento browniano*, que es el análogo en tiempo continuo.

1.1. Recordatorios sobre la esperanza condicional

Recordamos que para dos eventos A y B en $\mathcal{F}$, con $\mathbb{P}(B) \neq 0$, por definición

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Sean dos variables aleatorias reales definidas sobre $(\Omega, \mathcal{F})$, X y Y. Suponemos que X es integrable. En un primer momento, vamos a suponer que Y es una variable aleatoria discreta, con valores en $\mathbb{N}$. Suponemos también que para todo $y \in \mathbb{N}$, $\mathbb{P}(Y = y) > 0$. Entonces, para todo $A \in \mathcal{F}$,

$$\mathbb{P}(A \mid Y = y) = \frac{\mathbb{P}(A \cap \{Y = y\})}{\mathbb{P}(Y = y)}$$

lo que define una medida de probabilidad sobre $(\Omega, \mathcal{F})$, llamémosla $\mathbb{P}_y$, y es natural definir

$$\mathbb{E}(X \mid Y = y) := \int_\Omega X(\omega)\mathbb{P}_y(\,\mathrm{d}\omega) = \int X \,\mathrm{d}\mathbb{P}_y. \tag{1.1}$$

Se verifica directamente que

$$\mathbb{E}(X \mid Y = y) = \frac{\int X 1_{\{Y=y\}} \mathrm{d}\mathbb{P}}{\mathbb{P}(Y = y)} \tag{1.2}$$

Lo importante es que se ve, ya sea en (1.1) o (1.2), que $\mathbb{E}(X \mid Y = y)$ es una función de y, llamémosla $f(y)$. Entonces, simplemente se define $\mathbb{E}(X \mid Y) := f(Y)$. Atención: $\mathbb{E}(X)$ es un número real, determinista, mientras que $\mathbb{E}(X \mid Y) = f(Y)$ es una variable aleatoria, función de la variable Y, es decir, $\sigma(Y)$-medible.

El problema de esta construcción intuitiva es que no se puede extender al caso donde Y es una variable continua. Por lo tanto, hay que encontrar una caracterización de $\mathbb{E}(X \mid Y)$ que no implique dividir por $\mathbb{P}(Y = y)$ y pueda extenderse al caso general. Vamos a demostrar la siguiente propiedad: para toda variable Z acotada que sea además $\sigma(Y)$-medible, es decir, aquí, $Z = g(Y)$, se tiene

$$\mathbb{E}[\mathbb{E}(X \mid Y)Z] = \mathbb{E}(XZ). \tag{1.3}$$

En efecto:

$$\begin{aligned}
\mathbb{E}[\mathbb{E}(X \mid Y)Z] &= \mathbb{E}[\mathbb{E}(X \mid Y)g(Y)] \\
&= \mathbb{E}\left[\sum_{y=0}^{\infty} \mathbb{E}(X \mid Y = y)g(y)\mathbf{1}_{Y=y}\right] \\
&= \sum_{y=0}^{\infty} g(y)\mathbb{E}(X \mid Y = y)\mathbb{P}(Y = y) \\
&= \sum_{y=0}^{\infty} g(y)\mathbb{E}\left(X\mathbf{1}_{\{Y=y\}}\right) \\
&= \mathbb{E}\left[\sum_{y=0}^{\infty} X1_{\{Y=y\}}g(y)\right] \\
&= \mathbb{E}[Xg(Y)] \\
&= \mathbb{E}(XZ).
\end{aligned}$$

El siguiente teorema dice justamente que la propiedad (1.3) puede extenderse al caso general.

Teorema 1.9. *Sea $\mathcal{A} \subset \mathcal{F}$ una σ-álgebra y X una variable aleatoria real, positiva o integrable. Entonces existe una variable aleatoria $\mathcal{A}$-medible, positiva o integrable, única salvo en un p.c. (casi seguro), denotada $\mathbb{E}(X \mid \mathcal{A})$, tal que para toda variable aleatoria Z, $\mathcal{A}$-medible, positiva o acotada, se tiene*

$$\mathbb{E}[\mathbb{E}(X \mid \mathcal{A})Z] = \mathbb{E}(XZ). \tag{1.4}$$

En el caso donde $\mathcal{A} = \sigma(Y)$ para una variable aleatoria Y, simplemente se escribe $\mathbb{E}(X \mid Y)$ en vez de $\mathbb{E}(X \mid \sigma(Y))$.

Demostración. Idea general: se comienza por tratar el caso donde $X \in \mathcal{L}_2$, el conjunto de variables de cuadrado integrable. Entonces

1. Se verifica que $\langle X, Y\rangle := \mathbb{E}(XY)$ es un producto escalar que convierte a $\mathcal{L}^2$ en un espacio de Hilbert.
2. Se verifica que el conjunto de variables $\mathcal{A}$-medibles y en $\mathcal{L}_2$ es un subespacio vectorial cerrado de $\mathcal{L}_2$. Se puede entonces definir la proyección ortogonal sobre este subespacio, Π.
3. La variable $X - \Pi(X)$ es entonces ortogonal a toda variable Z $\mathcal{A}$-medible y en $\mathcal{L}_2$. Esto se traduce por

$$\langle X - \Pi(X), Z\rangle = \mathbb{E}\{[X - \Pi(X)]Z\}$$

es decir,

$$\mathbb{E}(XZ) = \mathbb{E}[\Pi(X)Z].$$

Por lo tanto, tomando $\mathbb{E}(X \mid \mathcal{A}) := \Pi(X)$, la relación (1.4) se satisface inmediatamente.

Resta luego trabajar un poco para extender esta definición al caso donde X es simplemente integrable ($X \in \mathcal{L}_1$) o positiva.

□

Recordemos algunas propiedades que se utilizarán constantemente a lo largo de este curso. Sean X y Y dos variables aleatorias reales, λ y μ dos números reales, $\mathcal{A}$ y $\mathcal{B}$ dos σ-álgebras, $(X_n)_{n\in\mathbb{N}}$ una sucesión de variables aleatorias, y f una función $\mathbb{R} \to \mathbb{R}$ (c.s para nota convergencia casi segura (*presque surement* en francés)) [5].

1. $\mathbb{E}(\lambda \mid \mathcal{A}) = \lambda$.
2. $\mathbb{E}(\lambda X + \mu Y \mid \mathcal{A}) = \lambda\mathbb{E}(X \mid \mathcal{A}) + \mu\mathbb{E}(Y \mid \mathcal{A})$.

3. X independiente de $\mathcal{A} \Rightarrow \mathbb{E}(X \mid \mathcal{A}) = \mathbb{E}(X)$.
4. X $\mathcal{A}$-medible $\Rightarrow \mathbb{E}(X \mid \mathcal{A}) = X$ y $\mathbb{E}(XY \mid \mathcal{A}) = X\mathbb{E}(Y \mid \mathcal{A})$.
5. $\mathbb{E}[\mathbb{E}(X \mid \mathcal{A})] = \mathbb{E}(X)$.
6. $\mathcal{B} \subset \mathcal{A} \Rightarrow \mathbb{E}[\mathbb{E}(X \mid \mathcal{A}) \mid \mathcal{B}] = \mathbb{E}(X \mid \mathcal{B})$.
7. $X \leq Y$ c.s. $\Rightarrow \mathbb{E}(X \mid \mathcal{A}) \leq \mathbb{E}(Y \mid \mathcal{A})$ c.s.
8. (Jensen condicional) f convexa $\Rightarrow f(\mathbb{E}(X \mid \mathcal{A})) \leq \mathbb{E}(f(X) \mid \mathcal{A})$.
9. (TCD condicional) $X_n \xrightarrow{\text{c.s.}} X, |X_n| \leq Z, Z \in L_1 \Rightarrow \mathbb{E}(X_n \mid \mathcal{A}) \xrightarrow{\text{c.s.}} \mathbb{E}(X \mid \mathcal{A})$.
10. (TCM condicional) $X_n \xrightarrow{\nearrow,\text{c.s.}} X, X_n \geq 0 \Rightarrow \mathbb{E}(X_n \mid \mathcal{A}) \xrightarrow{\nearrow,\text{c.s.}} \mathbb{E}(X \mid \mathcal{A})$.
11. (Fatou condicional) $X_n \geq 0 \Rightarrow \mathbb{E}\left(\liminf X_n \mid \mathcal{A}\right) \leq \liminf \mathbb{E}\left(X_n \mid \mathcal{A}\right)$.

Referencias

[1] P. Baldi, L. Mazliak y P. Priouret. *Martingales and Markov Chains*. Solved exercises and elements of theory, Translated from the 1998 French original. Boca Raton, FL: Chapman & Hall/CRC, 2002.

[2] P. Brémaud. *Markov Chains*. Vol. 31. Texts in Applied Mathematics. Gibbs fields, Monte Carlo simulation, and queues. New York: Springer-Verlag, 1999.

[3] Djalil Chafaï et al. *Processus stochastiques: Ébauche de notes de cours de M1*. Département de mathématiques et applications (DMA), École normale supérieure – PSL, Paris. Équipe pédagogique: Djalil Chafaï (cours) y Alexis Metz–Donnadieu (TD), revisado por Laëtitia Comminges y Arnaud Guyad. 2024.

[4] G. F. Lawler. *Introduction to Stochastic Processes*. 2nd. Boca Raton, FL: Chapman & Hall/CRC, 2006.

[5] J. R. Norris. *Markov Chains*. A fundamental reference for discrete-time Markov chains, ergodic theory, and applications. Cambridge: Cambridge University Press, 1998. ISBN: 978-0-521-55289-0. URL: https://doi.org/10.1017/CBO9780511810633.