

Ampliación de Teoría de la Probabilidad

Universidad de Málaga

Grado en Matemáticas

Curso 2024-2025

Contenidos

1. Introducción	5
1.1. Espacio de probabilidad	5
1.2. Variables aleatorias	7
1.3. Variables aleatorias discretas	10
1.4. Variables aleatorias absolutamente continuas	11
1.5. Vectores aleatorios	13
1.6. Independencia de sucesos y variables aleatorias	13
1.7. Esperanza	14
2. Convolución	17
2.1. Convolución de medidas de probabilidad	17
2.2. Convolución de funciones de distribución	17
2.3. Convolución y funciones de densidad	17
3. Función característica	21
3.1. Propiedades elementales y ejemplos	21
3.2. Reconocimiento de funciones características	24
3.3. Teoremas de inversión	26
3.4. Función característica y momentos	26
4. Convergencia de variables aleatorias	29
4.1. Convergencia débil y en distribución	29
4.2. Convergencia vaga	33
4.3. Ley 0-1 de Kolmogorov	34
4.4. Convergencia casi segura	36
4.5. Convergencia en probabilidad	37
4.6. Leyes de los grandes números	37

Introducción

Este tema no es más que un recordatorio de varias nociones elementales de la teoría de la probabilidad que se deben haber estudiado ya en unas cuantas asignaturas distintas.

1.1. Espacio de probabilidad

Definición 1.1.1. Sea Ω un conjunto no vacío y sea $\mathcal{A} \subset \mathcal{P}(\Omega)$. Se dice que \mathcal{A} es una σ -álgebra sobre Ω si se verifica lo siguiente:

- (a) $\emptyset \in \mathcal{A}$.
- (b) Si $A \in \mathcal{A}$, entonces $A^c \in \mathcal{A}$.
- (c) Si $\{A_i\}_{i=1}^{\infty}$ es una sucesión de elementos de \mathcal{A} , entonces $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

La dupla (Ω, \mathcal{A}) se denomina *espacio medible*, y los elementos de \mathcal{A} , *conjuntos medibles*.

Definición 1.1.2. Sea (Ω, \mathcal{A}) un espacio medible. Una aplicación $\mu: \mathcal{A} \rightarrow [0, \infty]$ se dice que es una *medida* si se verifica lo siguiente:

- (a) $\mu(\emptyset) = 0$.
- (b) Si $\{A_i\}_{i=1}^{\infty}$ es una sucesión de conjuntos medibles disjuntos dos a dos, entonces

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

La terna $(\Omega, \mathcal{A}, \mu)$ se denomina *espacio de medida*.

Proposición 1.1.3. Sea $(\Omega, \mathcal{A}, \mu)$ un espacio de medida.

- (a) Si $A, B \in \mathcal{A}$ y $A \subset B$, entonces $\mu(A) \leq \mu(B)$.
- (b) Si $A, B \in \mathcal{A}$, $A \subset B$ y $\mu(A) < \infty$, entonces $\mu(B \setminus A) = \mu(B) - \mu(A)$.
- (c) Si $\{A_i\}_{i=1}^{\infty}$ es una sucesión de conjuntos medibles tal que $A_i \subset A_{i+1}$ para todo $i \in \mathbb{N}$, entonces

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{i \rightarrow \infty} \mu(A_i).$$

- (d) Si $\{A_i\}_{i=1}^{\infty}$ es una sucesión de conjuntos medibles tal que $A_i \supset A_{i+1}$ para todo $i \in \mathbb{N}$ y $\mu(A_1) < \infty$, entonces

$$\mu\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{i \rightarrow \infty} \mu(A_i).$$

(e) Si $\{A_i\}_{i=1}^{\infty}$ es una sucesión de conjuntos medibles, entonces

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu(A_i).$$

Definición 1.1.4. Sea $(\Omega, \mathcal{A}, \mu)$ un espacio de medida.

- (a) Si existe una sucesión $\{A_i\}_{i=1}^{\infty}$ de elementos de \mathcal{A} con $\Omega = \bigcup_{i=1}^{\infty} A_i$ y $\mu(A_i) < \infty$ para cada $i \in \mathbb{N}$, se dice que μ es una *medida σ -finita* y que $(\Omega, \mathcal{A}, \mu)$ es un *espacio de medida σ -finita*.
- (b) Si $\mu(\Omega) < \infty$, se dice que μ es una *medida finita* y que $(\Omega, \mathcal{A}, \mu)$ es un *espacio de medida finita*.
- (c) Si $\mu(\Omega) = 1$, se dice que μ es una *medida de probabilidad* y que $(\Omega, \mathcal{A}, \mu)$ es un *espacio de probabilidad*.

Por motivos evidentes, en un espacio de probabilidad la medida se suele denotar por P en lugar de μ . Además, a los conjuntos medibles de un espacio de probabilidad también se les conoce como *sucesos*.

Ejemplo 1.1.5. Sea Ω un conjunto finito y sea $\mathcal{A} = \mathcal{P}(\Omega)$. La aplicación $\mu: \mathcal{A} \rightarrow [0, 1]$ definida por

$$\mu(A) = \frac{\#A}{\#\Omega}$$

es una medida de probabilidad, conocida como *medida uniforme*.

Ejemplo 1.1.6. Sea Ω un conjunto numerable, sea $\mathcal{A} = \mathcal{P}(\Omega)$, sea $\{p_{\omega}\}_{\omega \in \Omega}$ una sucesión de números reales no negativos y sea $\mu: \mathcal{A} \rightarrow [0, \infty]$ la función definida por

$$\mu(A) = \sum_{\omega \in A} p_{\omega}.$$

Se prueba fácilmente que μ es una medida (no necesariamente de probabilidad). En caso de que sea $p_{\omega} = 1$ para todo $\omega \in \Omega$, a μ se le conoce como *medida de conteo*.

Ejemplo 1.1.7. Sea Ω un conjunto no vacío, sea $\mathcal{A} = \mathcal{P}(\Omega)$, sea $\omega \in \Omega$ y sea $\delta_{\omega}: \mathcal{A} \rightarrow [0, 1]$ la función dada por

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

Se prueba fácilmente que δ_{ω} es una medida de probabilidad, conocida como *delta de Dirac*.

Definición 1.1.8. Sea Ω un conjunto no vacío y sea $\mathcal{E} \subset \mathcal{P}(\Omega)$. A la menor σ -álgebra que contiene a \mathcal{E} se le conoce como *σ -álgebra generada por \mathcal{E}* , y se le denota por $\sigma(\mathcal{E})$.

Se demuestra que la intersección de una familia arbitraria de σ -álgebras sobre un conjunto no vacío Ω es también una σ -álgebra sobre Ω , de donde se deduce que $\sigma(\mathcal{E})$ es la intersección de todas las σ -álgebras sobre Ω que contienen a \mathcal{E} .

Definición 1.1.9. Dado un espacio topológico (Ω, τ) , a la σ -álgebra generada por τ se le conoce como *σ -álgebra de Borel de Ω* , y se le denota por \mathcal{B}_{Ω} .

Siempre que se hable de la σ -álgebra de Borel de \mathbb{R} o de \mathbb{R}^n , se entenderá que la topología que hay detrás es la topología usual. A $\mathcal{B}_{\mathbb{R}}$ y $\mathcal{B}_{\mathbb{R}^n}$ se les suele denotar por \mathcal{B} y \mathcal{B}^n , respectivamente.

Proposición 1.1.10. *La σ -álgebra de Borel de \mathbb{R} junto con la topología usual está generada por cada una de las siguientes familias de intervalos:*

- (a) $\mathcal{E}_1 = \{(a, b): a, b \in \mathbb{R}\}.$
- (b) $\mathcal{E}_2 = \{[a, b): a, b \in \mathbb{R}\}.$
- (c) $\mathcal{E}_3 = \{(a, b]: a, b \in \mathbb{R}\}.$
- (d) $\mathcal{E}_4 = \{[a, b]: a, b \in \mathbb{R}\}.$
- (e) $\mathcal{E}_5 = \{(a, \infty): a \in \mathbb{R}\}.$
- (f) $\mathcal{E}_6 = \{[a, \infty): a \in \mathbb{R}\}.$
- (g) $\mathcal{E}_7 = \{(-\infty, b): b \in \mathbb{R}\}.$
- (h) $\mathcal{E}_8 = \{(-\infty, b]: b \in \mathbb{R}\}.$

Por la densidad de \mathbb{Q} en \mathbb{R} , este resultado sigue valiendo si todos los extremos de los intervalos se toman en \mathbb{Q} . En \mathbb{R}^n se tiene un resultado totalmente análogo.

En un espacio de medida $(\Omega, \mathcal{A}, \mu)$ verificando ciertas hipótesis, no va a ser necesario dar la imagen por μ de todos los elementos de \mathcal{A} para que μ quede totalmente determinada.

Definición 1.1.11. Sea Ω un conjunto no vacío y sea $\mathcal{E} \subset \mathcal{P}(\Omega)$. Se dice que \mathcal{E} es un π -sistema si para todos $A, B \in \mathcal{E}$ se tiene que $A \cap B \in \mathcal{E}$.

Teorema 1.1.12. Sea (Ω, \mathcal{A}) un espacio medible y sean $\mu_1, \mu_2: \mathcal{A} \rightarrow [0, \infty]$ dos medidas tales que $\mu_1(\Omega) = \mu_2(\Omega) < \infty$. Si \mathcal{E} es un π -sistema con $\sigma(\mathcal{E}) = \mathcal{A}$ y $\mu_1 = \mu_2$ en \mathcal{E} , entonces $\mu_1 = \mu_2$ en \mathcal{A} .

En particular, este resultado afirma que en un espacio de probabilidad (Ω, \mathcal{A}, P) , la medida de probabilidad queda totalmente determinada por los valores que toma en cualquier π -sistema que genere \mathcal{A} .

Ejemplo 1.1.13. Considérese el espacio medible $(\mathbb{R}, \mathcal{B})$, y sea $\mathcal{E} = \{(-\infty, x]: x \in \mathbb{R}\}$. Como \mathcal{E} es un π -sistema que genera \mathcal{B} , cualquier medida de probabilidad sobre \mathcal{B} queda totalmente determinada por los valores que toma en intervalos de la forma $(-\infty, x], x \in \mathbb{R}$.

El ejemplo que sigue muestra que la condición de que la familia \mathcal{E} que genera \mathcal{A} sea cerrada para intersecciones es fundamental para que el teorema anterior sea cierto.

Ejemplo 1.1.14. Sean $\Omega = \{1, 2, 3, 4\}$, $\mathcal{A} = \mathcal{P}(\Omega)$ y $\mathcal{E} = \{\{1, 2\}, \{2, 3\}\}$. Es fácil ver que $\sigma(\mathcal{E}) = \mathcal{A}$, y como $\{2\} \notin \mathcal{E}$, entonces \mathcal{E} no es cerrado para intersecciones. Considérense las medidas de probabilidad $\alpha = \frac{1}{2}\delta_2 + \frac{1}{2}\delta_4$ y $\beta = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_3$. Se tiene que $\alpha(\{1, 2\}) = \frac{1}{2}$, $\alpha(\{2, 3\}) = \frac{1}{2}$, $\beta(\{1, 2\}) = \frac{1}{2}$ y $\beta(\{2, 3\}) = \frac{1}{2}$. Por tanto, $\alpha = \beta$ en \mathcal{E} , pero $\alpha \neq \beta$ en \mathcal{A} .

1.2. Variables aleatorias

Definición 1.2.1. Sean (Ω, \mathcal{A}) y (Ω', \mathcal{A}') dos espacios medibles y sea $X: \Omega \rightarrow \Omega'$.

- (a) Se dice que X es medible con respecto a \mathcal{A} y \mathcal{A}' si $X^{-1}(A) \in \mathcal{A}$ para todo $A \in \mathcal{A}'$.
- (b) Si $\Omega' = \mathbb{R}$ y $\mathcal{A}' = \mathcal{B}$, se dice que X es una variable aleatoria.
- (c) Si $\Omega' = \mathbb{R}^n$ y $\mathcal{A}' = \mathcal{B}^n$, se dice que X es un vector aleatorio.
- (d) Si $\Omega = \mathbb{R}^m$, $\mathcal{A} = \mathcal{B}^m$, $\Omega' = \mathbb{R}^n$ y $\mathcal{A}' = \mathcal{B}^n$, se dice que X es una función de Borel.

El siguiente resultado muestra que las funciones medibles permiten, de alguna manera, transferir una medida de un espacio medible a otro.

Proposición 1.2.2. Sean (Ω, \mathcal{A}) y (Ω', \mathcal{A}') dos espacios medibles, sea $T: \Omega \rightarrow \Omega'$ una función medible con respecto a \mathcal{A} y \mathcal{A}' y sea μ una medida en (Ω, \mathcal{A}) .

(a) La función

$$\begin{aligned}\mu T^{-1}: \mathcal{A}' &\longrightarrow [0, \infty] \\ A &\longmapsto \mu(T^{-1}(A))\end{aligned}$$

es una medida en (Ω', \mathcal{A}') .

(b) Si $(\Omega, \mathcal{A}, \mu)$ es un espacio de probabilidad, entonces $(\Omega', \mathcal{A}', \mu T^{-1})$ también.

Lo habitual será trabajar con variables aleatorias en lugar de funciones medibles, y con espacios de probabilidad en lugar de espacios medibles. A veces, por ahorrar estritura, ni siquiera se hará referencia a espacios de probabilidad o espacios medibles. Cuando eso ocurra, debe entenderse que todas las variables aleatorias que aparezcan se consideran en un mismo espacio de probabilidad (Ω, \mathcal{A}, P) .

Proposición 1.2.3. Sean X e Y variables aleatorias en un espacio de probabilidad (Ω, \mathcal{A}, P) . Entonces:

- (a) $X + Y$ es una variable aleatoria.
- (b) λX es una variable aleatoria para todo $\lambda \in \mathbb{R}$.
- (c) XY es una variable aleatoria.
- (d) Si $Y(\omega) \neq 0$ para todo $\omega \in \Omega$, $\frac{X}{Y}$ es una variable aleatoria.
- (e) $\max\{X, Y\}$ es una variable aleatoria.
- (f) $\min\{X, Y\}$ es una variable aleatoria.

Este resultado se generaliza muy fácilmente por inducción a un número finito de variables aleatorias.

Definición 1.2.4. Sea X una variable aleatoria en un espacio de probabilidad (Ω, \mathcal{A}, P) . La *medida de probabilidad de X* es la función $P_X: \mathcal{B} \rightarrow [0, 1]$ dada por $P_X(B) = P(X^{-1}(B))$.

Por la Proposición 1.2.2, se tiene que P_X es una medida en $(\mathbb{R}, \mathcal{B})$ y que $(\mathbb{R}, \mathcal{B}, P_X)$ es un espacio de probabilidad.

Si $a \in \mathbb{R}$ y $B \in \mathcal{B}$, en contextos probabilísticos se suelen emplear las notaciones $P(X \in B) = P_X(B)$, $P(X = a) = P_X(\{a\})$, $P(X \leq a) = P_X((-\infty, a])$, $P(X \geq a) = P_X([a, \infty))$, $P(X < a) = P_X((-\infty, a))$ y $P(X > a) = P_X((a, \infty))$. Generalmente, también se omiten los puntos en los que se evalúan las variables aleatorias. Por ejemplo, si X e Y son variables aleatorias en un espacio de probabilidad (Ω, \mathcal{A}, P) , se denota $P(X = Y) = P(\{\omega \in \Omega: X(\omega) = Y(\omega)\})$.

Definición 1.2.5. Sea X una variable aleatoria en un espacio de probabilidad (Ω, \mathcal{A}, P) . La *función de masa de X* es la función $p_X: \mathbb{R} \rightarrow [0, 1]$ dada por $p_X(a) = P(X = a)$.

Definición 1.2.6. Sea X una variable aleatoria en un espacio de probabilidad (Ω, \mathcal{A}, P) . La *función de distribución de X* es la función $F_X: \mathbb{R} \rightarrow [0, 1]$ dada por $F_X(a) = P(X \leq a)$.

No debe confundirse *función de distribución* con *distribución*; este último término se suele utilizar para referirse a la medida de probabilidad o a la función de distribución de una variable aleatoria, dependiendo del contexto. Por ejemplo, dos variables aleatorias X e Y se dice que *siguen la misma distribución*, o que son *idénticamente distribuidas*, si $P_X = P_Y$.

Proposición 1.2.7. *Si X es una variable aleatoria y $g: \mathbb{R} \rightarrow \mathbb{R}$ es una función de Borel, entonces $Y = g \circ X$ es una variable aleatoria y $P_Y = P_X g^{-1}$.*

Se estudiarán a continuación las propiedades más relevantes sobre la función de distribución y la medida de probabilidad de una variable aleatoria. Antes de ello, se necesita introducir algo de notación: si $F: \mathbb{R} \rightarrow \mathbb{R}$ es una función cualquiera y $a \in \mathbb{R}$, se escribe

$$F(-\infty) = \lim_{x \rightarrow -\infty} F(x), \quad F(\infty) = \lim_{x \rightarrow \infty} F(x), \quad F(a^-) = \lim_{x \rightarrow a^-} F(x), \quad F(a^+) = \lim_{x \rightarrow a^+} F(x).$$

Proposición 1.2.8. *Sea X una variable aleatoria. Entonces:*

- (a) F_X es continua por la derecha.
- (b) F_X es creciente.
- (c) $F_X(-\infty) = 0$.
- (d) $F_X(\infty) = 1$.

Proposición 1.2.9. *Sea X una variable aleatoria. Si $a, b \in \mathbb{R}$ y $a < b$, entonces:*

- (a) $P_X((a, b]) = F_X(b) - F_X(a)$.
- (b) $P_X((a, b)) = F_X(b^-) - F_X(a)$.
- (c) $P_X([a, b)) = F_X(b^-) - F_X(a^-)$.
- (d) $P_X([a, b]) = F_X(b) - F_X(a^-)$.
- (e) $P_X((-\infty, a)) = F_X(a^-)$.
- (f) $P_X((a, \infty)) = 1 - F_X(a)$.
- (g) $P_X([a, \infty)) = 1 - F_X(a^-)$.
- (h) $p_X(a) = F_X(a) - F_X(a^-)$.
- (i) F_X es continua en a si y solo si $p_X(a) = 0$.

El conjunto de puntos de discontinuidad de una función $F: \mathbb{R} \rightarrow \mathbb{R}$ cualquiera se va a denotar por D_F , y el conjunto de puntos de continuidad, por C_F .

Así, si F_X es la función de distribución de una variable aleatoria X , por la proposición anterior se tiene $D_{F_X} = \{a \in \mathbb{R}: p_X(a) > 0\}$ y $C_{F_X} = \{a \in \mathbb{R}: p_X(a) = 0\}$.

Proposición 1.2.10. *Sea X una variable aleatoria. Entonces:*

- (a) D_{F_X} es finito o infinito numerable.
- (b) C_{F_X} es denso en \mathbb{R} .

Los siguientes resultados justifican que en muchas ocasiones se mencionen funciones de distribución o medidas de probabilidad sobre $(\mathbb{R}, \mathcal{B})$ sin hacer referencia a ninguna variable aleatoria.

Teorema 1.2.11. *Sea $F: \mathbb{R} \rightarrow \mathbb{R}$ una función verificando las siguientes propiedades:*

- (a) F es continua por la derecha.

(b) F es creciente.

(c) $F(-\infty) = 0$.

(d) $F(\infty) = 1$.

Entonces existe en algún espacio de probabilidad una variable aleatoria X tal que F es la función de distribución de X .

A toda función $F: \mathbb{R} \rightarrow \mathbb{R}$ verificando las cuatro propiedades anteriores se le llama *función de distribución*, a secas, sin variables aleatorias de por medio.

Si P es una medida de probabilidad en el espacio medible $(\mathbb{R}, \mathcal{B})$, es fácil probar que la función $F: \mathbb{R} \rightarrow \mathbb{R}$ dada por $F(x) = P((-\infty, x])$ es una función de distribución.

En resumen, una variable aleatoria X tiene asociadas una función de distribución y una medida de probabilidad sobre $(\mathbb{R}, \mathcal{B})$. Recíprocamente, una medida de probabilidad P en $(\mathbb{R}, \mathcal{B})$ permite definir una función de distribución F , que, a su vez, proporciona una variable aleatoria X en algún espacio de probabilidad que tiene a F por función de distribución y a P por medida de probabilidad.

1.3. Variables aleatorias discretas

Definición 1.3.1. Una variable aleatoria X en un espacio de probabilidad (Ω, \mathcal{A}, P) se dice que es *discreta* si existe un conjunto $B \in \mathcal{B}$ numerable tal que $P(X \in B) = 1$.

Si X es una variable aleatoria discreta y $B = \{b_i: i \in \mathbb{N}\} \in \mathcal{B}$ es tal que $P(X \in B) = 1$, entonces P_X queda totalmente determinada por los valores $P(X = b_i)$, $i \in \mathbb{N}$.

La siguiente definición recoge las variables aleatorias discretas a las que les suelen dar más importancia los probabilistas.

Definición 1.3.2. Sea X una variable aleatoria discreta en un espacio de probabilidad (Ω, \mathcal{A}, P) .

(a) Si $x_0 \in \mathbb{R}$, se dice que X sigue una *distribución degenerada en el punto x_0* , y se denota $X \sim D(x_0)$, cuando

$$P(X = x) = \begin{cases} 1 & \text{si } x = x_0, \\ 0 & \text{si } x \neq x_0. \end{cases}$$

(b) Si $\{x_1, \dots, x_n\} \subset \mathbb{R}$ y $x_i \neq x_j$ si $i \neq j$, se dice que X sigue una *distribución uniforme sobre los n puntos*, y se denota $X \sim U(\{x_1, \dots, x_n\})$, cuando

$$P(X = x) = \begin{cases} \frac{1}{n} & \text{si } x \in \{x_1, \dots, x_n\}, \\ 0 & \text{si } x \notin \{x_1, \dots, x_n\}. \end{cases}$$

(c) Si $p \in (0, 1)$, se dice que X sigue una *distribución de Bernoulli de parámetro p* , y se denota $X \sim \text{Ber}(p)$, cuando

$$P(X = x) = \begin{cases} 1 - p & \text{si } x = 0, \\ p & \text{si } x = 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Usualmente se escribe $q = 1 - p$.

- (d) Si $n \in \mathbb{N}$ y $p \in (0, 1)$, se dice que X sigue una *distribución binomial de parámetros n y p* , y se denota $X \sim \text{Bin}(n, p)$, cuando

$$P(X = k) = \begin{cases} \binom{n}{k} p^k q^{n-k} & \text{si } k \in \mathbb{N} \cup \{0\}, k \leq n, \\ 0 & \text{en otro caso.} \end{cases}$$

- (e) Si $p \in (0, 1)$, se dice que X sigue una *distribución geométrica de parámetro p* , y se denota $X \sim \text{Geo}(p)$, cuando

$$P(X = k) = \begin{cases} q^k p & \text{si } k \in \mathbb{N} \cup \{0\}, \\ 0 & \text{en otro caso.} \end{cases}$$

- (f) Si $\lambda > 0$, se dice que X sigue una *distribución de Poisson de parámetro λ* , y se denota $X \sim P(\lambda)$, cuando

$$P(X = k) = \begin{cases} e^{-\lambda} \frac{\lambda^k}{k!} & \text{si } k \in \mathbb{N} \cup \{0\}, \\ 0 & \text{en otro caso.} \end{cases}$$

- (g) Se dice que X sigue una *distribución de Rademacher*, y se denota $X \sim \text{Rad}$, cuando

$$P(X = k) = \begin{cases} \frac{1}{2} & \text{si } k \in \{-1, 1\}, \\ 0 & \text{en otro caso.} \end{cases}$$

1.4. Variables aleatorias absolutamente continuas

Definición 1.4.1. Una variable aleatoria X en un espacio de probabilidad (Ω, \mathcal{A}, P) se dice que es *absolutamente continua* si existe una función de Borel $f_X: \mathbb{R} \rightarrow \mathbb{R}$ no negativa tal que

$$P(X \in B) = \int_B f_X(t) dt$$

para cada $B \in \mathcal{B}$. A f_X se le denomina *función de densidad de X* .

Nótese que la integral anterior tiene perfecto sentido cualquiera que sea $B \in \mathcal{B}$ porque f_X es una función de Borel y no negativa.

Para el resultado que sigue, se recuerda que toda función creciente es derivable en casi todo punto (respecto de la medida de Lebesgue) y su derivada, allá donde exista, es no negativa.

Proposición 1.4.2. Sea X una variable aleatoria absolutamente continua.

- (a) Si $a, b \in \mathbb{R}$ y $a < b$,

$$F_X(b) - F_X(a) = \int_a^b f_X(t) dt.$$

- (b) Para todo $b \in \mathbb{R}$,

$$F_X(b) = \int_{-\infty}^b f_X(t) dt.$$

- (c) $F'_X(t) = f(t)$ para casi todo $t \in \mathbb{R}$.

(d) Se tiene que

$$\int_{\mathbb{R}} f_X(t) dt = 1.$$

Al igual que en el caso discreto, existe una lista de variables aleatorias absolutamente continuas especialmente relevantes en la teoría de la probabilidad.

Definición 1.4.3. Sea X una variable aleatoria absolutamente continua y sea f_X la función de densidad de X .

(a) Dado un intervalo $[a, b] \subset \mathbb{R}$, se dice que X sigue una *distribución uniforme en $[a, b]$* , y se denota $X \sim U([a, b])$, cuando

$$f_X(t) = \begin{cases} \frac{1}{b-a} & \text{si } t \in [a, b], \\ 0 & \text{en otro caso.} \end{cases}$$

(b) Si $\mu \in \mathbb{R}$ y $\sigma^2 > 0$, se dice que X sigue una *distribución normal de parámetros μ y σ^2* , y se denota $X \sim N(\mu, \sigma^2)$, cuando

$$f_X(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}.$$

Si $\mu = 0$ y $\sigma^2 = 1$, se habla de *distribución normal estándar*.

(c) Si $\lambda > 0$, se dice que X sigue una *distribución exponencial de parámetro λ* , y se denota $X \sim \text{Exp}(\lambda)$, cuando

$$f_X(t) = \begin{cases} \lambda e^{-\lambda t} & \text{si } t \geq 0, \\ 0 & \text{si } t < 0. \end{cases}$$

(d) Si $x_0 \in \mathbb{R}$ y $\gamma > 0$, se dice que X sigue una *distribución de Cauchy de parámetros x_0 y γ* , y se denota $X \sim C(x_0, \gamma)$, cuando

$$f_X(t) = \frac{1}{\pi\gamma(1 + (\frac{x-x_0}{\gamma})^2)}.$$

Si $x_0 = 0$ y $\gamma = 1$, se habla de *distribución de Cauchy estándar*.

(e) Si $a, b, c \in \mathbb{R}$ son tales que $a < c < b$, se dice que X sigue una *distribución triangular de parámetros a , b y c* , y se denota $X \sim T(a, b, c)$, cuando

$$f_X(t) = \begin{cases} 0 & \text{si } x < a, \\ \frac{2(x-a)}{(b-a)(c-a)} & \text{si } a \leq x < c, \\ \frac{2}{b-a} & \text{si } x = c, \\ \frac{2(b-x)}{(b-a)(b-c)} & \text{si } c \leq x < b, \\ 0 & \text{si } b < x. \end{cases}$$

Nótese que la clasificación de variables aleatorias realizada hasta ahora no es exhaustiva: existen variables aleatorias que no son ni discretas ni absolutamente continuas.

1.5. Vectores aleatorios

Se recuerda que los vectores aleatorios ya fueron definidos junto con las variables aleatorias. Las similitudes entre ambos son más que notables.

Proposición 1.5.1. Sea (Ω, \mathcal{A}, P) un espacio de probabilidad. Una función $(X_1, X_2, \dots, X_n): \Omega \rightarrow \mathbb{R}^n$ es un vector aleatorio si y solo si $X_i: \Omega \rightarrow \mathbb{R}$ es una variable aleatoria para cada $i = 1, 2, \dots, n$.

La mayoría de los resultados expuestos sobre variables aleatorias pueden ser adaptados convenientemente para vectores aleatorios, pero no se va a profundizar en ello. Solo se van a aportar las definiciones más elementales.

Definición 1.5.2. Sea $X = (X_1, X_2, \dots, X_n)$ un vector aleatorio. La *medida de probabilidad de X* es la función $P_X: \mathcal{B}^n \rightarrow [0, 1]$ dada por $P_X(B) = P(X^{-1}(B))$.

Definición 1.5.3. Sea $X = (X_1, X_2, \dots, X_n)$ un vector aleatorio. La *función de masa de X* es la función $p_X: \mathbb{R}^n \rightarrow [0, 1]$ dada por $p_X(a_1, a_2, \dots, a_n) = P(X_1 = a_1, X_2 = a_2, \dots, X_n = a_n)$.

Definición 1.5.4. Sea $X = (X_1, X_2, \dots, X_n)$ un vector aleatorio. La *función de distribución de X* es la función $F_X: \mathbb{R}^n \rightarrow [0, 1]$ dada por $F_X(a_1, a_2, \dots, a_n) = P(X_1 \leq a_1, X_2 \leq a_2, \dots, X_n \leq a_n)$.

A veces se utiliza el término *distribución conjunta* para referirse a la función de distribución o a la medida de probabilidad de un vector aleatorio.

1.6. Independencia de sucesos y variables aleatorias

Definición 1.6.1. Sea (Ω, \mathcal{A}, P) un espacio de probabilidad. Una colección $\{A_i\}_{i \in I}$ de elementos de \mathcal{A} se dice que es *independiente* si para toda subcolección finita $\{A_{i_1}, A_{i_2}, \dots, A_{i_k}\} \subset \{A_i\}_{i \in I}$ se tiene

$$P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \dots P(A_{i_k}).$$

Definición 1.6.2. Una colección $\{X_i\}_{i \in I}$ de variables aleatorias sobre un espacio de probabilidad (Ω, \mathcal{A}, P) se dice que es *independiente* si para toda subcolección finita $\{X_{i_1}, X_{i_2}, \dots, X_{i_k}\} \subset \{X_i\}_{i \in I}$ y para $B_1, B_2, \dots, B_k \in \mathcal{B}$ cualesquiera, se tiene

$$P(X_{i_1} \in B_1, X_{i_2} \in B_2, \dots, X_{i_k} \in B_k) = P(X_{i_1} \in B_1)P(X_{i_2} \in B_2) \dots P(X_{i_k} \in B_k).$$

Al ser I un conjunto arbitrario, también queda definida la independencia de un número finito de sucesos o variables aleatorias, así como la independencia de una sucesión de sucesos o variables aleatorias.

Proposición 1.6.3. Sea $X = (X_1, X_2, \dots, X_n)$ un vector aleatorio en un espacio de probabilidad (Ω, \mathcal{A}, P) . Entonces las variables aleatorias X_1, X_2, \dots, X_n son independientes si y solo si para todo $(a_1, a_2, \dots, a_n) \in \mathbb{R}^n$ se tiene que

$$F_X(a_1, a_2, \dots, a_n) = F_{X_1}(a_1)F_{X_2}(a_2) \dots F_{X_n}(a_n).$$

Proposición 1.6.4. Sea $X = (X_1, X_2, \dots, X_n)$ un vector aleatorio en un espacio de probabilidad (Ω, \mathcal{A}, P) . Si X_i es una variable aleatoria discreta para todo $i \in \{1, \dots, n\}$, entonces las variables aleatorias X_1, X_2, \dots, X_n son independientes si y solo si para todo $(a_1, a_2, \dots, a_n) \in \mathbb{R}^n$ se tiene que

$$p_X(a_1, a_2, \dots, a_n) = p_{X_1}(a_1)p_{X_2}(a_2)\dots p_{X_n}(a_n).$$

Proposición 1.6.5. Si X e Y son variables aleatorias independientes y $f, g: \mathbb{R} \rightarrow \mathbb{R}$ son funciones de Borel, entonces $f(X)$ y $g(Y)$ también son variables aleatorias independientes.

1.7. Esperanza

Definición 1.7.1. Sea Ω es un conjunto no vacío y sea $X: \Omega \rightarrow \mathbb{R}$ una función cualquiera.

- (a) Se define la *parte positiva de X* , y se denota por X^+ , como $X^+ = \max\{X, 0\}$.
- (b) Se define la *parte negativa de X* , y se denota por X^- , como $X^- = \max\{-X, 0\}$.

Es inmediato comprobar que $X = X^+ - X^-$ y que $|X| = X^+ + X^-$. Además, si X es una variable aleatoria, entonces X^+ , X^- y $|X|$ también lo son.

Definición 1.7.2. Sea X una variable aleatoria en un espacio de probabilidad (Ω, \mathcal{A}, P) . Se define la *esperanza de X* como

$$E(X) = \int_{\Omega} X dP,$$

siempre que dicha integral tenga sentido. Si $E(X) \in \mathbb{R}$, se dice que X es *integrable*.

Recuérdese que la esperanza de X tiene sentido siempre que $E(X^+) < \infty$ o $E(X^-) < \infty$, en cuyo caso se tiene que $E(X) = E(X^+) - E(X^-)$. Además, las esperanzas de X^+ y X^- siempre existen (pudiendo valer ∞) porque son variables aleatorias no negativas. El único caso en que una variable aleatoria no admite esperanza es cuando $E(X^+) = \infty$ y $E(X^-) = \infty$.

Por otra parte, si $1 \leq p < \infty$, se denota $\mathcal{L}^p(\Omega, \mathcal{A}, P)$, o simplemente \mathcal{L}^p , al conjunto de variables aleatorias $X: \Omega \rightarrow \mathbb{R}$ tales que

$$E(|X|^p) = \int_{\Omega} |X|^p d\mu < \infty.$$

Proposición 1.7.3. Si $X, Y \in \mathcal{L}^p$ y $\alpha, \beta \in \mathbb{R}$, entonces $\alpha X + \beta Y \in \mathcal{L}^p$ y $E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$.

El resultado anterior establece que \mathcal{L}^p junto con la suma de funciones y el producto por escalares reales tiene estructura de espacio vectorial sobre \mathbb{R} y, además, que la esperanza es una aplicación lineal sobre dicho espacio vectorial.

Proposición 1.7.4. Si X e Y son variables aleatorias con $X = Y$ en casi todo punto, entonces X es integrable si y solo si Y es integrable, y en ese caso, $E(X) = E(Y)$.

El resultado anterior sugiere introducir en el conjunto \mathcal{L}^p la relación \sim definida de la siguiente manera:

$$X \sim Y \text{ si y solo si } X = Y \text{ en casi todo punto.}$$

Es fácil probar que esta relación es de equivalencia. El conjunto cociente será denotado por L^p y, por comodidad, se llamará de la misma manera a las clases de equivalencia de L^p y a sus

representantes: en lugar de decir $[X] \in L^p$, se escribe simplemente $X \in L^p$. De nuevo, L^p es un espacio vectorial sobre \mathbb{R} .

Proposición 1.7.5. *Si $1 \leq p \leq q$, entonces $L^q \subset L^p$.*

Proposición 1.7.6. *Una variable aleatoria X es integrable si y solo si $E(|X|) < \infty$, y en ese caso, $|E(X)| \leq E(|X|)$.*

En otras palabras, esta proposición dice que una variable aleatoria X es integrable si y solo si $X \in L^1$.

Cuando haya que calcular esperanzas en casos concretos, lo más habitual será recurrir a los dos teoremas que siguen.

Teorema 1.7.7. *Sea X una variable aleatoria y sea $g: \mathbb{R} \rightarrow \mathbb{R}$ una función de Borel.*

(a) *Se tiene que*

$$E(g \circ X) = \int_{\mathbb{R}} g(t) dP_X(t),$$

siempre que la esperanza de $g \circ X$ tenga sentido. En particular,

$$E(X) = \int_{\mathbb{R}} t dP_X(t),$$

siempre que la esperanza de X tenga sentido.

(b) *Si X es discreta y $D_X = \{x \in \mathbb{R}: p_X(x) > 0\}$, entonces*

$$E(g \circ X) = \sum_{t \in D_X} g(t) p_X(t),$$

siempre que la esperanza de $g \circ X$ tenga sentido. En particular,

$$E(X) = \sum_{t \in D_X} t p_X(t),$$

siempre que la esperanza de X tenga sentido.

(c) *Si X es absolutamente continua, entonces*

$$E(g \circ X) = \int_{\mathbb{R}} g(t) f_X(t) dt,$$

siempre que la esperanza de $g \circ X$ tenga sentido. En particular,

$$E(X) = \int_{\mathbb{R}} t f_X(t) dt,$$

siempre que la esperanza de X tenga sentido.

Teorema 1.7.8. *Si $X \in L^1$, entonces*

$$E(X) = \int_0^\infty P(X > t) dt - \int_{-\infty}^0 P(X < t) dt.$$

Estos teoremas son útiles incluso cuando no se sabe si una variable aleatoria X admite esperanza, pues se puede aplicar primero a $|X|$ (o a X^+ y X^-), y si la esperanza sale finita, entonces ya se puede aplicar a X .

Ejemplo 1.7.9. Sea X una variable aleatoria absolutamente continua con función de densidad dada por $f_X(t) = \frac{1}{\pi(1+t^2)}$, $t \in \mathbb{R}$. Como la función $g: \mathbb{R} \rightarrow \mathbb{R}$ dada por $g(t) = \max\{t, 0\}$ es de Borel y $X^+ = \max\{X, 0\} = g \circ X$ es una variable aleatoria no negativa, por el teorema anterior,

$$E(X^+) = \int_{\mathbb{R}} \max\{t, 0\} f_X(t) dt = \int_0^{\infty} \frac{t}{\pi(1+t^2)} dt \geq \frac{1}{\pi} \int_1^{\infty} \frac{t}{1+t^2} dt \geq \frac{1}{\pi} \int_1^{\infty} \frac{1}{2t} dt = \infty,$$

donde se ha usado que $\frac{x}{x^2+1} \geq \frac{1}{2x}$ para todo $x \geq 1$ (se prueba fácilmente). De forma similar,

$$E(X^-) = \int_{\mathbb{R}} \max\{-t, 0\} f_X(t) dt = \int_{-\infty}^0 -\frac{t}{\pi(1+t^2)} dt = \int_0^{\infty} \frac{t}{\pi(1+t^2)} dt = \infty.$$

Por tanto, X no admite esperanza.

Ejemplo 1.7.10. Sea X una variable aleatoria con $X \sim \text{Exp}(\lambda)$, $\lambda > 0$. Se prueba fácilmente que la función de distribución de X es $F_X(t) = (1 - e^{-\lambda t}) \mathbb{1}_{[0, \infty)}(t)$, $t \in \mathbb{R}$. Como $P(X < 0) = 0$, entonces $X \geq 0$ en casi todo punto (respecto de la medida P) y por tanto

$$E(X) = E(|X|) = \int_0^{\infty} P(X > t) dt = \int_0^{\infty} (1 - F_X(t)) dt = \int_0^{\infty} e^{-\lambda t} dt = \frac{1}{\lambda}.$$

Proposición 1.7.11 (Desigualdad de Markov). Sea X una variable aleatoria. Para todo $\alpha > 0$ y todo $k \in \mathbb{N}$,

$$P(|X| \geq \alpha) \leq \frac{1}{\alpha^k} E(|X|^k).$$

Nótese que esta desigualdad tiene sentido incluso cuando X^k no es integrable, o sea, incluso cuando $E(|X|^k) = \infty$.

Al aplicar la desigualdad anterior a la variable aleatoria $X - E(X)$ (es necesario entonces que $X \in L^1$) y poner $k = 2$, se obtiene lo siguiente:

Corolario 1.7.12 (Desigualdad de Chebyshev). Sea X una variable aleatoria con $X \in L^1$. Para todo $\alpha > 0$,

$$P(|X - E(X)| \geq \alpha) \leq \frac{1}{\alpha^2} E((X - E(X))^2).$$

Definición 1.7.13. Si $X \in L^1$, se define la *varianza de X* como $\text{Var}(X) = E((X - E(X))^2)$.

Nótese que $\text{Var}(X)$ tiene sentido incluso cuando $X \notin L^2$, pues $(X - E(X))^2 = X^2 + E(X)^2 - 2E(X)X$, y se tiene que $X^2 + E(X)^2$ admite esperanza por ser no negativa y $2XE(X)$ tiene esperanza finita por ser $X \in L^1$. Así, la desigualdad de Chebyshev también puede escribirse de manera más abreviada como

$$P(|X - E(X)| \geq \alpha) \leq \frac{1}{\alpha^2} \text{Var}(X).$$

Proposición 1.7.14. Se verifican las siguientes propiedades:

- (a) Si $X \in L^1$, entonces $\text{Var}(X) = E(X^2) - E(X)^2$.
- (b) Si $a, b \in \mathbb{R}$ y $X \in L^1$, entonces $\text{Var}(aX + b) = a^2 \text{Var}(X)$.
- (c) Si $X, Y \in L^1$ y son independientes, entonces $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$.

Definición 1.7.15. Si $X \in L^k$ para algún $k \in \mathbb{N}$, se dice que $E(X^k)$ es el *momento de orden k de X* .

De la Proposición 1.7.5 se sigue que si existe el momento de orden k de X , también existe el de orden j para todo $j \leq k$.

Convolución

En este breve tema se estudiará una noción que en ciertas ocasiones facilita el cálculo de la distribución de la suma de dos variables aleatorias.

2.1. Convolución de medidas de probabilidad

Definición 2.1.1. Sean X e Y dos variables aleatorias. La *convolución de P_X y P_Y* es la función $P_X * P_Y: \mathcal{B} \rightarrow [0, 1]$ dada por

$$P_X * P_Y(B) = \int_{\mathbb{R}} P_Y(B - t) dP_X(t).$$

Proposición 2.1.2. Si X e Y son variables aleatorias independientes, entonces $P_X * P_Y$ es la medida de probabilidad de la variable aleatoria $X + Y$.

Como la suma de variables aleatorias es conmutativa y asociativa, lo mismo puede decirse de la convolución de medidas de probabilidad.

2.2. Convolución de funciones de distribución

Si X es una variable aleatoria cualquiera, en las integrales respecto de P_X a veces se escribe dF_X en lugar de dP_X .

Definición 2.2.1. Sean X e Y dos variables aleatorias. La *convolución de F_X y F_Y* es la función $F_X * F_Y: \mathbb{R} \rightarrow [0, 1]$ definida por

$$F_X * F_Y(z) = \int_{\mathbb{R}} F_Y(z - t) dF_X(t).$$

Proposición 2.2.2. Si X e Y son variables aleatorias independientes, entonces $F_X * F_Y$ es la función de distribución de $X + Y$.

2.3. Convolución y funciones de densidad

La convolución posee algunas propiedades interesantes cuando se trata con variables aleatorias absolutamente continuas.

Definición 2.3.1. Sean X e Y dos variables aleatorias.

- (a) Si Y es absolutamente continua, la *convolución de F_X y f_Y* no es más que la función $F_X * f_Y: \mathbb{R} \rightarrow \mathbb{R}$ dada por

$$F_X * f_Y(z) = \int_{\mathbb{R}} f_Y(z-t) dF_X(t).$$

- (b) Si además X es absolutamente continua, la *convolución de f_X y f_Y* no es más que la función $f_X * f_Y: \mathbb{R} \rightarrow \mathbb{R}$ dada por

$$f_X * f_Y(z) = \int_{\mathbb{R}} f_Y(z-t) f_X(t) dt.$$

Proposición 2.3.2. Sean X e Y dos variables aleatorias independientes.

- (a) Si Y es absolutamente continua, entonces $X+Y$ es absolutamente continua y $f_{X+Y} = F_X * f_Y$.
(b) Si X también es absolutamente continua, entonces $f_{X+Y} = f_X * f_Y$.

Ejemplo 2.3.3. Sean X e Y dos variables aleatorias independientes y absolutamente continuas con $X, Y \sim U([0, 1])$. Se comprueba fácilmente que sus respectivas funciones de distribución están definidas por

$$F_X(t) = F_Y(t) = \begin{cases} 0 & \text{si } t < 0, \\ t & \text{si } 0 \leq t < 1, \\ 1 & \text{si } 1 \leq t. \end{cases}$$

Sea $z \in \mathbb{R}$ y calculemos

$$F_X * F_Y(z) = \int_{\mathbb{R}} F_X(z-t) dF_Y(t) = \int_{\mathbb{R}} F_X(z-t) f_Y(t) dt,$$

donde f_Y es la función de densidad de Y . Como $\text{sop } f_Y = [0, 1]$, entonces

$$F_X * F_Y(z) = \int_0^1 F_X(z-t) f_Y(t) dt.$$

Se distinguen los siguientes casos:

- (a) Si $z < 0$, entonces $z-t < 0$ para todo $t \in [0, 1]$ y, en consecuencia, $F_X * F_Y(z) = 0$.
(b) Supongamos que $0 \leq z < 1$. Si $0 \leq z-t < 1$, es decir, si $0 < t \leq z$, entonces $F_X(z-t) = z-t$, y si $z-t < 0$, es decir, si $z < t \leq 1$, entonces $F_X(z-t) = 0$. Nótese que el caso $1 \leq z-t$ no puede darse. De todo esto se deduce que

$$F_X * F_Y(z) = \int_0^z (z-t) dt = \left[-\frac{(z-t)^2}{2} \right]_{t=0}^{t=z} = \frac{z^2}{2}.$$

- (c) Supongamos que $1 \leq z < 2$. Ahora el caso $z-t < 0$ no puede darse. Si $0 \leq z-t < 1$, es decir, si $z-1 < t \leq 1$, entonces $F_X(z-t) = z-t$. Y si $1 \leq z-t$, es decir, si $0 \leq t \leq z-1$, entonces $F_X(z-t) = 1$. En consecuencia,

$$F_X * F_Y(z) = \int_{z-1}^1 (z-t) dt + \int_0^{z-1} 1 dt = \left[-\frac{(z-t)^2}{2} \right]_{t=z-1}^{t=1} + z-1 = -\frac{z^2}{2} + 2z - 1.$$

- (d) Si $2 \leq z$, entonces $1 \leq z-t$ para todo $t \in [0, 1]$, luego $F_X(z-t) = 1$ y entonces

$$F_X * F_Y(z) = \int_0^1 1 dt = 1.$$

La función de distribución de $X + Y$ sería entonces

$$F_{X+Y}(z) = \begin{cases} 0 & \text{si } z < 0, \\ \frac{z^2}{2} & \text{si } 0 \leq z < 1, \\ -\frac{z^2}{2} + 2z - 1 & \text{si } 1 \leq z < 2, \\ 1 & \text{si } 2 \leq z. \end{cases}$$

Ejemplo 2.3.4. Sean X e Y dos variables aleatorias independientes y absolutamente continuas con $X, Y \sim U([0, 1])$. Entonces $X + Y$ es absolutamente continua con función de densidad dada por

$$f_{X+Y}(z) = \int_{\mathbb{R}} f_Y(z-t)f_X(t)dt.$$

Razonando como en el último ejemplo, se comprueba que

$$f_{X+Y}(z) = \begin{cases} 0 & \text{si } z < 0, \\ z & \text{si } 0 \leq z < 1, \\ 2-z & \text{si } 1 \leq z < 2, \\ 0 & \text{si } 2 \leq z. \end{cases}$$

Se observa que $F'_{X+Y} = f_{X+Y}$ en casi todo punto, como no podía ser de otra forma.

Función característica

Definición 3.0.1. Sea X una variable aleatoria. La *función característica de X* es la función $\varphi_X: \mathbb{R} \rightarrow \mathbb{C}$ definida por $\varphi_X(t) = E(e^{itX})$.

Lo primero que debe observarse es que φ_X está bien definida, es decir, e^{itX} es una variable aleatoria y su esperanza tiene sentido. Lo primero es porque la función $g: \mathbb{R} \rightarrow \mathbb{R}$ dada por $g(s) = e^{its}$ es de Borel (pues es continua); lo segundo, porque para todo $t \in \mathbb{R}$ se tiene

$$E(|e^{itX}|) = \int_{\Omega} |e^{itX(\omega)}| dP(\omega) = \int_{\Omega} dP(\omega) = P(\Omega) = 1 < \infty.$$

Además, por el Teorema 1.7.7,

$$\varphi_X(t) = \int_{\mathbb{R}} e^{its} dP_X(s).$$

3.1. Propiedades elementales y ejemplos

Proposición 3.1.1. Sea X una variable aleatoria. Se verifican las siguientes propiedades:

- (a) $\varphi_X(0) = 1$.
- (b) $|\varphi_X(t)| \leq 1$ para todo $t \in \mathbb{R}$.
- (c) $\overline{\varphi_X(t)} = \varphi_X(-t)$ para todo $t \in \mathbb{R}$.
- (d) φ_X es uniformemente continua.
- (e) Si $Y = aX + b$ con $a, b \in \mathbb{R}$, entonces $\varphi_Y(t) = e^{itb} \varphi_X(ta)$ para todo $t \in \mathbb{R}$.

Proposición 3.1.2. Si X e Y son variables aleatorias independientes, entonces para todo $t \in \mathbb{R}$ se verifica $\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$.

Proposición 3.1.3 (Identidad de Parseval). Sean X e Y dos variables aleatorias. Entonces

$$\int_{\mathbb{R}} \varphi_X(t) dP_Y(t) = \int_{\mathbb{R}} \varphi_Y(t) dP_X(t).$$

El resultado anterior resulta especialmente útil para hallar ciertas integrales evitando el cálculo de primitivas.

Distribución	Función característica
$Deg(a)$	e^{ita}
Rad	$\cos(t)$
$Ber(p)$	$1 - p + pe^{it}$
$Bin(n, p)$	$(1 - p + pe^{it})^n$
$Ge(p)$	$p(1 - (1 - p)e^{it})^{-1}$
$P(\lambda)$	$e^{\lambda(e^{it} - 1)}$
$U([a, b])$	$\frac{e^{itb} - e^{ita}}{it(b - a)}$
$Exp(\lambda)$	$\frac{\lambda}{\lambda - it}$
$N(\mu, \sigma^2)$	$e^{it\mu - \frac{\sigma^2 t^2}{2}}$
$C(0, 1)$	$e^{- t }$

Cuadro. Funciones características de algunas distribuciones.

En los ejemplos que siguen se calculan varias de las funciones características que aparecen en el cuadro anterior.

Ejemplo 3.1.4. Sea X una variable aleatoria absolutamente continua con $X \sim N(0, 1)$. Para cada $t \in \mathbb{R}$,

$$\varphi_X(t) = E(e^{itX}) = \int_{\mathbb{R}} e^{its} f_X(s) ds = \frac{1}{\sqrt{2\pi}} \left(\int_{\mathbb{R}} \cos(ts) e^{-\frac{s^2}{2}} ds + i \int_{\mathbb{R}} \operatorname{sen}(ts) e^{-\frac{s^2}{2}} ds \right).$$

Nótese que

$$\int_{\mathbb{R}} \operatorname{sen}(ts) e^{-\frac{s^2}{2}} ds = 0$$

por ser el integrando una función par. Por tanto,

$$\varphi_X(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \cos(ts) e^{-\frac{s^2}{2}} ds.$$

Como se tiene que $|\frac{\partial}{\partial t} \cos(ts) e^{-\frac{s^2}{2}}| = |-s \operatorname{sen}(ts) e^{-\frac{s^2}{2}}| \leq |s| e^{-\frac{s^2}{2}}$ y la función $s \mapsto |s| e^{-\frac{s^2}{2}}$, $s \in \mathbb{R}$ es integrable, entonces φ_X es derivable y para todo $t \in \mathbb{R}$,

$$\begin{aligned} \varphi'_X(t) &= -\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} s \operatorname{sen}(ts) e^{-\frac{s^2}{2}} ds \\ &= \frac{1}{\sqrt{2\pi}} \lim_{n \rightarrow \infty} \int_{-n}^n -s \operatorname{sen}(ts) e^{-\frac{s^2}{2}} ds \\ &= \frac{1}{\sqrt{2\pi}} \lim_{n \rightarrow \infty} \left(\left[\operatorname{sen}(ts) e^{-\frac{s^2}{2}} \right]_{s=-n}^{s=n} - t \int_{-n}^n \cos(ts) e^{-\frac{s^2}{2}} ds \right) \\ &= -\frac{t}{\sqrt{2\pi}} \int_{\mathbb{R}} \cos(ts) e^{-\frac{s^2}{2}} ds \\ &= -t \varphi_X(t), \end{aligned}$$

donde en la tercera igualdad se ha integrado por partes ($u(s) = \sin(ts)$, $v(s) = e^{-\frac{s^2}{2}}$). Por tanto, resolviendo la ecuación diferencial obtenida, existe $C \in \mathbb{R}$ tal que $\varphi_X(t) = Ce^{-\frac{t^2}{2}}$ para todo $t \in \mathbb{R}$. Como

$$\varphi_X(0) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{s^2}{2}} ds = \frac{1}{\sqrt{2\pi}} \sqrt{2\pi} = 1,$$

entonces $C = 1$ y se concluye que $\varphi_X(t) = e^{-\frac{t^2}{2}}$ para todo $t \in \mathbb{R}$.

Ejemplo 3.1.5. Sea X una variable aleatoria absolutamente continua con $X \sim N(\mu, \sigma^2)$. Entonces $Y = \frac{X-\mu}{\sigma} \sim N(0, 1)$, así que, por lo probado en el ejemplo anterior, $\varphi_Y(t) = e^{-\frac{t^2}{2}}$ para todo $t \in \mathbb{R}$. Y como $\varphi_Y(t) = e^{-\frac{it\mu}{\sigma}} \varphi_X\left(\frac{t}{\sigma}\right)$, entonces $\varphi_X(t) = e^{it\mu} e^{-\frac{\sigma^2 t^2}{2}}$ para todo $t \in \mathbb{R}$.

Ejemplo 3.1.6. Sea X una variable aleatoria absolutamente continua con $X \sim U([-1, 1])$. Para cada $t \in \mathbb{R}$ con $t \neq 0$,

$$\varphi_X(t) = E(e^{itX}) = \int_{\mathbb{R}} e^{its} f_X(s) ds = \frac{1}{2} \int_{-1}^1 e^{its} ds = \frac{1}{2} \left[\frac{1}{it} e^{its} \right]_{s=-1}^{s=1} = \frac{2i \sin t}{2it} = \frac{\sin t}{t},$$

mientras que $\varphi_X(0) = 1$.

Ejemplo 3.1.7. Sea X una variable aleatoria discreta con $X \sim Rad$. Para cada $t \in \mathbb{R}$,

$$\varphi_X(t) = E(e^{itX}) = e^{it \cdot (-1)} P(X = -1) + e^{it \cdot 1} P(X = 1) = \frac{e^{-it} + e^{it}}{2} = \cos(t).$$

Ejemplo 3.1.8. Sea X una variable aleatoria discreta con $X \sim Ber(p)$. Para cada $t \in \mathbb{R}$,

$$\varphi_X(t) = E(e^{itX}) = e^{it \cdot 1} P(X = 1) + e^{it \cdot 0} P(X = 0) = pe^{it} + 1 - p.$$

Ejemplo 3.1.9. Sea X una variable aleatoria discreta con $X \sim Bin(n, p)$. Se sabe que la distribución de probabilidad de X es la misma que la de $Y_1 + \dots + Y_n$, para ciertas variables aleatorias Y_1, \dots, Y_n independientes y con $Y_i \sim Ber(p)$ para cada $i \in \{1, \dots, n\}$. Por tanto, por el ejemplo anterior, para todo $t \in \mathbb{R}$ se tiene

$$\varphi_X(t) = \varphi_{Y_1 + \dots + Y_n}(t) = \varphi_{Y_1}(t) \dots \varphi_{Y_n}(t) = (pe^{it} + 1 - p)^n.$$

Ejemplo 3.1.10. Sea X una variable aleatoria discreta con $X \sim P(\lambda)$. Para cada $t \in \mathbb{R}$,

$$\varphi_X(t) = E(e^{itX}) = \sum_{k=0}^{\infty} e^{itk} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^{it})^k}{k!} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it} - 1)}.$$

Ejemplo 3.1.11. Sea X una variable aleatoria con función de distribución definida mediante $F_X(t) = (1 - \frac{1}{2}e^{-t}) \mathbb{1}_{[0, \infty)}(t)$, $t \in \mathbb{R}$. Se trata de hallar la función característica de X . Sea $t \in \mathbb{R}$. Como $P_X((-\infty, 0)) = 0$, entonces

$$\varphi_X(t) = \int_{\mathbb{R}} e^{its} dP_X(s) = \int_{[0, \infty)} e^{its} dP_X(s) = \int_{\{0\}} e^{its} dP_X(s) + \int_{(0, \infty)} e^{its} dP_X(s).$$

Examinemos cada uno de los sumandos. Por un lado,

$$\int_{\{0\}} e^{its} dP_X(s) = P(X = 0) = F_X(0) - F_X(0^-) = \frac{1}{2}.$$

Por otro lado, se observa que, para todo $t \in (0, \infty)$,

$$F_X(t) = 1 - \frac{1}{2}e^{-t}, \quad F'_X(t) = \frac{1}{2}e^{-t} = \frac{1}{2}f_Y(t),$$

donde Y es una variable absolutamente continua con $Y \sim \text{Exp}(1)$. Por tanto,

$$\int_{(0, \infty)} e^{its} dP_X(s) = \frac{1}{2} \int_0^\infty e^{its} f_Y(s) ds = \frac{1}{2} \int_0^\infty e^{(it-1)s} ds = \frac{1}{2} \lim_{n \rightarrow \infty} \left[\frac{e^{(it-1)s}}{it-1} \right]_{s=0}^{s=n}.$$

Como

$$|e^{(it-1)n}| = |e^{-n}| = e^{-n} \xrightarrow{n \rightarrow \infty} 0,$$

entonces

$$\frac{1}{2} \lim_{n \rightarrow \infty} \left[\frac{e^{(it-1)s}}{it-1} \right]_{s=0}^{s=n} = \frac{1}{2(1-it)},$$

concluyéndose que

$$\varphi_X(t) = \frac{1}{2} + \frac{1}{2(1-it)}$$

para todo $t \in \mathbb{R}$.

3.2. Reconocimiento de funciones características

Dada una función $\varphi: \mathbb{R} \rightarrow \mathbb{C}$, cabe preguntarse si existe una variable aleatoria X que tenga a φ por función característica.

Definición 3.2.1. Una función $\varphi: \mathbb{R} \rightarrow \mathbb{C}$ se dice que es una *función característica* si existe una variable aleatoria X en algún espacio de probabilidad que tenga a φ por función característica.

Proposición 3.2.2. Sea X una variable aleatoria. Entonces:

- (a) $\overline{\varphi_X}$ es una función característica.
- (b) $\text{Re}(\varphi_X)$ es una función característica.

Proposición 3.2.3. Sea $\{X_n\}_{n=1}^\infty$ una sucesión de variables aleatorias y sea $\{p_n\}_{n=1}^\infty$ una sucesión de números reales positivos verificando $\sum_{n=1}^\infty p_n = 1$. Entonces la función $\varphi: \mathbb{R} \rightarrow \mathbb{C}$ definida por $\varphi(t) = \sum_{n=1}^\infty p_n \varphi_{X_n}(t)$ es una función característica.

Nótese que la función φ de la proposición anterior está bien definida porque para todo $t \in \mathbb{R}$ se tiene $|\varphi_{X_n}(t)| \leq 1$, así que la serie $\sum_{n=1}^\infty p_n \varphi_{X_n}(t)$ es absolutamente convergente y, por tanto, convergente.

Ejemplo 3.2.4. Dado $n \in \mathbb{N}$, se trata de estudiar si la función $g: \mathbb{R} \rightarrow \mathbb{R}$ dada por $g(t) = \cos^n(t)$ es una función característica. Previamente se ha visto que si $X \sim \text{Rad}$, entonces $\varphi_X(t) = \cos(t)$. Por tanto, si X_1, \dots, X_n son variables aleatorias independientes con $X_i \sim \text{Rad}$ para cada $i \in \{1, \dots, n\}$, se tiene que $\varphi_{X_1 + \dots + X_n}(t) = \varphi_{X_1}(t) \dots \varphi_{X_n}(t) = \cos^n(t)$, así que $X_1 + \dots + X_n$ es una variable aleatoria cuya función característica es g .

Ejemplo 3.2.5. Dado $n \in \mathbb{N}$, se trata de estudiar si la función $g: \mathbb{R} \rightarrow \mathbb{R}$ dada por $g(t) = e^{2(\cos(t)-1)}$ es una función característica. Para todo $t \in \mathbb{R}$ se tiene

$$g(t) = e^{-2} e^{2\cos(t)} = e^{-2} \sum_{n=0}^{\infty} \frac{(2\cos(t))^n}{n!} = \sum_{n=0}^{\infty} \frac{1}{e^2} \frac{2^n}{n!} \cos^n(t).$$

Por el ejemplo anterior, la función $t \mapsto \cos^n(t)$, $t \in \mathbb{R}$ es la función característica de la variable aleatoria $X_1 + \dots + X_n$, donde $X_i \sim \text{Rad}$ para cada $i \in \{1, \dots, n\}$. Y como

$$\sum_{n=0}^{\infty} \frac{1}{e^2} \frac{2^n}{n!} = e^{-2} e^2 = 1,$$

entonces, por la proposición anterior, g es una función característica.

Ejemplo 3.2.6. Otra forma de resolver la cuestión planteada en el ejemplo anterior es la siguiente: para todo $t \in \mathbb{R}$,

$$g(t) = e^{2(\frac{e^{it}+e^{-it}}{2}-1)} = e^{e^{it}+e^{-it}-2} = e^{e^{it}-1} e^{-it-1}.$$

Si X e Y son variables aleatorias independientes con $X, Y \sim P(1)$, recordando el Ejemplo 3.1.10, se tiene $g(t) = \varphi_X(t)\varphi_Y(-t) = \varphi_X(t)\varphi_{-Y}(t) = \varphi_{X-Y}(t)$, así que g es la función característica de la variable aleatoria $X - Y$.

Ejemplo 3.2.7. Sea X una variable aleatoria con función de característica $\varphi_X: \mathbb{R} \rightarrow \mathbb{C}$ dada por $\varphi_X(t) = \log(g(t)) + 1$ para alguna función $g: \mathbb{R} \rightarrow (0, \infty)$. Se trata de probar que g es una función característica. En efecto, para todo $t \in \mathbb{R}$,

$$g(t) = e^{\varphi_X(t)-1} = e^{-1} \sum_{n=0}^{\infty} \frac{\varphi_X^n(t)}{n!} = \sum_{n=0}^{\infty} \frac{e^{-1}}{n!} \varphi_X^n(t).$$

Se tiene que

$$\sum_{n=0}^{\infty} \frac{e^{-1}}{n!} = 1$$

y además φ_X^n es la función característica de la suma de n variables aleatorias independientes con la misma distribución que X , así que, por la proposición anterior, g es una función característica.

Definición 3.2.8. Una función $g: \mathbb{R} \rightarrow \mathbb{C}$ se dice que es *definida positiva* si para todo $n \in \mathbb{N}$ y todos $t_1, \dots, t_n \in \mathbb{R}$, $\alpha_1, \dots, \alpha_n \in \mathbb{C}$, se verifica

$$\sum_{i=1}^n \sum_{j=1}^n \alpha_i \overline{\alpha_j} g(t_i - t_j) \geq 0.$$

Puede probarse que para cada $i, j \in \{1, \dots, n\}$ se tiene que $\alpha_i \overline{\alpha_j} g(t_i - t_j) \in \mathbb{R}$, así que tiene sentido preguntarse si la suma anterior es positiva o negativa.

Teorema 3.2.9 (Teorema de Bochner). Una función $\varphi: \mathbb{R} \rightarrow \mathbb{C}$ es una función característica si y solo si es definida positiva, continua en 0 y tal que $\varphi(0) = 1$.

Teorema 3.2.10 (Criterio de Pólya). Sea $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ una función verificando:

- (a) $\varphi(0) = 1$.
- (b) φ es no negativa, par y continua.
- (c) φ es convexa y decreciente en $[0, \infty)$.

Entonces φ es una función característica.

3.3. Teoremas de inversión

En esta sección se exponen algunos resultados que permiten obtener información de una variable aleatoria X a partir de su función característica.

Teorema 3.3.1 (Teorema de inversión). *Sea X una variable aleatoria y sean $a, b \in \mathbb{R}$ con $a < b$. Entonces*

$$P(a < X < b) + \frac{1}{2}P(X = a) + \frac{1}{2}P(X = b) = \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t) dt.$$

Como consecuencia de este resultado, se obtiene que la función característica de una variable aleatoria determina de forma única su función de distribución y su medida de probabilidad.

Teorema 3.3.2. *Dos variables aleatorias X e Y son idénticamente distribuidas si y solo si $\varphi_X = \varphi_Y$.*

Teorema 3.3.3 (Teorema de inversión para densidades). *Si X es una variable aleatoria con*

$$\int_{\mathbb{R}} |\varphi_X(s)| ds < \infty,$$

entonces X es absolutamente continua y para todo $t \in \mathbb{R}$,

$$f_X(t) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-its} \varphi_X(s) ds.$$

Ejemplo 3.3.4. Como aplicación del teorema anterior, vamos a calcular

$$\int_0^\infty \left(\frac{\sin x}{x} \right)^2 dx.$$

La comprobación de que esta integral es finita es fácil. En un ejemplo anterior se vio que si $X \sim U([-1, 1])$, entonces $\varphi_X(t) = \frac{\sin t}{t}$ para todo $t \neq 0$. En consecuencia, si X e Y son variables aleatorias independientes con $X, Y \sim U([-1, 1])$, entonces $\varphi_{X+Y}(t) = \left(\frac{\sin t}{t} \right)^2$ para todo $t \neq 0$. Por el teorema anterior, $X + Y$ es absolutamente continua con función de densidad dada por

$$f_{X+Y}(t) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-its} \frac{\sin^2 s}{s^2} ds$$

para todo $t \in \mathbb{R}$. En particular,

$$f_{X+Y}(0) = \frac{1}{2\pi} \int_{\mathbb{R}} \frac{\sin^2 s}{s^2} ds. \quad (*)$$

Por otro lado, la densidad de $X + Y$ puede hallarse por convolución, obteniéndose tras hacer un par de cuentas que $f_{X+Y}(t) = \frac{1}{2}(1 - \frac{1}{2}|t|)\mathbb{1}_{[-2, 2]}(t)$, $t \in \mathbb{R}$. Como $f_{X+Y}(0) = \frac{1}{2}$, igualando con (*) y usando que el integrando es una función par se concluye que

$$\int_0^\infty \left(\frac{\sin x}{x} \right)^2 dx = \frac{1}{2} \int_{\mathbb{R}} \left(\frac{\sin x}{x} \right)^2 dx = \frac{1}{2} \frac{2\pi}{2} = \frac{\pi}{2}.$$

3.4. Función característica y momentos

Teorema 3.4.1. *Sea X una variable aleatoria. Si para algún $k \in \mathbb{N}$ se tiene que $X \in L^k$, entonces φ_X es k veces derivable y $\varphi_X^{(j)}(t) = E((iX)^j e^{itX})$ para todo $j \leq k$ y todo $t \in \mathbb{R}$.*

Como consecuencia inmediata de este teorema se obtiene un resultado que permite calcular los momentos de una variable aleatoria (siempre que tengan sentido) a partir de la función característica.

Corolario 3.4.2. *Sea X una variable aleatoria. Si para algún $k \in \mathbb{N}$ se tiene que $X \in L^k$, entonces $\varphi_X^{(k)}(0) = i^k E(X^k)$.*

Ejemplo 3.4.3. Si X es una variable aleatoria absolutamente continua con $X \sim N(\mu, \sigma^2)$, se ha visto que $\varphi_X(t) = e^{it\mu - \frac{\sigma^2 t^2}{2}}$ para todo $t \in \mathbb{R}$, luego $\varphi'_X(t) = (i\mu - \sigma^2 t)e^{it\mu - \frac{\sigma^2 t^2}{2}}$, y por el corolario anterior, $E(X) = \frac{\varphi'_X(0)}{i} = \mu$.

En caso de que existan los momentos de todos los órdenes, la función característica de una variable aleatoria adopta una expresión sencilla y manejable.

Teorema 3.4.4. *Sea X una variable aleatoria. Si $X \in L^k$ para cada $k \in \mathbb{N}$, entonces para todo $t \in \mathbb{R}$ se tiene*

$$\varphi_X(t) = \sum_{k=0}^{\infty} \frac{(it)^k}{k!} E(X^k).$$

La existencia de los momentos de todos los órdenes es una condición bastante restrictiva. Con tener la existencia algún momento ya se puede obtener una estimación decente de φ_X .

Teorema 3.4.5. *Sea X una variable aleatoria. Si $X \in L^k$ para algún $k \in \mathbb{N}$, entonces para todo $t \in \mathbb{R}$ se tiene*

$$\varphi_X(t) = \sum_{j=0}^k \frac{(it)^j}{j!} E(X^j) + o(t^k).$$

Se recuerda que si $f, g: \mathbb{R} \rightarrow \mathbb{R}$ son funciones no nulas, cuando se escribe $f(t) = o(g(t))$ se está diciendo que

$$\lim_{t \rightarrow \infty} \frac{f(t)}{g(t)} = 0.$$

Convergencia de variables aleatorias

Al manejar variables aleatorias y funciones de distribución, la convergencia puntual presenta ciertos inconvenientes, pues, como consecuencia del mal comportamiento de la convergencia puntual frente a la continuidad, el límite puntual de funciones de distribución no tiene por qué ser una función de distribución. Es por esto que interesa estudiar otros tipos de convergencia.

Ejemplo 4.0.1. Para cada $n \in \mathbb{N}$, sea X_n una variable aleatoria discreta con $X_n \sim D(\frac{1}{n})$. La función de distribución de X_n viene dada por

$$F_{X_n}(x) = \begin{cases} 0 & \text{si } x < \frac{1}{n}, \\ 1 & \text{si } \frac{1}{n} \leq x. \end{cases}$$

Para todo $x \in \mathbb{R}$,

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = \begin{cases} 0 & \text{si } x \leq 0, \\ 1 & \text{si } 0 < x, \end{cases}$$

luego el límite puntual de $\{F_{X_n}\}_{n=1}^{\infty}$ no es una función de distribución, pues no es continua por la derecha.

4.1. Convergencia débil en distribución

Definición 4.1.1. Sea X una variable aleatoria y sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias.

- (a) Se dice que $\{F_{X_n}\}_{n=1}^{\infty}$ converge débilmente a F_X , y se denota $F_{X_n} \xrightarrow{d} F_X$, si para todo $t \in C_F$ se tiene que

$$\lim_{n \rightarrow \infty} F_{X_n}(t) = F_X(t).$$

- (b) Se dice que $\{X_n\}_{n=1}^{\infty}$ converge en distribución a X , y se denota $X_n \xrightarrow{d} X$, si $\{F_{X_n}\}_{n=1}^{\infty}$ converge débilmente a F_X .

- (c) Se dice que $\{P_{X_n}\}_{n=1}^{\infty}$ converge débilmente a P_X , y se denota $P_{X_n} \xrightarrow{d} P_X$, si para todo $t \in \mathbb{R}$ con $p_X(t) = 0$ se tiene que

$$\lim_{n \rightarrow \infty} P(X_n \leq t) = P(X \leq t).$$

Proposición 4.1.2. Sea X una variable aleatoria y sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias. Entonces

$$F_{X_n} \xrightarrow{d} F_X \iff P_{X_n} \xrightarrow{d} P_X.$$

Ya no hace falta especificar si con *límite débil* se está hablando de funciones de distribución o de medidas de probabilidad, pues ambas nociones son equivalentes.

Proposición 4.1.3. El límite débil, si existe, es único.

Teorema 4.1.4. Sea X una variable aleatoria y sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias. Si $P_{X_n} \xrightarrow{d} P_X$ y $g: \mathbb{R} \rightarrow \mathbb{R}$ es una función de Borel tal que $P_X(D_g) = 0$, entonces $P_{X_n} g^{-1} \xrightarrow{d} P_X g^{-1}$.

Nótese que $P_X(D_g)$ tiene perfecto sentido porque $D_g \in \mathcal{B}$, como es fácil demostrar (esto sigue siendo cierto para una función $g: \mathbb{R} \rightarrow \mathbb{R}$ arbitraria, no necesariamente de Borel).

Por otra parte, se recuerda que $P_{X_n} g^{-1}$ es la medida de probabilidad de la variable aleatoria $g \circ X_n$ para cada $n \in \mathbb{N}$, y $P_X g^{-1}$ es la medida de probabilidad de $g \circ X$.

Corolario 4.1.5. Sea X una variable aleatoria y sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias. Si $X_n \xrightarrow{d} X$ y $g: \mathbb{R} \rightarrow \mathbb{R}$ es una función de Borel tal que $P_X(D_g) = 0$, entonces $g \circ X_n \xrightarrow{d} g \circ X$.

Antes de enunciar el próximo teorema, se aclara que la frontera de un subconjunto $B \subset \mathbb{R}^n$ se va a denotar por ∂B . También se denota $\mathcal{C}_b(\mathbb{R}) = \{f: \mathbb{R} \rightarrow \mathbb{R}: f \text{ es continua y acotada}\}$, mientras que $\mathcal{C}_c(\mathbb{R}) = \{f: \mathbb{R} \rightarrow \mathbb{R}: f \text{ es continua y } \overline{\text{sop } f} \text{ es compacto}\}$.

Teorema 4.1.6. Sea X una variable aleatoria y sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias. Son equivalentes:

(a) $P_{X_n} \xrightarrow{d} P_X$.

(b) Para toda $f \in \mathcal{C}_b(\mathbb{R})$,

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} f dP_{X_n} = \int_{\mathbb{R}} f dP_X.$$

(c) Para toda $f \in \mathcal{C}_c(\mathbb{R})$,

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} f dP_{X_n} = \int_{\mathbb{R}} f dP_X.$$

(d) Para todo $B \in \mathcal{B}$ con $P(X \in \partial B) = 0$,

$$\lim_{n \rightarrow \infty} P(X_n \in B) = P(X \in B).$$

Gracias al último apartado, en la definición de convergencia débil de medidas de probabilidad puede sustituirse $P(X_n \leq t)$ (siendo $t \in \mathbb{R}$ tal que $p_X(t) = 0$) por $P(X_n \in B)$, para cualquier $B \in \mathcal{B}$ con $P(X \in \partial B) = 0$. Si $P(X \in \partial B) \neq 0$, el resultado no es cierto. Veámoslo.

Ejemplo 4.1.7. En el espacio medible $(\mathbb{R}, \mathcal{B})$, considérese la medida de probabilidad δ_0 y la sucesión de medidas de probabilidad $\{\delta_{\frac{1}{n}}\}_{n=1}^{\infty}$. Si $f: \mathbb{R} \rightarrow \mathbb{R}$ es continua y acotada,

$$\int_{\mathbb{R}} f d\delta_{\frac{1}{n}} = f\left(\frac{1}{n}\right).$$

Como f es continua,

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} f d\delta_{\frac{1}{n}} = \lim_{n \rightarrow \infty} f\left(\frac{1}{n}\right) = f(0).$$

Por tanto,

$$\int_{\mathbb{R}} f d\delta_0 = f(0) = \lim_{n \rightarrow \infty} \int_{\mathbb{R}} f d\delta_{\frac{1}{n}}.$$

Por el teorema anterior, $\delta_{\frac{1}{n}} \xrightarrow{d} \delta_0$. Sin embargo, tomando $B = (-\infty, 0]$, se tiene que

$$\lim_{n \rightarrow \infty} \delta_{\frac{1}{n}}(B) = \lim_{n \rightarrow \infty} 0 = 0,$$

pero $\delta_0(B) = 1$, luego

$$\lim_{n \rightarrow \infty} \delta_{\frac{1}{n}}(B) \neq \delta_0(B).$$

La hipótesis que falla es que $\delta_0(\partial B) = \delta_0(\{0\}) = 1 \neq 0$.

Si se trabaja con variables aleatorias absolutamente continuas, se dispone de una condición suficiente para la convergencia en distribución que puede ser útil en la práctica.

Proposición 4.1.8. *Sea X una variable aleatoria absolutamente continua y sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias absolutamente continuas. Si para casi todo $t \in \mathbb{R}$ se tiene que*

$$\lim_{n \rightarrow \infty} f_{X_n}(t) = f_X(t),$$

entonces $X_n \xrightarrow{d} X$.

Ejemplo 4.1.9. Sea $\{U_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias independientes y absolutamente continuas con $U_n \sim U([0, 1])$ para todo $n \in \mathbb{N}$. Consideremos la sucesión de variables aleatorias $\{X_n\}_{n=1}^{\infty}$, donde $X_n = \max\{U_1, U_2, \dots, U_n\}$ para cada $n \in \mathbb{N}$. Se trata de estudiar la convergencia en distribución de $\{X_n\}_{n=1}^{\infty}$. Dado $n \in \mathbb{N}$,

$$F_{X_n}(t) = P(X_n \leq t) = P(U_1 \leq t, U_2 \leq t, \dots, U_n \leq t) = \prod_{i=1}^n P(U_i \leq t) = \prod_{i=1}^n F_{U_i}(t),$$

donde, para cada $i \in \{1, \dots, n\}$,

$$F_{U_i}(t) = \begin{cases} 0 & \text{si } t < 0, \\ t & \text{si } 0 \leq t < 1, \\ 1 & \text{si } 1 \leq t. \end{cases}$$

Por tanto,

$$F_{X_n}(t) = \begin{cases} 0 & \text{si } t < 0, \\ t^n & \text{si } 0 \leq t < 1, \\ 1 & \text{si } 1 \leq t. \end{cases}$$

Tomando límite,

$$\lim_{n \rightarrow \infty} F_{X_n}(t) = \begin{cases} 0 & \text{si } t < 1, \\ 1 & \text{si } 1 \leq t. \end{cases}$$

Se observa que $\{F_{X_n}\}_{n=1}^{\infty}$ converge débilmente a F_X , siendo F_X la función de distribución de una variable aleatoria discreta X con $X \sim D(1)$. Se concluye que la sucesión de variables aleatorias $\{X_n\}_{n=1}^{\infty}$ converge en distribución a X , o, como suele decirse, $\{X_n\}_{n=1}^{\infty}$ converge en distribución a $D(1)$.

Ejemplo 4.1.10. Siguiendo con el ejemplo anterior, sea $Y_n = n(1 - X_n)$ para cada $n \in \mathbb{N}$. Se trata de estudiar la convergencia en distribución de $\{Y_n\}_{n=1}^{\infty}$. Se tiene que

$$F_{Y_n}(t) = P(Y_n \leq t) = P\left(X_n \geq 1 - \frac{t}{n}\right) = 1 - F_{X_n}\left(\left(1 - \frac{t}{n}\right)^-\right).$$

Como F_{X_n} es continua,

$$F_{Y_n}(t) = 1 - F_{X_n}\left(1 - \frac{t}{n}\right) = \begin{cases} 1 & \text{si } 1 - \frac{t}{n} < 0, \\ 1 - \left(1 - \frac{t}{n}\right)^n & \text{si } 0 \leq 1 - \frac{t}{n} < 1, \\ 0 & \text{si } 1 \leq 1 - \frac{t}{n}. \end{cases} = \begin{cases} 0 & \text{si } t \leq 0, \\ 1 - \left(1 - \frac{t}{n}\right)^n & \text{si } 0 < t \leq n, \\ 1 & \text{si } n < t. \end{cases}$$

Por tanto,

$$\lim_{n \rightarrow \infty} F_{Y_n}(t) = \begin{cases} 0 & \text{si } t \leq 0, \\ 1 - e^{-t} & \text{si } 0 < t. \end{cases}$$

En consecuencia, $\{F_{Y_n}\}_{n=1}^{\infty}$ converge débilmente a la función de distribución de una variable aleatoria Y absolutamente continua y con $Y \sim \text{Exp}(1)$, luego $\{Y_n\}_{n=1}^{\infty}$ converge en distribución a Y , o, como suele decirse, $\{Y_n\}_{n=1}^{\infty}$ converge en distribución a $\text{Exp}(1)$.

Ejemplo 4.1.11. Para cada $n \in \mathbb{N}$, sea $F_n : \mathbb{R} \rightarrow [0, 1]$ la función definida por

$$F_n(x) = \left(\frac{1}{2} + n\frac{x}{2}\right) \mathbb{I}_{(-\frac{1}{n}, \frac{1}{n})}(x) + \mathbb{I}_{[\frac{1}{n}, \infty)}(x).$$

Es claro que F_n es una función de distribución para cada $n \in \mathbb{N}$. Además,

$$\lim_{n \rightarrow \infty} F_n(x) = \begin{cases} 0 & \text{si } x < 0, \\ \frac{1}{2} & \text{si } x = 0, \\ 1 & \text{si } 0 < x. \end{cases}$$

Se observa que el límite puntual de F_n no es una función de distribución (no es continuo por la derecha en 0), así que no puede ser límite débil de $\{F_n\}_{n=1}^{\infty}$. Sin embargo, si se define $\tilde{F} : \mathbb{R} \rightarrow [0, 1]$ mediante

$$\tilde{F}(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 & \text{si } 0 \leq x, \end{cases}$$

entonces \tilde{F} es una función de distribución y

$$\lim_{n \rightarrow \infty} F_n(x) = \tilde{F}(x)$$

para todo $x \in C_{\tilde{F}}$. Por tanto, $\{F_n\}_{n=1}^{\infty}$ converge en distribución a \tilde{F} .

Para simplificar la escritura, en el resultado que sigue se hará referencia a una función característica φ asociada a una función de distribución F . Esto quiere decir que si X es una variable aleatoria con función de distribución F (que se sabe que existe), entonces φ es la función característica de X .

Recuérdese que si Y fuese otra variable aleatoria con función característica φ , entonces X e Y son idénticamente distribuidas, así que esto de hablar de funciones características asociadas a funciones de distribución en lugar de variables aleatorias no presenta ninguna ambigüedad.

Teorema 4.1.12 (Teorema de continuidad de Lévy). Sea $\{X_n\}_{n=1}^\infty$ una sucesión de variables aleatorias.

- (a) Si existe una función de distribución F tal que $F_{X_n} \xrightarrow{d} F$, entonces $\varphi_{X_n}(t) \xrightarrow{n \rightarrow \infty} \varphi(t)$ para todo $t \in \mathbb{R}$, donde φ es la función característica asociada a F .
- (b) Recíprocamente, si existe una función $\varphi: \mathbb{R} \rightarrow \mathbb{C}$ continua en 0 y con $\varphi_{X_n}(t) \xrightarrow{n \rightarrow \infty} \varphi(t)$ para todo $t \in \mathbb{R}$, entonces φ es la función característica asociada a una función de distribución F verificando $F_{X_n} \xrightarrow{d} F$.

Ejemplo 4.1.13. Considérense n variables aleatorias independientes X_1, \dots, X_n con $X_i \sim \text{Rad}$ para cada $i \in \{1, \dots, n\}$. Se sabe que $\varphi_{X_i}(t) = \cos(t)$ para todo $i \in \{1, \dots, n\}$ y todo $t \in \mathbb{R}$. Estudiemos la convergencia en distribución de $\{Y_n\}_{n=1}^\infty$, donde, para cada $n \in \mathbb{N}$, $Y_n = \frac{X_1 + \dots + X_n}{\sqrt{n}}$. Si $n \in \mathbb{N}$ y $t \in \mathbb{R}$,

$$\varphi_{Y_n}(t) = \varphi_{X_1 + \dots + X_n} \left(\frac{t}{\sqrt{n}} \right) = \varphi_{X_1} \left(\frac{t}{\sqrt{n}} \right) \dots \varphi_{X_n} \left(\frac{t}{\sqrt{n}} \right) = \cos^n \left(\frac{t}{\sqrt{n}} \right).$$

Se tiene que

$$\lim_{n \rightarrow \infty} \cos^n \left(\frac{t}{\sqrt{n}} \right) = \lim_{n \rightarrow \infty} e^{\log(\cos^n(\frac{t}{\sqrt{n}}))} = \lim_{n \rightarrow \infty} e^{n \log(\cos(\frac{t}{\sqrt{n}}))}. \quad (*)$$

Se halla primero el límite del exponente:

$$\begin{aligned} \lim_{x \rightarrow \infty} x \log \left(\cos \left(\frac{t}{\sqrt{x}} \right) \right) &= \lim_{x \rightarrow \infty} \frac{\log(\cos(\frac{t}{\sqrt{x}}))}{\frac{1}{x}} = \lim_{x \rightarrow \infty} \frac{\frac{1}{\cos(\frac{t}{\sqrt{x}})} \text{sen}(\frac{t}{\sqrt{x}}) \cdot \frac{t}{2x^{3/2}}}{-\frac{1}{x^2}} = -\frac{t}{2} \lim_{x \rightarrow \infty} \sqrt{x} \frac{\text{sen}(\frac{t}{\sqrt{x}})}{\cos(\frac{t}{\sqrt{x}})} \\ &= -\frac{t}{2} \lim_{x \rightarrow \infty} \frac{\text{sen}(\frac{t}{\sqrt{x}})}{\frac{t}{\sqrt{x}}} \cdot \frac{t}{\cos(\frac{t}{\sqrt{x}})} = -\frac{t}{2} \cdot 1 \cdot \frac{t}{\cos(0)} = -\frac{t^2}{2}, \end{aligned}$$

utilizándose en la segunda igualdad la regla de L'Hôpital, y en la última igualdad que

$$\lim_{x \rightarrow 0} \frac{\text{sen } x}{x} = 1.$$

Volviendo a (*), para todo $t \in \mathbb{R}$ se verifica

$$\lim_{n \rightarrow \infty} \varphi_{Y_n}(t) = \lim_{n \rightarrow \infty} \cos^n \left(\frac{t}{\sqrt{n}} \right) = e^{-\frac{t^2}{2}} = \varphi(t).$$

Se observa que φ , además de ser continua en 0, es la función característica de una variable aleatoria Y absolutamente continua con $Y \sim N(0, 1)$, como ya se ha visto en algún ejemplo anterior. Por el teorema de continuidad de Lévy, se concluye que $\{Y_n\}_{n=1}^\infty$ converge en distribución a $N(0, 1)$.

4.2. Convergencia vaga

Definición 4.2.1. Sea μ una medida sobre $(\mathbb{R}, \mathcal{B})$ y sea $\{\mu_n\}_{n=1}^\infty$ una sucesión de medidas sobre $(\mathbb{R}, \mathcal{B})$. Se dice que $\{\mu_n\}_{n=1}^\infty$ converge vagamente a μ y se denota $\mu_n \xrightarrow{v} \mu$, si para toda $f \in \mathcal{C}_c(\mathbb{R})$ se tiene que

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} f d\mu_n = \int_{\mathbb{R}} f d\mu.$$

En caso de que todas las medidas que aparecen en la definición anterior sean de probabilidad, el Teorema 4.1.6 permite afirmar que la convergencia vaga y la convergencia débil son equivalentes.

Definición 4.2.2. Una familia $\{\mu_i\}_{i \in I}$ de medidas de probabilidad sobre $(\mathbb{R}, \mathcal{B})$ se dice que es *ajustada* si para todo $\varepsilon > 0$ existe $a > 0$ tal que $\mu_i([-a, a]) \geq 1 - \varepsilon$ para todo $i \in I$.

Al ser I un conjunto arbitrario, queda claro qué quiere decir que una sucesión de medidas de probabilidad sobre $(\mathbb{R}, \mathcal{B})$ sea ajustada.

Proposición 4.2.3. Sea μ una medida de probabilidad sobre $(\mathbb{R}, \mathcal{B})$ y sea $\{\mu_n\}_{n=1}^\infty$ una sucesión de medidas de probabilidad sobre $(\mathbb{R}, \mathcal{B})$. Si $\mu_n \xrightarrow{d} \mu$, entonces $\{\mu_n\}_{n=1}^\infty$ es ajustada.

En el ejemplo que sigue se muestra una sucesión de medidas de probabilidad sobre $(\mathbb{R}, \mathcal{B})$ que no es ajustada.

Ejemplo 4.2.4. Para cada $n \in \mathbb{N}$, sea $\mu_n = \frac{1}{2}(\delta_0 + \delta_n)$. Es claro que μ_n es una medida de probabilidad sobre $(\mathbb{R}, \mathcal{B})$ para cada $n \in \mathbb{N}$. Veamos que la sucesión $\{\mu_n\}_{n=1}^\infty$ no es ajustada, es decir, que existe $\varepsilon_0 > 0$ tal que para todo $a > 0$ existe $n_0 \in \mathbb{N}$ verificando $\mu_{n_0}([-a, a]) < 1 - \varepsilon_0$. En efecto, si $a > 0$, tomamos $n_0 \in \mathbb{N}$ con $a < n_0$ y entonces se tiene

$$\mu_{n_0}([-a, a]) = \frac{1}{2}(\delta_0([-a, a]) + \delta_{n_0}([-a, a])) = \frac{1}{2}.$$

Basta tomar $\varepsilon_0 = \frac{1}{2}$.

Teorema 4.2.5. Sea $\{\mu_n\}_{n=1}^\infty$ una sucesión de medidas de probabilidad sobre $(\mathbb{R}, \mathcal{B})$ y sea μ una medida sobre $(\mathbb{R}, \mathcal{B})$ y Si $\mu_n \xrightarrow{v} \mu$, entonces $\mu(\mathbb{R}) = 1$ si y solo si la sucesión $\{\mu_n\}_{n=1}^\infty$ es ajustada, y en ese caso, $\mu_n \xrightarrow{d} \mu$.

Definición 4.2.6. Una familia $\mathcal{M} = \{\mu_i\}_{i \in I}$ de medidas de probabilidad sobre $(\mathbb{R}, \mathcal{B})$ se dice que es *relativamente compacta* si toda sucesión en \mathcal{M} tiene una subsucesión que converge débilmente.

Si una familia $\mathcal{M} = \{\mu_i\}_{i \in I}$ de medidas de probabilidad sobre $(\mathbb{R}, \mathcal{B})$ es relativamente compacta y $\{\mu_n\}_{n=1}^\infty$ es una sucesión en \mathcal{M} , el límite débil de una subsucesión cualquiera de $\{\mu_n\}_{n=1}^\infty$ (en caso de que exista) no tiene por qué estar en \mathcal{M} .

Teorema 4.2.7 (Teorema de Prohorov). Una familia $\{\mu_i\}_{i \in I}$ de medidas de probabilidad sobre $(\mathbb{R}, \mathcal{B})$ es relativamente compacta si y solo si es ajustada.

Corolario 4.2.8. Una sucesión $\{\mu_n\}_{n=1}^\infty$ de medidas de probabilidad sobre $(\mathbb{R}, \mathcal{B})$ es ajustada si y solo si existe una subsucesión de $\{\mu_n\}_{n=1}^\infty$ que converge débilmente.

4.3. Ley 0-1 de Kolmogorov

Definición 4.3.1. Sea (Ω, \mathcal{A}) un espacio medible y sea $\{A_n\}_{n=1}^\infty$ una sucesión de elementos de \mathcal{A} .

(a) Se define el *límite inferior* de $\{A_n\}_{n=1}^\infty$ como

$$\liminf_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k.$$

(b) Se define el *límite superior* de $\{A_n\}_{n=1}^\infty$ como

$$\limsup_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k.$$

Es claro que el límite superior y el límite inferior son elementos de \mathcal{A} , pues no son más que intersecciones o uniones numerables de elementos de \mathcal{A} . Además,

$$\omega \in \liminf_{n \rightarrow \infty} A_n \iff \text{existe } n \in \mathbb{N} \text{ tal que para todo } k \geq n \text{ se tiene que } \omega \in A_k,$$

es decir,

$$\omega \in \liminf_{n \rightarrow \infty} A_n \iff \omega \text{ pertenece a todos los } A_k \text{ excepto un número finito.}$$

Por otro lado,

$$\omega \in \limsup_{n \rightarrow \infty} A_n \iff \text{para todo } n \in \mathbb{N} \text{ existe } k \geq n \text{ tal que } \omega \in A_k,$$

o, dicho de otra manera,

$$\omega \in \limsup_{n \rightarrow \infty} A_n \iff \omega \text{ pertenece a infinitos } A_k.$$

Proposición 4.3.2. Sea (Ω, \mathcal{A}, P) un espacio de probabilidad y sea $\{A_n\}_{n=1}^{\infty}$ una sucesión de elementos de \mathcal{A} . Entonces

$$P\left(\liminf_{n \rightarrow \infty} A_n\right) \leq \liminf_{n \rightarrow \infty} P(A_n) \leq \limsup_{n \rightarrow \infty} P(A_n) \leq P\left(\limsup_{n \rightarrow \infty} A_n\right).$$

Un resultado en el que se afirma que un determinado suceso tiene probabilidad 0 o probabilidad 1, se conoce como *ley 0-1*. El teorema que sigue es una de estas leyes.

Teorema 4.3.3 (Lemas de Borel-Cantelli). Sea (Ω, \mathcal{A}, P) un espacio de probabilidad y sea $\{A_n\}_{n=1}^{\infty}$ una sucesión de elementos de \mathcal{A} .

(a) Si $\sum_{n=1}^{\infty} P(A_n) < \infty$, entonces

$$P\left(\limsup_{n \rightarrow \infty} A_n\right) = 0.$$

(b) Si los sucesos $\{A_n\}_{n=1}^{\infty}$ son independientes y $\sum_{n=1}^{\infty} P(A_n) = \infty$, entonces

$$P\left(\limsup_{n \rightarrow \infty} A_n\right) = 1.$$

El primer apartado del resultado anterior se conoce como *primer lema de Borel-Cantelli*, y el segundo, como *segundo lema de Borel-Cantelli*.

Definición 4.3.4. Sea (Ω, \mathcal{A}, P) un espacio de probabilidad y sea $\{A_n\}_{n=1}^{\infty}$ una sucesión de elementos de \mathcal{A} . Se define la σ -álgebra cola de $\{A_n\}_{n=1}^{\infty}$ como $\mathcal{T} = \bigcap_{n=1}^{\infty} \sigma(\{A_k : k \geq n\})$. Los elementos de \mathcal{T} se denominan *sucesos cola*.

Si (Ω, \mathcal{A}, P) es un espacio de probabilidad y $\{A_n\}_{n=1}^{\infty}$ es una sucesión de elementos de \mathcal{A} , se observa que

$$\liminf_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k \in \sigma(\{A_k : k \geq m\}), \quad \limsup_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k \in \sigma(\{A_k : k \geq m\}),$$

para todo $m \in \mathbb{N}$, así que

$$\liminf_{n \rightarrow \infty} A_n \in \mathcal{T}, \quad \limsup_{n \rightarrow \infty} A_n \in \mathcal{T}.$$

Teorema 4.3.5 (Ley 0-1 de Kolmogorov). Sea (Ω, \mathcal{A}, P) un espacio de probabilidad y sea $\{A_n\}_{n=1}^{\infty}$ una sucesión de elementos de \mathcal{A} . Si los sucesos $\{A_n\}_{n=1}^{\infty}$ son independientes, entonces para todo $A \in \mathcal{T}$ se tiene que $P(A) \in \{0, 1\}$.

En consecuencia, si $\{A_n\}_{n=1}^{\infty}$ son sucesos independientes en un espacio de probabilidad (Ω, \mathcal{A}, P) , entonces

$$P\left(\limsup_{n \rightarrow \infty} A_n\right) \in \{0, 1\}.$$

Posteriormente será de utilidad disponer de una ley 0-1 de Kolmogorov en términos de variables aleatorias en lugar de sucesos. Se introducen para ello las definiciones siguientes:

Definición 4.3.6. Sea X una variable aleatoria. Se define la σ -álgebra generada por X como $\sigma(X) = \{X^{-1}(B) : B \in \mathcal{B}\}$.

Definición 4.3.7. Sea $\{X_i\}_{i \in I}$ una familia de variables aleatorias. Se define la σ -álgebra generada por $\{X_i\}_{i \in I}$ como $\sigma(X_i : i \in I) = \sigma(\bigcup_{i \in I} \sigma(X_i))$.

En otros términos, la σ -álgebra generada por una variable aleatoria X es la menor σ -álgebra respecto de la cual X es una variable aleatoria, mientras que la σ -álgebra generada por una familia de variables aleatorias $\{X_i\}_{i \in I}$ es la menor σ -álgebra respecto de la cual X_i es una variable aleatoria para todo $i \in I$.

Definición 4.3.8. Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias. Se define la σ -álgebra cola de $\{X_n\}_{n=1}^{\infty}$ como $\mathcal{T} = \bigcap_{n=1}^{\infty} \sigma(X_k : k \geq n)$.

Teorema 4.3.9 (Ley 0-1 de Kolmogorov). Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias independientes. Entonces $P(A) \in \{0, 1\}$ para todo $A \in \mathcal{T}$.

4.4. Convergencia casi segura

Definición 4.4.1. Sea X una variable aleatoria y sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias. Se dice que $\{X_n\}_{n=1}^{\infty}$ converge casi seguro a X , y se denota $X_n \xrightarrow{cs} X$, si

$$P(\{\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}) = 1.$$

Es bien sabido que $A = \{\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}$ es un conjunto medible, y para probarlo basta tener en cuenta que

$$A = \bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} \left\{ \omega \in \Omega : |X_m(\omega) - X(\omega)| < \frac{1}{k} \right\}.$$

Esto también muestra que $A \in \mathcal{T} = \bigcap_{n=1}^{\infty} \sigma(X_k : k \geq n)$, y de la ley 0-1 de Kolmogorov se deduce lo siguiente:

Corolario 4.4.2. Sea X una variable aleatoria y sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias. Si las variables aleatorias $\{X_n\}_{n=1}^{\infty}$ son independientes, entonces $X_n \xrightarrow{cs} X$ si y solo si

$$P(\{\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}) \neq 0.$$

Teorema 4.4.3 (Caracterización de la convergencia casi segura). Sea X una variable aleatoria y sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias. Entonces $X_n \xrightarrow{cs} X$ si y solo si para todo $\varepsilon > 0$ se tiene que

$$\lim_{n \rightarrow \infty} P \left(\bigcap_{m=n}^{\infty} \{\omega \in \Omega : |X_m(\omega) - X(\omega)| < \varepsilon\} \right) = 1,$$

o, equivalentemente,

$$\lim_{n \rightarrow \infty} P \left(\bigcup_{m=n}^{\infty} \{\omega \in \Omega : |X_m(\omega) - X(\omega)| \geq \varepsilon\} \right) = 0.$$

4.5. Convergencia en probabilidad

Definición 4.5.1. Sea X una variable aleatoria y sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias. Se dice que $\{X_n\}_{n=1}^{\infty}$ converge en probabilidad a X , y se denota $X_n \xrightarrow{p} X$, si para todo $\varepsilon > 0$ se tiene que

$$\lim_{n \rightarrow \infty} P(\{\omega \in \Omega : |X_n(\omega) - X(\omega)| \geq \varepsilon\}) = 0,$$

o, equivalentemente,

$$\lim_{n \rightarrow \infty} P(\{\omega \in \Omega : |X_n(\omega) - X(\omega)| < \varepsilon\}) = 1.$$

Se enuncian a continuación las propiedades más destacadas de la convergencia en probabilidad, entre ellas su relación con la convergencia casi segura y la convergencia en distribución.

Proposición 4.5.2. Se verifican las siguientes propiedades:

- (a) Si $X_n \xrightarrow{p} X$ y $X_n \xrightarrow{p} Y$, entonces $P(\{\omega \in \Omega : X(\omega) = Y(\omega)\}) = 1$.
- (b) Si $f: \mathbb{R} \rightarrow \mathbb{R}$ es continua y $X_n \xrightarrow{p} X$, entonces $f(X_n) \xrightarrow{p} f(X)$.
- (c) Si $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ es continua, $X_n \xrightarrow{p} X$ e $Y_n \xrightarrow{p} Y$, entonces $f(X_n, Y_n) \xrightarrow{p} f(X, Y)$.
- (d) Si $X_n \xrightarrow{cs} X$, entonces $X_n \xrightarrow{p} X$.
- (e) Si $X_n \xrightarrow{p} X$, entonces $X_n \xrightarrow{d} X$.
- (f) Si $X_n \xrightarrow{d} c$ para alguna constante $c \in \mathbb{R}$, entonces $X_n \xrightarrow{p} c$.
- (g) $X_n \xrightarrow{p} X$ si y solo si para cada subsucesión $\{X_{n_k}\}_{k=1}^{\infty}$ de $\{X_n\}_{n=1}^{\infty}$ hay una subsucesión $\{X_{n_{k_j}}\}_{j=1}^{\infty}$ de $\{X_{n_k}\}_{k=1}^{\infty}$ tal que $X_{n_{k_j}} \xrightarrow{cs} X$.

Como consecuencia del tercer apartado, se obtiene que la convergencia en probabilidad se comporta bien frente a la suma y el producto.

Corolario 4.5.3. Supóngase que $X_n \xrightarrow{p} X$ e $Y_n \xrightarrow{p} Y$. Entonces:

- (a) $aX_n + bY_n \xrightarrow{p} aX + bY$ para $a, b \in \mathbb{R}$ cualesquiera.
- (b) $X_n Y_n \xrightarrow{p} XY$.
- (c) Si $P(\{\omega \in \Omega : Y_n(\omega) \neq 0\}) = 1$ para todo $n \in \mathbb{N}$, entonces $\frac{X_n}{Y_n} \xrightarrow{p} \frac{X}{Y}$.

4.6. Leyes de los grandes números

Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias y sean $\{a_n\}_{n=1}^{\infty}$ y $\{b_n\}_{n=1}^{\infty}$ dos sucesiones de números reales. Si $b_n \neq 0$ para todo $n \in \mathbb{N}$ y $b_n \xrightarrow{n \rightarrow \infty} \infty$, las leyes de los grandes números estudian la convergencia en algún sentido de la sucesión de variables aleatorias $\{\frac{S_n - a_n}{b_n}\}_{n=1}^{\infty}$, donde

$S_n = \sum_{i=1}^n X_i$. En concreto, en esta sección se estudiará la convergencia en probabilidad y la convergencia casi segura de la sucesión $\{\frac{S_n - E(S_n)}{n}\}_{n=1}^\infty$.

Definición 4.6.1. Una sucesión de variables aleatorias $\{X_n\}_{n=1}^\infty$ se dice que verifica la *ley débil de los grandes números* si

$$\frac{S_n - E(S_n)}{n} \xrightarrow{p} 0,$$

donde $S_n = \sum_{i=1}^n X_i$ para cada $n \in \mathbb{N}$.

Definición 4.6.2. Una sucesión de variables aleatorias $\{X_n\}_{n=1}^\infty$ se dice que verifica la *ley fuerte de los grandes números* si

$$\frac{S_n - E(S_n)}{n} \xrightarrow{cs} 0,$$

donde $S_n = \sum_{i=1}^n X_i$ para cada $n \in \mathbb{N}$.

Teorema 4.6.3. Si $\{X_n\}_{n=1}^\infty$ es una sucesión de variables aleatorias independientes e idénticamente distribuidas con esperanza $\mu \in \mathbb{R}$, entonces

$$\frac{S_n}{n} \xrightarrow{p} \mu,$$

es decir, $\{X_n\}_{n=1}^\infty$ verifica la ley débil de los grandes números.

Teorema 4.6.4 (Criterio de convergencia de Kolmogorov). Sea $\{X_n\}_{n=1}^\infty$ una sucesión de variables aleatorias independientes con $X_n \in L^1$ para todo $n \in \mathbb{N}$. Si $\sum_{n=1}^\infty \text{Var}(X_n) < \infty$, entonces

$$P\left(\sum_{n=1}^\infty (X_n - E(X_n)) < \infty\right) = 1.$$

Ejemplo 4.6.5. Sea $\{X_n\}_{n=1}^\infty$ una sucesión de variables aleatorias independientes tales que

$$P\left(X_n = \frac{1}{n^\alpha}\right) = P\left(X_n = -\frac{1}{n^\alpha}\right) = \frac{1}{2}$$

para cada $n \in \mathbb{N}$, con $\alpha > 0$ fijo. Entonces $E(X_n) = 0$ y

$$\text{Var}(X_n) = E(X_n^2) - E(X_n)^2 = \frac{1}{n^{2\alpha}} \frac{1}{2} + \frac{1}{n^{2\alpha}} \frac{1}{2} = \frac{1}{n^{2\alpha}}$$

para todo $n \in \mathbb{N}$. Por tanto, si $\alpha > \frac{1}{2}$, se tiene que $\sum_{n=1}^\infty \text{Var}(X_n) < \infty$ y el teorema anterior permite afirmar que

$$P\left(\sum_{n=1}^\infty X_n < \infty\right) = 1.$$

Teorema 4.6.6 (Lema de Kronecker). Sea $\{X_n\}_{n=1}^\infty$ una sucesión de variables aleatorias y sea $\{a_n\}_{n=1}^\infty$ una sucesión de números reales con $a_n > 0$ para todo $n \in \mathbb{N}$ y $a_n \xrightarrow{n \rightarrow \infty} \infty$. Si

$$P\left(\sum_{n=1}^\infty \frac{X_n}{a_n} < \infty\right) = 1,$$

entonces

$$\frac{1}{a_n} \sum_{k=1}^n X_k \xrightarrow{cs} 0.$$

Teorema 4.6.7. Si $\{X_n\}_{n=1}^{\infty}$ es una sucesión de variables aleatorias independientes e idénticamente distribuidas con esperanza $\mu \in \mathbb{R}$, entonces

$$\frac{S_n}{n} \xrightarrow{cs} \mu,$$

es decir, $\{X_n\}_{n=1}^{\infty}$ verifica la ley fuerte de los grandes números.

Si las variables aleatorias son independientes pero no idénticamente distribuidas, bajo ciertas hipótesis se sigue verificando la ley fuerte de los grandes números.

Teorema 4.6.8 (Condición suficiente de Kolmogorov). Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias independientes con $X_n \in L^1$ para todo $n \in \mathbb{N}$. Si $\sum_{n=1}^{\infty} \frac{\text{Var}(X_n)}{n^2} < \infty$, entonces

$$\frac{S_n - E(S_n)}{n} \xrightarrow{cs} 0,$$

es decir, $\{X_n\}_{n=1}^{\infty}$ verifica la ley fuerte de los grandes números.

Definición 4.6.9. Dos sucesiones de variables aleatorias $\{X_n\}_{n=1}^{\infty}$ e $\{Y_n\}_{n=1}^{\infty}$ se dicen *equivalentes en convergencia* si $\sum_{n=1}^{\infty} P(X_n \neq Y_n) < \infty$.

Por ejemplo, si $\{X_n\}_{n=1}^{\infty}$ es una sucesión de variables aleatorias y $\{c_n\}_{n=1}^{\infty}$ es una sucesión de números reales tal que $\sum_{n=1}^{\infty} P(|X_n| > c_n) < \infty$, entonces $\{X_n\}_{n=1}^{\infty}$ e $\{X_n \mathbb{1}_{|X_n| \leq c_n}\}_{n=1}^{\infty}$ son equivalentes en convergencia.

Proposición 4.6.10. Si dos sucesiones de variables aleatorias $\{X_n\}_{n=1}^{\infty}$ e $\{Y_n\}_{n=1}^{\infty}$ son equivalentes en convergencia, entonces $\{X_n\}_{n=1}^{\infty}$ verifica la ley fuerte de los grandes números si y solo si $\{Y_n\}_{n=1}^{\infty}$ verifica la ley fuerte de los grandes números.

Si $\{X_n\}_{n=1}^{\infty}$ es una sucesión de variables aleatorias tal que $\sum_{n=1}^{\infty} \frac{\text{Var}(X_n)}{n^2} = \infty$ y quiere estudiarse la convergencia casi segura de la sucesión $\{\frac{S_n - E(S_n)}{n}\}_{n=1}^{\infty}$, puede construirse una sucesión $\{Y_n\}_{n=1}^{\infty}$ que sea equivalente en convergencia a $\{X_n\}_{n=1}^{\infty}$ y que verifique $\sum_{n=1}^{\infty} \frac{\text{Var}(Y_n)}{n^2} < \infty$. En estas circunstancias, la proposición anterior y la condición suficiente de Kolmogorov permitirán afirmar que $\{X_n\}_{n=1}^{\infty}$ verifica la ley fuerte de los grandes números.

Por otro lado, si X es una variable aleatoria en un espacio de probabilidad (Ω, \mathcal{A}, P) y $c > 0$, se denota $X^{(c)} = X \mathbb{1}_{|X| \leq c}$.

Teorema 4.6.11 (Teorema de las tres series de Kolmogorov). Sea $\{X_n\}_{n=1}^{\infty}$ una sucesión de variables aleatorias independientes. Si existe $c > 0$ tal que

$$\sum_{n=1}^{\infty} P(|X_n| > c) < \infty, \quad \sum_{n=1}^{\infty} E(X_n^{(c)}) < \infty, \quad \sum_{n=1}^{\infty} \text{Var}(X_n^{(c)}) < \infty,$$

entonces

$$P\left(\sum_{n=1}^{\infty} X_n < \infty\right) = 1.$$

Definición 4.6.12. Una sucesión de variables aleatorias $\{X_n\}_{n=1}^{\infty}$ se dice que verifica el *teorema central del límite* si

$$\frac{S_n - E(S_n)}{\sqrt{\text{Var}(S_n)}} \xrightarrow{d} N(0, 1),$$

donde $S_n = \sum_{i=1}^n X_i$ para cada $n \in \mathbb{N}$.

Teorema 4.6.13 (Teorema de Lindeberg-Lévy). *Considérese una sucesión $\{X_n\}_{n=1}^{\infty}$ de variables aleatorias independientes e idénticamente distribuidas con esperanza $\mu \in \mathbb{R}$ y varianza $\sigma^2 \in (0, \infty)$. Entonces*

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} N(0, 1),$$

es decir, $\{X_n\}_{n=1}^{\infty}$ verifica el teorema central del límite.