Probabilidad II
Tema 1: Espacios de probabilidad

José R. Berrendero

Departamento de Matemáticas
Universidad Autónoma de Madrid

Estructura de este tema

- Álgebras, σ-álgebras.
- Espacios de probabilidad.
- Propiedades elementales.
- Límites de sucesiones de conjuntos.
- π-sistemas, λ-sistemas. Teorema de Dynkin.
- Función de distribución.
- Probabilidad condicionada.
- Variables y vectores aleatorios.
- σ-álgebra generada por una variable aleatoria.

σ-álgebras

Sea Ω un conjunto no vacío.

Un álgebra $\mathcal{F}$ es una clase de subconjuntos de Ω que verifica

- $\Omega \in \mathcal{F}$.
- Si $A \in \mathcal{F}$, entonces $A^c \in \mathcal{F}$, donde $A^c := \Omega - A$.
- Si $A, B \in \mathcal{F}$, entonces $A \cup B \in \mathcal{F}$.

Una σ-álgebra $\mathcal{F}$ es una clase de subconjuntos de Ω que verifica

- $\Omega \in \mathcal{F}$.
- Si $A \in \mathcal{F}$, entonces $A^c \in \mathcal{F}$.
- Si $A_n \in \mathcal{F}$, $n \geq 1$, entonces $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$.

σ-álgebras

Ejemplos

- El conjunto de las partes de Ω, $\mathcal{F} = \mathcal{P}(\Omega)$.
- $\mathcal{F} = \{\emptyset, \Omega\}$.
- $\mathcal{F} = \{\emptyset, \Omega, A, A^c\}$, donde $A \subset \Omega$.
- $\mathcal{F} = \{A \subset \mathbb{R} :\ A \text{ numerable}\} \bigcup \{A \subset \mathbb{R} :\ A^c \text{ numerable}\}$.

Observación: La intersección de σ-álgebras es una σ-álgebra. En general la unión de σ-álgebras **no es** una σ-álgebra.

σ-álgebra generada por una clase de conjuntos
Sea $\mathcal{C} \subset \mathcal{P}(\Omega)$. Se define la σ-álgebra generada por $\mathcal{C}$ como la intersección de todas las σ-álgebras que contienen a $\mathcal{C}$ (la "mínima" σ-álgebra que contiene a $\mathcal{C}$).

σ-álgebra de Borel
Sea (Ω, τ) un espacio topológico. La σ-álgebra de Borel $\mathcal{B} := \mathcal{B}(\Omega)$ es la generada por los conjuntos abiertos de Ω.

σ-álgebra de Borel en $\mathbb{R}$

$$\begin{aligned}\mathcal{B}(\mathbb{R}) =& \sigma((a,b], a, b \in \mathbb{R}) = \sigma((a,b), a, b \in \mathbb{R}) = \sigma([a,b), a, b \in \mathbb{R}) \\ =& \sigma((-\infty, b], b \in \mathbb{R}) = \sigma((-\infty, b), b \in \mathbb{R}) \\ =& \sigma((a, \infty], a \in \mathbb{R}) = \sigma([a, \infty), a \in \mathbb{R}).\end{aligned}$$

Espacio de probabilidad

Un espacio de probabilidad es una terna $(\Omega, \mathcal{F}, \mathrm{P})$, donde Ω es un conjunto no vacío, $\mathcal{F}$ es una σ-álgebra de subconjuntos de Ω y P es una medida de probabilidad sobre $\mathcal{F}$, es decir, $\mathrm{P} : \mathcal{F} \to [0, 1]$ tal que

- $\mathrm{P}(\Omega) = 1$.
- Si $A_1, A_2, \ldots$ son sucesos en $\mathcal{F}$ tales que $A_i \cap A_j = \emptyset$ si $i \neq j$, entonces

$$\mathrm{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathrm{P}(A_n).$$

Propiedades elementales

- Paso al complementario: $\mathrm{P}(A^c) = 1 - \mathrm{P}(A)$.
- Probabilidad del vacío: $\mathrm{P}(\emptyset) = 0$.
- Monotonía: Si $A \subset B$, entonces $\mathrm{P}(A) \leq \mathrm{P}(B)$.
- Subaditividad:
$$\mathrm{P}\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mathrm{P}(A_n).$$
- Fórmula de inclusión-exclusión:
$$\mathrm{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathrm{P}(A_i) - \sum_{i<j} \mathrm{P}(A_i \cap A_j)$$
$$+ \sum_{i<j<k} \mathrm{P}(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n+1} \mathrm{P}(A_1 \cap \cdots \cap A_n).$$

Límites de sucesiones de conjuntos

Límite inferior de una sucesión de conjuntos:

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

Límite superior de una sucesión de conjuntos:

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

- Comprueba $\liminf A_n \subset \limsup A_n$.
- Expresa de forma alternativa $(\liminf A_n)^c$ y $(\limsup A_n)^c$.

Límites de sucesiones de conjuntos

Límite de una sucesión de conjuntos
La sucesión A_n tiene límite A ($A_n \to A$) si $\liminf A_n = \limsup A_n = A$.

Límites de sucesiones monótonas

- Si $A_1 \subset A_2 \subset \cdots$, entonces $A_n \uparrow A = \cup_{n=1}^{\infty} A_n$.
- Si $A_1 \supset A_2 \supset \cdots$, entonces $A_n \downarrow A = \cap_{n=1}^{\infty} A_n$.

Continuidad para sucesiones monótonas
Si $A_n \uparrow A$, entonces $\mathrm{P}(A_n) \uparrow \mathrm{P}(A)$. Si $A_n \downarrow A$, entonces $\mathrm{P}(A_n) \downarrow \mathrm{P}(A)$.

Relación entre la probabilidad del límite inferior y superior

$$\mathrm{P}(\liminf A_n) \le \liminf \mathrm{P}(A_n) \le \limsup \mathrm{P}(A_n) \le \mathrm{P}(\limsup A_n).$$

Continuidad
Si $A_n \to A$, entonces $\mathrm{P}(A_n) \to \mathrm{P}(A)$.

π-sistemas y λ-sistemas

Un π-**sistema** $\mathcal{C}$ es una clase de subconjuntos de Ω tal que si $A, B \in \mathcal{C}$, entonces $A \cap B \in \mathcal{C}$.

Un λ-**sistema** es una clase $\mathcal{L}$ de subconjuntos de Ω tal que:
(a) $\Omega \in \mathcal{L}$, (b) si $A \in \mathcal{L}$, entonces $A^c \in \mathcal{L}$, (c) si $A_n \in \mathcal{L}$, $n \geq 1$, y $A_i \cap A_j = \emptyset$ para $i \neq j$, entonces $\cup_{n=1}^{\infty} A_n \in \mathcal{L}$.

Observaciones

- Un álgebra es un π-sistema.
- Una σ-álgebra es un λ-sistema.
- $\mathcal{C}$ es una σ-álgebra si y solo si $\mathcal{C}$ es π-sistema y λ-sistema.

Teorema de Dynkin

Teorema de Dynkin
Sea $\mathcal{C}$ un π-sistema y $\mathcal{L}$ un λ-sistema. Si $\mathcal{C} \subset \mathcal{L}$ entonces $\sigma(\mathcal{C}) \subset \mathcal{L}$.

Aplicación habitual de este teorema
Supongamos que $\mathcal{L}$ contiene los conjuntos que satisfacen cierta propiedad. Para demostrar que la propiedad se cumple para todos los conjuntos de una σ-álgebra ($\sigma(\mathcal{C}) \subset \mathcal{L}$) basta probarla para los conjuntos de un π-sistema que la genera ($\mathcal{C} \subset \mathcal{L}$).

Teorema de Dynkin: esquema de la demostración

- Basta demostrar que el mínimo λ-sistema que contiene a $\mathcal{C}$, $\lambda(\mathcal{C})$, es un π-sistema.
- Dado $A \subset \Omega$, definimos $\mathcal{G}_A := \{B \subset \Omega : A \cap B \in \lambda(\mathcal{C})\}$.
- Si $A \in \lambda(\mathcal{C})$, entonces $\mathcal{G}_A$ es λ-sistema.
- Si $A \in \mathcal{C}$, entonces $\lambda(\mathcal{C}) \subset \mathcal{G}_A$. Para esto basta ver que $\mathcal{C} \subset \mathcal{G}_A$. (Por tanto, si $A \in \mathcal{C}$ y $B \in \lambda(\mathcal{C})$, entonces $A \cap B \in \lambda(\mathcal{C})$).
- Cambiando los papeles de A y B tenemos que si $A \in \lambda(\mathcal{C})$, entonces $\mathcal{C} \subset \mathcal{G}_A$. Como consecuencia, $\lambda(\mathcal{C}) \subset \mathcal{G}_A$ y hemos terminado.

Función de distribución

Sea P una medida de probabilidad sobre $(\mathbb{R}, \mathcal{B})$, donde $\mathcal{B}$ es la σ-álgebra de Borel. La **función de distribución** $F : \mathbb{R} \to [0, 1]$ correspondiente a P se define como

$$F(x) = \mathrm{P}((-\infty, x]), \quad x \in \mathbb{R}.$$

Observaciones

(a) Una función de distribución F tiene las tres propiedades siguientes:
- Es continua por la derecha.
- Es monótona no decreciente.
- $F(\infty) := \lim_{x\to\infty} F(x) = 1$; $F(-\infty) := \lim_{x\to-\infty} F(x) = 0$.

(b) Si F es una función de distribución, $\mathrm{Cont}(F)^c$ es numerable, donde $\mathrm{Cont}(F)$ es el conjunto de puntos en los que F es continua.

(c) Si F_1 y F_2 son dos funciones de distribución tales que $F_1(x) = F_2(x)$, para todo $x \in \mathrm{Cont}(F_1) \cap \mathrm{Cont}(F_2)$, entonces $F_1(x) = F_2(x)$, para todo $x \in \mathbb{R}$.

Extensión de medidas

Teorema
Sea $F : \mathbb{R} \to [0,1]$ una función continua por la derecha, monótona no decreciente, con $F(\infty) = 1$, $F(-\infty) = 0$. Definamos $\mathrm{P}((a,b]) = F(b) - F(a)$, para $a < b$. Existe una única medida de probabilidad que extiende P a $\mathcal{B}(\mathbb{R})$.

Teorema de extensión de Carathèodory
Sea μ una medida σ-finita sobre un álgebra $\mathcal{F}_0 \subset \mathcal{P}(\Omega)$. Entonces existe una extensión única de μ a la σ-álgebra $\mathcal{F}$ generada por $\mathcal{F}_0$.

La unicidad se puede deducir directamente de:
Teorema
Sean P_1 y P_2 medidas de probabilidad sobre $\sigma(\mathcal{C})$, donde $\mathcal{C}$ es un π-sistema. Si $\mathrm{P}_1(A) = \mathrm{P}_2(A)$ para todo $A \in \mathcal{C}$, entonces $\mathrm{P}_1(A) = \mathrm{P}_2(A)$ para todo $A \in \sigma(\mathcal{C})$.

Probabilidad condicionada

Sea $(\Omega, \mathcal{F}, \mathrm{P})$ un espacio de probabilidad y A, $B \in \mathcal{F}$ con $\mathrm{P}(B) > 0$.

Se llama probabilidad de A condicionada a B a

$$\mathrm{P}(A|B) := \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}.$$

La aplicación:

$$\begin{aligned} \mathrm{P}(\cdot|B) : \mathcal{F} &\longrightarrow [0,1] \\ A &\longmapsto \mathrm{P}(A|B) \end{aligned}$$

es una medida de probabilidad sobre $(\Omega, \mathcal{F})$.

Probabilidad condicionada

Fórmula del producto
Si $A_1, \dots, A_n \in \mathcal{F}$ con $\mathrm{P}(A_1 \cap \cdots \cap A_n) > 0$, se tiene

$$\mathrm{P}(A_1 \cap \cdots \cap A_n) = \mathrm{P}(A_1)\mathrm{P}(A_2|A_1)\mathrm{P}(A_3|A_1 \cap A_2) \cdots \mathrm{P}(A_n|A_1 \cap \cdots \cap A_{n-1}).$$

Una colección de sucesos disjuntos $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$ es una partición de Ω si:

(a) $\mathrm{P}(A_i) > 0$, $i \geq 1$.

(b) $\Omega = \cup_{i=1}^{\infty} A_i$.

Fórmula de la probabilidad total
Sea $(\Omega, \mathcal{F}, \mathrm{P})$ un espacio de probabilidad y $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$ una partición.
Entonces, para cualquier $B \in \mathcal{F}$,

$$\mathrm{P}(B) = \sum_{i=1}^{\infty} \mathrm{P}(A_i)\mathrm{P}(B|A_i).$$

Probabilidad condicionada

Fórmula de Bayes
Sea $(\Omega, \mathcal{F}, \mathrm{P})$ un espacio de probabilidad, $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$ una partición y $B \in \mathcal{F}$ con $\mathrm{P}(B) > 0$, entonces

$$\mathrm{P}(A_j|B) = \frac{\mathrm{P}(A_j)\mathrm{P}(B|A_j)}{\sum_{i=1}^{\infty} \mathrm{P}(A_i)\mathrm{P}(B|A_i)}.$$

$\mathrm{P}(A_n)$ se llaman **probabilidades a priori**
$\mathrm{P}(A_n|B)$ se llaman **probabilidades a posteriori**

Variables aleatorias

Una **variable aleatoria** sobre un espacio de probabilidad $(\Omega, \mathcal{F}, \mathrm{P})$ es una función $X : \Omega \to \mathbb{R}$ medible respecto a $\mathcal{F}$ y la σ-álgebra de Borel $\mathcal{B}(\mathbb{R})$.

$$X^{-1}(B) \in \mathcal{F}, \quad \text{para todo} \quad B \in \mathcal{B}(\mathbb{R}).$$

- Intuitivamente, $X(\omega)$ representa una cantidad numérica relacionada con el resultado $\omega \in \Omega$ de un experimento aleatorio.
- Queremos calcular probabilidades de sucesos de la forma

$$\{a < X \leq b\} := \{\omega : \, a < X(\omega) \leq b\} = X^{-1}(a, b].$$

- Es necesario $X^{-1}(a, b] \in \mathcal{F}$, para todo $a, b \in \mathbb{R}$, lo que implica $X^{-1}(B) \in \mathcal{F}$ para todo $B \in \mathcal{B}(\mathbb{R})$.
- A veces hace falta considerar v.a. extendidas $X : \Omega \to \overline{\mathbb{R}}$, donde $\overline{\mathbb{R}} = [-\infty, \infty]$. Aquí, $\mathcal{B}(\overline{\mathbb{R}})$ es generada por intervalos $(a, b]$, con $-\infty \leq a < b \leq \infty$ y $[-\infty, b]$, con $-\infty \leq b \leq \infty$.

Distribución inducida por una v.a.

Una **variable aleatoria** X induce una medida de probabilidad P_X sobre $(\mathbb{R}, \mathcal{B})$:

$$\mathrm{P}_X(B) := \mathrm{P}(X \in B) = \mathrm{P}\{\omega : X(\omega) \in B\}, \quad B \in \mathcal{B}.$$

- Es fácil comprobar que P_X es en efecto una medida de probabilidad sobre $(\mathbb{R}, \mathcal{B})$.
- La función de distribución de una v.a. X es la función de distribución de la medida de probabilidad P_X:

$$F_X(x) = \mathrm{P}_X((-\infty, x]) = \mathrm{P}(X \leq x), \quad x \in \mathbb{R}.$$

- Dada una función F monótona no decreciente y continua por la derecha con $F(\infty) = 1$ y $F(-\infty) = 0$, siempre existe una v.a. X tal que $F = F_X$.

Observaciones

- Si $X_1, \ldots, X_n$ son v.a. y $g : \mathbb{R}^n \to \mathbb{R}$ es medible (Borel), entonces $g(X_1, \ldots, X_n)$ es una v.a. Por ejemplo: $X_1 + \ldots + X_n$ es una v.a.

- Si $X_1, X_2, \cdots$ son v.a. entonces $\sup_n X_n$ e $\inf_n X_n$ son v.a.

- Si $X_1, X_2, \cdots$ son v.a. entonces $\limsup_n X_n$ y $\liminf_n X_n$ son v.a.

- Si $X_1, X_2, \cdots$ son v.a. tales que $X_n(\omega)$ converge para todo $\omega \in \Omega$, entonces $\lim_n X_n$ es una v.a.

Algunos tipos de variables aleatorias

(1) Una v.a. X es **degenerada** en $c \in \mathbb{R}$ si $\mathrm{P}(X = c) = 1$. ¿Cuál es su función de distribución?

(2) Una v.a. X es **discreta** si el conjunto de valores que toma X es finito o numerable.

(3) Una v.a. X es **absolutamente continua** si existe una función medible (Borel) y no negativa $f : \mathbb{R} \to \mathbb{R}$ tal que $F_X(x) = \int_{-\infty}^{x} f(t)dt$. Se dice que f es la **función de densidad** de X. Se verifica:

- $\int_{-\infty}^{\infty} f(t)dt = 1$.
- $P(X \in B) = \int_B f(t)dt$, para todo $B \in \mathcal{B}$.

(4) Una v.a. X es **singular** si existe $B \in \mathcal{B}$ tal que $\lambda(B) = 0$ y $\mathrm{P}_X(B) = 1$, donde λ es la medida de Lebesgue.

Vectores aleatorios

Un **vector aleatorio** sobre un espacio de probabilidad $(\Omega, \mathcal{F}, \mathrm{P})$ es una función $X : \Omega \to \mathbb{R}^n$ medible respecto a $\mathcal{F}$ y la σ-álgebra de Borel $\mathcal{B}(\mathbb{R}^n)$.

Observación
$X = (X_1, \ldots, X_n)$ es un vector aleatorio si y solo si X_i es una variable aleatoria para todo $i = 1, \ldots, n$.

Distribución y función de distribución de un vector aleatorio
Distribución:

$$\mathrm{P}_X(B) := \mathrm{P}(X \in B) = \mathrm{P}\{\omega : X(\omega) \in B\}, \quad B \in \mathcal{B}(\mathbb{R}^n).$$

Función de distribución: $F : \mathbb{R}^n \to \mathbb{R}$ tal que

$$F_X(x) = F(x_1, \ldots, x_n) := \mathrm{P}(X_i \leq x_i, \quad i = 1, \ldots, n).$$

Las funciones de distribución de las v.a. X_i, $i = 1, \ldots, n$ se llaman **funciones de distribución marginales**.

Vectores aleatorios absolutamente continuos

Una vector aleatorio X es **absolutamente continuo** si existe una función medible (Borel) y no negativa $f : \mathbb{R}^n \to \mathbb{R}$ tal que

$$F_X(x_1, \ldots, x_n) = \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} f(t_1, \ldots, t_n) dt_1 \cdots dt_n.$$

Se dice que f es la **función de densidad** de X. Se verifica:

$$P(X \in B) = \int_B f(t_1, \ldots, t_n) dt_1 \cdots dt_n, \quad \text{para todo} \quad B \in \mathcal{B}(\mathbb{R}^n).$$

σ-álgebra generada por una variable aleatoria

Sea $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B})$ una v.a. La clase de conjuntos

$$\sigma(X) := \{X^{-1}(B) : \ B \in \mathcal{B}\}$$

es una sub σ-álgebra de $\mathcal{F}$ denominada **σ-álgebra generada por** X.

$\sigma(X)$ es la mínima σ-álgebra que hace medible a X.

Determina $\sigma(X)$ para las siguientes v.a.:

- $X(\omega) = a$ para todo $\omega \in \Omega$.
- $X = I_A$, donde $A \in \mathcal{F}$.
- X es una v.a. simple, es decir, X toma un número finito de valores $\{a_1, \ldots, a_k\} \subset \mathbb{R}$.

σ-álgebra generada por una variable aleatoria

Intuitivamente, $\sigma(X)$ recoge la información que se obtiene al observar la v.a. X.

Dada una familia de v.a. $\{X_i : i \in I\}$, se define $\sigma(X_i, i \in I)$ como la mínima σ-álgebra que hace medibles a todas las X_i simultáneamente, es decir,

$$\sigma(X_i, i \in I) := \sigma\left(\bigcup_{i \in I} \sigma(X_i)\right).$$

Un proceso en un intervalo de tiempo $[0, T]$, determina una familia de v.a. $\{X_t : t \in [0, T]\}$. Entonces, $\sigma(X_t, t \in [0, T])$ representa la información obtenida al observar el proceso.

σ-álgebra generada por una variable aleatoria

Proposición

Si X es una v.a. y $\mathcal{C}$ es una clase de subconjuntos de $\mathbb{R}$

$$X^{-1}[\sigma(\mathcal{C})] = \sigma[X^{-1}(\mathcal{C})].$$

Demostración:

- $X^{-1}[\sigma(\mathcal{C})] \supset \sigma[X^{-1}(\mathcal{C})]$ ya que $X^{-1}[\sigma(\mathcal{C})]$ es σ-álgebra que contiene a $X^{-1}(\mathcal{C})$.

- $\mathcal{A} = \{B \subset \mathbb{R} : X^{-1}(B) \in \sigma[X^{-1}(\mathcal{C})]\}$ es una σ-álgebra tal que $\mathcal{C} \subset \mathcal{A}$. Por tanto $\sigma(\mathcal{C}) \subset \mathcal{A}$.

- Esta inclusión implica $X^{-1}[\sigma(\mathcal{C})] \subset \sigma[X^{-1}(\mathcal{C})]$ por definición de $\mathcal{A}$.

Corolario

$\sigma(X) = \sigma(\{\{X \leq x\} : x \in \mathbb{R})\} = \sigma(\{X^{-1}((-\infty, x]) : x \in \mathbb{R}\}).$