PROBABILIDAD II

Grado en Matemáticas

Tema 1
Repaso de teoría de la medida

Jesús Munárriz

Departamento de Matemáticas
Universidad Autónoma de Madrid
jesus.munarriz@uam.es

1. σ-álgebras
2. Espacios de medida
3. Teorema de extensión
4. Medida de Lebesgue y Borel-Stieljes
5. Funciones medibles
6. Integrales
7. Paso al límite bajo signo integral
8. Continuidad absoluta de medidas
9. El Teorema de Radon-Nikodym
10. Espacios L^p. Desigualdades importantes.

Sea Ω un conjunto no vacío. Una colección de conjuntos $\mathcal{F} \subset \mathcal{P}(\Omega)$ (partes de Ω) se dice que es una σ-**álgebra** si

(1) $\Omega \in \mathcal{F}$.

(2) $\mathcal{F}$ es cerrada o estable para la complementación:

$$\text{Si } A \in \mathcal{F}, \text{ entonces } A^c \in \mathcal{F}.$$

(3) $\mathcal{F}$ es estable para la unión numerable:

$$\text{Si } \{A_i\}_{i=1}^{\infty} \subset \mathcal{F}, \text{ entonces } \cup_{i=1}^{\infty} A_i \in \mathcal{F}.$$

El par $(\Omega, \mathcal{F})$ se denomina **espacio medible** y los elementos de $\mathcal{F}$ **conjuntos medibles**.

Si $\{\mathcal{F}_i\}_{i\in I} \subset \mathcal{P}(\Omega)$ es una colección de σ-álgebras, entonces $\bigcap_{i\in I} \mathcal{F}_i$ es σ-álgebra.

Dado $\mathcal{C} \subset \mathcal{P}(\Omega)$, se define

$$\sigma(\mathcal{C}) = \bigcap_{\mathcal{F}} \{\mathcal{F} : \mathcal{F} \supset \mathcal{C} \text{ y } \mathcal{F}\ \sigma\text{-álgebra}\}.$$

El conjunto $\mathcal{C}$ se denomina **generador de la σ-álgebra** $\sigma(\mathcal{C})$.

Si (Ω, τ) es un espacio topológico, a la σ-álgebra $\sigma(\tau)$ se denomina **σ-álgebra Boreliana o de Borel asociada a** τ.

De interés especial para nosotros serán:

- $\Omega = \mathbb{R}$ o $\overline{\mathbb{R}}$, $\tau = \tau_u$ (topología usual).
- $\Omega = \mathbb{R}^k$ o $\overline{\mathbb{R}}^k$, $\tau = \tau_u$ (topología usual).

Sea $(\Omega, \mathcal{F})$ un espacio medible. Se dice que μ es una **medida (positiva)** en Ω si $\mu : \mathcal{F} \longrightarrow [0, \infty]$ satisface:

(1) $\mu(\emptyset) = 0$.

(2) σ-**aditividad** o **aditividad numerable**: si $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$ es una colección de conjuntos disjuntos dos a dos (i.e., $A_i \cap A_j = \emptyset$, $i \neq j$), entonces:

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

El triplete $(\Omega, \mathcal{F}, \mu)$ se llama **espacio de medida**.

$(\Omega, \mathcal{F}, \mu)$ es un **espacio de medida finita** si $\mu(\Omega) < \infty$.

$(\Omega, \mathcal{F}, \mu)$ es un **espacio de medida** σ-**finita** si existe $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$ tal que $\Omega = \bigcup_{i=1}^{\infty} A_i$ y $\mu(A_i) < \infty$, para todo i.

Si $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$, decimos que A_n **crece** hasta A, $A_n \uparrow A$, si $A_1 \subset A_2 \subset A_3 \subset \cdots$ y $A = \bigcup_{i=1}^{\infty} A_i$.

Si $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$, decimos que A_n **decrece** hasta A, $A_n \downarrow A$, si $A_1 \supset A_2 \supset A_3 \supset \cdots$ y $A = \bigcap_{i=1}^{\infty} A_i$.

1. $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$ y $A_n \uparrow A$, entonces $\mu(A_n) \uparrow \mu(A)$.
2. $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$ con $A_n \downarrow A$ y $\mu(A_1) < \infty$, entonces $\mu(A_n) \downarrow \mu(A)$.
3. $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$, $\mu(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu(A_i)$.

Una colección $\mathcal{A} \subset \mathcal{P}(\Omega)$ se dice que es una **álgebra** si

(1) $\Omega \in \mathcal{A}$.

(2) Si $A \in \mathcal{A}$, entonces $A^c \in \mathcal{A}$.

(3) Si $A, B \in \mathcal{A}$, entonces $A \cup B \in \mathcal{A}$.

Teorema de extensión de Caratheodory (1948)

Sea $\mu_0 : \mathcal{A} \longrightarrow [0,\infty]$ una medida sobre el álgebra $\mathcal{A}$. Existe una medida $\mu : \sigma(\mathcal{A}) \longrightarrow [0,\infty]$ que es una extensión de μ_0, es decir, $\mu_{|\mathcal{A}} = \mu_0$.

Los conjuntos $B \in \mathcal{F}$ con $\mu(B) = 0$ se llaman **conjuntos μ-nulos**
Un espacio de medida se dice **completo** si para todo conjunto μ-nulo B y para todo $A \subset B$, se tiene que $A \in \mathcal{F}$.

Teorema de completación

Sea $(\Omega, \mathcal{F}, \mu)$ un espacio de medida. Existe un espacio medida $(\Omega, \overline{\mathcal{F}}, \overline{\mu})$ completo tal que $\mathcal{F} \subset \overline{\mathcal{F}}$ y $\overline{\mu}$ es una extensión de μ.

Además, este espacio de medida $(\Omega, \overline{\mathcal{F}}, \overline{\mu})$ es único y se dice que es la **completación** de $(\Omega, \mathcal{F}, \mu)$.

Medida de Borel-Stieljes asociada a g

Sea $I \subset \mathbb{R}$ (intervalo) y $g : I \longrightarrow \mathbb{R}$ funcion no decreciente y continua por la derecha.

Existe una única medida $m_g : \mathcal{B}(I) \longrightarrow [0, \infty]$ tal que $m_g((a, b]) = g(b) - g(a)$. A m_g se le llama la **medida de Borel-Stieljes asociada a** g.

Nota: Se tiene $m_g(\{x_0\}) = g(x_0) - g(x_0^-)$.

Nota: Si g, h son dos funciones continuas por la derecha y no decrecientes tales que $m_g = m_h$, entonces $g - h =$constante.

Función de distribución de una medida

Dada $\mu : \mathcal{B}(I) \longrightarrow [0, \infty]$ medida finita, existe g no decreciente y continua por la derecha tal que $\mu = \mu_g$. Además, podemos tomar $g(x) = \mu(I \cap (-\infty, x])$ que se llama **función de distribución de** μ.

Sean $(\Omega, \mathcal{F})$, $(\Omega', \mathcal{F}')$ dos espacios medibles. Se dice que una función $f : (\Omega, \mathcal{F}) \longrightarrow (\Omega', \mathcal{F}')$ es $\mathcal{F} - \mathcal{F}'$-**medible** si para todo $A \in \mathcal{F}'$, se tiene que $f^{-1}(A) \in \mathcal{F}$.

$f^{-1}(A) = \{\omega \in \Omega : f(\omega) \in A\}$ (**pre-imagen** de A por f).

Si $f : \Omega \longrightarrow \Omega'$ es constante, entonces es medible.

Sea $\mathbb{K} = \mathbb{R}$ ó $\mathbb{C}$, con $\mathcal{F}'$ los respectivos conjuntos de Borel.

Si $A \subset \Omega$, el **indicador** (o función indicatriz) de A es $\mathbf{1}_A : \Omega \longrightarrow \{0, 1\}$, definida como $\mathbf{1}_A(\omega) = 1$ si $\omega \in A$ y $\mathbf{1}_A(\omega) = 0$ si $\omega \notin A$.

- ¿Cuándo es $\mathbf{1}_A$ una función medible?

Una función $f : (\Omega, \mathcal{F}) \longrightarrow \mathbb{R}$ se dice **simple** si es medible y $f(\Omega)$ es finito.

Nota: $f : (\Omega, \mathcal{F}) \longrightarrow \mathbb{R}$ es simple si y solo si $f = \sum_{i=1}^{k} a_i \mathbf{1}_{A_i}$, con $A_i \in \mathcal{F}$ disjuntos.

Sea $(\Omega, \mathcal{F}, \mu)$ un espacio de medida. Si $s = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i}$ es función simple positiva (es decir, no no negativa), se define la **integral de** s **con respecto a la medida** μ como

$$\int_{\Omega} s \, d\mu := \sum_{i=1}^{n} a_i \mu(A_i).$$

Si $f : (\Omega, \mathcal{F}) \longrightarrow [0, \infty]$ es medible, entonces existe una sucesión $\{s_n\}_{n=1}^{\infty}$ simples con $0 \leq s_n \uparrow f$. Se define la **integral de f con respecto a la medida μ** como

$$\int_\Omega f\, d\mu = \lim_{n\to\infty} \int_\Omega s_n\, d\mu = \sup_{0\leq s\leq f,\ s \text{ simple}} \int_\Omega s\, d\mu.$$

Si $f : (\Omega, \mathcal{F}) \longrightarrow [-\infty, \infty]$ es medible, entonces podemos escribir $f = f^+ - f^-$. Si $\int_\Omega f^+ \, d\mu < \infty$ ó $\int_\Omega f^- \, d\mu < \infty$, definimos la **integral de** f como

$$\int_\Omega f \, d\mu := \int_\Omega f^+ \, d\mu - \int_\Omega f^- \, d\mu.$$

Si $f : (\Omega, \mathcal{F}) \longrightarrow \mathbb{C}$ es medible, entonces podemos escribir $f = f_1 + if_2$, donde $f_1, f_2 : (\Omega, \mathcal{F}) \longrightarrow \mathbb{R}$. Si $\int_\Omega f_1 \, d\mu < \infty$ y $\int_\Omega f_2 \, d\mu < \infty$, definimos

$$\int_\Omega f \, d\mu := \int_\Omega f_1 \, d\mu + i \int_\Omega f_2 \, d\mu.$$

Si $\int_\Omega |f| \, d\mu < \infty$, f se dice μ-**integrable** ($f \in \mathcal{L}^1(\mu)$).

Para $f, g \in \mathcal{L}^1(\mu)$, $a, b \in \mathbb{K}$, se tiene:

1. **Linealidad:** $af + bg \in \mathcal{L}^1(\mu)$ y
$$\int_\Omega (af + bg)\, d\mu = a \int_\Omega f\, d\mu + b \int_\Omega g\, d\mu.$$
2. **Positividad:** Si $f \geq 0$ entonces $\int_\Omega f\, d\mu \geq 0$; o equivalentemente, si $f \leq g$ (luego f y g toman valores reales) entonces
$$\int_\Omega f\, d\mu \leq \int_\Omega g\, d\mu \quad \left(\text{en particular,} \quad \left|\int_\Omega f\, d\mu\right| \leq \int_\Omega |f|\, d\mu.\right).$$
3. $\int_\Omega |f|\, d\mu = 0$ si y solo si $f = 0$ en casi todo punto.
4. Si $f \geq 0$ y $\int_\Omega f\, d\mu < \infty$, entonces $\mu(\{\omega \in \Omega : f(\omega) = \infty\}) = 0$, es decir, $f \neq \infty$ c.s.

Dada f medible, para $A \in \mathcal{F}$ se define

$$\int_A f \, d\mu := \int_\Omega f \, \mathbf{1}_A \, d\mu.$$

Proposición: Sea $f : (\Omega, \mathcal{F}) \longrightarrow [0, \infty]$ medible. La aplicación

$$\begin{aligned} \nu : \mathcal{F} &\longrightarrow [0, \infty] \\ A &\longmapsto \nu(A) = \int_A f \, d\mu, \end{aligned}$$

es una medida sobre $(\Omega, \mathcal{F})$. Se denota $d\nu = f \, d\mu$.

Teorema de la convergencia monótona

Hacia arriba: Si $f_n \uparrow f$ y existe k tal que $f_k^- \in \mathcal{L}^1(\mu)$, entonces

$$\int_\Omega f\, d\mu = \lim_{n\to\infty} \int_\Omega f_n\, d\mu.$$

Hacia abajo: Si $f_n \downarrow f$ y existe k tal que $f_k^+ \in \mathcal{L}^1(\mu)$, entonces

$$\int_\Omega f\, d\mu = \lim_{n\to\infty} \int_\Omega f_n\, d\mu.$$

Lema de Fatou-Lebesgue

(a) Si $f_n \leq f$ y $f^+ \in \mathcal{L}^1(\mu)$, entonces

$$\int_\Omega \limsup_{n\to\infty} f_n \, d\mu \geq \limsup_{n\to\infty} \int_\Omega f_n \, d\mu.$$

(b) Si $f_n \geq f$ y $f^- \in \mathcal{L}^1(\mu)$, entonces

$$\int_\Omega \liminf_{n\to\infty} f_n \, d\mu \leq \liminf_{n\to\infty} \int_\Omega f_n \, d\mu.$$

Teorema de la convergencia dominada

Si $f_n \to f$ en casi todo punto, y existe $g \in \mathcal{L}^1(\mu)$ tal que $|f_n| \leq g$ para todo n, entonces
$$\int_\Omega f\, d\mu = \lim_{n\to\infty} \int_\Omega f_n\, d\mu.$$

Sea $(\Omega, \mathcal{F})$ un espacio medible y ν, μ dos medidas. Se dice que ν es **absolutamente continua con respecto a** μ y se denota $\nu << \mu$ si para todo $A \in \mathcal{F}$ tal que $\mu(A) = 0$, entonces $\nu(A) = 0$.

Una medida μ se dice que se **concentra** en un conjunto $A \in \mathcal{F}$ si $\mu(A^c) = 0$, o equivalentemente, si para todo $B \in \mathcal{F}$, se tiene que $\mu(B) = \mu(A \cap B)$.

Se dice que ν es **singular con respecto a** μ y se denota $\nu \perp \mu$ si existe $A \in \mathcal{F}$ tal que $\mu(A) = 0$ y ν se concentra en A.

Teorema de descomposición de Lebesgue

Sean ν, μ dos medidas σ-finitas sobre $(\Omega, \mathcal{F})$. Existen dos únicas medidas ν_a, ν_s tales que $\nu_a << \mu$ y $\nu_s \perp \mu$ con $\nu = \nu_a + \nu_s$. Esta descomposición se denomina **descomposición de Lebesgue** de la medida ν (con respecto a μ).

Sean μ y ν dos medidas (positivas) finitas. Decimos que τ es una medida con valores reales o con signo si es de la forma $\tau = \mu - \nu$. Decimos que γ es una medida con valores complejos si es de la forma $\gamma = \tau_1 + i\tau_2$, donde τ_1 y τ_2 son medidas reales.

En el caso de las medidas con signo, se puede admitir que una de las medidas sea infinita.

Teorema de Radon-Nikodym

Sean ν, μ dos medidas σ-finitas con signo, sobre $(\Omega, \mathcal{F})$, tales que $\nu << \mu$. Existe una única función medible $f : (\Omega, \mathcal{F}) \longrightarrow [0, \infty]$ tal que

$$\nu(A) = \int_A f \, d\mu, \quad A \in \mathcal{F}.$$

Es decir, $d\nu = f d\mu$.

A la función $f = \frac{d\nu}{d\mu}$ se le llama la **derivada de Radon-Nikodym de ν con respecto a μ** (f es la μ-**densidad de ν**.)

Espacios $\mathcal{L}_p$, $1 \leq p \leq \infty$: Sea $(\Omega, \mathcal{F}, \mu)$ un espacio de medida. Definimos

$$\mathcal{L}_p = \left\{ f^{\text{ medible}} : \Omega \longrightarrow \mathbb{K} : \int_\Omega |f|^p d\mu < \infty \right\}$$

si $1 \leq p < \infty$, y

$$\mathcal{L}_\infty = \{ f^{\text{ medible}} : \Omega \longrightarrow \mathbb{K} : \exists a \geq 0 \text{ tal que } \mu(|f|^{-1}(a, \infty)) = 0 \}.$$

Puede demostrarse que los espacios $\mathcal{L}_p$ son espacios vectoriales.

Definimos, para $1 \leq p < \infty$,

$$\|f\|_p = \left(\int_\Omega |f|^p d\mu\right)^{1/p},$$

y

$$\|f\|_\infty = \text{ess sup}\,|f| := \text{mín}\{a \geq 0 \text{ tal que } \mu(|f|^{-1}(a,\infty)) = 0\}.$$

Como

1. $\|f+g\|_p \leq \|f\|_p + \|g\|_p$ (desigualdad triangular o de Minkowski).
2. $\|\lambda f\|_p = |\lambda| \|f\|_p$.

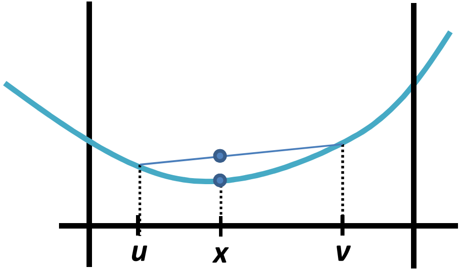
tenemos que $\|\cdot\|_p$ es una **seminorma**. No es una norma ya que $\|f\|_p = 0$ cuando $f = 0$ c. p. t.

Definiendo la relación de equivalencia "$f \sim g$ si y solo si $f = g$ en casi todo punto", tenemos que $\mathcal{L}_p / \sim = L_p$ es un espacio vectorial normado y completo (toda sucesión de Cauchy es convergente). En otras palabras, los espacios L^p para $1 \leq p \leq \infty$ son espacios de Banach.

Funciones convexas

Definición: Sea I un intervalo. Una función $f : I \longrightarrow \mathbb{R}$ se dice **convexa** si para todo $u, v \in I$, para todo $\lambda \in [0, 1]$, se tiene que

$$f(\lambda u + (1 - \lambda)v) \leq \lambda f(u) + (1 - \lambda)f(v).$$



Ejemplos: $f(x) = e^x$; $f(x) = |x|^\alpha$, $\alpha \geq 1$; $f(x) = 1/x$ $(x > 0)$.

Propiedades de las funciones convexas

1. Sea $f : I \longrightarrow \mathbb{R}$ derivable. f es convexa si y sólo si f' es no decreciente.
2. Sea $f : I \longrightarrow \mathbb{R}$ dos veces derivable. f es convexa si y sólo si $f'' \geq 0$.
3. Si f convexa, entonces f es continua salvo quizá en ∂I (extremos del intervalo I).
4. Si f es convexa, para todo $(u, f(u))$ tal que $u \in I^{\circ}$, existe (al menos) una recta que pasa por $(u, f(u))$ (recta soporte) tal que su gráfica está totalmente por debajo de f. Es decir,

$$\forall u \in I^{\circ}, \exists a \in \mathbb{R} \quad : \quad a(x-u) + f(u) \leq f(x), \quad \forall x \in I.$$

Teorema: Desigualdad de Jensen

Sea I un intervalo, y sea $(\Omega, \mathcal{F}, \mu)$ un espacio de medida con $\mu(\Omega) = 1$. Si $f : \Omega \longrightarrow I$ pertenece a $L^1(\Omega, \mathcal{F}, \mu)$ y $\varphi : I \longrightarrow \mathbb{R}$ es convexa, entonces

$$\varphi\left(\int_\Omega f d\mu\right) \leq \int_\Omega \varphi(f) d\mu.$$

Consecuencia: φ es convexa si y sólo si para todos los puntos $u_1, \ldots, u_n \in I$, y para todos los pesos $\lambda_1, \ldots, \lambda_n \geq 0$ con $\lambda_1 + \cdots + \lambda_n = 1$, se tiene que

$$\varphi(\lambda_1 u_1 + \cdots + \lambda_n u_n) \leq \lambda_1 \varphi(u_1) + \cdots + \lambda_n \varphi(u_n).$$

Nota: Una función φ es **cóncava** si $-\varphi$ es convexa. La desigualdad de Jensen para funciones cóncavas es $\varphi\left(\int_\Omega f d\mu\right) \geq \int_\Omega \varphi(f) d\mu.$

Se dice que p y q son exponentes conjugados si $1/p + 1/q = 1$, o equivalentemente, si $q = p/(p-1)$. El conjugado de 1 es ∞.

El siguiente lema es un caso especial de la desigualdad aritmético-geométrica; con frecuencia se le denomina "desigualdad de Young".

Lema: Sean $p, q > 1$ tales que $1/p + 1/q = 1$. Entonces para todo $a, b \geq 0$ se tiene que

$$a \cdot b \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

En general, en las demostraciones los casos $p = 1$ y $p = \infty$ deben tratarse aparte. Pero suelen ser triviales.

Teorema: Desigualdad de Hölder

Sean $p, q \geq 1$ tales que $1/p + 1/q = 1$. Entonces

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

Por tanto, $fg \in \mathcal{L}_1$ si $f \in \mathcal{L}_p$ y $g \in \mathcal{L}_q$.

Corolario: Desigualdad de Cauchy-Schwarz

Si $f, g \in \mathcal{L}_2$ y , entonces $\|fg\|_1 \leq \|f\|_2 \|g\|_2$.

Corolario

Si $f \in \mathcal{L}_p$, entonces $\|f\|_p = \sup_{\{g : \|g\|_q = 1\}} \int fg$.

Teorema: Desigualdad de Minkowski

Sea $1 \leq p \leq \infty$. Si $f, g \in \mathcal{L}_p$, entonces $\|f+g\|_p \leq \|f\|_p + \|g\|_p$.

Dem: Por el corolario anterior.

Desigualdad Jensen o de Hölder, y espacios de probabilidad

• En espacios de probabilidad, y más generalmente en espacios de medida finita, la desigualdad de Jensen, o la de Hölder, implican que cuanto mayor es el exponente p, más pequeño es el espacio.

Teorema

Sean $0 < r < s \leq \infty$, y sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad. Entonces para toda variable aleatoria X se verifica que

$$\|X\|_r \leq \|X\|_s.$$

Por tanto, $\mathcal{L}_s \subset \mathcal{L}_r$.

Dem: El caso $s = \infty$ es trivial. Si $s < \infty$, o bien escribimos $1 = s/s$ y usamos Jensen, o bien escribimos $X = X\mathbf{1}_\Omega$, tomamos $p = s/r > 1$, y sabiendo que $\mathbf{1}_\Omega \in \mathcal{L}_q$, usamos Hölder.

Comentario: cuando $0 < r < 1$, $\|\cdot\|_r := \left(\int |\cdot|^r\right)^{1/r}$ no satisface la desigualdad triangular, luego no es una norma (puede demostrarse que es una cuasi norma).

Teorema: Desigualdad de Markov

Sea $(\Omega, \mathcal{F}, \mu)$ un espacio de medida, y sea $f \geq 0$ medible. Entonces para todo $t > 0$ se verifican las desigualdades

$$\mu\{f \geq t\} \leq \frac{1}{t} \int_{\{f \geq t\}} f d\mu \leq \frac{1}{t} \int_{\Omega} f d\mu.$$

Dem: trivial.