# NOTAS DE TEORÍA DE LA MEDIDA

HÉCTOR A. CHANG-LARA

## Índice

## 1. Introducción

La teoría de la medida puede ser motivada desde tres perspectivas distintas:

**1.1. Perspectiva geométrica.** El objetivo es definir una medida, es decir una función sobre los subconjuntos del espacio euclídeo

$$\mu\colon \mathcal{P}_{\mathbb{R}^d} := \{E \mid E \subseteq \mathbb{R}^d\} \to [0,\infty]$$

la cual refleje el tamaño de cada subconjunto de $\mathbb{R}^d$. Para ello requerimos algunas propiedades naturales. Por ejemplo, el tamaño de la unión de partes disjuntas debe ser la suma de los tamaños de las partes.

1. $\mu(\varnothing) = 0$
2. **Aditiva**: Para $E, F \subseteq \mathbb{R}^d$ disjuntos
$$\mu(E \cup F) = \mu(E) + \mu(F).$$
3. **Continua (por abajo)**: Para toda sucesión creciente de subconjuntos ($E_k \subseteq E_{k+1} \subseteq \mathbb{R}^d$)
$$\mu\left(\bigcup_{k=1}^{\infty} E_k\right) = \lim_{k\to\infty} \mu(E_k).$$
4. **Continua (por arriba)**: Para toda sucesión decreciente de subconjuntos ($E_k \supseteq E_{k+1} \subseteq \mathbb{R}^d$) tal que $\mu(E_1) < \infty$
$$\mu\left(\bigcup_{k=1}^{\infty} E_k\right) = \lim_{k\to\infty} \mu(E_k).$$
5. **Invariante por traslaciones**: Para todo $E \subseteq \mathbb{R}^d$ y $h \in \mathbb{R}^d$
$$\mu(E + h) = \mu(E).$$
6. **Compatible con la fórmula del rectángulo**:
$$\mu([a_1, b_1] \times \ldots \times [a_d, b_d]) = (b_1 - a_1) \ldots (b_d - a_d).$$

Desafortunadamente esto no es posible ni siquiera en el caso unidimensional como lo muestra el siguiente ejemplo de Vitali.

Consideramos una clase de equivalencia en $\mathbb{R}$ para la cual $x \sim y$ si y solo si $x - y \in \mathbb{Q}$. Gracias al axioma de elección, existe un conjunto $N \subseteq [0,1]$ que contiene exactamente un representante de cada clase de equivalencia. Por tanto

$$[0,1] \subseteq \bigcup_{r\in\mathbb{Q}\cap[-1,1]} (N + r) \subseteq [-1,2].$$

En caso de que tal medida exista, y dado que las traslaciones de $N$ por racionales son disjuntas por construcción, obtendríamos que

$$1 \leqslant \sum_{r \in Q \cap [-1,1]} \mu(N) \leqslant 3.$$

Esto es una contradicción dado que la desigualdad por abajo implica que $\mu(N) > 0$, mientras que la desigualdad por arriba solamente es posible si $\mu(N) = 0$.

Para poder construir la **medida de Lebesgue** en $\mathbb{R}^d$ debemos abandonar alguna de las hipótesis previas. Una opción es definir la medida en una subcolección de $\mathcal{P}_\mathbb{R}$, otra es abandonar la aditividad por una condición más débil. De hecho, **ambas ideas terminarán siendo útiles para nuestro propósito.**

**1.2. Perspectiva funcional.** El conjunto de funciones reales sobre un intervalo $E \subseteq \mathbb{R}$ dado admiten estructuras similares a las que conocemos en el espacio euclídeo. Forman un espacio vectorial pero de dimensión infinita. Es justamente la inconmensurabilidad de la dimensión lo que distingue drásticamente la topología entre estos espacios. Por ejemplo:

1. El teorema de equivalencia de las normas solamente se vale en dimensión finita.
2. **Compacidad**: Las sucesiones acotadas tienen alguna subsucesión convergentes solamente en dimensión finita.

Existen distintos espacios de funciones los cuales se caracterizan por su topología o noción de límite. Entre ellas tenemos los siguientes. A continuación consideramos la covergencia de una sucesión de funciones $\{f_k\}$ de $E$ en $\mathbb{R}$ a una dada función $f\colon E \to \mathbb{R}$:

1. **Puntual**: Para todo $x \in E$,
$$\lim_{k\to\infty} f_k(x) = f(x).$$
2. **Uniforme**:
$$\lim_{k\to\infty} \|f_k - f\|_{L^\infty(E)} = 0, \qquad \|f\|_{L^\infty(E)} := \sup_E |f|.$$
3. $L^p$:
$$\lim_{k\to\infty} \|f_k - f\|_{L^p(E)} = 0, \qquad \|f\|_{L^p(E)} := \left(\int_E |f|^p dx\right)^{1/p}.$$

La convergencia puntual es la primera que se aprende en un curso de cálculo. Sin embargo, en general esta no puede ser caracterizada por una métrica, como es el caso de las otras dos.

La convergencia uniforme es la noción más fuerte de convergencia. Siempre implica la convergencia puntual, y también la convergencia en $L^p$ si $E$ es un intervalo acotado. La métrica $\|\cdot\|_{L^\infty(E)}$ es completa cuando nos restringimos al espacio de funciones acotadas sobre $E$. Más aún, también es completa cuando nos restringimos al espacio de funciones continuas sobre un intervalo compacto.

Para $p \geqslant 1$, la función $\|\cdot\|_{L^p(E)}$ es una seminorma sobre el espacio de funciones Riemann integrables. A diferencia de una norma, $\|f\|_{L^p(E)} = 0$ no implica que $f = 0$ como puede ser el caso de $f = \mathbb{1}_{x_0}$ para cualquier $x_0 \in E$. Sin embargo, cualquier seminorma permite definir una norma dada en el espacio cociente definido por la siguiente relación de equivalencia: Decimos que $f \sim g$ si y solo si $\|f - g\|_{L^p(E)} = 0$.

Una desventaja de la integral de Riemann es que la norma $L^p$ no es completa. La completitud de un espacio métrico es una propiedad que permite demostrar la existencia de soluciones de ecuaciones, por ejemplo el teorema de Weierstrass, el teorema del punto fijo, y el teorema de la función inversa dependen esencialmente de esta propiedad.

La **integral de Lebesgue** nos permitirá definir la norma $L^p$ de modo que los espacios de funciones sean completos. Esta completitud es también un requisito para probar la existencia de soluciones a ecuaciones en espacios de funciones, como por ejemplo las ecuaciones diferenciales.

**1.3. Perspectiva probabilista.** Las **variables aleatorias** son unos de los objetos fundamentales en la teoría de probabilidad. Estas caracterizan el resultado de algún evento aleatorio. Una variable aleatoria es descrita en términos de su **distribución** la cual define la probabilidad de que la variable tome diversos valores.

La formalización matemática de esta idea consiste en dar un espacio de probabilidad abstracto $\Omega$ y una medida de probabilidad $\mathbb{P}\colon \mathcal{P}_\Omega \to [0,1]$ (aditiva y continua). Las variables aleatorias con valores en algún conjunto $S$ están definidas a través funciones $X : \Omega \to S$. La distribución de $X$ es la medida de probabilidad $\mu$ sobre $S$ tal que para $E \subseteq S$

$$\mu(E) := \mathbb{P}(X \in E) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in E\}).$$

El desarrollo de la **integral de Lebesgue-Stieljes** permitió definir rigurosamente las distribuciones sobre $S = \mathbb{R}$. Quedaba por otro lado abierta la pregunta sobre la posible realización de dichas distribuciones a partir de un espacio abstracto de probabilidad $\Omega$. Los casos más interesantes son aquellos donde existen más de una variable aleatoria con diferentes interacciones entre ellas, también denominados como **acoplamientos**.

**1.4. Algunos libros y recomendaciones.** De acuerdo con las perspectivas que hemos discutido, podríamos clasificar tres tipos de libros:

1. **Euclídeanos**: Hacen mayor énfasis en las construcciones de $\mathbb{R}^d$, en particular la medida e integral de Lebesgue. Por ejemplo, Wheeden-Zygmund, Royden-Fitzpatrick (Parte 1), Stein-Shakarchi, Tao, Evans-Gariepy.
2. **Abstractos**: Desarrollan la teoría general de la medida. Folland y papa Rudin.
3. **Probabilísticos**: Durret y Billingsley.

Las recomendaciones varían dependiendo si lo que buscan un libro para aprender la materia desde cero o de referencia. En mi opinión, Tao es muy bueno para aprender la materia de forma independiente, Stein-Shakarchi y Durret son muy buenos para dictar la materia puesto que tiene problemas interesantes. Papa Rudin, Billingsley, y Evans-Gariepy son libros de referencia ampliamente conocidos y citados. Por su puesto que existen otros libros muy buenos que desconozco.

Algunos comentarios con más detalle:

**An Introduction to Measure Theory** por Terence Tao.

Son las notas del curso de teoría de la medida que suele dictar durante un trimestre. Se basan en el libro de Stein-Shakarchi y también en el libro de Folland. Buena opción para autodidactas, muchos resultados se dejan como ejercicios al lector. Hace énfasis en el desarrollo intuitivo de conceptos y teoría. **En particular recomiendo leer el capítulo Problem solving strategies**.

Comienza con la construcción de la medida de Jordan como motivación de la medida de Lebesgue. La medida abstracta aparece en el último capítulo.

No muestra la organización típica de un libro de texto y puede ser difícil de navegar. El teorema de Radon-Nikodym, espacios de Lebesgue, y el teorema de Riesz aparecen en otro libro: **An Epsilon of Room I**.

Errata: https://terrytao.wordpress.com/books/an-introduction-to-measure-theory/

**Real Analysis** por Elias Stein y Rami Shakarchi.

Tercer volumen de la serie de análisis de Stein y Shakarchi. Altamente recomendado por sus ejercicios y problemas. Incluye además varias secciones con tópicos interesantes que suelen estar fuera del temario básico. **Recomiendo leer el capítulo introductorio, a mi parecer es una de las mejores presentaciones para motivar la materia.**

Incluye dos capítulos sobre espacios de Hilbert con énfasis en análisis armónico. Los espacios de Lebesgue aparecen en otro volumen: **Functional Analysis**.

Errata: https://www.princeton.edu/releases/m8008_errata.pdf

**Real Analysis** por Gerald Folland.

Hace énfasis desde el principio en el desarrollo abstracto de la teoría de la medida. Tiene un capítulo sobre teoría de probabilidad y la construcción de la medida de Wiener.

Errata: https://sites.math.washington.edu//~folland/Homepage/reals.pdf

**2020-2021 Real Analysis (MTH-RA)** por Emanuel Carneiro.

Serie de 15 videos de 90 minutos cada uno con el curso de teoría de la medida para el programa de postgrado del ICTP. Basado en Stein-Shakarchi y Folland.

https://youtube.com/playlist?list=PLp0hSY2uBeP8hajKOVGZ9oIPjG3HKMfoY&si=z-wMfGT-2FZVphA8

**Measure Theory and Fine Properties of Functions** por Craig Evans y Ronald Gariepy.

Al igual que Stein-Shakarchi y Tao es un libro de texto primordialmente sobre medidas en $\mathbb{R}^d$. Presenta algunos de los resultados modernos que no aparecen en otro libros de un nivel similar: Fórmula de área y coárea, espacios de Sobolev y funciones de variación acotada.

No incluye ejercicios.

Errata: https://math.berkeley.edu/~evans/errata.Evans-Gariepybook2.pdf

### 1.5. Ejercicios.

1. Demuestra que una función $\mu\colon \mathcal{P}_X \to [0,\infty]$ tal que $\mu(\varnothing) = 0$, es aditiva, y es continua por abajo cumple las siguientes propiedades:
   *a*) **Monótona**: Para $E \subseteq F \subseteq X$, se tiene que $\mu(E) \leqslant \mu(F)$.

*b*) **$\sigma$-aditiva**: Para toda sucesión de subconjuntos disjuntos a pares

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

*c*) **$\sigma$-subaditiva**: Para toda sucesión de subconjuntos, no necesariamente disjuntos,

$$\mu\left(\bigcup_{k=1}^{\infty} E_k\right) \leqslant \sum_{k=1}^{\infty} \mu(E_k).$$

*d*) **Continua por arriba**: Para toda sucesión decreciente de subconjuntos, tal que $\operatorname{ínf}_k \mu(E_k) < \infty$,

$$\mu\left(\bigcap_{k=1}^{\infty} E_k\right) = \lim_{k\to\infty} \mu(E_k).$$

2. Sea $E \subseteq \mathbb{R}$:
   *a*) Demuestra que $\|\cdot\|_{L^\infty(E)}$ es una norma en el conjunto de funciones acotadas sobre $E$.
   *b*) Demuestra que la norma $\|\cdot\|_{L^\infty(E)}$ define una métrica completa en el conjunto de funciones acotadas sobre $E$.
   *c*) Demuestra que la norma $\|\cdot\|_{L^\infty([a,b])}$ define una métrica completa en el conjunto de funciones continuas sobre $[a, b]$.
3. Demuestra que $\|\cdot\|_{L^1([a,b])}$ es una seminorma sobre el conjunto de funciones Riemann integrables sobre $[a, b]$.
4. Sea $\{r_k\}$ una enumeración de los racionales y

$$E_k = \bigcup_{j=1}^{k} (r_j - 2^{-j}, r_j + 2^{-j}).$$

   *a*) Demuestra que $\{\mathbb{1}_{E_k}\}$ es una sucesión de Cauchy con respecto a $\|\cdot\|_{L^1(\mathbb{R})}$.
   *b*) Demuestra que $\{\mathbb{1}_{E_k}\}$ converge puntualmente a $\mathbb{1}_E$ donde

$$E = \bigcup_{j=1}^{\infty} (r_j - 2^{-j}, r_j + 2^{-j}).$$

   *c*) Demuestra que $\mathbb{1}_E$ no es Riemann integrable.

   Queda pendiente ver que el límite de $\{\mathbb{1}_{E_k}\}$, pero con respecto a $\|\cdot\|_{L^1(\mathbb{R})}$, es esencialmente igual a $\mathbb{1}_E$.
5. Dados $\{x_k\}, \{y_k\} \subseteq \mathbb{R}$ y $\{p_k\}, \{q_k\} \subseteq [0, 1]$ tales que $\sum_{k=1}^{\infty} p_k = \sum_{k=1}^{\infty} q_k = 1$, definimos $\mu, \nu \colon \mathcal{P}_{\mathbb{R}} \to [0, 1]$ tal que

$$\mu(E) := \sum_{k=1}^{\infty} p_k \mathbb{1}_E(x_k), \qquad \nu(E) := \sum_{k=1}^{\infty} q_k \mathbb{1}_E(y_k).$$

   Sea $\gamma \colon \mathcal{P}_{\mathbb{R}^2} \to [0, 1]$ tal que

$$\gamma(E \times \mathbb{R}) = \mu(E), \qquad \gamma(\mathbb{R} \times F) = \nu(F).$$

   *a*) Demuestra que $\mu$ y $\nu$ son aditivas y continuas.
   *b*) Construye un conjunto $\Omega$, una función $\mathbb{P} \colon \mathcal{P}_\Omega \to [0, 1]$ aditiva y continua, y una función $X \colon \Omega \to \mathbb{R}$ tal que para todo $E \subseteq \mathbb{R}$ se cumpla que $\mu(E) = \mathbb{P}(X \in E)$.

*c*) Demuestra que mientras $\mu$ y $\nu$ están únicamente definidas por las relaciones dadas, en general existe más de una posible $\gamma$ que satisface las relaciones dadas para ella.

*d*) Construye un conjunto $\Omega$, una función $\mathbb{P}\colon \mathcal{P}_\Omega \to [0,1]$ aditiva y continua, y un par de funciones $X\colon \Omega \to \mathbb{R}$ e $Y\colon \Omega \to \mathbb{R}$ tal que para todo $E, F \subseteq \mathbb{R}$ se cumpla que $\gamma(E \times F) = \mathbb{P}(X \in E, Y \in F)$.

*e*) Demuestra que las funciones $X$ e $Y$ dadas arriba satisfacen necesariamente que $\mu(E) = \mathbb{P}(X \in E)$ y $\nu(F) = \mathbb{P}(Y \in F)$.

## 2. Medida

Una $\sigma$-álgebra es una colección no vacía de conjuntos que es cerrada bajo las operaciones básicas (unión, intersección, diferencias y límites). Basta exigir unos pocos axiomas.

**Definición 2.1.** *Una colección de subconjuntos $\mathcal{M} \subseteq \mathcal{P}_X$ es una $\sigma$-álgebra si cumple que:*

- $\varnothing \in \mathcal{M}$.
- ***Cerrada bajo el complemento:*** *Para todo $E \in \mathcal{M}$, se tiene que $X \backslash E \in \mathcal{M}$.*
- ***Cerrada bajo la unión:*** *Para todo $E, F \subseteq \mathcal{M}$, se tiene que $E \cup F \in \mathcal{M}$.*
- ***Continua por abajo:*** *Para toda sucesión creciente $\{E_k\} \subseteq \mathcal{M}$, es decir que $E_k \subseteq E_{k+1}$; se tiene que $\bigcup_{k=1}^{\infty} E_k \in \mathcal{M}$.*

*En este caso denominamos al par $(X, \mathcal{M})$ un **espacio medible**. Los miembros de $\mathcal{M}$ se conocen como los **conjuntos medibles**.*

**Definición 2.2.** *Dado un espacio medible $(X, \mathcal{M})$, una **medida** es una función $\mu\colon \mathcal{M} \to [0, \infty]$ $\sigma$-aditiva tal que $\mu(\varnothing) = 0$. En este caso denominamos a la tripleta $(X, \mathcal{M}, \mu)$ un **espacio de medida**.*

Si $(X, \mathcal{M}, \mu)$ satisface que $\mu(X) < \infty$, entonces se dice que $X$ es de medida finita. Un caso particular ocurre cuando $\mu(X) = 1$, en cuyo caso se dice que $(X, \mathcal{M}, \mu)$ es un espacio de probabilidad. Probalistas suelen preferir la notación $(\Omega, \mathcal{F}, \mathbb{P})$, donde el conjunto $\Omega$ se denomina el espacio **muestral**. Las $\sigma$-álgebras también se conocen en probabilidad como **filtraciones**, y una filtración $\mathcal{F} \subseteq \mathcal{P}_\Omega$ dada sobre $\Omega$ se denomina como los **eventos**. La medida $\mathbb{P}\colon \mathcal{F} \to [0, 1]$ es la **medida de probabilidad**.

Dada una propiedad sobre puntos de $X$, se dice esta se cumple en **casi todo punto** si esta se cumple fuera de un conjunto excepcional de medida cero.

Un espacio de medida $(\mathbb{R}, \mathcal{M}, \mu)$ sobre los reales se dice invariante por traslaciones si para todo $E \in \mathcal{M}$ y $h \in \mathbb{R}$, $E + h \in \mathcal{M}$ y $\mu(E + h) = \mu(E)$.

El primer resultado fundamental del curso será la construcción de la medida de Lebesgue sobre los números reales.

**Teorema 1.** *Los reales admiten estructura de espacio de medida $(\mathbb{R}, \mathcal{L}_\mathbb{R}, m)$ invariante por traslaciones tal que $\mathcal{L}_\mathbb{R}$ contiene los intervalos de $\mathbb{R}$ y $m([0, 1]) = 1$.*

### 2.1. Ejercicios.

1. Dado un espacio medible $(X, \mathcal{M})$ y $E \subseteq X$ se define la $\sigma$-álgebra inducida $\mathcal{M}|_E := \{A \cap E \mid A \in \mathcal{M}\}$. Demuestra que $\mathcal{M}|_E$ es efectivamente una $\sigma$-álgebra.
2. Demuestra que la intersección de una familia de $\sigma$-álgebras es una $\sigma$-álgebra. Dado $\mathcal{E} \subseteq \mathcal{P}_X$, la $\sigma$-álgebra que se obtiene al intersectar todas las $\sigma$-álgebras que contienen a $\mathcal{E}$ se conoce como la $\sigma$-álgebra generada por $\mathcal{E}$.
3. Si $X$ es un espacio topológico, se define la $\sigma$-álgebra generada por los conjuntos abiertos como la $\sigma$-álgebra de **Borel**. En el caso de $\mathbb{R}^d$ se denotará por $\mathcal{B}_{\mathbb{R}^d}$. Demuestra que $\mathcal{B}_\mathbb{R} \subseteq \mathcal{L}_\mathbb{R}$.

4. Dado $T : X \to Y$ y $\mathcal{N} \subseteq \mathcal{P}_Y$ definimos el pullback $T^{-1}(\mathcal{N}) = \{T^{-1}(E) \mid E \in \mathcal{N}\}$. Dada $\mu\colon T^{-1}(\mathcal{N}) \to \mathbb{R}$, definimos el pushforward $T_\#\mu\colon \mathcal{N} \to \mathbb{R}$ tal que $T_\#\mu(E) := \mu(T^{-1}(E))$.
   *a*) Demuestra que si $\mathcal{N}$ es una $\sigma$-álgebra, entonces $T^{-1}(\mathcal{N})$ en una $\sigma$-álgebra.
   *b*) Demuestra que si $X \subseteq Y$ y $T(x) = x$, entonces $T^{-1}(\mathcal{N}) = \mathcal{N}|_X$.
   *c*) Demuestra que si $\mu$ es una medida, entonces $T_\#\mu$ es una medida.
   *d*) Demuestra que el pullback satisface que $(S \circ T)^{-1} = T^{-1} \circ S^{-1}$.
   *e*) Demuestra que el pushforward satisface que $(S \circ T)_\# = S_\# \circ T_\#$.
5. Sea $\{(X_\alpha, \mathcal{M}_\alpha)\}$ una colección de espacios medibles. Sean $\pi_\alpha\colon \prod_\beta X_\beta \to X_\alpha$ los mapas de coordenadas. Se define la $\sigma$-álgebra producto $\bigotimes_\alpha \mathcal{M}_\alpha$ como aquella que generada por $\bigcup_\alpha \pi_\alpha^{-1}(\mathcal{M}_\alpha)$. Demuestra que
$$\mathcal{B}_{\mathbb{R}^d} = \mathcal{B}_{\mathbb{R}}^{\otimes d} = \underbrace{\mathcal{B}_{\mathbb{R}} \otimes \ldots \otimes \mathcal{B}_{\mathbb{R}}}_{d \text{ veces}}.$$

## 3. La construcción de Carathéodory

Las ideas detrás de la construcción de la medida de Lebesgue son útiles en las construcciones de otras medidas: Lebesgue-Stieljes, producto, Hausdorff, Wiener, etc. Por esta razón optamos por presentar el enfoque abstracto desde el comienzo.

**Definición 3.1.** *Una **medida exterior** es una función $\mu_*\colon \mathcal{P}_X \to [0,\infty]$ $\sigma$-subaditiva, monótona y tal que $\mu_*(\varnothing) = 0$.*

La medida exterior de Lebesgue en $\mathbb{R}$ se construye a partir de la longitud de los intervalos. Un intervalo es cualquier conjunto convexo de $\mathbb{R}$. La longitud de un intervalo $I \neq \varnothing$ se calcula por

$$\ell(I) = \sup_{[a,b]\subseteq I} (b-a).$$

En el caso del conjunto vacío se define $\ell(\varnothing) = 0$.

Tenemos así la medida exterior de Lebesgue

$$m_*(A) := \text{ínf} \left\{ \sum_{k=1}^{\infty} \ell(I_k) \mid A \subseteq \bigcup_{k=1}^{\infty} I_k \text{ intervalos} \right\}.$$

En este caso se tiene que $m_*$ extiende la longitud, es decir que $m_*(I) = \ell(I)$ para todo intervalo $I$.

**Definición 3.2** (Criterio de Carathédory)**.** *Dada una medida exterior $\mu_*\colon \mathcal{P}_X \to [0,\infty]$ se define la colección $\mathcal{M}(\mu_*) \subseteq \mathcal{P}_X$ de conjuntos medibles tal que $E \in \mathcal{M}(\mu_*)$ si y solo si para todo $B \subseteq X$ se tiene que*

$$\mu_*(B) = \mu_*(B \cap E) + \mu_*(B \backslash E).$$

Los conjuntos $B$ que son usados en el criterio de Carathédory se denominan conjuntos de prueba. Su propósito es determinar localmente la aditividad del conjunto $E$ con respecto a su complemento.

**Teorema 2.** *Dada una medida exterior $\mu_*\colon \mathcal{P}_X \to [0,\infty]$ se tiene que $\mathcal{M}(\mu_*)$ es una $\sigma$-álgebra y $\mu := \mu_*|_{\mathcal{M}(\mu_*)}$ es una medida sobre $\mathcal{M}(\mu_*)$.*

### 3.1. Ejercicios.

1. Dado $\mathcal{E} \subseteq \mathcal{P}_X$ tal que $X, \varnothing \in \mathcal{E}$, y $\ell\colon \mathcal{E} \to [0,\infty]$ tal que $\ell(\varnothing) = 0$, sea $\mu_*\colon \mathcal{P}_X \to [0,\infty]$ tal que
$$\mu_*(A) = \text{ínf} \left\{ \sum_{k=1}^{\infty} \ell(E_k) \mid \{E_k\} \subseteq \mathcal{E}, A \subseteq \bigcup_{k=1}^{\infty} E_k \right\}.$$
   *a*) Demuestra que $\mu_*$ una medida exterior.
   *b*) Demuestra que si $\mathcal{E} = \mathcal{P}_X$ y $\ell$ es una medida exterior, entonces $\mu_* = \ell$.
2. $\star$ Demuestra que la longitud $\ell$ es $\sigma$-aditiva. Es decir si $\{I_k\}$ es una sucesión de intervalos disjuntos y $I = \bigcup_{k=1}^{\infty} I_k$ también es un intervalo, entonces $\ell(I) = \sum_{k=1}^{\infty} \ell(I_k)$.
3. Demuestra que la misma medida exterior de Lebesgue también se obtiene si en vez de usar todos los posibles intervalos de la recta real, usamos exclusivamente los siguientes conjuntos:

   *a*) intervalos abiertos,
   *b*) intervalos cerrados.
4. Demuestra que para todo intervalo $I$ se tiene que $m_*(I) = \ell(I)$.
5. Demuestra que para todo $E \subseteq \mathbb{R}$ se tiene que
$$m_*(E) = \text{ínf}\{m_*(U) \mid E \subseteq U \text{ abierto}\}.$$
6. Demuestra que si $A, B \subseteq \mathbb{R}$ son tales que
$$\operatorname{dist}(A, B) := \text{ínf}\{|x - y| \mid x \in A, y \in B\} > 0,$$
entonces $m_*(A \cup B) = m_*(A) + m_*(B)$.

## 4. Funciones medibles

En comparación con la teoría de medida, donde los protagonistas son subconjuntos de $X$ a los que se les asigna una medida, la teoría de integración presenta una perspectiva distinta: los objetos de interés son funciones con valores reales a las que se les asigna una integral.

**Definición 4.1.** *Sean $(X, \mathcal{M})$ e $(Y, \mathcal{N})$ espacios medibles. Decimos que $f : X \to Y$ es medible si $f^{-1}(\mathcal{N}) \subseteq \mathcal{M}$.*

Esta definición hace ver que la composición de funciones medibles es medible. Sin embargo, en el caso de que el codominio esté dado por $Y \subseteq \mathbb{R}$, asumimos por defecto que $\mathcal{N} = \mathcal{B}_{\mathbb{R}}|_Y$. Mientras que si el dominio está dado por $X \subseteq \mathbb{R}$, asumimos que $\mathcal{M} = \mathcal{L}_{\mathbb{R}}|_X$. Esta discrepancia hace que la composición de dos funciones $f, g\colon \mathbb{R} \to \mathbb{R}$ no sea necesariamente medible si $f$ y $g$ son medibles. Por otro lado, si $f$ es medible y $g$ es continua, entonces sí se tiene que $g \circ f$ es medible.

Para verificar que $f\colon X \to \mathbb{R}$ es medible basta ver que para todo $\lambda \in \mathbb{R}$ se tiene que $\{f > \lambda\} \in \mathcal{M}$. Esto se sigue porque la $\sigma$-álgebra de Borel está generada por intervalos de la forma $(\lambda, \infty)$. Criterios similares se valen si reemplazamos $>$ por $<$, $\geqslant$, ó $\leqslant$. Este simple criterio no se cumpliría si asumimos que $\mathcal{N}$ es una $\sigma$-álgebra más fina que $\mathcal{B}_{\mathbb{R}}|_Y$, como por ejemplo $\mathcal{L}_{\mathbb{R}}|_Y$.

Algunas construcciones que preservan medibilidad de funciones reales son la suma, el producto, supremos, ínfimos, límites superior e inferior, y el límite en caso de que esté definido.

Dado $E \subseteq X$ se tiene que la función indicadora $\mathbb{1}_E\colon X \to \mathbb{R}$ dada por

$$\mathbb{1}_E(x) := \begin{cases} 1 \text{ si } x \in E, \\ 0 \text{ en cualquier otro caso,} \end{cases}$$

es medible si y solo si $E \in \mathcal{M}$.

Las combinaciones lineales de funciones indicadoras de conjuntos medibles se conocen como las **funciones simples**. Esta clase de funciones son extremadamente útiles a la hora de entender las funciones medibles.

### 4.1. Modos de convergencia.

Las funciones medibles de $X$ en los reales presentan distintos tipos de convergencia. Entre ellas ya hemos mencionado la convergencia uniforme y la convergencia puntual.

Dada la noción de conjuntos nulos, tenemos que la convergencia puntual puede ser extendida a la convergencia en casi todo punto. Esta última sucede cuando una sucesión de $\{f_k\}$ funciones medibles convergen a una función $f$ medible, salvo en un conjunto de medida cero. Equivalentemente tenemos la siguiente definición.

**Definición 4.2.** *Una sucesión de funciones $\{f_k\}$ medibles converge en casi todo punto a una función $f$ si para todo $\varepsilon > 0$ se tiene que*

$$\mu\left(\limsup_{k\to\infty}\{|f_k - f| > \varepsilon\}\right) = 0.$$

Si intercambiamos el límite y la medida obtenemos la noción de **convergencia en medida**.

**Definición 4.3.** *Una sucesión de funciones* $\{f_k\}$ *medibles converge en medida a* $f$ *si para todo* $\varepsilon > 0$

$$\lim_{k\to\infty} \mu(\{|f_k - f| > \varepsilon\}) = 0.$$

Dados todos estos modos de convergencia, es natural preguntarse las relaciones entre ellos.

La convergencia uniforme es sin duda la más fuerte de todas. Implica la convergencia puntual y en medida. En la dirección recíproca y sobre espacios de medida finita, es posible promover la convergencia puntual a convergencia uniforme fuera de un conjunto excepcional arbitrariamente pequeño.

**Teorema 3** (Egorov)**.** *Sea* $(X, \mathcal{M}, \mu)$ *un espacio de medida finita y* $\{f_k\}$ *una sucesión de funciones medibles de* $X$ *en* $\mathbb{R}$ *que convergen en casi todo punto a una función* $f$*. Dado* $\varepsilon > 0$*, existe un conjunto* $\Sigma_\varepsilon \in \mathcal{M}$ *tal que* $\mu(\Sigma_\varepsilon) < \varepsilon$ *y* $\{f_k\}$ *converge uniformemente a* $f$ *en* $X \backslash \Sigma_\varepsilon$*.*

El **teorema de Borel-Cantelli** junto con el argumento diagonal de Cantor permite demostrar que toda sucesión de funciones que convergen en medida tiene una subsucesión que converge puntualmente. Esta idea será útil para demostrar que los espacios de Lebesgue son completos.

### 4.2. Ejercicios.

1. Sean $(X, \mathcal{M})$, $(Y_1, \mathcal{N}_1)$, y $(Y_2, \mathcal{N}_2)$ espacios medibles. Demuestra que $(f, g) : X \to Y_1 \times Y_2$ es medible si y solo si $f : X \to Y_1$ y $g \colon X \to Y_2$ son funciones medibles.
2. Sea $(X, \mathcal{M})$ un espacio medible y $f, g \colon X \to \mathbb{R}$ funciones medibles
   *a*) Dadas $f, g \colon X \to \mathbb{R}$ medibles, demuestra que $f + g$ y $fg$ son medibles.
   *b*) Dada una sucesión $\{f_k\}$ de funciones medibles de $X$ a $\mathbb{R}$, demuestra que $\sup f_k$, $\inf f_k$, $\liminf f_k$ y $\limsup f_k$ son medibles.
   *c*) Dada una sucesión $\{f_k\}$ de funciones medibles de $X$ a $\mathbb{R}$, demuestra que el conjunto donde $\lim_{k\to\infty} f_k$ no está definido es medible.
3. Sea $(X, \mathcal{M}, \mu)$ un espacio de medida, $\{f_k\}$ una sucesión de funciones medibles de $X$ en $\mathbb{R}$ y $f \colon X \to \mathbb{R}$ medible. Demuestra que
   $$\mu\left(\left\{\lim_{k\to\infty} f_k \text{ no está definido}\right\} \cup \left\{\lim_{k\to\infty} f_k \text{ está definido y es distinto de } f\right\}\right) = 0,$$
   si y solo si para todo $\varepsilon > 0$,
   $$\mu\left(\limsup_{k\to\infty} \{|f_k - f| > \varepsilon\}\right) = 0.$$
4. Da un ejemplo de una sucesión de funciones que converjan puntualmente pero no en medida.
5. Demuestra que en espacios de medida finita la convergencia en casi todo punto implica la convergencia en medida.

6. Demuestra el teorema de Borel-Cantelli: Dada una sucesión de conjuntos medibles $\{E_k\}$

$$\sum_{k=1}^{\infty} \mu(E_k) < \infty \qquad \Rightarrow \qquad \mu\left(\limsup_{k\to\infty} E_k\right) = 0.$$

7. Demuestra que toda sucesión de funciones que convergen en medida tiene una subsucesión que converge en casi todo punto.
8. Se define la sucesión de funciones $f_k\colon \mathbb{R} \to \mathbb{R}$ (también conocida como typewriter) como

$$f_k = \mathbb{1}_{[(k-1)2^{-n}, k2^{-n}]}, \qquad 2^n \leqslant k < 2^{n+1}$$

   *a*) Demuestra que estas convergen en medida pero no en casi todo punto.
   *b*) Da una subsucesión que converga en casi todo punto.

## 5. Integral

Dado un espacio de medida $(X, \mathcal{M}, \mu)$, la integral es un mapa que asigna a cada función medible $f : X \to [0, \infty)$ un número $\int f d\mu \in [0, \infty]$ tal que:

1. **Compatibilidad con la medida**: Dado $E \in \mathcal{M}$, se tiene que $\int \mathbb{1}_E d\mu = \mu(E)$.
2. **Linealidad:**
   *a*) **Aditividad:** $\int (f + g) d\mu = \int f d\mu + \int g d\mu$.
   *b*) **Homogeneidad:** Para toda constante $c > 0$, $\int c f d\mu = c \int f d\mu$.
3. **Convergencia monótona:** Dada una sucesión $\{f_k\}_{k \geqslant 1}$ de funciones medibles que crecen a $f$, se tiene que
$$\lim_{n \to \infty} \int f_k d\mu = \int f d\mu.$$

Las dos primeras propiedades definen de forma única la integral en funciones simples no-negativas. Es decir
$$\int \sum_{j=1}^{k} c_j \mathbb{1}_{E_j} d\mu = \sum_{j=1}^{k} c_j \mu(E_j).$$

La convergencia monótona sugiere una forma de extender la integral a las demás funciones medibles no-negativas
$$\int f d\mu := \sup \left\{ \int \varphi d\mu \mid 0 \leqslant \varphi \leqslant f \text{ es una función simple} \right\}.$$

**Teorema 4.** *Dado un espacio de medida $(X, \mathcal{M}, \mu)$, existe una única integral que satisface las axiomas dados.*

Para definir la integral de una función medible $f \colon X \to \mathbb{R}$, que ahora podría tomar valores negativos, usamos la descomposición en su parte positiva y negativa
$$f = f_+ - f_-, \qquad f_\pm(x) := \max\{\pm f(x), 0\}.$$
Si por lo menos una de las integrales $\int f_\pm d\mu$ es finita decimos que $f$ es **integrable** y definimos
$$\int f d\mu := \int f_+ d\mu - \int f_- d\mu.$$
El resultado de la integral bien podría ser $\pm\infty$.

Dado que $|f| = f_+ + f_-$, ambas integrales $\int f_\pm d\mu < \infty$ si y solo si $\int |f| d\mu < \infty$. En este caso decimos que $f$ es **absolutamente integrable** y se denota por $L^1(X, \mathcal{M}, \mu)$ ó $L^1(X)$ al conjunto de todas estas funciones. El conjunto $L^1(X)$ admite la estructura de un espacio vectorial. Más aún,
$$\|f\|_{L^1(X)} := \int |f| d\mu$$
define norma con la salvedad que $\|f\|_{L^1(X)} = 0$ no implica que $f = 0$, sino que en cambio esta igualdad se cumple en casi todo punto.

**5.1. Distribuciones.** En probabilidad se denominan las funciones medibles $X : \Omega \to \mathbb{R}$ como **variables aleatorias**. El pushforward $\mu = X_{\#}\mathbb{P}$ se conoce como la **medida de distribución de la variable aleatoria**. La **función de distribución acumulada** se define como

$$F_X(\lambda) := \mathbb{P}(X \leqslant \lambda) = \mu((-\infty, \lambda]).$$

La integral con respecto de $\mathbb{P}$ se denomina la esperanza y se denota por

$$\mathbb{E}(X) := \int X d\mathbb{P}.$$

La fórmula de cambio de variable (ver ejercicio en esta sección) permite calcular esperanzas a partir de la distribución dada en $\mathbb{R}$. Dada $f\colon \mathbb{R} \to \mathbb{R}$ integrable con respecto a $\mu = X_{\#}\mathbb{P}$

$$\mathbb{E}(f(X)) = \int_\Omega f(X) d\mathbb{P} = \int_{\mathbb{R}} f d\mu.$$

**5.2. Ejercicios.**

1. Sea $(X, \mathcal{M})$ un espacio medible y $f\colon X \to [0, \infty)$ medible. Demuestra que

$$\varphi_k := \sum_{j=1}^{2^{2k}} c_{j,k} \mathbb{1}_{E_{j,k}}, \qquad c_{j,k} := j2^{-k}, \qquad E_{j,k} := \{j2^{-k} \leqslant f < (j+1)2^{-k}\},$$

   define una sucesión funciones simples no-negativas que crecen a $f$.
2. Demuestra que $L^1(X)$ es un espacio vectorial y que

$$\|f\|_{L^1(X)} := \int |f| d\mu$$

   define una seminorma para la cual $\|f\|_{L^1(X)} = 0$ implica que $f = 0$ en casi todo punto.
3. Demuestra las propiedades de la integral a partir de sus axiomas:
   *a*) **Monotonicidad**: Para $f \leqslant g$ se tiene que $\int f d\mu \leqslant \int g d\mu$.
   *b*) **Desigualdad de Markov**: Para $\lambda > 0$

$$\mu(\{|f| > \lambda\}) \leqslant \frac{1}{\lambda} \int |f| d\mu.$$

4. Demuestra el **lema de Fatou:** Dada una sucesión $\{f_k\}_{k\geqslant 1}$ de funciones medibles no negativas que convergen puntualmente a $f$, se tiene que

$$\int \liminf_{k\to\infty} f_k d\mu \leqslant \liminf_{k\to\infty} \int f_k d\mu.$$

5. Demuestra el teorema de **convergencia dominada:** Sea $\{f_k\}$ una sucesión de funciones medibles de $X$ en $\mathbb{R}$ que convergen puntualmente a una función $f$. Demuestra que si $\sup_k |f_k| \in L^1(X)$ entonces se tiene que

$$\lim_{k\to\infty} \int f_k d\mu = \int f d\mu.$$

6. Demuestra la fórmula de **cambio de variables:** Sean $(X, \mathcal{M}, \mu)$ un espacio de medida, $(Y, \mathcal{N})$ un espacio medible, y sea $T : X \to Y$ medible. Para toda función medible $f : Y \to [0, \infty)$ se tiene que
$$\int_X f \circ T d\mu = \int_Y f dT_{\#}\mu.$$
Extiende el resultado a funciones $f : Y \to \mathbb{R}$ integrables con respecto a $T_{\#}\mu$.

## 6. Teorema de Fubini

Dados $(X, \mathcal{M}, \mu)$ e $(Y, \mathcal{N}, \nu)$ espacios de medida, buscamos definir una medida $\gamma$ en $(X \times Y, \mathcal{M} \otimes \mathcal{N})$ que sea compatible con la estructura producto. En general, decimos que $\gamma$ es un **acoplamiento** entre estos espacios si

$$(\pi_X)_\# \gamma = \mu, \qquad (\pi_Y)_\# \gamma = \nu.$$

Estas medidas de tipo pushforward también se conocen como las **marginales** de $\gamma$.

Por ejemplo, asumamos que $T\colon X \to Y$ es medible y $\nu = T_\# \mu$. En este caso tenemos que el mapeo en la gráfica $G\colon X \to X \times Y$ tal que $G(x) := (x, T(x))$ empuja a $\mu$ en la medida $\gamma := G_\# \mu$, la cual acopla $\mu$ y $\nu$.

Otro acoplamiento posible es la **medida producto** $\gamma = \mu \otimes \nu$ la cual se manifiesta si para todo $E \in \mathcal{M}$ y $F \in \mathcal{N}$ se tiene que $(\mu \otimes \nu)(E \times F) = \mu(E)\nu(F)$. La existencia y unicidad de esta medida no es trivial.

**Teorema 5.** *Dados dos espacios de medida $(X, \mathcal{M}, \mu)$ e $(Y, \mathcal{N}, \nu)$, existe un espacio de medida producto $(X \times Y, \mathcal{M} \otimes \mathcal{N}, \mu \otimes \nu)$. Si además los factores son espacios de medida finita, entonces el producto es único.*

Cuando tomamos la perspectiva de las funciones medibles obtenemos en cambio el teorema de Fubini.

Dada $f\colon X \times Y \to Z$ y $(x_0, y_0) \in X \times Y$ denotamos

$$f_{x_0}(y) := f(x_0, y), \qquad f^{y_0}(x) := f(x, y_0).$$

**Teorema 6.** *Sean $(X, \mathcal{M}, \mu)$ e $(Y, \mathcal{N}, \nu)$ dos espacios de medida finita y $f\colon X \times Y \to [0, \infty)$ medible:*

1. *Para todo $(x_0, y_0) \in X \times Y$, las funciones $f_{x_0}$ y $f^{y_0}$ son medibles.*
2. *Las funciones $x \mapsto \int_Y f_x d\nu$ e $y \mapsto \int_X f^y d\mu$ son medibles.*
3. *Se tiene que*

$$\int_{X \times Y} f d(\mu \otimes \nu) = \int_Y \left( \int_X f^y d\mu \right) d\nu = \int_X \left( \int_Y f_x d\nu \right) d\mu.$$

### 6.1. Independencia.

En probabilidad, el acoplamiento describe la relación entre las variables aleatorias. Dadas $X, Y\colon \Omega \to \mathbb{R}$, tenemos que la medida $\gamma = (X, Y)_\# \mathbb{P}$ en $\mathbb{R}^2$ acopla las medidas $\mu = X_\# \mathbb{P}$ y $\nu = Y_\# \mathbb{P}$ en $\mathbb{R}$. Se dice que $X$ e $Y$ son **independientes** si $\gamma = \mu \otimes \nu$. En el otro extremo, se dice que $Y$ depende de $X$ si existe una función $T\colon \mathbb{R} \to \mathbb{R}$ tal que $Y = T(X)$.

### 6.2. Ejercicios.

1. Sean $(X, \mathcal{M}, \mu)$ e $(Y, \mathcal{N}, \nu)$ dos espacios de medida y $T : X \to Y$ tal que $\nu = T_\# \mu$. Dado $G\colon X \to X \times Y$ tal que $G(x) := (x, T(x))$, demuestra que $\gamma := G_\# \mu$ tiene marginales $\mu$ y $\nu$.
2. Sean $(X, \mathcal{M})$ e $(Y, \mathcal{N})$ espacios medibles:

*a*) Demuestra que si $E \in \mathcal{M} \otimes \mathcal{N}$, entonces para todo $(x_0, y_0) \in X \times Y$

$$E_{x_0} := \{y \in Y \mid (x_0, y) \in E\} \in \mathcal{N}, \qquad E^{y_0} := \{x \in X \mid (x, y_0) \in E\} \in \mathcal{M}.$$

*b*) Demuestra que si $f: X \times Y \to \mathbb{R}$ es medible, entonces $f_{x_0}$ y $f^{y_0}$ son medibles.

*c*) Demuestra que si $g: X \to \mathbb{R}$ es medible, entonces $G: X \times Y \to \mathbb{R}$ tal que $G(x, y) = g(x)$, es medible.

3. Sea $(X, \mathcal{M}, \mu)$ un espacio de medida finita y $f : X \to [0, \infty)$:

   *a*) Demuestra que $f$ es medible si y solo si

   $$A := \{(x, y) \in X \times [0, \infty) \mid 0 \leqslant y \leqslant f(x)\} \in \mathcal{M} \otimes \mathcal{L}_{\mathbb{R}}|_{[0,\infty)}.$$

   *b*) En caso de que $f$ sea medible, demuestra que

   $$(\mu \otimes m)(A) = \int_X f d\mu = \int_{[0,\infty)} \mu(\{f > \lambda\}) dm(\lambda).$$

4. Dada una sucesión de reales positivos $\{a_k\}$ tales que $s := \sum_{k=1}^{\infty} a_k < \infty$, sean

   *a*) $s_k = \sum_{j=1}^{k} a_k$,

   *b*) $r_0 = 0$ y $r_{k+1} = r_k + 2^{-k}$,

   *c*) $A_k = [r_{k-1}, r_k] \times [r_{k-1}, r_k]$ y $B_k = [r_{k-1}, r_k] \times [r_k, r_{k+1}]$,

   *d*) $\alpha_k = s_k/m(A_k)$ y $\beta_k = s_k/m(B_k)$,

   $$f = \sum_{k=1}^{\infty} (\alpha_k \mathbb{1}_{A_k} - \beta_k \mathbb{1}_{B_k}).$$

   Demuestra que el teorema de Fubini no se puede satisfacer para esta función dado que

   $$\int_0^1 \left( \int_0^1 f dx \right) dy \neq \int_0^1 \left( \int_0^1 f dy \right) dx.$$

   ¿Cuál de las hipótesis no se cumple en este caso?

5. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y $X, Y: \Omega \to \mathbb{R}$ variables aleatorias independientes. Dados $f, g: \mathbb{R} \to \mathbb{R}$ tales que $f(X), g(Y) \in L^1(\Omega)$, demuestra que

   $$\mathbb{E}(f(X)g(Y)) = \mathbb{E}(f(X))\mathbb{E}(g(Y)).$$

## 7. Diferenciación

Dado un espacio de medida $(X, \mathcal{M}, \mu)$, podemos definir medidas en $(X, \mathcal{M})$ a partir de $\mu$ y las funciones medibles no negativas. Dada $f: X \to [0, \infty)$ medible tenemos que

$$\mu_f(E) := \int_E f d\mu = \int f \mathbb{1}_E d\mu$$

es una medida sobre $\mathcal{M}$.

No todas las medidas sobre $\mathcal{M}$ se pueden obtener por medio de esta contrucción. En particular, observamos que para cualquier conjunto $E \in \mathcal{M}$ tal que $\mu(E) = 0$ se tiene que $\mu_f(E) = 0$. Equivalentemente encontramos la siguiente definición la cual guarda la misma estructura de la definición de continuidad uniforme en un espacio métrico.

**Definición 7.1.** *Una medida $\nu: \mathcal{M} \to [0, \infty]$ se dice **absolutamente continua** con respecto a $\mu$ si para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que $\nu(E) < \varepsilon$ para todo conjunto $E \in \mathcal{M}$ tal que $\mu(E) < \delta$. Esta noción se denota por $\nu \ll \mu$.*

La primera parte del teorema de Radon-Nikodym indica que toda medida absolutamente continua sí se puede obtener a partir de la construcción que hemos descrito al principio en espacios de medida finita.

**Teorema 7.** *Sea $(X, \mathcal{M}, \mu)$ un espacio de medida finita. Para toda medida $\nu \ll \mu$, también finita, existe una función $f \in L^1(X, \mathcal{M}, \mu)$ no negativa tal que $\nu = \mu_f$.*

La función $f$ es única salvo conjuntos de medida cero. Esta se conoce como la derivada de Radon-Nikodym y se denota por

$$\frac{d\nu}{d\mu} = f.$$

No todas las medidas definidas sobre $\mathcal{M}$ son absolutamente continuas. Por ejemplo, una medida $\nu$ podría concentrar su masa en un conjunto nulo de $\mu$.

**Definición 7.2.** *Una medida $\mu: \mathcal{M} \to [0, \infty]$ se dice soportada en $E \in \mathcal{M}$ si $\mu(X \backslash E) = 0$.*

**Definición 7.3.** *Dos medidas $\mu, \nu: \mathcal{M} \to [0, \infty]$ se dicen **mutualmente singulares** si existe $E \in \mathcal{M}$ tal que $\mu$ está soportada en $E$ y $\nu$ está soportada en $X \backslash E$. Esta noción se denota por $\nu \perp \mu$.*

La segunda parte del teorema de Radon-Nikodym muestra que toda medida se puede descomponer en una parte singular y otra absolutamente continua, una vez más bajo una hipótesis de finitud.

**Teorema 8.** *Sea $(X, \mathcal{M}, \mu)$ un espacio de medida finita. Para toda medida $\nu$, también finita, existe una función $f$ medible y no-negativa tal que $(\nu - \mu_f) \perp \mu$.*

### 7.1. Esperanza condicional.

Una aplicación del teorema de Radon-Nikodym en probabilidad permite definir la **esperanza condicional**. La idea es proyectar una variable aleatoria definida en un espacio $(\Omega, \mathcal{F})$ a un espacio con menor información, es decir una filtración ($\sigma$-álgebra) más gruesa.

Consideremos un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ y una filtración $\mathcal{G} \subseteq \mathcal{F}$. Una variable aleatoria $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ define un par de medidas $\mu_\pm \colon \mathcal{G} \to [0, \infty)$ tales que

$$\mu_\pm(A) = \mathbb{E}(X_\pm \mathbb{1}_A) = \int_A X_\pm d\mathbb{P}.$$

Ambas medidas son absolutamente continuas y finitas con respecto a $\mathcal{G}$, por lo tanto existen variables aleatorias $Y_\pm \in L^1(\Omega, \mathcal{G}, \mathbb{P})$ tales que

$$\mu_\pm(A) = \mathbb{E}(Y_\pm \mathbb{1}_A).$$

El punto es que las variables aleatorias $Y_\pm$ no son necesariamente $X_\pm$ dado que $X$ es medible con respecto a $\mathcal{F}$, pero no necesariamente con respecto a $\mathcal{G}$. La variable aleatoria $Y = Y_+ - Y_-$ se conoce como la esperanza condicional de $X$ respecto de $\mathcal{G}$ y se denota por $\mathbb{E}(X|\mathcal{G}) = Y$.

**7.2. Teorema de diferenciación de Lebesgue.** El teorema de Radon-Nikodym plantea la existencia de la derivada entre dos medidas bajo una hipótesis de continuidad absoluta. Sin embargo no propone una construcción de esta densidad. El teorema de diferenciación de Lebesgue describe como calcular la derivada cuando $\mu = m$ es la medida de Lebesgue en $\mathbb{R}^d$.

Si $f \colon \mathbb{R}^d \to \mathbb{R}$ es una función continua se tiene que para todo $x \in \mathbb{R}^d$

$$f(x) = \lim_{\varepsilon \to 0^+} \frac{1}{m(B_\varepsilon)} \int_{B_\varepsilon(x)} f dm.$$

Este resultado admite la siguiente generalización en funciones integrables.

**Teorema 9.** *Sea $f \in L^1(\mathbb{R}^d)$. Para casi todo $x \in \mathbb{R}^d$ se tiene que*

$$f(x) = \lim_{\varepsilon \to 0^+} \frac{1}{m(B_\varepsilon)} \int_{B_\varepsilon(x)} f dm.$$

**7.3. Ejercicios.**

1. Sea $(X, \mathcal{M}, \mu)$ un espacio de medida, $f \colon X \to [0, \infty)$ medible, y $E \in \mathcal{M}$. Demuestra que si $\mu(E) = 0$, entonces $\int_E f d\mu = 0$.
2. Sea $(X, \mathcal{M}, \mu)$ un espacio de medida y $\nu \colon \mathcal{M} \to [0, \infty]$ una medida. Demuestra que las siguientes aseveraciones son equivalentes:
   *a*) $\nu(E) = 0$ para todo $E \in \mathcal{M}$ tal que $\mu(E) = 0$.
   *b*) Para todo $\varepsilon > 0$ existe un $\delta > 0$ tal que $\nu(E) < \varepsilon$ para todo $E \in \mathcal{M}$ tal que $\mu(E) < \delta$.
3. Sea $(X, \mathcal{M}, \mu)$ un espacio de medida finita y $f, g \colon X \to [0, \infty)$ medibles
   *a*) Demuestra que si $\mu_f = \mu_g$, entonces $f = g$ en casi todo punto.
   *b*) Da un contraejemplo que muestre que este resultado no se cumple necesariamente si el espacio no es de medida finita.
4. Demuestra que si $\nu \ll \mu$ y $\nu \perp \mu$, entonces $\nu = 0$.
5. Sea $(X, \mathcal{M}, \mu)$ un espacio de medida y $N \in \mathcal{M} \backslash \{\varnothing\}$ tal que $\mu(N) = 0$. Definimos $\nu \colon \mathcal{M} \to [0, \infty]$ tal que

$$\nu(E) = \begin{cases} \infty \text{ si } E \cap N \neq \varnothing, \\ 0 \text{ si } E \cap N = \varnothing. \end{cases}$$

   *a*) Demuestra que $\nu$ es una medida.

*b*) Demuestra que $\mu \perp \nu$.

*c*) Demuestra que no existe una función $f$ medible y no-negativa tal que $(\nu - m_f) \perp \mu$.

6. Sean $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \mathcal{F}_3$ filtraciones sobre $\Omega$, y $X \in L^1(\Omega, \mathcal{F}_3, \mathbb{P})$. Demuestra que
$$\mathbb{E}(\mathbb{E}(X|\mathcal{F}_2)|\mathcal{F}_1) = \mathbb{E}(X|\mathcal{F}_1).$$

7. Demuestra que si $f\colon \mathbb{R}^d \to \mathbb{R}$ es una función continua, entonces se tiene que para todo $x \in \mathbb{R}^d$
$$f(x) = \lim_{\varepsilon \to 0^+} \frac{1}{m(B_\varepsilon)} \int_{B_\varepsilon(x)} f dm.$$

## 8. Espacios de Lebesgue

La integral de Lebesgue permite definir una familia de seminormas. Dado $p \in [1, \infty)$ y $f \colon X \to \mathbb{R}$ medible

$$\|f\|_{L^p(X)} := \left( \int |f|^p d\mu \right)^{1/p}.$$

Mientras que para $p = \infty$

$$\|f\|_{L^\infty(X)} := \text{ínf}\{M > 0 \mid \mu(\{|f| > M\}) = 0\}.$$

El conjunto de las funciones medibles para las cuales $\|f\|_{L^p(X)} < \infty$ forman el espacio vectorial $L^p(X)$. Más aún, la función dada arriba define una seminorma para la cual $\|f\|_{L^p(X)} = 0$ implica $f = 0$ en casi todo punto.

Para poder definir un espacio normado a partir de $L^p(X)$ consideramos la relación de equivalencia $f \sim g$ cuando $f = g$ en casi todo punto. En este caso se tiene que $\|f\|_{L^p(X)} = \|g\|_{L^p(X)}$, por lo cual se tiene que la norma de una clase de equivalencia $[f]$

$$\|[f]\|_{L^p(X)} = \|f\|_{L^p(X)}$$

está bien definida.

La completitud de $L^p(X)$ se sigue del teorema de Borel-Cantelli. Los espacios normados completos se conocen como **espacios de Banach**.

Cuando $p = 2$ se tiene que $L^2(X)$ también adquiere un producto definido por

$$(f, g) := \int fg d\mu.$$

Los espacios con producto interno y completos se conocen como **espacios de Hilbert**.

Uno de los objetivos de esta unidad será presentar algunas de las desigualdades clásicas del análisis:

1. **La desigualdad de triangular**: Para $p \in [1, \infty]$ y $f, g \in L^p(X)$
$$\|f + g\|_{L^p(X)} \leqslant \|f\|_{L^p(X)} + \|f\|_{L^p(X)}.$$
2. **Hölder:** Para $p, q \in [1, \infty]$ tales que $\frac{1}{p} + \frac{1}{q} = 1$
$$\int |fg| d\mu \leqslant \|f\|_{L^p(X)} \|g\|_{L^q(X)}$$
El caso $p = q = 2$ se conoce como la desigualdad de **Cauchy-Schwarz**.
3. **Interpolación**: Dado $\theta \in [0, 1]$, $1 \leqslant p_0 \leqslant p_\theta \leqslant p_1 \leqslant \infty$ tales que
$$\frac{1}{p_\theta} = \frac{1 - \theta}{p_0} + \frac{\theta}{p_1},$$
se tiene que para toda $f \in L^{p_0}(X) \cap L^{p_1}(X)$
$$\|f\|_{L^{p_\theta}(X)} \leqslant \|f\|_{L^{p_0}(X)}^{1-\theta} \|f\|_{L^{p_1}(X)}^{\theta}$$

Cada una de de las desigualdades describe de forma cuantitativa diferentes aspectos de los espacios de Lebesgue: La desigualdad triangular es la propiedad básica de un espacio métrico, junto con la propiedad de homogeneidad establece que la norma es una función convexa. La desigualdad de interpolación permite ver que $L^{p_0}(X) \cap L^{p_1}(X) \subseteq L^p(X)$ para todo $p \in [p_0, p_1]$.

En la siguiente sección discutiremos la relación de la desigualdad de Hölder con los espacios duales.

Muchos de los resultados conocidos de la norma en $\mathbb{R}^d$ pueden ser extendidos a espacios de dimensión infinita. Sin embargo, se debe prestar especial cuidado a la noción de compacidad. Por ejemplo, un resultado que falla drásticamente en dimensión infinita es la compacidad de los conjuntos cerrados y acotados. En particular, existen sucesiones de funciones acotadas que no admiten subsucesiones convergentes.

Cuando $(X, \mathcal{M}, \mu) = (\mathbb{R}^d, \mathcal{L}_{\mathbb{R}^d}, m)$ tenemos el siguiente criterio de compacidad.

Dada $f\colon \mathbb{R}^d \to \mathbb{R}$ y $h \in \mathbb{R}^d$ denotamos $\tau_h f(x) := f(x - \tau)$.

**Teorema 10** (Fréchet–Kolmogorov). *Una sucesión $\{f_k\} \subseteq L^p(\mathbb{R}^d)$ admite una subsucesión convergente en $L^p$ si y solo si es*

1. ***Equicontinua:***
$$\lim_{|h|\to 0^+} \sup_k \int_{\mathbb{R}^d} |\tau_h f_k - f_k|^p = 0.$$
2. ***Uniformemente tensa:***
$$\lim_{R\to\infty} \sup_k \int_{\mathbb{R}^d \setminus B_R} |f_k|^p = 0.$$

## 8.1. Espacio de variación total.

Otro espacio de Banach importante en análisis y probabilidad es espacio de medidas signadas, es decir combinaciones lineales de medidas finitas sobre una dada $\sigma$-álgebra $\mathcal{M}$. Estas las podemos pensar como una generalización del espacio $L^1$.

La **variación** se define como
$$|\mu|(E) = \sup\left\{ \sum_{k=1}^{\infty} |\mu(E_k)| \mid \{E_k\} \subseteq \mathcal{M} \text{ disjuntos a pares, } \bigcup_{k=1}^{\infty} E_k = E \right\}$$
Esta construcción resulta en una medida no-negativa. La norma de **variación total** se obtiene como $\|\mu\| := |\mu|(X)$.

## 8.2. Ejercicios.

1. Dado $p \in [1, \infty)$ sea $|\cdot|_p\colon \mathbb{R}^d \to \mathbb{R}$ tal que
$$|x|_p = \left( \sum_{k=1}^{d} |x_k|^p \right)^{1/p},$$
mientras que para $p = \infty$,
$$|x|_\infty = \max_k |x_k|.$$
   a) Demuestra que $|\cdot|_p$ es 1-homogénea y convexa.
   b) Demuestra la desigualdad triangular
$$|x + y|_p \leqslant |x|_p + |y|_p.$$

2. Sea $x \in \mathbb{R}^d$
   *a*) Demuestra que $f\colon [0,1] \to \mathbb{R}$ tal que $f(1/p) = \ln |x|_p$ es una función convexa.
   *b*) Demuestra que para $\theta \in [0,1]$, $1 \leqslant p_0 \leqslant p_\theta \leqslant p_1 \leqslant \infty$ tales que
   $$\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1},$$
   se tiene que
   $$|x|_{p_\theta} \leqslant |x|_{p_0}^{1-\theta} |x|_{p_1}^{\theta}.$$
3. Demuestra que la convergencia en $L^p$ implica la convergencia en medida.
4. Sea $\{f_k\}$ una sucesión de funciones medibles de $X$ en $\mathbb{R}$ que convergen en casi todo punto a una función $f$. Demuestra que si $\sup_k |f_k| \in L^p(X)$, entonces la sucesión también converge en la norma $L^p$.
5. Sea $\{f_k\} \subseteq L^p(X)$ y $f \in L^p(X)$
   *a*) Demuestra que si $\sum_{k=1}^{\infty} \|f_k - f\|_{L^p(X)} < \infty$, entonces $\{f_k\}$ converge a $f$ en $L^p$ y también en casi todo punto.
   *b*) Demuestra que si $\{f_k\}$ converge a $f$ en $L^p$, entonces existe una subsucesión que converge a $f$ en casi todo punto.
   *c*) Demuestra que si $\sum_{k=1}^{\infty} \|f_{k+1} - f_k\|_{L^p(X)} < \infty$, entonces $\{f_k\}$ converge en $L^p$ y en casi todo punto a alguna función.
   *d*) Demuestra que $L^p(X)$ es un espacio completo.
6. Sea $(X, \mathcal{M}, \mu)$ un espacio de medida. Demuestra que para toda $f \in L^1(X, \mathcal{M}, \mu)$ se tiene que $\|\mu_f\| = \|f\|_{L^1(X)}$.
7. Demuestra que la variación define una medida no negativa y la variación total define una norma completa.
8. Sea $(X, \mathcal{M})$ un espacio medible y $\mu, \nu\colon \mathcal{M} \to [0,1]$ dos medidas de probabilidad. Demuestra que
   $$\|\mu - \nu\| = 2 \sup\{|\mu(E) - \nu(E)| \mid E \in \mathcal{M}\}.$$

# 9. Dualidad

Dualidad es una idea que aparece en diversas construcciones matemáticas. En general, esta consiste en definir un emparejamiento $(\cdot,\cdot)\colon W \times V \to \mathbb{R}$ entre dos conjuntos $V$ y $W$ dados. Por ejemplo, un emparejamiento que ha estado presente desde el principio en estas notas consiste de aquel para el cual $V = \mathcal{M}$, $W$ el conjunto de medidas definidas en $\mathcal{M}$, y $(\mu, E) = \mu(E)$.

En general, los emparejamientos permiten describir el conjunto $V$ a través de la acción de $W$. Por ejemplo, puede verse que dos elementos $v_1, v_2 \in V$ son distintos si para algún $w \in W$ se tiene que $(w, v_1) \neq (w, v_2)$.

Si $V$ es un subconjunto del espacio de funciones de $X$ en $\mathbb{R}$ podemos tomar el emparejamiento $(f, x) = f(x)$ entre $V$ y $X$. En el caso de los espacios de Lebesgue esta noción es de poca utilidad, a pesar de que estos sean espacios de funciones. Tengamos en cuenta que los elementos de $L^p(X)$ son de forma precisa clases de equivalencias de funciones. Si $\mu$ es una medida para la cual los puntos de $X$ tienen medida cero, el valor de $f(x_0)$ para un punto $x_0 \in X$ dado no está bien definido en términos de la clase de equivalencia de $f$.

Para enfatizar, la distinción de funciones en $L^p(X)$ no debe hacerse con la evaluación puntual de las misma. Una alternativa es tomar $W = \mathcal{M}$ y

$$(E, f) = \int_E f d\mu = \int \mathbb{1}_E f dm.$$

Más aún, esta construcción se generaliza si reemplazamos $\mathbb{1}_E$ por cualquier función $g$ tal que $fg \in L^1(X)$.

Sean $p, q \in [1, \infty]$ tal que $\frac{1}{p} + \frac{1}{q} = 1$. La desigualdad de Hölder establece que el emparejamiento

$$(f, g) = \int fg d\mu$$

entre $L^p(X)$ y $L^q(X)$ está bien definido.

Las siguientes definiciones ilustrarán cómo esta construcción forma parte de una idea abstracta en análisis funcional.

Dado un espacio normado $(V, \|\cdot\|)$, se tiene que una funcional lineal $w\colon V \to \mathbb{R}$ es continua si y solo si es acotada en el siguiente sentido

$$\|w\| := \sup_{v \in V \setminus \{0\}} \frac{|(w, v)|}{\|v\|} < \infty.$$

El emparejamiento en este caso está definido como $(w, v) = w(v)$. Nótese que $(\cdot,\cdot)$ es multilineal y junto con las normas se tiene la desigualdad de Cauchy-Schwarz generalizada

$$|(w, v)| \leqslant \|w\|\|v\|.$$

Dado un espacio normado $(V, \|\cdot\|)$, se define su espacio dual $(V', \|\cdot\|)$ como el espacio de las funcionales lineales y acotadas.

Gracias a la desigualdad de Hölder tenemos que cada $g \in L^q(X)$ define un funcional lineal acotado sobre $L^p(X)$. Más aún, se sigue que la norma de esta funcional lineal es justamente $\|g\|_{L^q(X)}$.

Observamos que en general, cada elemento $v \in V$ también define una funcional lineal acotada sobre $V'$ dado por $w \in V' \mapsto (w, v)$. Más aún

$$\|v\| = \sup_{w \in V' \setminus \{0\}} \frac{|(w, v)|}{\|w\|}.$$

La desigualdad $\geqslant$ se sigue de la desigualdad de Cauchy-Schwarz. La desigualdad opuesta suele demostrarse usando el teorema de Hahn-Banach.

Es posible que $V$ no contenga todos los funcionales lineales acotados sobre $V'$. Sin embargo, cuando esto sucede se dice que $V$ es **reflexivo**. El siguiente teorema indica que los espacios $L^p(X)$ son reflexivos para $p \in (1, \infty)$.

**Teorema 11.** *Para $p, q \in [1, \infty)$ tales que $\frac{1}{p} + \frac{1}{q} = 1$ se tiene que $L^q(X)$ es el espacio dual de $L^p(X)$.*

### 9.1. Topología débil.

La dualidad define dos topologías importantes.

**Definición 9.1.** *Una sucesión $\{v_k\} \subseteq V$ converge en el sentido débil a $v \in V$ si para todo $w \in V'$ se tiene que la sucesión numérica $(w, v_k) \to (w, v)$. Se denota por $v_k \rightharpoonup v$.*

*Por otro lado, una sucesión $\{w_k\} \subseteq V$ converge en el sentido débil-estrella a $w \in V'$ si para todo $v \in V$ se tiene que la sucesión numérica $(w_k, v) \to (w, v)$. Se denota por $v_k \overset{*}{\rightharpoonup} v$.*

Dado que $v_k \rightharpoonup v$, se sigue del teorema de cota uniforme que la sucesión $\{\|v_k\|\}$ es acotada. A partir de acá se sigue la continuidad inferior de la norma con respecto a la convergencia débil

$$\|v\| \leqslant \liminf_{k \to \infty} \|v_k\|.$$

Los mismos resultados también se cumplen para la convergencia débil-estrella.

Una de las bondades de la convergencia débil es que permite recuperar la compacidad de una sucesión acotada.

**Teorema 12.** *Sea $(V, \|\cdot\|)$ separable. Toda sucesión $\{w_k\} \subseteq V'$ acotada, tiene una subsucesión que converge en el sentido débil-estrella.*

En el caso de los espacios de Lebesgue se tiene que para $p \in [1, \infty)$, los espacios $L^p(\mathbb{R}^d)$ son separables. Por lo tanto, para $q \in (1, \infty]$ y toda sucesión $\{f_k\} \subseteq L^q(\mathbb{R}^d)$ tal que $\sup_k \|f_k\|_{L^q(\mathbb{R}^d)} < \infty$, existe un límite $f \in L^q(\mathbb{R}^d)$ y una subsucesión $\{f_{k_j}\}$ tal que para todo $g \in L^p(\mathbb{R}^d)$ se tiene que

$$\lim_{j \to \infty} \int f_{k_j} g dm = \int f g dm.$$

### 9.2. Teorema de Riesz.

Otro ejemplo de dualidad ocurre en el espacio de las funciones continuas en $\mathbb{R}^d$.

Dado $K \subseteq \mathbb{R}^d$ compacto, sea $C(K)$ el espacio de las funciones continuas. Este espacio es de Banach con la norma uniforme. Cualquier $f \in L^1(K)$ define una funcional lineal acotada sobre $C(K)$ tal que

$$\varphi \in C(K) \mapsto \int_K \varphi f dm.$$

Sin embargo, el dual de $C(K)$ admite otros funcionales lineales.

**Definición 9.2.** *Una medida $\mu\colon \mathcal{B}_{\mathbb{R}^d}|_K \to \mathbb{R}$ es Borel regular si para todo $E \in \mathcal{B}_{\mathbb{R}^d}|_K$ se tiene que*

$$\mu(E) = \inf\{\mu(U) \mid E \subseteq U \text{ abierto}\} = \sup\{\mu(F) \mid F \subseteq E \text{ compacto}\}.$$

**Teorema 13.** *Dado $K \subseteq \mathbb{R}^d$ compacto, el dual de $(C(K), \|\cdot\|_{L^\infty(\mathbb{R}^d)})$ consiste de las medidas regulares de variación acotada con la norma de variación total.*

Dado que $C(K)$ es separable se tiene que para toda sucesión $\{\mu_k\}$ de medidas regulares tales que $\sup_k |\mu_k|(K) < \infty$, existe un límite $\mu$ y una subsucesión $\{\mu_{k_j}\}$ tal que para toda función $\varphi \in C(K)$

$$\lim_{j\to\infty} \int \varphi d\mu_{k_j} = \int \varphi d\mu.$$

En el caso de una sucesión $\{f_k\} \subseteq L^1(K)$ tal que $\sup_k \|f_k\|_{L^1(K)} < \infty$ obtenemos la existencia del límite, pero si admitimos que este sea una medida.

### 9.3. Ejercicios.

1. Sean $p, q \in [1,\infty]$ tales que $\frac{1}{p} + \frac{1}{q} = 1$
   *a*) Demuestra la desigualdad de Young: Para $x, y \in \mathbb{R}$ y $p, q \in (1,\infty)$
   $$|xy| \leqslant \frac{|x|^p}{p} + \frac{|y|^q}{q}.$$
   *b*) Demuestra la desigualdad de Hölder: Para $x, y \in \mathbb{R}^d$ y $p, q \in [1,\infty]$
   $$x \cdot y \leqslant |x|_p |y|_q$$
   Pista: Considera el caso $|x|_p = |y|_q = 1$.
   *c*) Determina bajo cuales condiciones se cumple la igualdad en la desigualdad de Hölder.
2. Sea $(V, \|\cdot\|)$ un espacio normado y $w\colon V \to \mathbb{R}$ lineal, demuestra que $w$ es continua si y solo si $\|w\| < \infty$.
3. Demuestra que la norma dual efectivamente define una norma.
4. Demuestra que la norma es débilmente semicontinua por abajo
5. Da una sucesión de funciones $\{f_k\} \in L^1([0,1])$ tal que $\sup_k \|f_k\|_{L^1([0,1])} < \infty$, pero no exista un límite $f \in L^1([0,1])$ ni ninguna subsucesión $\{f_{k_j}\}$ tal que para todo $g \in L^\infty([0,1])$ se tenga que
   $$\lim_{j\to\infty} \int_0^1 f_{k_j} g dm = \int_0^1 f g dm.$$

Departamento de Matemáticas, CIMAT, Guanajuato, México