Capítulo 2

Espacios de Medida

2.1. Funciones Aditivas de Conjunto

Definición 2.1 Sea $\mu : \mathcal{C} \to \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ una función definida sobre una colección de conjuntos $\mathcal{C}$. Decimos que μ es *finitamente aditiva* si

1. $\mu(\emptyset) = 0$.
2. Para toda colección finita $(A_n)_{1 \leq n \leq m}$ de conjuntos de $\mathcal{C}$, disjuntos dos a dos, tales que $\cup_{i=1}^m A_i \in \mathcal{C}$ se tiene
$$\mu(\cup_{n=1}^m A_n) = \sum_{n=1}^m \mu(A_n). \tag{2.1}$$

Observación 2.1 Todas las sumas en (2.1) deben estar bien definidas, de modo que no puede ocurrir que $\mu(A_i) = -\infty$ y $\mu(A_j) = \infty$ para algunos índices i, j. Además hay que observar que no estamos suponiendo que $\mathcal{C}$ sea cerrada bajo uniones finitas, aunque el dominio natural para una función de este tipo es un álgebra.

Por otro lado la condición 1 es prácticamente redundante, pues basta con que haya un conjunto $A \in \mathcal{C}$ con $\mu(A)$ finita para que sea cierta:
$$\mu(A) = \mu(A \cup \emptyset) = \mu(A) + \mu(\emptyset) \Rightarrow \mu(\emptyset) = 0.$$

Ejemplos 2.1

1. Sea Ω un conjunto cualquiera y sea $\mathcal{C} = \mathcal{P}(\Omega)$. Definimos $\mu(A)$ como el número de puntos de A. Si A tiene n elementos, $n = 0, 1, 2, \dots$, entonces $\mu(A) = n$; si A es un conjunto con infinitos elementos, $\mu(A) = \infty$.
2. Sea Ω un conjunto cualquiera con al menos dos puntos, x, y y sea $\mathcal{C} = \mathcal{P}(\Omega)$. Definimos
$$\mu(A) = \begin{cases} 0 & \text{si } x \notin A, y \notin A \text{ o } x \in A, y \in A \\ 1 & \text{si } x \in A, y \notin A, \\ -1 & \text{si } x \notin A, y \in A, \end{cases}$$
Esta función es aditiva en $\mathcal{C}$.
3. Sea $\Omega = (0, 1]$, $\mathcal{A}$ el álgebra que definimos en el ejemplo 1.1.4, es decir, los elementos de $\mathcal{A}$ son de la forma $\cup_{i=1}^n (a_i, b_i]$, donde los intervalos son disjuntos. Definimos
$$\mu(\cup_{i=1}^n (a_i, b_i]) = \sum_{i=1}^n (b_i - a_i)$$
Es fácil verificar que μ es finitamente aditiva sobre $\mathcal{A}$.

4. Sea $\Omega = (0, 1]$, $\mathcal{C}$ la clase de los intervalos semiabiertos de la forma $(a, b]$ con $0 \leq a \leq b \leq 1$. Definimos μ por

$$\mu(a, b] = b - a \text{ si } a \neq 0,$$
$$\mu(0, b] = \infty.$$

Teorema 2.1 *Sea μ una función aditiva sobre el álgebra $\mathcal{A}$.*

(a) $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$ para todo $A, B \in \mathcal{A}$.

(b) Si $A, B \in \mathcal{F}$ y $B \subset A$, entonces $\mu(A) = \mu(B) + \mu(A - B)$. Como consecuencia, si $\mu(B)$ es finita, $\mu(A - B) = \mu(A) - \mu(B)$ y si $\mu(A - B) \geq 0$, $\mu(B) \leq \mu(A)$.

(c) Si μ es no-negativa, $\mu(\cup_{i=1}^n A_i) \leq \sum_{i=1}^n \mu(A_i)$.

Demostración. (a) Por aditividad tenemos

$$\mu(A) = \mu(A \cap B) + \mu(A - B),$$
$$\mu(B) = \mu(A \cap B) + \mu(B - A).$$

Sumando las dos ecuaciones anteriores obtenemos

$$\begin{aligned}\mu(A) + \mu(B) &= \mu(A \cap B) + \big[\mu(A - B) + \mu(B - A) + \mu(A \cap B)\big] \\ &= \mu(A \cap B) + \mu(A \cup B).\end{aligned}$$

(b) Podemos escribir $A = B \cup (A - B)$, y entonces $\mu(A) = \mu(B) + \mu(A - B)$.
(c) Tenemos que

$$\cup_{i=1}^n A_i = A_1 \cup (A_1^c \cap A_2) \cup (A_1^c \cap A_2^c \cap A_3) \cup \cdots \cup (A_1^c \cap \cdots \cap A_{n-1}^c \cap A_n).$$

Los conjuntos que aparecen al lado derecho de esta relación son disjuntos y por lo tanto

$$\mu(\cup_{i=1}^n A_i) = \mu(A_1) + \mu(A_1^c \cap A_2) + \mu(A_1^c \cap A_2^c \cap A_3) + \cdots + \mu(A_1^c \cap \cdots \cap A_{n-1}^c \cap A_n).$$

Por otro lado tenemos que $A_1^c \cap \cdots \cap A_{n-1}^c \cap A_n \subset A_n$ y como μ es no-negativa el aparte (b) implica el resultado. ■

2.2. Medidas

Definición 2.2 Sea $\mu : \mathcal{F} \to \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ una función definida sobre una colección de conjuntos $\mathcal{F}$. Decimos que μ es *σ-aditiva* o *numerablemente aditiva* si

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

2. Para toda colección $(A_n)_{n \geq 1}$ de conjuntos de $\mathcal{F}$, disjuntos dos a dos, tales que $\cup_{i=1}^\infty A_i \in \mathcal{F}$ se tiene

$$\mu(\cup_{n \geq 1} A_n) = \sum_{n \geq 1} \mu(A_n). \tag{2.2}$$

Las observaciones 2.1 también son válidas en este caso, con las modificaciones obvias.

Definición 2.3 Si $\mathcal{F}$ es una σ-álgebra de subconjuntos del conjunto Ω, decimos que $(\Omega, \mathcal{F})$ es un *espacio medible*. Una *medida* μ sobre $\mathcal{F}$ es una función σ-aditiva que toma únicamente valores positivos: $\mu : \mathcal{F} \to [0, \infty]$.

Si $\mu(\Omega) < \infty$ decimos que la medida es *finita*. Si $\mu(\Omega) = 1$ decimos que μ es una *(medida de) probabilidad*. Si el conjunto Ω se puede descomponer como la unión numerable de subconjuntos de medida finita, decimos que la medida es *σ-finita*.

Finalmente, decimos que $(\Omega, \mathcal{F}, \mu)$ es un *espacio de medida* y es un *espacio de probabilidad* en el caso particular $\mu(\Omega) = 1$. Los conjuntos de $\mathcal{F}$ se llaman conjuntos *medibles* (o *eventos* en el caso de espacios de probabilidad).

Ejemplos 2.2

1. La función definida en el ejemplo 2.1.1 es una medida, conocida como la medida de contar. Esta medida es σ-finita si y sólo si Ω es numerable.

2. Otra medida similar se define de la siguiente manera: Sea $\Omega = \{x_1, x_2, \dots\}$ un conjunto finito o numerable y sea $p_1, p_2, \dots$ una sucesión de números no-negativos. De nuevo, sea $\mathcal{F} = \mathcal{P}(\Omega)$ y definimos
$$\mu(A) = \sum_{i : x_i \in A} p_i.$$
Con esta definición μ es una medida sobre $\mathcal{F}$ y $\mu\{x_i\} = p_i$, $i = 1, 2, \dots$. μ es una probabilidad sii $\sum_i p_i = 1$. Si todos los p_i valen 1, μ es la medida del ejemplo anterior.

3. La función definida en el ejemplo 2.1.4 es aditiva pero no σ-aditiva.

Dada una función $\mu : \mathcal{A} \to \overline{\mathbb{R}}$, donde $\mathcal{A}$ es un álgebra usualmente es sencillo verificar si es aditiva ya que basta con verificar (2.1) para $n = 2$. Para poder verificar si además es σ-aditiva es útil tener una caracterización de la σ-aditividad en términos de la continuidad de μ respecto a sucesiones monótonas.

Teorema 2.2 *Sea μ una función σ-aditiva definida sobre la σ-álgebra $\mathcal{F}$.*

(a) Si $A_n \in \mathcal{F}$ para $n \geq 1$ y $A_n \uparrow A$, entonces $\mu(A_n) \to \mu(A)$ cuando $n \to \infty$.

(b) Si $A_n \in \mathcal{F}$ para $n \geq 1$, $A_n \downarrow A$ y $\mu(A_1)$ es finita, entonces $\mu(A_n) \to \mu(A)$ cuando $n \to \infty$.

Demostración. (a) Si $\mu(A_n) = \infty$ para algún n, entonces $\mu(A) = \mu(A_n) + \mu(A - A_n) = \infty + \mu(A - A_n) = \infty$. De manera similar, reemplazando A por A_k vemos que $\mu(A_k) = \infty$ para todo $k \geq n$ y esto demuestra el resultado en este caso. De igual manera se puede tratar el caso $\mu(A_n) = -\infty$ para todo n. Por lo tanto podemos suponer que todos los $\mu(A_n)$ son finitos.

Como los conjuntos A_n forman una sucesión creciente podemos escribir
$$A = A_1 \cup (A_2 - A_1) \cup (A_3 - A_2) \cup \cdots \cup (A_n - A_{n-1}) \cup \cdots$$
donde los conjuntos que aparecen a la derecha son disjuntos. Por el teorema 2.1 (b) tenemos
$$\begin{aligned}
\mu(A) &= \mu(A_1) + \sum_{n=2}^{\infty} \mu(A_n - A_{n-1}) \\
&= \mu(A_1) + \lim_{m \to \infty} \sum_{n=2}^{m} \mu(A_n - A_{n-1}) \\
&= \mu(A_1) + \lim_{m \to \infty} \big[\mu(A_2) - \mu(A_1) + \mu(A_3) - \mu(A_2) + \cdots + \mu(A_m) - \mu(A_{m-1})\big] \\
&= \lim_{m \to \infty} \mu(A_m).
\end{aligned}$$

(b) Si $A_n \downarrow A$, entonces $A_1 - A_n \uparrow A_1 - A$, de modo que $\mu(A_1 - A_n) \to \mu(A_1 - A)$ por (a). El resultado sigue ahora del teorema 2.1 (b). ■

Definición 2.4 Sea μ una función aditiva definida sobre una colección de conjuntos $\mathcal{C}$. Decimos que μ es *continua por debajo* en $A \in \mathcal{C}$ si $A_n \in \mathcal{C}$ para $n \geq 1$, y $A_n \uparrow A$, implican que $\mu(A_n) \to \mu(A)$ cuando $n \to \infty$. Decimos que μ es *continua por arriba* en el conjunto A si $A_n \in \mathcal{C}$ para $n \geq 1$, $A_n \downarrow A$ implica que $\mu(A_n) \to \mu(A)$ cuando $n \to \infty$.

Teorema 2.3 *Sea μ una función aditiva definida sobre el álgebra $\mathcal{A}$.*

(a) Supongamos que μ es continua por debajo en todo $A \in \mathcal{A}$, entonces μ es σ-aditiva en $\mathcal{A}$.

(b) Supongamos que μ es continua por arriba en el conjunto vacío, entonces μ es σ-aditiva en $\mathcal{A}$

Demostración. (a) Sean $A_1, A_2, \dots$ conjuntos disjuntos en $\mathcal{A}$ cuya unión está en $\mathcal{A}$. Si $B_n = \cup_{i=1}^n A_i$ entonces $B_n \uparrow A$, y en consecuencia $\mu(B_n) \to \mu(A)$, por hipótesis. Pero por aditividad, $\mu(B_n) = \sum_{i=1}^n \mu(A_i)$, y en consecuencia $\mu(A) = \lim_{n\to\infty} \sum_{i=1}^n \mu(A_i)$.

(b) Sean $A_1, A_2, \dots$ conjuntos disjuntos en $\mathcal{A}$ cuya unión está en $\mathcal{A}$ y sea $B_n = \cup_{i=1}^n A_i$. Por el teorema 2.1 (b) $\mu(A) = \mu(B_n) + \mu(A - B_n)$, pero $A - B_n \downarrow \emptyset$, y por hipótesis $\mu(A - B_n) \to 0$. Por lo tanto $\mu(B_n) \to \mu(A)$ y el resultado sigue igual que en el aparte anterior. ■

Corolario 2.1 *Sea $\mathcal{F}$ una σ-álgebra y sea $P : \mathcal{F} \to [0,1]$ que satisface $P(\Omega) = 1$ y es aditiva. Entonces las siguientes proposiciones son equivalentes:*

1. *P es una probabilidad.*
2. *$A_n \in \mathcal{F},\ A_n \downarrow \emptyset \Rightarrow P(A_n) \downarrow 0$.*
3. *$A_n \in \mathcal{F},\ A_n \downarrow A \Rightarrow P(A_n) \downarrow P(A)$.*
4. *$A_n \in \mathcal{F},\ A_n \uparrow \Omega \Rightarrow P(A_n) \uparrow 1$.*
5. *$A_n \in \mathcal{F},\ A_n \uparrow A \Rightarrow P(A_n) \uparrow P(A)$.*

Demostración. Ejercicio.

Teorema 2.4 *Sea $(\Omega, \mathcal{F}, \mu)$ un espacio de medida finita.*

(a) Para toda sucesión $A_n \in \mathcal{F}$, $n \geq 1$ se tiene que

$$\mu(\liminf_{n\to\infty} A_n) \leq \liminf_{n\to\infty} \mu(A_n) \leq \limsup_{n\to\infty} \mu(A_n) \leq \mu(\limsup_{n\to\infty} A_n). \tag{2.3}$$

(b) Si $A_n \to A$ entonces $\mu(A_n) \to \mu(A)$.

Demostración. (b) es consecuencia de (a) ya que si $A_n \to A$ entonces

$$\liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n = A,$$

y por lo tanto los extremos de (2.3) son ambos iguales a $\mu(A)$.

Para demostrar (a) sea $B_n = \cap_{k=n}^\infty A_k$, $C_n = \cup_{k=n}^\infty A_k$, entonces $B_n \uparrow \liminf A_n$ y $C_n \downarrow \limsup A_n$. Por los teoremas 2.2 y 2.3 tenemos

$$\mu(A_n) \geq \mu(B_n) \to \mu(\liminf A_n)$$
$$\mu(A_n) \leq \mu(C_n) \to \mu(\limsup A_n)$$

y haciendo $n \to \infty$ obtenemos (2.3). ■

Otras propiedades para medidas finitas.

1. $\mu(A^c) = \mu(\Omega) - \mu(A)$.
2. El principio de inclusión-exclusión: Si $A_1, \dots, A_n$ son conjuntos medibles entonces

$$\begin{aligned}\mu(\cup_{i=1}^n A_i) = &\sum_{i=1}^n \mu(A_i) - \sum_{1\le i<j\le n} \mu(A_i \cap A_j)\\ &+ \sum_{1\le i<j<k\le n} \mu(A_i \cap A_j \cap A_k)\\ &\vdots\\ &+ (-1)^{n+1}\mu(A_1 \cap \cdots \cap A_n).\end{aligned}$$

   El caso $n = 2$ corresponde al teorema 2.1(a). Esta fórmula se puede demostrar por inducción. Como los términos del lado derecho alternan en signo, es posible obtener desigualdades cuando truncamos las sumas. Estas desigualdades se conocen como desigualdades de Bonferroni. Por ejemplo,

$$\mu(\cup_{i=1}^n A_i) \le \sum_{i=1}^n \mu(A_i),$$
$$\mu(\cup_{i=1}^n A_i) \ge \sum_{i=1}^n \mu(A_i) - \sum_{1\le i<j\le n} \mu(A_i \cap A_j).$$

3. Subaditividad numerable: Si $A_n, n \ge 1$ son conjuntos medibles

$$\mu(\cup_{n\ge1} A_n) \le \sum_{n\ge1} \mu(A_n)$$

   La demostración es similar a la del teorema 2.1(c).

2.3. El Teorema de Dynkin

Definición 2.5 Una colección $\mathcal{C}$ de subconjuntos de Ω es un *sistema* π si es cerrada bajo intersecciones:

$$A, B \in \mathcal{C} \quad \Rightarrow \quad A \cap B \in \mathcal{C}.$$

Una colección $\mathcal{L}$ de subconjuntos de Ω es un *sistema* λ o una *clase de Dynkin* si satisface las siguientes condiciones:

(a) $\Omega \in \mathcal{L}$.

(b) $A \in \mathcal{L} \Rightarrow A^c \in \mathcal{L}$.

(c) $\mathcal{L}$ es cerrada bajo uniones disjuntas: Si $A_n \in \mathcal{L}$ para $n \ge 1$, $A_i \cap A_j = \emptyset$ cuando $i \ne j$, entonces $\cup_{n\ge1} A_n \in \mathcal{L}$.

Alternativamente, las condiciones (b) y (c) pueden reemplazarse por las siguientes:

(b') $A, B \in \mathcal{L},\ A \subset B \Rightarrow B - A \in \mathcal{L}$.

(c') $\mathcal{L}$ es cerrada bajo límites crecientes: Si $A_n \in \mathcal{L}$ para $n \ge 1$, $A_n \uparrow A$ cuando $n \to \infty$, entonces $\cup_{n\ge1} A_n = A \in \mathcal{L}$.

Una colección $\mathcal{M}$ de subconjuntos de Ω es una *clase monótona* si es cerrada bajo límites monótonos: Si $A_n \in \mathcal{M}$ para $n \geq 1$, $A_n \uparrow A$ o $A_n \downarrow A$ cuando $n \to \infty$, entonces $A \in \mathcal{M}$.

Una clase que es a la vez un sistema π y un sistema λ es un álgebra porque es cerrada bajo complementos e intersecciones. Pero un álgebra que satisface la condición (c') es una σ-álgebra. En consecuencia, una clase que es a la vez un sistema π y un sistema λ es una σ-álgebra.

Hemos definido varias colecciones de conjuntos: álgebra, σ-álgebra, sistema π, sistema λ y clase monótona. Vamos a llamar *estructura* a cualquiera de estas clases de conjuntos. Tenemos la siguiente definición, que generaliza la de σ-álgebra generada. Fijemos un tipo de estructura y llamémosla $\mathcal{S}$.

Definición 2.6 Sea $\mathcal{C}$ una colección de subconjuntos de Ω. La menor estructura $\mathcal{S}$ generada por $\mathcal{C}$ es la estructura que satisface las siguientes condiciones:

(a) $\mathcal{C} \subset \mathcal{S}$.

(b) Si $\mathcal{S}'$ es otra estructura que contiene a $\mathcal{C}$ entonces $\mathcal{S} \subset \mathcal{S}'$.

Al igual que en el caso de la σ-álgebra generada, es posible ver que la estructura generada por $\mathcal{C}$ es única. La demostración es similar.

Teorema 2.5 (Dynkin) *(a) Si $\mathcal{C}$ es un sistema π y $\mathcal{L}$ es un sistema λ tal que $\mathcal{C} \subset \mathcal{L}$ entonces $\sigma(\mathcal{C}) \subset \mathcal{L}$.*
(b) Si $\mathcal{C}$ es un sistema π

$$\sigma(\mathcal{C}) = \mathcal{L}(\mathcal{C}),$$

es decir, la σ-álgebra generada por $\mathcal{C}$ coincide con el sistema λ generado por $\mathcal{C}$.

Demostración. (a) Sea $\mathcal{L}(\mathcal{C})$ el sistema λ generado por $\mathcal{C}$, es decir, la intersección de todos los sistemas λ que contienen a $\mathcal{C}$. Si $\mathcal{L}(\mathcal{C})$ fuese un sistema π entonces sería una σ-álgebra, y en consecuencia $\sigma(\mathcal{C}) \subset \mathcal{L}(\mathcal{C}) \subset \mathcal{L}$. Por lo tanto basta mostrar que $\mathcal{L}(\mathcal{C})$ es un sistema π.

Para cada conjunto A definimos $\mathcal{G}_A$ como la clase de los conjuntos B tales que $A \cap B \in \mathcal{L}(\mathcal{C})$. Si $A \in \mathcal{C}$ entonces $\mathcal{G}_A$ es un sistema λ:

(i) $\Omega \in \mathcal{G}_A$ porque $\Omega \cap A = A \in \mathcal{C} \subset \mathcal{L}(\mathcal{C})$.

(ii) Supongamos que $B \subset C,\ B, C \in \mathcal{G}_A$, es decir, $B \cap A \in \mathcal{L}(\mathcal{C})$, $C \cap A \in \mathcal{L}(\mathcal{C})$. Entonces $(C - B) \cap A = (C \cap A) - (B \cap A)$ y como esto es una diferencia propia, está en $\mathcal{L}(\mathcal{C})$. Por lo tanto $C - B \in \mathcal{G}_A$.

(iii) Sea $B_n \uparrow B,\ B_n \in \mathcal{G}_A$, entonces $B_n \cap A \in \mathcal{L}(\mathcal{C}),\ B_n \cap A \uparrow B \cap A$ y como $\mathcal{L}(\mathcal{C})$ es cerrada bajo límites crecientes, $B \cap A \in \mathcal{L}(\mathcal{C})$ y $B \in \mathcal{G}_A$.

Veamos ahora que $\mathcal{L}(\mathcal{C})$ es un sistema π. Si $A \in \mathcal{C}$, $B \in \mathcal{C}$ entonces $A \cap B \in \mathcal{C} \subset \mathcal{L}(\mathcal{C})$. Por lo tanto si $A \in \mathcal{C}$, tenemos que $\mathcal{C} \subset \mathcal{G}_A$ y como $\mathcal{G}_A$ es un sistema λ, $\mathcal{L}(\mathcal{C}) \subset \mathcal{G}_A$. Por lo tanto $A \in \mathcal{C} \Rightarrow \mathcal{L}(\mathcal{C}) \subset \mathcal{G}_A$, es decir,

$$A \in \mathcal{C},\ B \in \mathcal{L}(\mathcal{C}) \Rightarrow A \cap B \in \mathcal{L}(C).$$

Pero esto quiere decir que si $B \in \mathcal{L}(\mathcal{C})$ entonces $\mathcal{C} \subset \mathcal{G}_B$, y como $\mathcal{G}_B$ es un sistema λ, tenemos que $\mathcal{L}(\mathcal{C}) \subset \mathcal{G}_B$. En consecuencia

$$B \in \mathcal{L}(\mathcal{C}),\ C \in \mathcal{L}(\mathcal{C}) \Rightarrow B \cap C \in \mathcal{L}(C)$$

y $\mathcal{L}(\mathcal{C})$ es un sistema π.

(b) Es consecuencia de (a). Como $\mathcal{C} \subset \mathcal{L}(\mathcal{C})$ tenemos que $\sigma(\mathcal{C}) \subset \mathcal{L}(\mathcal{C})$. Por otro lado $\sigma(\mathcal{C})$ es una σ-álgebra y por lo tanto un sistema λ que contiene a $\mathcal{C}$. En consecuencia $\mathcal{L}(\mathcal{C}) \subset \sigma(\mathcal{C})$. ■

Hay un teorema similar al teorema de Dynkin, que históricamente es anterior, conocido como el Teorema de la Clase Monótona, que enunciamos a continuación. Su demostración es similar a la del teorema de Dynkin y la omitiremos.

Teorema 2.6 (Clase Monótona) *Sea $\mathcal{A}$ un álgebra de subconjuntos de Ω y sea $\mathcal{M}$ una clase monótona de subconjuntos de Ω. Si $\mathcal{A} \subset \mathcal{M}$ entonces $\sigma(\mathcal{A}) \subset \mathcal{M}$.*

Para dar una idea de cómo se usan estos dos resultados presentamos el siguiente 'Principio General':

(a) Supongamos que cierta propiedad vale para un sistema de Dynkin $\mathcal{L}$ de subconjuntos de Ω. Si $\mathcal{A}$ es un álgebra que genera a la σ-álgebra $\mathcal{F}$ y $A \subset \mathcal{L}$, entonces $\mathcal{F} \subset \mathcal{L}$, es decir, todos los conjuntos de $\mathcal{F}$ tienen la propiedad.

(b) Supongamos que cierta propiedad vale para una clase monótona $\mathcal{M}$ de subconjuntos de Ω. Si $\mathcal{A}$ es un álgebra que genera a la σ-álgebra $\mathcal{F}$ y $A \subset \mathcal{M}$, entonces $\mathcal{F} \subset \mathcal{M}$, es decir, todos los conjuntos de $\mathcal{F}$ tienen la propiedad.

Para ver por qué funciona este principio sea

$$\mathcal{E} = \{E : \text{ la propiedad vale para } E\};$$

entonces para (a), por hipótesis y el teorema de Dynkin tenemos que

$$\mathcal{F} = \sigma(\mathcal{A}) = \mathcal{L}(\mathcal{A}) \subset \mathcal{E}.$$

De manera similar, para (b), por el teorema de la clase monótona,

$$\mathcal{F} = \sigma(\mathcal{A}) = \mathcal{M}(\mathcal{A}) \subset \mathcal{E}$$

■

Proposición 2.1 *Sean P_1 y P_2 dos probabilidades sobre $(\Omega, \mathcal{F})$. La clase*

$$\mathcal{L} = \{A \in \mathcal{F} : P_1(A) = P_2(A)\}$$

es un sistema λ.

Demostración.

1. $\Omega \in \mathcal{L}$ porque $P_1(\Omega) = 1 = P_2(\Omega)$.
2. Supongamos $A, B \in \mathcal{L},\ A \subset B$ entonces
$$P_1(B - A) = P_1(B) - P_1(A) = P_2(B) - P_2(A) = P_2(B - A)$$
y $B - A \in \mathcal{L}$.
3. Supongamos $A_n \in \mathcal{L},\ n \geq 1,\ A_n \uparrow A$, entonces
$$P_1(A) = \lim_{n\to\infty} P_1(A_n) = \lim_{n\to\infty} P_2(A_n) = P_2(A)$$
de modo que $A \in \mathcal{L}$. ■

Corolario 2.2 *Si P_1 y P_2 son dos probabilidades sobre $(\Omega, \mathcal{F})$ y $\mathcal{C}$ es un sistema π tal que para todo $A \in \mathcal{C},\ P_1(A) = P_2(A)$, entonces P_1 y P_2 coinciden en $\sigma(\mathcal{C})$.*

Demostración.

$$\mathcal{L} = \{A \in \mathcal{F} : P_1(A) = P_2(A)\}$$

es un sistema λ y contiene a $\mathcal{C}$. Por el teorema de Dynkin, $\sigma(\mathcal{C}) \subset \mathcal{L}$. ■

2.4. El Teorema de Extensión

Nuestro objetivo ahora es demostrar el teorema de extensión que permite pasar de una función μ positiva y σ-aditiva definida sobre el álgebra $\mathcal{A}$ a una medida sobre la σ-álgebra $\mathcal{F} = \sigma(\mathcal{A})$ generada por $\mathcal{A}$. Comenzamos con algunas definiciones.

Definición 2.7 Dada una función μ positiva y σ-aditiva definida sobre un álgebra $\mathcal{A}$ y dado cualquier conjunto $E \subset \Omega$ definimos la medida exterior μ^* de E por

$$\mu^*(E) = \text{ínf}\{\sum_{n\geq 1} \mu(A_n) : A_n \in \mathcal{A},\ E \subset \cup_{n\geq 1} A_n\}$$

con la convención usual de que el ínfimo de un conjunto vacío es $+\infty$.

Veamos algunas propiedades de la medida exterior.

Lema 2.1 *Si $E, E_n, n \geq 1$ son subconjuntos de Ω y $E \subset \cup_{n\geq 1} E_n$ entonces $\mu^*(E) \leq \sum_n \mu^*(E_n)$.*

Demostración. Si la suma vale ∞, el resultado es cierto. En caso contrario, dado $\varepsilon > 0$ para cada n tomamos $A_{nm} \in \mathcal{A}$ tal que $E_n \subset \cup_m A_{nm}$ y $\sum_m \mu(A_{nm}) \leq \mu^*(E_n) + \varepsilon 2^{-n}$. Entonces $E \subset \cup_n E_n \subset \cup_n \cup_m A_{nm}$ y por lo tanto,

$$\mu^*(E) \leq \sum_{n\geq 1} \mu^*(E_n) + \frac{\varepsilon}{2^n} \leq \varepsilon + \sum_{n\geq 1} \mu^*(E_n).$$

Haciendo $\varepsilon \downarrow 0$ obtenemos el resultado. ■

Lema 2.2 *Para todo $A \in \mathcal{A}$ se tiene $\mu^*(A) = \mu(A)$.*

Demostración. Si $A \subset \cup_n A_n,\ A_n \in \mathcal{A}$ entonces por monotonía de μ, $\mu(A) \leq \mu(\cup_n A_n)$ y por subaditividad $\mu(\cup_n A_n) \leq \sum_n \mu(A_n)$ de modo que $\mu(A) \leq \mu^*(A)$. Recíprocamente, tomando $A_1 = A,\ A_n = \emptyset$ para $n > 1$ obtenemos $\mu^*(A) \leq \mu(A)$. En consecuencia $\mu^*(A) = \mu(A)$. ■

Definición 2.8 Un conjunto $E \subset \Omega$ es medible respecto a μ^* o μ^*-medible si para todo $F \subset \Omega$ se tiene

$$\mu^*(F) = \mu^*(F \cap E) + \mu^*(F \cap E^c) \tag{2.4}$$

Usaremos la notación $\mathcal{M}(\mu^*)$ para la clase de los conjuntos medibles respecto a μ^*.

Es importante resaltar que el concepto de medibilidad depende de la medida exterior μ^*. Un conjunto E puede ser medible respecto a μ_1^* y no serlo respecto a μ_2^*. La ecuación (2.4) dice que para cualquier $F \subset \Omega$, el conjunto E divide a F en dos partes, $F \cap E$ y $F \cap E^c$, tales que la suma de sus medidas exteriores es la medida exterior de F. Por lo tanto la medibilidad de E depende de lo que el conjunto E hace a todos los subconjuntos de Ω.

Observamos que por la subaditividad siempre se tiene que

$$\mu^*(F) \leq \mu^*(F \cap E) + \mu^*(F \cap E^c)$$

para cualesquiera F, E. En consecuencia, E es medible si y sólo si

$$\mu^*(F) \geq \mu^*(F \cap E) + \mu^*(F \cap E^c). \tag{2.5}$$

Como esta desigualdad es válida si $\mu^*(F) = \infty$, E es medible si y sólo si (2.5) es válida para todo $F \subset \Omega$ con $\mu^*(F) < \infty$.

Lema 2.3 *Todos los conjuntos de $\mathcal{A}$ son medibles: $\mathcal{A} \subset \mathcal{M}(\mu^*)$*

Demostración. Sea $A \in \mathcal{A}$ y $F \subset \Omega$ con $\mu^*(F) < \infty$. Dado $\varepsilon > 0$ sea $A_n \in \mathcal{A}$ con $F \subset \cup_n A_n$ y $\sum_n \mu(A_n) \leq \mu^*(F) + \varepsilon$. Entonces

$$F \cap A \subset \cup_n (A_n \cap A), \qquad F \cap A^c \subset \cup_n (A_n \cap A^c),$$

y por lo tanto

$$\mu^*(F \cap A) + \mu^*(F \cap A^c) \leq \sum_n \mu^*(A_n \cap A) + \mu^*(A_n \cap A^c) \leq \sum_n \mu(A_n) \leq \mu^*(F) + \varepsilon.$$

Haciendo $\varepsilon \downarrow 0$ se obtiene la desigualdad (2.5). ■

Lema 2.4 *$\mathcal{M}(\mu^*)$ es una σ-álgebra y μ^* es una medida sobre $\mathcal{M}$.*

Demostración. Por la definición de medibilidad es claro que $A \in \mathcal{M}(\mu^*)$ si y sólo si $A^c \in \mathcal{M}(\mu^*)$. Si $A, B \in \mathcal{M}(\mu^*)$, para cualquier $F \subset \Omega$,

$$\begin{aligned}\mu^*(F) &= \mu^*(F \cap A) + \mu^*(F \cap A^c) \\ &= \mu^*(F \cap A \cap B) + \mu^*((F \cap A) \cap B^c) + \mu^*(F \cap A^c)\end{aligned}$$

Veamos ahora que

$$\mu^*((F \cap A) \cap B^c) + \mu^*(F \cap A^c) = \mu^*(F \cap (A \cap B)^c) \tag{2.6}$$

con lo cual obtendríamos que para cualquier $F \subset \Omega$,

$$\mu^*(F) = \mu^*(F \cap (A \cap B)) + \mu^*(F \cap (A \cap B)^c)$$

y esto muestra que $A \cap B \in \mathcal{M}(\mu^*)$ y por lo tanto $\mathcal{M}(\mu^*)$ es un álgebra.

Observemos que

$$F \cap (A \cap B)^c = F \cap (A^c \cup B^c) = (F \cap A^c) \cup (F \cap A \cap B^c).$$

Como A es medible, tenemos para este conjunto

$$\begin{aligned}\mu^*(F \cap (A \cap B)^c) &= \mu^*(F \cap (A \cap B)^c \cap A) + \mu^*(F \cap (A \cap B)^c \cap A^c) \\ &= \mu(F \cap A \cap B^c) + \mu^*(F \cap A^c).\end{aligned}$$

con lo cual se demuestra (2.6). Por lo tanto $\mathcal{M}(\mu^*)$ es un álgebra. Consideremos ahora $A_n \in \mathcal{M}(\mu^*)$ para $n \geq 1$, y veamos la medibilidad de $A = \cup_{j \geq 1} A_j$. Sea $B_n = \cup_{j=1}^n A_j \in \mathcal{M}(\mu^*)$, como $\mathcal{M}(\mu^*)$ es un álgebra tenemos $A_n \cap B_{n-1}^c \in \mathcal{M}(\mu^*)$ para todo n, y podemos suponer que los A_n son disjuntos para probar la medibilidad de A. Sea $F \subset \Omega$ entonces

$$\begin{aligned}\mu^*(F) &= \mu^*(F \cap B_n^c) + \mu^*(F \cap B_n) \\ &= \mu^*(F \cap B_n^c) + \mu^*(F \cap A_n) + \mu^*(F \cap B_{n-1}).\end{aligned}$$

Haciendo inducción en n,

$$\mu^*(F) = \mu^*(F \cap B_n^c) + \sum_{j=1}^n \mu^*(F \cap A_j) \geq \mu^*(F \cap A^c) + \sum_{j=1}^n \mu^*(F \cap A_j).$$

Haciendo $n \to \infty$ obtenemos

$$\mu^*(F) \geq \mu^*(F \cap A^c) + \sum_{j=1}^\infty \mu^*(F \cap A_j) \geq \mu^*(F \cap A^c) + \mu^*(F \cap A)$$

por el lema 2.1. Por lo tanto $A \in \mathcal{M}(\mu^*)$ y de nuevo por el lema 2.1

$$\mu^*(F) = \mu^*(F \cap A^c) + \sum_{j=1}^{\infty} \mu^*(F \cap A_j).$$

Poniendo $F = A$ muestra que μ^* es σ-aditiva en $\mathcal{M}(\mu^*)$ y esto demuestra el lema. ■

Uniendo los resultados de los últimos cuatro lemas y las definiciones que hemos visto obtenemos el siguiente teorema de extensión.

Teorema 2.7 (de extensión) *Dados cualquier conjunto Ω, cualquier álgebra $\mathcal{A}$ de subconjuntos de Ω y cualquier función μ positiva y σ-aditiva, definida sobre $\mathcal{A}$, existe una extensión de μ a la σ-álgebra generada por $\mathcal{A}$, es decir, existe una medida μ' sobre $\sigma(\mathcal{A})$ tal que para todo $A \in \mathcal{A}$, $\mu(A) = \mu'(A)$.*

Los conjuntos de medida exterior nula siempre son medibles:

Proposición 2.2 *Si $\mu^*(E) = 0$ entonces $E \in \mathcal{M}(\mu^*)$.*

Demostración. Para cualquier $F \subset \Omega$,

$$\mu^*(F) \geq \mu^*(F \cap E^c) = \mu^*(F \cap E^c) + \mu^*(E) \geq \mu^*(F \cap E^c) + \mu^*(F \cap E).$$

■

Aun nos falta considerar el problema de unicidad de la extensión de medidas. Para ver que en general la extensión no es única consideremos la colección de los intervalos $(a, b]$ en $\mathbb{R}$. Sea $\mathcal{A}$ el álgebra generada por estos intervalos y sea $\mu(A) = \infty$ para todos los conjuntos de $\mathcal{A}$. Sea ahora $\mathcal{F}$ la σ-álgebra generada por $\mathcal{A}$. La medida de contar sobre $\mathcal{F}$ coincide con μ en $\mathcal{A}$, pero también lo hace la función que asigna valor $+\infty$ a todos los conjuntos no vacíos de $\mathcal{F}$. Como veremos en el siguiente teorema, la propiedad clave para que la extensión sea única es que la μ sea σ-finita.

Teorema 2.8 (de unicidad) *Sea μ una función positiva y σ-aditiva sobre un álgebra $\mathcal{A}$. Sea ν una medida sobre la σ-álgebra $\mathcal{F}$ generada por $\mathcal{A}$ con $\mu = \nu$ en $\mathcal{A}$. Entonces para cualquier $A \in \mathcal{F}$ con $\mu^*(A) < \infty$, $\nu(A) = \mu^*(A)$. Si μ es σ-finita, entonces la extensión de μ a la σ-álgebra $\mathcal{F}$ que ella genera es única, y la extensión coincide con ν en $\mathcal{F}$.*

Demostración. Para cualquier $A \in \mathcal{F}$ y $A_n \in \mathcal{A}$ con $A \subset \cup_n A_n$ tenemos por la demostración del lema 2.3 que $\nu(A) \leq \sum_n \nu(A_n) = \sum_n \mu(A_n)$. Tomando ínfimo obtenemos que $\nu(A) \leq \mu^*(A)$. Si $\mu^*(A) < \infty$, dado $\varepsilon > 0$, escogemos A_n con $A \subset \cup_{n\geq 1} A_n$ y $\sum_n \mu(A_n) < \mu^*(A) + \varepsilon/3$. Sea $B_k = \cup_{m=1}^k A_m$. Entonces $B_k \in \mathcal{A}$ para k finito y $B_\infty \in \mathcal{F}$. Para k suficientemente grande, $\mu^*(B_\infty - B_k) < \varepsilon/3$. Como $A \subset B_\infty$,

$$\mu^*(A) \leq \mu^*(B_\infty) < \mu^*(A) + \varepsilon/3$$

y en consecuencia

$$\nu(B_k \triangle A) \leq \mu^*(B_k \triangle A) = \mu^*(A - B_k) + \mu^*(B_k - A) \leq \mu^*(B_\infty - B_k) + \mu^*(B_\infty - A) < 2\varepsilon/3$$

Por lo tanto, teniendo en cuenta que $B_k \subset A \cup (B_k - A) \subset A \cup (B_k \triangle A)$,

$$\nu(A) \geq \nu(B_k) - 2\varepsilon/3 = \mu^*(B_k) - 2\varepsilon/3 \geq \mu^*(B_\infty) - \varepsilon \geq \mu^*(A) - \varepsilon.$$

Haciendo $\varepsilon \downarrow 0$ obtenemos $\nu(A) \geq \mu^*(A)$, de modo que $\nu(A) = \mu^*(A)$. ■

2.5. Semiálgebras

Definición 2.9 Una clase $\mathcal{S}$ de subconjuntos de Ω es una semiálgebra si

(a) $\emptyset, \Omega \in \mathcal{S}$.

(b) $\mathcal{S}$ es un sistema π.

(c) Si $A \in \mathcal{S}$ para algún n finito existen conjuntos $C_1, \dots, C_n$ en $\mathcal{S}$ tales que $A^c = \cup_{i=1}^n C_i$.

Ejemplos 2.3
1. $\Omega = \mathbb{R}$, $\mathcal{S} = \{(a, b] : -\infty \le a \le b \le \infty\}$. $\mathcal{S}$ es una semiálgebra.

2. $\Omega = \mathbb{R}^k$, $\mathcal{S}_k$ es la colección de rectángulos.

Proposición 2.3 *Sea $\mathcal{S}$ una semiálgebra de subconjuntos de Ω y definimos*

$$\mathcal{K} = \{\cup_{i\in I} S_i,\ I \text{ finito},\ S_i \text{ disjuntos dos a dos},\ S_i \in \mathcal{S} \text{ para } i \in I\} \tag{2.7}$$

Entonces $\mathcal{K}$ es el álgebra generada por $\mathcal{S}$.

Demostración. Es claro que $\mathcal{S} \subset \mathcal{K}$. Veamos ahora que $\mathcal{K}$ es un álgebra.

(i) $\Omega \in \mathcal{K}$ porque $\Omega \in \mathcal{S}$.

(ii) Si $\cup_{i\in I} S_i$ y $\cup_{j\in J} S'_j$ están en $\mathcal{K}$ entonces

$$\big(\cup_{i\in I} S_i\big) \cap \big(\cup_{j\in J} S'_j\big) = \cup_{(i,j)\in I\times J} S_i \cap S'_j$$

y este conjunto está en $\mathcal{K}$ porque $\{S_i \cap S'_j, (i, j) \in I \times J\}$ es una colección finita y disjunta de elementos del π-sistema $\mathcal{S}$.

(iii) Para ver que es cerrada bajo complementos sea $\cup_{i\in I} S_i \in \mathcal{K}$, entonces

$$\big(\cup_{i\in I} S_i\big)^c = \cap_{i\in I} S_i^c.$$

A partir de los axiomas de una semiálgebra, $S_i \in \mathcal{S}$ implica que $S_i^c = \cup_{j\in J_i} S_{ij}$ para una colección finita de conjuntos disjuntos $\{S_{ij}, j \in J_i\}$ en $\mathcal{S}$. Ahora observamos que por (ii), $\cap_{i\in I} S_i^c \in \mathcal{K}$.

En consecuencia $\mathcal{K}$ es un álgebra, $\mathcal{S} \subset \mathcal{K}$ y por lo tanto $\mathcal{A}(\mathcal{S}) \subset \mathcal{K}$. Como además

$$\cup_{i\in I} S_i \in \mathcal{K} \Rightarrow \cup_{i\in I} S_i \in \mathcal{A}(\mathcal{S})$$

obtenemos que $\mathcal{K} \subset \mathcal{A}(\mathcal{S})$ y por lo tanto $\mathcal{K} = \mathcal{A}(\mathcal{S})$. ■

Teorema 2.9 *Sea $\mathcal{S}$ una semiálgebra y α una función aditiva de $\mathcal{S}$ en $[0, \infty]$. Si $A = \cup_{i\in I} S_i \in \mathcal{A}(\mathcal{S})$ definimos*

$$\mu(A) = \sum_{i\in I} \alpha(S_i)$$

Entonces μ está bien definida en $\mathcal{A}$. Si α es σ-aditiva en $\mathcal{S}$, también lo es μ en $\mathcal{A}$, y entonces podemos extender μ a una medida en la σ-álgebra generada por $\mathcal{S}$ o $\mathcal{A}$.

Demostración. Supongamos que $\{C_i, 1 \leq i \leq n\}$ son conjuntos disjuntos en $\mathcal{S}$, y que también lo son $\{D_j, 1 \leq j \leq m\}$ y además

$$\cup_{i=1}^{n} C_i = \cup_{j=1}^{m} D_j.$$

Entonces para cada i, $C_i = \cup_{j=1}^{m} C_i \cap D_j$, es una unión disjunta de conjuntos en $\mathcal{S}$. Por aditividad

$$\sum_{i=1}^{n} \alpha(C_i) = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha(C_i \cap D_j) = \sum_{j=1}^{m} \alpha(D_j)$$

de modo que μ está bien definida. Si $B_1, \dots, B_n$ son disjuntos en $\mathcal{A}$, sea $B_i = \cup_{j=1}^{m(i)} B_{ij}, \ i = 1, \dots n$, donde para cada i, los B_{ij} son disjuntos, están en $\mathcal{S}$ y $m(i) < \infty$. Por lo tanto todos los B_{ij} son disjuntos. Sea $B = \cup_i \cup_j B_{ij}$. Como μ está bien definida

$$\mu(B) = \sum_{i=1}^{n} \sum_{j=1}^{m(i)} \alpha(B_{ij}) = \sum_{i=1}^{n} \mu(B_i),$$

de modo que μ es aditiva. Supongamos ahora que α es σ-aditiva, $B \in \mathcal{A}$, tal que $B = \cup_{i=1}^{n} C_i$ para ciertos conjuntos disjuntos $C_i \in \mathcal{S}$ y $B = \cup_{j=1}^{\infty} A_j$ para ciertos conjuntos disjuntos $A_j \in \mathcal{A}$. Sea $A_j = \cup_{k=1}^{m(j)} A_{jk}$ para $A_{jk} \in \mathcal{S}$ disjuntos. Entonces

$$B = \cup_{i=1}^{n} \cup_{j=1}^{\infty} \cup_{k=1}^{m(j)} C_i \cap A_{jk}$$

que es la unión de conjuntos disjuntos en $\mathcal{S}$. Cada A_j es la unión disjunta y finita de los conjuntos $C_i \cap A_{jk}$, para $i = 1, \dots, n; \ k = 1, \dots, m(j)$. Por σ-aditividad de α en $\mathcal{S}$ y reordenando la suma de términos positivos

$$\begin{aligned} \mu(B) = \sum_{i=1}^{n} \alpha(C_i) &= \sum_{i=1}^{n} \sum_{j=1}^{\infty} \sum_{k=1}^{m(j)} \alpha(C_i \cap A_{jk}) \\ &= \sum_{j=1}^{\infty} \sum_{i=1}^{n} \sum_{k=1}^{m(j)} \alpha(C_i \cap A_{jk}) = \sum_{j=1}^{\infty} \mu(A_j), \end{aligned}$$

de modo que μ es σ-aditiva en $\mathcal{A}$. Por el teorema 2.7 podemos extenderla a una medida sobre la σ-álgebra generada por $\mathcal{A}$. ■

Corolario 2.3 *Sea $\mathcal{S}$ una semiálgebra de subconjuntos de Ω y P una función σ-aditiva de $\mathcal{S}$ a $[0,1]$ con $P(\Omega) = 1$. Entonces hay una única probabilidad sobre $\sigma(\mathcal{S})$ que extiende a μ.*

2.6. Completitud y Aproximación

En el teorema de extensión, ¿Qué diferencia hay entre $\sigma(\mathcal{A})$ y $\mathcal{M}$? Para hallar la respuesta tenemos que considerar los conjuntos nulos. Un conjunto N es nulo si su medida exterior es 0: $\mu^*(N) = 0$.

Definición 2.10 Una medida μ en una σ-álgebra $\mathcal{F}$ es *completa* si $A \in \mathcal{F}$ y $\mu(A) = 0$ implican que para todo $B \subset A$ se tiene que $B \in \mathcal{F}$.

La *completación* de un espacio de medida $(\Omega, \mathcal{F}, \mu)$ se define de la siguiente manera: Sea

$$\mathcal{F}_\mu = \{A \cup N : A \in \mathcal{F}, N \subset M \in \mathcal{F}, \mu(M) = 0\}$$

$\mathcal{F}_\mu$ es una σ-álgebra que incluye a $\mathcal{F}$:

- Si $A \cup N \in \mathcal{F}_\mu$, $N \subset M \in \mathcal{F}$, $\mu(M) = 0$ entonces

$$(A \cup N)^c = A^c \cap N^c = (A^c \cap N^c) \cap (M^c \cup M) = (A^c \cap M^c) \cup (A^c \cap N^c \cap M)$$

  pero $A^c \cap N^c \cap M) \subset M$, de modo que $(A \cup N)^c \in \mathcal{F}_\mu$.

- Si $A_k \cup N_k \in \mathcal{F}_\mu$ para $k \geq 1$ con $N_k \subset M_k \in \mathcal{F}$ y $\mu(M_k) = 0$ entonces $\cup_{k\geq 1}(A_k \cap N_k) = (\cup_{k\geq 1} A_k) \cup (\cup_{k\geq 1} N_k)$ con $(\cup_{k\geq 1} N_k) \subset (\cup_{k\geq 1} M_k) \in \mathcal{F}$ y $\mu(\cup_{k\geq 1} M_k) = \sum_{k\geq 1} \mu(M_k) = 0$.

Extendemos μ a $\mathcal{F}_\mu$ definiendo $\mu(A \cup N) = \mu(A)$. Esta definición es consistente porque si $A_1 \cup N_1 = A_2 \cup N_2 \in \mathcal{F}_\mu$ tenemos

$$\mu(A_1) = \mu(A_1 \cap A_2) + \mu(A_1 \cap A_2^c) = \mu(A_1 \cap A_2)$$

porque $A_1 \cap A_2 \subset N_2$. Por lo tanto $\mu(A_1) \leq \mu(A_2)$ y por simetría obtenemos que $\mu(A_1) = \mu(A_2)$. El espacio de medida $(\Omega, \mathcal{F}_\mu, \mu)$ se conoce como la *completación* de $(\Omega, \mathcal{F}, \mu)$ y $\mathcal{F}_\mu$ es la *completación* de $\mathcal{F}$ respecto de μ.

La completación es, en efecto, completa, ya que si $E \subset A \cup N \in \mathcal{F}_\mu$, con $A \in \mathcal{F}$, $\mu(A) = 0$, $N \subset M \in \mathcal{F}$, $\mu(M) = 0$ entonces $E \subset A \cup M \in \mathcal{F}$, $\mu(A \cup M) = 0$. Por lo tanto $E \in \mathcal{F}_\mu$.

Teorema 2.10 *En el teorema 2.7,* $(\Omega, \mathcal{M}, \mu^*)$ *es la completación de* $(\Omega, \mathcal{F} = \sigma(\mathcal{A}), \mu)$.

Demostración. Tenemos que demostrar que $\mathcal{M} = \mathcal{F}_{\mu^*}$. Si $A \in \mathcal{M}$, por definición de $\mu^*(A)$ y $\mu^*(A^c)$ podemos hallar conjuntos G_n, $H_n \in \mathcal{F}$ para $n \geq 1$ tales que $G_n \subset A \subset H_n$ y $\mu^*(G_n) \to \mu^*(A)$, $\mu^*(H_n) \to \mu^*(A)$. Sea $G = \cup_{n\geq 1} G_n$, $H = \cap_{n\geq 1} H_n$. Entonces $A = G \cup (A - G)$, $G \in \mathcal{F}$, $A - G \subset H - G \in \mathcal{F}$, $\mu^*(H - G) \leq \mu^*(H_n - G_n) \to 0$ de modo que $\mu^*(H - G) = 0$. Por lo tanto $A \in \mathcal{F}_{\mu^*}$.

Recíprocamente, si $B \in \mathcal{F}_{\mu^*}$ entonces $B = A \cup N$, $A \in \mathcal{F}$, $N \subset M \in \mathcal{F}$, $\mu^*(M) = 0$. Como $\mathcal{F} \subset \mathcal{M}$ tenemos que $A \in \mathcal{M}$, y como $(\Omega, \mathcal{M}, \mu^*)$ es completo entonces que $N \in \mathcal{M}$. Por lo tanto $B \in \mathcal{M}$. ■

El siguiente resultado muestra que si $\mathcal{F} = \sigma(\mathcal{A})$, donde $\mathcal{A}$ es una álgebra, los conjuntos de $\mathcal{F}$ pueden ser aproximados por conjuntos de $\mathcal{A}$ en algún sentido.

Teorema 2.11 (Aproximación.) *Sea* $(\Omega, \mathcal{F}, \mu)$ *un espacio de medida y sea* $\mathcal{A}$ *un álgebra en* Ω *tal que* $\sigma(\mathcal{A}) = \mathcal{F}$. *Supongamos que* μ *es* σ*-finita en* $\mathcal{A}$ *y sea* $\varepsilon > 0$ *dado. Si* $A \in \mathcal{F}$ *y* $\mu(A) < \infty$, *existe un conjunto* $B \in \mathcal{A}$ *tal que* $\mu(A \triangle B) < \varepsilon$.

Demostración. Haremos la demostración sólo en el caso de una medida finita. Sea $\mathcal{G}$ la clase de las uniones numerables de conjuntos de $\mathcal{A}$. La conclusión del teorema vale para cualquier $A \in \mathcal{G}$ por continuidad de la medida. Si μ es finita, por el teorema de unicidad y la definición de la medida exterior, cualquier $A \in \mathcal{F}$ puede ser aproximado por un conjunto de $\mathcal{G}$, y por lo tanto el teorema vale para μ finita. ■

Ejemplo 2.4
Sea $\Omega = \mathbb{Q}$, $\mathcal{A}$ el álgebra de uniones finitas y disjuntas de intervalos $(a, b]$ con $a, b \in \mathbb{Q}$ (incluimos (a, ∞) y Ω en esta clase). Sea $\mathcal{F} = \sigma(\mathcal{A})$. Entonces

(a) $\mathcal{F} = \mathcal{P}(\Omega)$.

(b) Si $\mu(A)$ es el número de puntos de A, entonces μ es σ-finita en $\mathcal{F}$ pero no en $\mathcal{A}$.

(c) Hay conjuntos $A \in \mathcal{F}$ de medida finita que no pueden ser aproximados por conjuntos de $\mathcal{A}$, es decir, no existe una sucesión $A_n \in \mathcal{A}$ con $\mu(A \triangle A_n) \to 0$.

(d) Si $\lambda = 2\mu$ entonces $\lambda = \mu$ en $\mathcal{A}$ pero no en $\mathcal{F}$.

Así, tanto el teorema de extensión como el de aproximación fallan en este caso. Veamos la demostración de estas afirmaciones.

(a) Tenemos $\{x\} = \cap_{n\geq 1}(x - (1/n), x]$, y por lo tanto todos los conjuntos de un solo elemento están en $\mathcal{F}$. Pero entonces todos los subconjuntos de Ω están en $\mathcal{F}$ porque Ω es numerable.

(b) Como Ω es la unión numerable de conjuntos de un solo elemento, μ es σ-finita en $\mathcal{F}$. Pero todo conjunto no vacío en $\mathcal{A}$ tiene medida infinita, de modo que μ no es σ-finita en $\mathcal{A}$.

(c) Si A es cualquier subconjunto finito no vacío de Ω, entonces $\mu(A\triangle B) = \infty$ para todo $B \in \mathcal{A}$, $B \neq \emptyset$, porque los conjuntos no vacíos en $\mathcal{A}$ deben contener infinitos puntos que no están en A.

(d) Como $\lambda(\{x\}) = 2$ y $\mu(\{x\}) = 1$, $\lambda \neq \mu$ en $\mathcal{F}$. Pero $\lambda(A) = \mu(A) = \infty$ para todo $A \in \mathcal{A}$, $A \neq \emptyset$.

2.7. La Medida de Lebesgue

2.7.1. La Medida de Lebesgue en $(0, 1]$

Sea $\mathcal{S} = \{(a, b] : 0 \leq a \leq b \leq 1\}$ y definimos en $\mathcal{S}$ la función $\lambda : \mathcal{S} \to [0, 1]$ por

$$\lambda(\emptyset) = 0, \qquad \lambda(a, b] = b - a.$$

Observemos que $\lambda(A) \geq 0$. Para demostrar que λ tiene una única extensión tenemos que mostrar que λ es σ-aditiva en $\mathcal{S}$, pues claramente es σ-finita. Veamos primero que λ es aditiva. Sea $(a, b] \in \mathcal{S}$ y supongamos que

$$(a, b] = \cup_{i=1}^{n}(a_i, b_i],$$

donde los intervalos de la derecha son disjuntos. Si los intervalos han sido indexados convenientemente tenemos

$$a = a_1 \leq b_1 = a_2 \leq b_2 = a_3 \leq \cdots \leq b_{n-1} = a_n \leq b_n = b.$$

Entonces $\lambda(a, b] = b - a$ y $\sum_{i=1}^{n} \lambda(a_i, b_i] = \sum_{i=1}^{n}(b_i - a_i) = b - a$. Esto muestra que λ es aditiva.

Veamos ahora que λ es σ-aditiva. Sea $(a, b] = \cup_{n\geq 1}(a_i, b_i]$, donde los intervalos en la unión son disjuntos dos a dos. Demostraremos primero que

$$b - a \leq \sum_{i=1}^{\infty}(b_i - a_i). \tag{2.8}$$

Sea $\varepsilon < b - a$ y observemos que

$$[a + \varepsilon, b] \subset \cup_{i=1}^{\infty}(a_i, b_i + \varepsilon 2^{-i}) \tag{2.9}$$

El intervalo a la izquierda de (2.9) es cerrado y acotado y por lo tanto compacto y el lado derecho representa un cubrimiento abierto. Por compacidad hay un subcubrimiento finito, es decir, existe un entero N tal que

$$[a + \varepsilon, b] \subset \cup_{i=1}^{N}(a_i, b_i + \varepsilon 2^{-i}). \tag{2.10}$$

Basta probar

$$b - a - \varepsilon \leq \sum_{i=1}^{N}(b_i - a_i + \varepsilon 2^{-i}) \tag{2.11}$$

ya que entonces tendríamos

$$b - a \leq \sum_{i}^{\infty}(b_i - a_i) + 2\varepsilon.$$

Como ε es arbitrario esto implica

$$b - a \leq \sum_{1}^{\infty}(b_i - a_i).$$

Reescribiendo las relaciones (2.10) y (2.11), tenemos que probar que

$$[a,b] \subset \cup_1^N (a_i, b_i) \tag{2.12}$$

implica

$$b - a \leq \sum_1^N (b_i - a_i). \tag{2.13}$$

Vamos a probar esto por inducción. En primer lugar observamos que (2.12) implica que (2.13) vale para $N = 1$. Ahora hacemos la hipótesis inductiva de que si (2.12) vale para $N-1$ entonces (2.13) también vale para $N-1$. Veamos que lo mismo es cierto para N.

Supongamos que $a_N = \sup_i a_i$ y

$$a_N < b \leq b_N. \tag{2.14}$$

Es posible usar un argumento similar si esto es falso. Supongamos que la ecuación (2.12) vale. Consideramos dos casos

Caso 1: $a_N \leq a$ entonces

$$b - a \leq b_N - a_N \leq \sum_1^N (b_i - a_i).$$

Caso 2: $a_N > a$. Entonces si (2.12) vale

$$[a, a_N] \subset \cup_1^{N-1} (a_i, b_i)$$

y por la hipótesis inductiva

$$a_N - a \leq \sum_{i=1}^{N-1} (b_i - a_i)$$

de modo que

$$\begin{aligned} b - a = b - a_N + a_N - a &\leq b - a_N + \sum_{i=1}^{N-1} (b_i - a_i) \\ &\leq b_N - a_N + \sum_{i=1}^{N-1} (b_i - a_i) = \sum_{i=1}^{N} (b_i - a_i) \end{aligned}$$

que es (2.13). Esto demuestra (2.8).

Ahora obtendremos una desigualdad en el sentido inverso a (2.8). Veremos que si $(a,b] = \cup_{i=1}^{\infty} (a_i, b_i]$ donde los intervalos son disjuntos, entonces para todo n

$$\lambda(a,b] = b - a \geq \sum_{i=1}^{n} \lambda(a_i, b_i] = \sum_{i=1}^{n} (b_i - a_i). \tag{2.15}$$

Esto es fácil de verificar porque sabemos que λ es aditiva en $\mathcal{S}$. Para todo n, $\cup_{i=1}^{n} (a_i, b_i]$ es la unión finita y disjunta de intervalos, y también lo es

$$(a,b] - \cup_{i=1}^{n} (a_i, b_i] = \cup_{j=1}^{m} I_j.$$

Por aditividad

$$\begin{aligned} \lambda(a,b] &= \lambda\big(\cup_{i=1}^{n} (a_i, b_i] \cup \cup_{j=1}^{m} I_j \big) \\ &= \sum_{i=1}^{n} \lambda(a_i, b_i] + \sum_{j=1}^{m} \lambda(I_j) \geq \sum_{i=1}^{n} \lambda(a_i, b_i]. \end{aligned}$$

Haciendo ahora $n \to \infty$ obtenemos

$$\lambda(a,b] \geq \sum_{i=1}^{\infty} \lambda(a_i, b_i].$$

Esto completa la demostración. ■

2.7.2. La Medida de Lebesgue en $\mathbb{R}^k$

Antes de considerar este caso hacemos una observación técnica. Al igual que introducimos la noción de anillo en una lista de ejercicios, introducimos ahora los semianillos. Una colección $\mathcal{C}$ de subconjuntos de Ω es un semianillo si $\emptyset \in \mathcal{C}$, $\mathcal{C}$ es un sistema π y la diferencia entre dos conjuntos cualquiera de $\mathcal{C}$ se puede escribir como la unión finita y disjunta de conjuntos de $\mathcal{C}$. La diferencia entre un semianillo y una semiálgebra es que para un semianillo no se pide que $\Omega \in \mathcal{C}$. Toda semiálgebra es un semianillo, pero lo contrario no siempre es cierto.

Para desarrollar el teorema de extensión es posible comenzar con una función μ positiva y σ-aditiva, definida en un semianillo $\mathcal{C}$, extenderla al anillo generado por $\mathcal{C}$ y luego de allí a la σ-álgebra generada. El procedimiento para demostrar este tipo de extensión es similar al que hemos llevado a cabo y se puede encontrar, por ejemplo, en el libro de R. M. Dudley. Usaremos este resultado para la definición de la medida de Lebesgue en $\mathbb{R}^k$.

En $\mathbb{R}^k$ definimos los intervalos semiabiertos de la siguiente manera: Sean $\mathbf{a} = (a_1, \dots, a_k)$, $\mathbf{b} = (b_1, \dots, b_k)$ con $a_i \leq b_i$ para $i = 1, \dots, k$. Entonces

$$(\mathbf{a}, \mathbf{b}] = \prod_{i=1}^{k} (a_i, b_i] = \{x_1, x_2, \dots, x_k) : a_i < x_i \leq b_i,\ 1 \leq i \leq k\}$$

Esta colección de intervalos forma un semianillo $\mathcal{S}_k$ que genera el anillo $\mathcal{R}(\mathcal{S}_k) = \mathcal{E}_k$.

Definimos la función $\lambda : \mathcal{S}_k \to [0, \infty]$ por

$$\lambda(\mathbf{a}, \mathbf{b}] = \prod_{i=1}^{k} (b_i - a_i).$$

Es posible demostrar que esta función es σ-aditiva en $\mathcal{S}_k$, es decir, si $(\mathbf{a}, \mathbf{b}] = \cup_{n\geq 1} (\mathbf{a}_n, \mathbf{b}_n]$ donde los intervalos $(\mathbf{a}_n, \mathbf{b}_n]$ son disjuntos dos a dos, entonces $\lambda(\mathbf{a}, \mathbf{b}]) = \sum_{n\geq 1} \lambda(\mathbf{a}_n, \mathbf{b}_n])$. Usando el teorema de extensión obtenemos una medida λ en la σ-álgebra $\mathcal{B}_k = \sigma(\mathcal{S}_k)$ generada por $\mathcal{S}_k$, que es la σ-álgebra de los borelianos. Como λ es σ-finita, la extensión es única.

El procedimiento de extensión usando el teorema de Carathéodory produce una extensión de μ a la colección de los conjuntos medibles $\mathcal{M}$, que en este caso se conocen como los conjuntos medibles (según) Lebesgue, y esta clase contiene a los conjuntos borelianos.

2.7.3. Propiedades de la Medida de Lebesgue

Hemos definido la medida de Lebesgue en los espacios euclídeos. Vamos a estudiar en esta sección sus propiedades geométricas. En primer lugar para cualquier punto $x \in \mathbb{R}^k$, el conjunto $\{x\}$ puede ser cubierto por intervalos de longitud arbitrariamente pequeña, y por lo tanto

$$\lambda(\{\mathbf{x}\}) = 0 \qquad \text{para todo } \mathbf{x} \in \mathbb{R}^k$$

Como consecuencia,

$$\lambda([\mathbf{a}, \mathbf{b}]) = \lambda((\mathbf{a}, \mathbf{b}]) = \lambda([\mathbf{a}, \mathbf{b})) = \lambda((\mathbf{a}, \mathbf{b})) = \prod_{n=1}^{k} (b_i - a_i)$$

Cualquier conjunto numerable en $\mathbb{R}^k$ tiene medida 0. En particular el conjunto de puntos de $\mathbb{R}^k$ con coordenadas racionales forman un conjunto de medida 0.

En $\mathbb{R}^k,\ k \geq 2$, cualquier segmento de longitud ℓ de una recta puede ser cubierto por $\lceil n\ell \rceil$ cubos de lado $1/n$, de modo que la medida de Lebesgue del segmento debe ser menor que

$$\big(\frac{1}{n}\big)^k \lceil n\ell \rceil = O\big(\frac{1}{n^{k-1}}\big) \quad \text{cuando } n \to \infty,$$

y por lo tanto $\lambda(L) = 0$ para cualquier segmento L de longitud finita. Toda recta es la unión numerable de segmentos de longitud finita, de modo que $\lambda(L) = 0$ para cuaquier recta L en $\mathbb{R}^k,\ k \geq 2$. Como consecuencia, si estamos calculando el área de cualquier figura geométrica en el plano que esté acotada por una colección numerable de rectas, entonces el área no cambia si se incluyen todas, algunas o ninguna de las rectas en la figura.

El resultado anterior muestra que hay conjuntos $E \subset \mathbb{R}^k,\ k \geq 2$ que no son numerables pero $\lambda(E) = 0$. Esto también es cierto para $\mathbb{R}$ como lo muestra el conjunto de Cantor. Para construir este conjunto partimos del intervalo $C_0 = [0, 1]$ y le quitamos un intervalo abierto centrado de longitud $1/3$, obteniendo el conjunto

$$C_1 = [0, 1/3] \cup [2/3, 1].$$

En el siguiente paso, a cada uno de los intervalos cerrados que componen C_1 le quitamos un intervalo abierto centrado de longitud igual a $1/3$ de la longitud del intervalo, y obtenemos

$$C_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1].$$

Continuamos inductivamente este procedimiento: En cada paso, a cada intervalo cerrado de C_n le quitamos un intervalo abierto centrado cuya longitud sea igual a $1/3$ de la longitud del intervalo cerrado, y obtenemos C_{n+1}. El conjunto de Cantor se define como

$$C = \cap_{n=0}^{\infty} C_n.$$

El procedimiento muestra que

$$\lambda(C_n) = \frac{2}{3}\lambda(C_{n-1}) = \big(\frac{2}{3}\big)^n \lambda(C_0) = \big(\frac{2}{3}\big)^n,$$

de modo que $\lambda(C) = \lim_n \lambda(C_n) = 0$.

Para ver que el conjunto de Cantor que hemos construido no es numerable observamos que cualquier número real en el intervalo $[0, 1]$ se puede escribir usando un desarrollo en potencias de 3

$$x = \sum_{n=1}^{\infty} a_i 3^{-n},$$

donde $a_i = a_i(x)$ puede tomar los valores $0, 1$ ó 2. Este desarrollo no siempre es único, por ejemplo, $1/3$ se puede escribir de dos maneras:

$$\frac{1}{3} = 1 \cdot 3^{-1} + 0 \cdot 3^{-2} + 0 \cdot 3^{-3} + \cdots = 0 \cdot 3^{-1} + 1 \cdot 3^{-2} + 1 \cdot 3^{-3} + \cdots$$

En estos casos ambiguos tomamos el desarrollo infinito, es decir, escribimos $1/3 = 0.0111\ldots$. Con esta convención vemos que el primer intervalo que quitamos, $(1/3, 2/3)$, corresponde a los números que tienen $a_1(x) = 1$, de modo que para cualquier punto x en C se tiene que $a_1(x) \neq 1$. De manera similar el segundo conjunto que quitamos C_2 se puede describir como el conjunto de puntos de C_1 para los cuales el segundo término en el desarrollo en base 3 es igual a 1: $a_2(x) = 1$. Por lo tanto, para todos los puntos de C se tiene que $a_2(x) \neq 1$. Continuando este argumento vemos que los puntos de C se pueden caracterizar como aquellos puntos de $[0, 1]$ cuyo desarrollo en base tres no tiene 1's: $a_n(x) \neq 1$ para todo $n \geq 1$.

Con esta caracterización es fácil ver que C no es numerable porque podemos hacer una biyección de C a $[0,1]$. Dado un punto $x \in C$ con desarrollo $x = \sum_{n=1}^{\infty} a_n(x)3^{-n}$ le asignamos el número $t(x) = \sum_{n=1}^{\infty} b_n(x)2^{-n}$ donde $b_n(x) = a_n(x)/2$, es decir, $b_n(x) = 0$ si $a_n(x) = 0$, $b_n(x) = 1$ si $a_n(x) = 2$. Es fácil ver que t es una biyección y por lo tanto hay una cantidad no numerable de puntos en C.

Veamos ahora qué ocurre a la medida de Lebesgue cuando se realizan trasformaciones elementales del espacio.

Traslaciones

Sea $\mathbf{x} \in \mathbb{R}^k$ y $E \subset \mathbb{R}^k$. Definimos $E(\mathbf{x}) = \{\mathbf{z} : \mathbf{z} = \mathbf{x} + \mathbf{y},\ \mathbf{y} \in E\}$. Para los intervalos $I \subset \mathbb{R}^k$ es inmediato que $\lambda(I(\mathbf{x})) = \lambda(I)$ y en consecuencia la medida exterior λ^* es invariante bajo traslaciones. Por lo tanto la medida de Lebesgue también lo será si la medibilidad se conserva bajo traslaciones. Supongamos que E es medible y $A \subset \Omega$. Entonces

$$\lambda^*(A(-\mathbf{x})) = \lambda^*(A(-\mathbf{x}) \cap E) + \lambda^*(A(-\mathbf{x}) \cap E^c)$$

y en consecuencia

$$\lambda^*(A) = \lambda^*(A \cap E(\mathbf{x})) + \lambda^*(A \cap E^c(\mathbf{x})),$$

de modo que $E(\mathbf{x})$ es medible.

Reflección Respecto a un Plano Perpendicular a los Ejes

Para $k = 1$ esto es reflección respecto a un punto y para $k = 2$ es reflección respecto a una recta paralela a algún eje. Es fácil ver que λ^* es invariante bajo una reflección de este tipo porque una reflección de un cubrimiento produce intervalos de igual medida. Un argumento similar al anterior muestra que que la medibilidad se conserva.

Dilatación Uniforme

Para $c > 0$ la transformación de $\mathbb{R}^k$ que se obtiene poniendo $\mathbf{y} = c \cdot \mathbf{x}$ es una dilatación por el factor c, y cE denota el resultado de aplicar esta dilatación a los puntos del conjunto E. Si I es un intervalo entonces es inmediato que cI también es un intervalo y que $\lambda(cI) = c^k \lambda(I)$. Por lo tanto

$$\lambda^*(cE) = c^k \lambda^*(E)$$

para todos los conjuntos E. Un argumento similar al que usamos para las traslaciones muestra que la medibilidad se conserva bajo dilataciones, de modo que si E es medible Lebesgue, también los es cE y $\lambda(cE) = c\lambda(E)$.

Rotaciones Respecto al Origen

La medida de Lebesgue también es invariante bajo rotaciones, pero es un poco más trabajoso demostrarlo. La idea fundamental es que cualquier esfera abierta con centro O es invariante bajo rotaciones respecto a O. Supongamos que I es un intervalo fijo en $\mathbb{R}^k$. Entonces para cualquier $\mathbf{x} \in \mathbb{R}^k$ y $c > 0$, $cI(\mathbf{x})$ es un intervalo de $\mathbb{R}^k$ similar a I y con lados paralelos a los de I. Si Γ denota la transformación de $\mathbb{R}^k$ que consiste de una rotación dada respecto a O, entonces

$$\Gamma(cI)(\mathbf{x}) = (c\Gamma(I))(\Gamma(\mathbf{x}))$$

Usando las propiedades anteriores

$$\lambda(\Gamma(cI)(\mathbf{x})) = c^k \lambda(\Gamma(I)), \qquad \lambda((cI)(\mathbf{x})) = c^k \lambda(I),$$

de modo que

$$\lambda(\Gamma(cI)(\mathbf{x})) = \frac{\lambda(\Gamma(I))}{\lambda(I)} \lambda((cI)(\mathbf{x}))$$

para todo $c > 0$, $\mathbf{x} \in \mathbb{R}^k$. Esto significa que, dados Γ e I, el efecto sobre la medida de Lebesgue es el mismo para todos los intervalos de la forma $(cI)(\mathbf{x})$.

Ahora cualquier conjunto abierto G se puede escribir como unión numerable de conjuntos disjuntos de la forma $(cI)(\mathbf{x})$. En particular la esfera unitaria abierta S centrada en el origen puede escribirse

$$S = \cup_{i=1}^{\infty} (c_i I)(\mathbf{x}_i),$$

y

$$\lambda(S) = \sum_{i=1}^{\infty} \lambda((c_i I)(\mathbf{x}_i)).$$

Pero $\Gamma(S) = S$, de modo que

$$\sum_{i=1}^{\infty} \lambda((c_i I)(\mathbf{x}_i)) = \lambda(S) = \lambda(\Gamma(S)) = \sum_{i=1}^{\infty} \lambda(\Gamma(c_i I)(\mathbf{x}_i)) = \frac{\lambda(\Gamma(I))}{\lambda(I)} \sum_{i=1}^{\infty} \lambda((c_i I)(\mathbf{x}_i))$$

y esto implica que $\lambda(\Gamma(I)) = \lambda(I)$. Este argumento es válido para cualquier intervalo I.

Ahora podemos usar los argumentos similares a los anteriores para probar que para cualquier conjunto $E \subset \mathbb{R}^k$

$$\lambda^*(\Gamma(E)) = \lambda^*(E)$$

y la medibilidad se conserva bajo Γ. Por lo tanto si E es medible, $\Gamma(E)$ también lo es y $\lambda(\Gamma(E)) = \lambda(E)$.

Observamos además que una reflección respecto a cualquier plano se puede obtener combinando las operaciones anteriores. Por lo tanto hemos demostrado el siguiente

Teorema 2.12 *La clase $\mathcal{M}_k$ de los conjuntos medibles Lebesgue de $\mathbb{R}^k$ y la medida de Lebesgue son invariantes bajo traslaciones, reflecciones y rotaciones. Si E y F son dos subconjuntos de $\mathbb{R}^k$ que son congruentes y E es medible, también lo es F y $\lambda(E) = \lambda(F)$. Para $c > 0$, entonces cE es medible si E lo es y $\lambda(cE) = c^k \lambda(E)$.*

2.7.4. Las Medidas de Lebesgue-Stieltjes

Hay otras medidas en $\mathbb{R}^k$ que son de gran importancia en la teoría de probabilidad y que se construyen de manera similar a la medida de Lebesgue. Vamos a comenzar por considerar el caso de las medidas sobre $\mathbb{R}$. Sea $F : \mathbb{R} \to \mathbb{R}$ una función no-decreciente que es continua por la derecha. A estas funciones las llamaremos funciones de Stieltjes. Definimos

$$\mu_F(a, b] = F(b) - F(a)$$

para cualquier intervalo $(a, b] \subset \mathbb{R}$. Esta función es positiva y al igual que en el caso de la medida de Lebesgue, es posible demostrar que es σ-aditiva en el semianillo $\mathcal{S}$ de los intervalos de la forma $(a, b]$. De hecho, la medida de Lebesgue corresponde a tomar la función $F(x) = x$. Usando el teorema de extensión sabemos que existe una extensión de μ_F a la clase $\mathcal{B}$ de los borelianos de $\mathbb{R}$, que se conoce como la medida de Lebesgue-Stieltjes asociada a la función F.

La colección de los conjuntos medibles depende de la función μ_F y por lo tanto también de F, y la denotaremos por $\mathcal{M}_F$. En cualquier caso sabemos, por el teorema de extensión, que es una σ-álgebra y que contiene a los borelianos. Al igual que en el caso de la medida de Lebesgue, es posible ver que la clase de los conjuntos medibles $\mathcal{M}_F$ es completa respecto a la medida μ_F.

Es fácil ver que hay infinitas funciones de Stieltjes que generan la misma medida de Lebesgue-Stieltjes: si tomamos $F'(x) = F(x) + k$ para cualquier constante $k \in \mathbb{R}$ entonces

$$\mu_{F'}(a, b] = F'(b) - F'(a) = F(b) - F(a)$$

y en consecuencia las medidas generadas por ambas funciones son iguales.

Un caso particular de las funciones de Stieltjes son las funciones de distribución, que juegan un papel central en la Teoría de Probabilidad.

Definición 2.11 Una función $F : \mathbb{R} \to \mathbb{R}$ es una *función de distribución* si

(a) F es no-decreciente y continua por la derecha.

(b) $F(x) \to 0$ cuando $x \to -\infty$, $F(x) \to 1$ cuando $x \to \infty$.

Toda función de distribución define una medida de Lebesgue-Stieltjes μ_F en $\mathcal{M}_F \supset \mathbb{B}$ y se tiene que $\mu_F(\mathbb{R}) = 1$, de modo que toda función de distribución en $\mathbb{R}$ determina una medida de probabilidad y $(\mathbb{R}, \mathcal{M}_F, \mu_F)$ es un espacio de probabilidad completo. $(\mathbb{R}, \mathcal{B}, \mu_F)$ también es un espacio de probabilidad pero no necesariamente es completo.

Veamos ahora que el proceso inverso también es cierto: toda probabilidad en $\mathbb{R}$ genera una función de distribución asociada.

Teorema 2.13 *Sea $\mathcal{F}$ una σ-álgebra en $\mathbb{R}$ que contiene a los conjuntos abiertos y sea $\mu : \mathcal{F} \to [0, \infty]$ una medida completa que es finita en conjuntos acotados de $\mathcal{F}$. Entonces existe una función de Stieltjes $F : \mathbb{R} \to \mathbb{R}$ tal que $\mathcal{M}_F \subset \mathcal{F}$ y μ coincide con μ_F en $\mathcal{M}_F$. Si $(\mathbb{R}, \mathcal{F}, \mu)$ es un espacio de probabilidad, podemos tomar a F como una función de distribución.*

Demostración. Como $\mathcal{F}$ contiene a los conjuntos abiertos y es una σ-álgebra, tiene que contener a la clase de los borelianos $\mathcal{B}$ y en particular a los intervalos de la forma $(a, b]$. Definimos F por

$$F(x) = \begin{cases} \mu(0, x] & \text{para } x \geq 0, \\ -\mu(x, 0] & \text{para } x < 0. \end{cases}$$

Entonces $F : \mathbb{R} \to \mathbb{R}$ está bein definida y es no-decreciente con $F(0) = 0$. Si (x_n) es cualquier sucesión decreciente con $x_n \downarrow x$ entonces

$$\begin{aligned} &\text{si } x \geq 0, \quad \lim_n (0, x_n] = (0, x], \\ &\text{si } x < 0, \quad \lim_n (x_n, 0] = (x, 0]. \end{aligned}$$

Por el teorema de continuidad de medidas obtenemos que $\lim_n F(x_n) = F(x)$, de modo que F es continua por la derecha, y por lo tanto es una función de Stieltjes.

Ahora

$$\begin{aligned} &\text{si } a \geq 0, && \mu(a, b] = \mu(0, b] - \mu(0, a] = F(b) - F(a), \\ &\text{si } a < 0 \leq b, && \mu(a, b] = \mu(a, 0] + \mu(0, b] = F(b) - F(a), \\ &\text{si } b < 0, && \mu(a, b] = \mu(a, 0] - \mu(b, 0] = F(b) - F(a), \end{aligned}$$

de modo que μ coincide con μ_F en $\mathcal{S}$. Por unicidad de la extensión de una medida a la σ-álgebra generada y a su comletación, tenemos $\mu = \mu_F$ en $\mathcal{M}_F$ y $\mathcal{M}_F \subset \mathcal{F}$.

Si μ es una medida de probabilidad en $\mathcal{F}$ necesariamente

$$\lim_{x \to +\infty} F(x) - \lim_{x \to -\infty} F(x) = \lim_{n \to \infty} \mu(-n, n] = 1,$$

de modo que

$$F_1(x) = F(x) - \lim_{s \to -\infty} F(x)$$

es una función de distribución que genera la misma medida de Stieltjes que F. ■

Si F es una función de distribución y μ es la medida de Lebesgue-Stieltjes asociada, sabemos que $\mu(a, b] = F(b) - F(a)$ para $a < b$. La medida de cualquier intervalo se puede expresar en términos de F. Usaremos la notación $F(x^-) = \lim_{y \uparrow x} F(y)$.

$$\begin{aligned} \mu(a, b] &= F(b) - F(a) \\ \mu(a, b) &= F(b^-) - F(a) \\ \mu[a, b) &= F(b^-) - F(a^-) \\ \mu[a, b] &= F(b) - F(a^-) \end{aligned}$$

Si F es continua en a y b, las cuatro expresiones valen lo mismo. Veamos, por ejemplo, la demostración de la segunda de estas relaciones:

$$\mu(a, b) = \lim_{n \to \infty} \mu(a, b - \frac{1}{n}] = \lim_{n \to \infty} (F(b - \frac{1}{n}) - F(a)) = F(b^-) - F(a).$$

Si ponemos $a = b = x$ en la tercera de estas relaciones obtenemos que $\mu\{x\} = F(x) - F(x^-)$, es decir, F es continua en x si y sólo si $\mu\{x\} = 0$. El tamaño de la discontinuidad de F en x coincide con la medida del conjunto $\{x\}$.

De manera similar se pueden obtener las siguientes relaciones. Recordamos que para una función de distribución se tiene que $F(-\infty) = 0$ y $F(\infty) = 1$.

$$\begin{aligned} \mu(-\infty, x] &= F(x) - F(-\infty), \\ \mu(-\infty, x) &= F(x^-) - F(-\infty), \\ \mu(x, \infty] &= F(\infty) - F(x), \\ \mu[x, \infty] &= F(\infty) - F(x^-), \\ \mu(\mathbb{R}) &= F(\infty) - F(-\infty) \end{aligned}$$

Con la definición de las medidas de Lebesgue-Stieltjes podemos definir una gran cantidad de medidas en $\mathcal{B}$. Por ejemplo, si $f : \mathbb{R} \to \mathbb{R}$, $f \geq 0$ y f es integrable en cualquier intervalo finito, fijamos el valor de $F(0)$ y definimos

$$\begin{aligned} F(x) - F(0) &= \int_0^x f(t)\, dt, \qquad x > 0 \\ F(0) - F(x) &= \int_x^0 f(t)\, dt, \qquad x < 0, \end{aligned}$$

entonces F es una función de Stieltjes continua y genera a una medida de Lebesgue-Stieltjes que satisface

$$\mu(a, b] = \int_a^b f(t)\, dt$$

En particular si $f(x) = 1$ para todo x, $F(x) = x$, $\mu(a, b] = b - a$ y obtenemos la medida de Lebesgue. Si

$$\int_{-\infty}^{\infty} f(t)\, dt = 1$$

obtenemos una medida de probabilidad. En este caso decimos que f es la densidad de la medida μ o de la función F.

Probabilidades Discretas

Sean $p_i,\ i \geq 1$ números reales con $p_i \geq 0$ para todo i y $\sum_i p_i = 1$. Sea $S = \{x_1, x_2, \dots\}$ un subconjunto numerable de $\Omega = \mathbb{R}$. Podemos definir una probabilidad P concentrada en los puntos de S poniendo, para cualquier $A \subset \Omega$,

$$P(A) = \sum_{i:x_i \in A} p_i.$$

Esta probabilidad está definida para todos los subconjuntos de $\mathbb{R}$, no solo para los borelianos. La función de distribución F que corresponde a esta medida es continua en $\mathbb{R} - S$. Si $\mu\{x_n\} = p_n > 0$, F tiene un salto en x_n de magnitud p_n. Si $x, y \in S$ y no hay ningún punto de S entre x e y, entonces F es constante en $[x, y)$: Si $x \leq b < y$, $F(b) - F(x) = \mu(x, b] = 0$.

2.8. Conjuntos no Medibles

Como hemos mencionado anteriormente, la medibilidad de un conjunto depende de la función μ que estemos considerando. Veamos un par de ejemplos sencillos antes de considerar el caso de la medida de Lebesgue.

Ejemplo 2.5
Sea $\Omega = \{0, 1\},\ \mathcal{F} = \{\emptyset, \Omega\},\ \mu(\Omega) = 1$. Entonces $\mu^*(\{0\}) = \mu^*(\{1\}) = 1$ y por lo tanto ni $\{0\}$ ni $\{1\}$ son medibles para μ.

Ejemplo 2.6
Sea Ω un conjunto no numerable y $\mathcal{F} = \{A \subset \Omega : A \text{ o } A^c \text{ es numerable }\}$ Sea m la función

$$m(A) = \begin{cases} 0 & \text{si } A \text{ es numerable,} \\ 1 & \text{si } A^c \text{ es numerable.} \end{cases}$$

Cualquier B no numerable con B^c no numerable no es medible porque

$$m^*(B) + m^*(B^c) = 1 + 1 \neq m^*(\Omega) = m(\Omega) = 1.$$

Regresemos ahora a considerar la medida de Lebesgue. Decimos que un conjunto $E \subset \mathbb{R}$ es *medible según Lebesgue* o *Lebesgue medible* si $E \in \mathcal{M}(\lambda^*)$. Vamos a demostrar que no todos los subconjuntos de $\mathbb{R}$ son Lebesgue medibles. Una idea posible para demostrar esto sería ver que hay más subconjuntos de $\mathbb{R}$ que conjuntos medibles, en términos de cardinalidad. Sabemos que la cardinalidad de $[0, 1]$ es c y que $\mathcal{P}([0, 1])$ tiene cardinalidad 2^c. ¿Cuál es la cardinalidad de $\mathcal{M}(\lambda^*)$?

Hemos visto que el conjunto de Cantor C es medible y nulo, y como la σ-álgebra $\mathcal{M}$ es completa, todos los subconjuntos de C están en $\mathcal{M}$. Pero también vimos que C tiene la misma cardinalidad de $[0, 1]$, es decir c, y por lo tanto $\mathcal{P}(C)$ tiene cardinalidad 2^c y está contenida en $\mathcal{M}$. En consecuencia $\mathcal{M}$ tiene la misma cardinalidad de $\mathcal{P}([0, 1])$.

Así que para demostrar que hay subconjuntos de $[0, 1]$ no medibles es necesario usar otro argumento. Comenzamos por definir la operación de suma módulo 1 en $(0, 1]$: Para $x, y \in (0, 1]$ definimos

$$x \oplus y = \begin{cases} x + y & \text{si } x + y \in (0, 1], \\ x + y - 1 & \text{si } x + y \notin (0, 1], \end{cases}$$

y $A \oplus x = \{a \oplus x : a \in A\}$. Por las propiedades de la medida de Lebesgue, si $E \in \mathcal{M}$ entonces $E \oplus x \in \mathcal{M}$ y $\lambda(E \oplus x) = \lambda(E)$.

Decimos que x e y son equivalentes ($x \sim y$) si $x \oplus r = y$ para algún racional $r \in (0, 1]$. Es fácil ver que esta es, efectivamente, una relación de equivalencia y por lo tanto define una partición de $(0, 1]$ en clases de equivalencia. Sea H un subconjunto de $(0, 1]$ que contiene exactamente un representante de cada clase de equivalencia. Este conjunto existe por el Axioma de Elección. Consideramos ahora los conjuntos $H \oplus r$ para r racional. Hay una cantidad numerable de ellos.

- Estos conjuntos son disjuntos porque no puede haber en H un par de puntos que sean equivalentes: Si $H \oplus r$ y $H \oplus s$ tienen un punto común $h_1 \oplus r = h_2 \oplus s$, entonces $h_1 \sim h_2$, lo cual es imposible a menos que $h_1 = h_2$, en cuyo caso $r = s$.
- Cada punto de $(0, 1]$ está en alguno de estos conjuntos porque H tiene un representante de cada clase de equivalencia: Si $x \sim h \in H$, entonces $x = h \oplus r \in H \oplus r$ para algún racional r.

Por lo tanto

$$(0, 1] = \bigcup_r H \oplus r$$

que es una unión numerable. Si H fuese medible entonces tendríamos que

$$\lambda(0, 1] = \sum_r \lambda(H \oplus r)$$

lo cual es imposible porque $\lambda(H \oplus r)$ vale lo mismo para todo r. Por lo tanto H no es medible.