

Índice

1. Modelos probabilísticos continuos	2
1.1. Introducción	2
1.2. El problema de la aguja de Buffon	3
1.3. Otro problema de lanzamientos	9
1.4. Ejercicios	11
2. El problema de la medida	13
2.1. Introducción	13
2.2. La longitud en $\mathbb{R}$	14
2.3. σ -álgebras. σ -álgebra de Borel en $\mathbb{R}$	19
2.4. Los conjuntos con longitud	24
2.5. El área en $\mathbb{R}^2$	26
2.6. El volumen en $\mathbb{R}^k$	33
2.7. Ejercicios	34
3. Espacios de probabilidad	36
3.1. Espacio muestral y σ -álgebra de sucesos	36
3.2. Medidas de probabilidad	38
3.3. Conjuntos de probabilidad cero o uno	44
3.4. Probabilidad condicionada	45
3.5. Independencia	49
3.6. Ejercicios	53
4. Probabilidades en $\mathbb{R}$	55
4.1. Introducción	55
4.2. Funciones de distribución	55
4.3. Distribuciones en un subconjunto de $\mathbb{R}$	60
4.4. Tipos de distribuciones	64
4.4.1. Distribuciones discretas	66
4.4.2. Distribuciones absolutamente continuas	68
4.4.3. Distribuciones singulares	77
4.4.4. Distribuciones mixtas	79
4.5. Ejercicios	80
5. Variables aleatorias	84
5.1. Definición	84
5.2. Distribución de una variable aleatoria	88
5.3. Tipos de variables aleatorias	94

5.4. Cambios de variable	95
5.4.1. Cambios de variable discreta	97
5.4.2. Cambios de variable absolutamente continua	98
5.5. Ejercicios	103
6. Independencia de variables aleatorias	106
6.1. Introducción	106
6.2. Variables aleatorias independientes	107
6.3. Espacios producto	111
6.4. Suma de variables aleatorias independientes	117
6.5. Ejercicios	121
7. Probabilidades en $\mathbb{R}^2$	124
7.1. Introducción	124
7.2. Funciones de distribución bidimensionales	124
7.3. Distribuciones en un subconjunto de $\mathbb{R}^2$	131
7.4. Tipos de distribuciones bidimensionales	132
7.4.1. Distribuciones discretas	132
7.4.2. Distribuciones singulares	134
7.4.3. Distribuciones absolutamente continuas	136
7.4.4. Distribuciones mixtas	142
7.5. Ejercicios	143
8. Variables aleatorias bidimensionales	145
8.1. Definición	145
8.2. Distribución de probabilidad	145
8.3. Cambios de variable	149
8.4. Ejercicios	155
9. Distribuciones marginales y condicionadas	157
9.1. Introducción	157
9.2. Distribuciones marginales	157
9.3. Distribuciones condicionadas	161
9.4. La cuestión inversa	168
9.5. El caso de independencia	179
9.6. Ejercicios	182
10. Probabilidades y variables aleatorias en $\mathbb{R}^k$	185
10.1. Introducción	185
10.2. Distribuciones de probabilidad en $\mathbb{R}^k$	185
10.2.1. Distribuciones discretas	187
10.2.2. Distribuciones absolutamente continuas	188
10.2.3. Distribuciones singulares	189
10.2.4. Distribuciones mixtas	190
10.3. Variables aleatorias k -dimensionales	190
10.4. Distribuciones marginales	193

10.5. Distribuciones condicionadas	195
10.6. Composición de marginales y condicionadas	198
10.7. El caso de independencia	200
10.8. Ejercicios	201
11. Esperanza matemática	204
11.1. Valor esperado de una variable aleatoria	204
11.2. Propiedades de la esperanza matemática	212
11.3. Esperanza condicionada	215
11.4. Ejercicios	220
12. Análisis descriptivo de las distribuciones en $\mathbb{R}$	223
12.1. Introducción	223
12.2. Momentos de una distribución	224
12.3. Medidas de posición y dispersión	230
12.3.1. Indicadores de posición	230
12.3.2. Indicadores de dispersión	231
12.4. Ejercicios	233
13. Análisis descriptivo de las distribuciones en $\mathbb{R}^k$	234
13.1. Introducción	234
13.2. Momentos de una distribución k -dimensional	235
13.3. Regresión bidimensional	240
13.4. Regresión k -dimensional	246
13.5. Ejercicios	252
14. Función característica	254
14.1. Introducción	254
14.2. Variables aleatorias complejas	255
14.3. Función característica	255
14.3.1. Teoremas de inversión	258
14.3.2. Relación con los momentos	261
14.3.3. Tipos de funciones características	264
14.4. Función característica k -dimensional	267
14.5. Ejercicios	269
15. Distribución Normal k-dimensional	271
15.1. Introducción	271
15.2. Densidad normal y normal singular	271
15.3. Marginales, independencia y condicionales	275
15.4. Ejercicios	276
16. Convergencia de variables aleatorias	278
16.1. Introducción	278
16.2. Convergencia casi segura y en probabilidad	278
16.3. Convergencia en media de orden p	281

16.4. Convergencia en distribución	284
16.5. Variables aleatorias k -dimensionales	293
16.6. Ejercicios	294
17. Teorema central del límite	297
17.1. Introducción	297
17.2. El caso de sumandos igualmente distribuidos	299
17.3. El caso general	304
17.4. El caso k -dimensional	312
17.5. Ejercicios	314
18. Leyes de los grandes números	315
18.1. Introducción	315
18.2. Leyes débiles	317
18.3. Leyes fuertes	319
18.4. Series de variables aleatorias independientes	325
18.4.1. El caso simétrico	325
18.4.2. El caso general	330
18.4.3. El orden de crecimiento de las sumas	333
18.5. Ejercicios	336
A. Solución de los ejercicios	340
Capítulo 1	340
Capítulo 2	350
Capítulo 3	357
Capítulo 4	370
Capítulo 5	391
Capítulo 6	410
Capítulo 7	425
Capítulo 8	437
Capítulo 9	447
Capítulo 10	460
Capítulo 11	477
Capítulo 12	487
Capítulo 13	494
Capítulo 14	506
Capítulo 15	515
Capítulo 16	519
Capítulo 17	530
Capítulo 18	535
Tabla 1: Distribución Normal(0,1)	550
Tabla 2: Distribuciones notables	551
Índice alfabético	554

Capítulo 2

El problema de la medida

2.1. Introducción

Longitud, área y volumen son conceptos análogos cuya finalidad es evaluar el tamaño de conjuntos de una, dos o tres dimensiones, es decir de conjuntos que pueden considerarse contenidos en $\mathbb{R}^1$, $\mathbb{R}^2$ y $\mathbb{R}^3$ respectivamente. En los tres casos, los propósitos y las ideas son análogos y permiten además tratar en forma similar el caso de un espacio euclídeo arbitrario, $\mathbb{R}^k$.

Las nociones de área y volumen que se remontan a la matemática griega en el caso de figuras geométricas, se vieron pronto ligadas con la idea de integral, surgida precisamente para calcular el área bajo una curva o el volumen limitado por una superficie. Así continuaron hasta finales del siglo XIX, cuando el desarrollo del Análisis matemático hizo necesaria la integración de funciones con una gran cantidad de discontinuidades para las que no sirven las definiciones de integral dadas por Cauchy o por Riemann. Ello motivó un cambio radical, a partir del cual el área y el volumen, junto con la longitud (tradicionalmente ignorada por no requerir técnicas de cálculo específicas), se revelaron como conceptos básicos que son imprescindibles para perfeccionar el concepto de integral, en vez de estar supeditados a él. Estas innovaciones, introducidas por Borel y a las que dio forma definitiva Lebesgue en la primera década del siglo XX, constituyen el núcleo de lo que se denomina *teoría de la medida* y, en la actualidad, longitud, área y el volumen se denominan *medidas de Lebesgue* en los espacios $\mathbb{R}^1$, $\mathbb{R}^2$ y $\mathbb{R}^3$ respectivamente.

Sobre espacios muestrales continuos, la probabilidad comparte muchas de las propiedades comunes a los conceptos de longitud, área o volumen; incluso en ocasiones (en modelos uniformes del tipo considerado en el capítulo anterior) la probabilidad puede coincidir, a menos de una constante de proporcionalidad, con la medida de Lebesgue sobre el espacio muestral. De ahí el interés, en este momento, de prestar atención a algunos de los problemas propios de la *teoría de la medida* para prever las dificultades que surgen en los modelos probabilísticos continuos.

2.2. La longitud en $\mathbb{R}$

Puesto que las complicaciones crecen al aumentar la dimensión, consideremos en primer lugar la idea de longitud. En esencia, ¿qué es la longitud? Simplemente, una forma de asignar un número no negativo a ciertos conjuntos de $\mathbb{R}$ (a todos si fuese posible) con tres propiedades fundamentales:

- (a) La longitud de un intervalo de extremos a y b , independientemente de que sea abierto, cerrado o semiabierto, es $|b - a|$.
- (b) Si A_1 y A_2 son dos subconjuntos disjuntos de $\mathbb{R}$ con longitud definida, su unión tiene como longitud la suma de las longitudes de cada uno de ellos. Es decir, si $\lambda(A)$ designa la longitud de un conjunto A , se cumple

$$\lambda(A_1 \cup A_2) = \lambda(A_1) + \lambda(A_2)$$

supuesto que $A_1 \cap A_2 = \emptyset$.

- (c) Si $\{A_n\}$ es una sucesión creciente de subconjuntos de $\mathbb{R}$, con longitud definida, y $A = \bigcup_n A_n$ entonces

$$\lambda(A) = \lim_{n \rightarrow \infty} \lambda(A_n).$$

La segunda propiedad, denominada *aditividad* de la longitud, se extiende de manera inmediata al caso de un número finito de conjuntos *disjuntos*, para los que se cumple:

$$\lambda(A_1 \cup A_2 \cup \dots \cup A_n) = \lambda(A_1) + \lambda(A_2) + \dots + \lambda(A_n). \quad (2.1)$$

La aditividad y la continuidad para sucesiones crecientes que expresa la condición (c), son propiedades que la longitud comparte con cualquier función de probabilidad. Aunque aquí interesa considerarlas por separado, pueden agruparse en una sola condición: que la igualdad (2.1) se verifique también para cualquier sucesión de conjuntos disjuntos (cf. proposiciones 3.3 y 3.4).

La longitud tiene una cuarta propiedad, muy relacionada con (b), que resulta de gran utilidad:

- (d) Dados dos conjuntos $A_2 \subset A_1$ con longitudes $\lambda(A_2)$ y $\lambda(A_1)$ respectivamente, el conjunto $A_1 - A_2$ tiene longitud $\lambda(A_1) - \lambda(A_2)$ siempre que la longitud $\lambda(A_2)$ no sea infinita¹. Es decir

$$\lambda(A_1 - A_2) = \lambda(A_1) - \lambda(A_2). \quad (2.2)$$

Nótese que si $A_1 - A_2$ tiene longitud $\lambda(A_1 - A_2)$, como $A_1 = A_2 \cup (A_1 - A_2)$, se sigue de (b) que $\lambda(A_1) = \lambda(A_2) + \lambda(A_1 - A_2)$ de donde resulta (2.2). De esta manera, la condición (d) sólo exige atribuir a $A_1 - A_2$ (cuando $\lambda(A_2) < \infty$) la única longitud permitida por (b).

Sin más, estas simples observaciones sobre lo que significa la longitud, permiten construir ejemplos de gran interés.

¹ Hay que exceptuar el caso en que $\lambda(A_2) = \infty$, puesto que entonces $\lambda(A_1) = \infty$ y la afirmación $\lambda(A_1 - A_2) = \infty - \infty$ carece de sentido. Tal es el caso, por ejemplo, si $A_2 = [a, \infty)$.

Ejemplo 2.1

El conjunto de Cantor. La intensa actividad matemática de Cantor (1845 – 1918) estuvo relacionada sobre todo con la cardinalidad de los conjuntos infinitos. Probó, por ejemplo, que el conjunto $\mathcal{Q}$ de los números racionales es numerable, mientras que no lo es el conjunto $\mathbb{R}$ de los números reales. Estableció también que $\mathbb{R}$, $\mathbb{R}^2$ y, en general, $\mathbb{R}^k$ son biyectivos y tienen, por tanto, el mismo cardinal: el cardinal del continuo, $\aleph_1$.

En el curso de estas investigaciones, Cantor ideó un subconjunto del intervalo $[0, 1]$ de la recta real con propiedades notables. Su construcción se muestra en la figura 2.1 y consiste en suprimir del intervalo $[0, 1]$ el tercio central, conservando los dos tercios extremos; a continuación, de ambos intervalos se elimina el tercio central para formar en total cuatro intervalos; en el siguiente paso, se detrae el tercio central de estos cuatro segmentos y así sucesivamente.

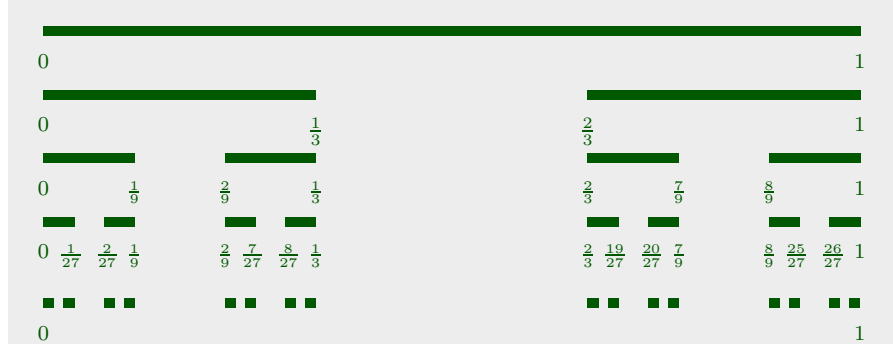


Figura 2.1: El conjunto de Cantor

Más exactamente, si

$$\begin{aligned} A_1 &= \left(\frac{1}{3}, \frac{2}{3} \right) \\ A_2 &= \left(\frac{1}{9}, \frac{2}{9} \right) \cup \left(\frac{7}{9}, \frac{8}{9} \right) \\ A_3 &= \left(\frac{1}{27}, \frac{2}{27} \right) \cup \left(\frac{7}{27}, \frac{8}{27} \right) \cup \left(\frac{19}{27}, \frac{20}{27} \right) \cup \left(\frac{25}{27}, \frac{26}{27} \right) \dots \end{aligned}$$

sea $A = \bigcup_{n=1}^{\infty} A_n$. El conjunto de Cantor es la diferencia

$$C = [0, 1] - A = [0, 1] - \bigcup_{n=1}^{\infty} A_n.$$

Como unión de intervalos abiertos contenidos en $[0, 1]$, A es un conjunto abierto distinto del intervalo cerrado $[0, 1]$. De manera que C es un conjunto cerrado y no vacío. Por lo que se refiere a las longitudes, es

$$\lambda(A_1) = \frac{1}{3}, \quad \lambda(A_2) = \frac{2}{9}, \quad \lambda(A_3) = \frac{4}{27}, \quad \dots \quad \text{y, en general, } \lambda(A_n) = \frac{2^{n-1}}{3^n}.$$

Al crecer N , $\bigcup_{n=1}^N A_n$ crece hacia A luego, de acuerdo con (b) y (c), se tiene

$$\lambda(A) = \lim_{N \rightarrow \infty} \lambda\left(\bigcup_{n=1}^N A_n\right) = \lim_{N \rightarrow \infty} \sum_{n=1}^N \lambda(A_n) = \sum_{n=1}^{\infty} \frac{2^{n-1}}{3^n} = \frac{1/3}{1 - 2/3} = 1.$$

En consecuencia, según (d),

$$\lambda(C) = 1 - \lambda(A) = 0.$$

El hecho de que la longitud de C sea cero indica que C no contiene ningún intervalo abierto, pues si contuviese al intervalo abierto I tendría que ser

$$\lambda(C) = \lambda(I) + \lambda(C - I) \geq \lambda(I) > 0.$$

Así, A es denso en el intervalo $[0, 1]$, puesto que cualquier subintervalo I , por pequeño que sea, contiene puntos de A . Pero esto no significa que C tenga pocos puntos. En efecto, cada número real $x \in [0, 1]$ admite un desarrollo en base 3

$$x = 0'a_1a_2a_3 \dots a_i \dots = \sum_{i=1}^{\infty} \frac{a_i}{3^i}$$

donde cada cifra a_i puede ser 0, 1 ó 2. Así por ejemplo, en base 3,

$$\frac{1}{3} = 0'1, \quad \frac{2}{3} = 0'2, \quad \frac{2}{9} = 0'02, \quad \frac{7}{9} = 0'21, \quad \frac{8}{27} = 0'022, \quad \text{etc.}$$

si bien este tipo de números que admiten un desarrollo finito, son precisamente aquellos que admiten otro infinito:

$$\frac{1}{3} = 0'0222 \dots, \quad \frac{2}{9} = 0'01222 \dots, \quad \frac{7}{9} = 0'20222 \dots, \quad \frac{8}{27} = 0'021222 \dots, \quad \text{etc.}$$

El intervalo A_1 está compuesto por todos los números x cuya expresión en base 3 es de la forma $0'1 \dots$; A_2 contiene los números x cuya desarrollo comienza por $0'01 \dots$ o por $0'21 \dots$; las cuatro componentes de A_3 se caracterizan respectivamente por desarrollos de la forma $0'001 \dots$, $0'021 \dots$, $0'201 \dots$ y $0'221 \dots$; y así sucesivamente. Por consiguiente, A está formado por todos los números x cuyo desarrollo en base 3 contiene algún 1, mientras que $C = [0, 1] - A$ está constituido por los números que tienen un desarrollo en el que sólo intervienen las cifras 0 y 2. Esto es

$$C = \left\{ x \in [0, 1] \mid x = \sum_{i=1}^{\infty} \frac{a_i}{3^i} \text{ donde } a_i = 0 \text{ ó } 2 \right\}.$$

Según ello, es inmediato establecer una correspondencia biyectiva entre C y todo el intervalo $[0, 1]$. Basta asociar a cada elemento:

$$x = \sum_{i=1}^{\infty} \frac{a_i}{3^i} \in C \longrightarrow \varphi(x) = \sum_{i=1}^{\infty} \frac{a_i/2}{2^i}$$

es decir sustituir, en el desarrollo en base 3 de x , los doses por unos e interpretarlo como un desarrollo en base 2. Desde luego, φ es una aplicación inyectiva y sobre el intervalo $[0, 1]$, ya que cualquier elemento $x \in [0, 1]$ admite un desarrollo en base 2

$$x = \sum_{i=1}^{\infty} \frac{b_i}{2^i} \quad \text{donde } b_i = 0 \text{ ó } 1$$

que proviene del punto de C

$$\varphi^{-1}(x) = \sum_{i=1}^{\infty} \frac{2b_i}{3^i}.$$

Queda así establecido que C tiene el mismo cardinal $\aleph_1$ que todo el intervalo $[0, 1]$. La conclusión, aparentemente paradójica, es que todos los puntos de un intervalo de longitud 1, se pueden reubicar dentro del intervalo (según la transformación φ^{-1}) de manera que sus imágenes ocupen un conjunto C de longitud cero.

Tal hecho puede resultar menos sorprendente si se observa que todos los intervalos de $\mathbb{R}$ tienen el mismo cardinal, dado que la aplicación lineal

$$\psi(x) = \frac{d-c}{b-a}(x-a) + c$$

transforma de forma biunívoca el intervalo (a, b) en el intervalo (c, d) , pese a que las longitudes de ambos pueden ser muy diferentes. De forma similar, la aplicación $\tau(x) = ae^{-x}$ es una biyección del intervalo infinito $[0, \infty)$ en el intervalo $(0, a]$, donde a puede ser tan pequeño como se quiera.

Es inmediato que los conjuntos numerables tienen longitud cero; así se establece en el siguiente ejemplo para el caso concreto de los racionales del intervalo $[0, 1]$.

Ejemplo 2.2

Es bien sabido que, dentro del intervalo $[0, 1]$, el conjunto de los números racionales $\mathcal{Q} \cap [0, 1]$ es un conjunto numerable, que puede escribirse en forma de sucesión $\{r_n\}_{n \in \mathbb{N}}$. Puesto que cada racional r es un intervalo cerrado de longitud cero, se sigue de (b) y (c) que

$$\lambda(\mathcal{Q} \cap [0, 1]) = \lim_{N \rightarrow \infty} \lambda(\{r_n \mid n \leq N\}) = \lim_{N \rightarrow \infty} \sum_{n=1}^N \lambda(\{r_n\}) = 0$$

de manera que el conjunto de los puntos racionales no ocupa ninguna longitud. El argumento anterior puede reforzarse mostrando que, para cualquier $\varepsilon > 0$ arbitrariamente pequeño, existe un subconjunto abierto A_ε de $[0, 1]$ que contiene a $\mathcal{Q} \cap [0, 1]$ y cuya longitud es inferior a ε . En efecto, se puede suponer que ε es irracional, porque en caso contrario puede sustituirse por un irracional menor que ε . Sea $\varepsilon_1 = \varepsilon/4$ y consideremos el intervalo abierto $I_1 = (r_1 - \varepsilon_1, r_1 + \varepsilon_1)$.

Sea r'_2 el primer racional de la sucesión no incluido en I_1 , su distancia d_2 a I_1 tiene un valor irracional (no nulo), al igual que $\varepsilon_2 = \min\{\varepsilon/2^3, d_2\}$; formemos entonces el intervalo abierto $I_2 = (r'_2 - \varepsilon_2, r'_2 + \varepsilon_2)$.

Existirá después un primer racional r'_3 no incluido en $I_1 \cup I_2$, situado a una distancia irracional d_3 de dicho conjunto, sea $\varepsilon_3 = \min\{\varepsilon/2^4, d_3\}$ y considérese el intervalo $I_3 = (r'_3 - \varepsilon_3, r'_3 + \varepsilon_3)$.

Esta regla de formación, continuada indefinidamente, produce una sucesión $\{I_n\}$ de intervalos abiertos, disjuntos (puesto que el centro de cada uno dista de los anteriores una cantidad superior a su radio). Su unión

$$A_\varepsilon = \bigcup_{n=1}^{\infty} I_n$$

es un conjunto abierto que contiene a todos los racionales de $[0, 1]$ (porque el racional r_n o bien está en uno de los intervalos I_k con $k < n$ o es centro de su propio intervalo). Además, A_ε tiene longitud total

$$\lambda(A_\varepsilon) = \sum_{n=1}^{\infty} \lambda(I_n) = \sum_{n=1}^{\infty} 2\varepsilon_n \leq 2 \sum_{n=1}^{\infty} \frac{\varepsilon}{2^{n+1}} = \varepsilon.$$

El resultado es sorprendente habida cuenta que $\mathcal{Q} \cap [0, 1]$ es denso en $[0, 1]$, pero el razonamiento no deja lugar a dudas: existe un recubrimiento mediante intervalos abiertos de los racionales del intervalo $[0, 1]$ cuyo complementario, compuesto exclusivamente por números irracionales, tiene una longitud superior a $1 - \varepsilon$. Dado que ε puede ser arbitrariamente pequeño, queda confirmado que

$$\lambda(\mathcal{Q} \cap [0, 1]) = 0 \quad \text{y} \quad \lambda([0, 1] - \mathcal{Q}) = 1.$$

Los ejemplos anteriores muestran la utilidad del concepto de longitud; pero también apuntan la complejidad que pueden tener ciertos subconjuntos de la recta real. Cabe entonces poner en duda la idea ingenua de que todos los subconjuntos de $\mathbb{R}$ deben tener una longitud definida.

De hecho, es posible construir un conjunto al que no se puede asignar longitud, si se admite que *al trasladar un conjunto de longitud definida su longitud no se altera*. Esta propiedad intuitivamente evidente puede deducirse de (a), (b) y (c), una vez asignada una longitud a todos aquellos conjuntos para los que es posible hacerlo (cf. Prop. 2.9). En cualquier caso, tal afirmación da sentido a la siguiente construcción.

Ejemplo 2.3

Un conjunto sin longitud. Entre los puntos del intervalo $[0, 1]$ puede definirse la relación de equivalencia

$$x \sim y \quad \text{si y sólo si} \quad x - y \in \mathcal{Q}$$

que relaciona dos números x e y cuando su diferencia es un número racional. La clase de equivalencia de cada valor x es

$$A_x = \{y \in [0, 1] \mid y = x + r \text{ con } r \in \mathcal{Q}\}$$

y contiene un número numerable de puntos. Como se sabe, la colección $\{A_i\}_{i \in I}$ de todas las clases de equivalencia distintas es una partición de $[0, 1]$; esto es

$$[0, 1] = \bigcup_{i \in I} A_i.$$

De ello, puesto que $[0, 1]$ no es numerable, se deduce que tiene que haber una cantidad no numerable de clases diferentes (es decir, que I no es numerable). Ahora bien, de todas formas, el *axioma de elección* asegura la existencia un conjunto $W \subset [0, 1]$ formado por un único punto de cada A_i .

Si se somete el conjunto W a traslaciones de cada magnitud $r \in \mathcal{Q} \cap [-1, 1]$ se obtiene una colección numerable de conjuntos

$$W_r = \{r + w \mid w \in W\}$$

contenidos en $[-1, 2]$ (pues $-1 + 0 \leq r + w \leq 1 + 1$), de la misma longitud y disjuntos entre sí (ya que si W_r y $W_{r'}$ tuviesen un punto en común, x , sería $x = r + w = r' + w'$ para ciertos $w, w' \in W$; con lo cual $w - w' = r' - r$ sería un número racional y $w \sim w'$ en contra de la hipótesis de que W sólo contiene un punto de cada A_i).

Además, cualquier $x \in [0, 1]$ es equivalente al único elemento $w \in W$ escogido de la clase de equivalencia de x ; de forma que $x = r + w$ para un cierto $r \in \mathcal{Q} \cap [-1, 1]$ (como $r = x - w$ es, sin duda, $-1 \leq r \leq 1$). En resumen, el conjunto

$$\mathcal{W} = \bigcup_{r \in \mathcal{Q} \cap [-1, 1]} W_r \quad \text{verifica} \quad [0, 1] \subset \mathcal{W} \subset [-1, 2].$$

Si se atribuye a W longitud no nula, $\lambda(W) = l$, resulta $\lambda(\mathcal{W}) = \infty$ (puesto que $\mathcal{W}$ es unión de un número numerable de conjuntos disjuntos, todos ellos de longitud l). Lo cual conduce a la contradicción de que el intervalo $[-1, 2]$, de longitud 3, contiene un conjunto de longitud infinita.

En cambio, si se toma $\lambda(W) = 0$, es $\lambda(\mathcal{W}) = 0$ y el intervalo $[0, 1]$, de longitud 1, resulta estar contenido en un conjunto de longitud nula.

En definitiva, no es posible asignar al conjunto W ninguna longitud, ni nula ni positiva, sin caer en contradicción con las propiedades que debe tener la longitud.

Puede entenderse ahora cual es el verdadero *problema de la medida* resuelto por Borel y Lebesgue. Puesto que entre los subconjuntos de $\mathbb{R}$ hay muchos *medibles*, en el sentido de que se les puede asignar longitud, y otros que no lo son, ¿qué caracterización puede darse de los primeros? y ¿cuál es el mecanismo adecuado para determinar su medida?

Dar respuesta a ambas preguntas es uno de los objetivos básicos de la *teoría de la medida* y supone un trabajo considerable. Aquí, sólo se describirán algunos de los aspectos más importantes de la solución lo cual requiere, pese a todo, de cierta terminología introducida en la próxima sección.

2.3. σ -álgebras. σ -álgebra de Borel en $\mathbb{R}$

El objetivo de esta sección es analizar la estructura que pueden tener algunas familias de subconjuntos de un conjunto y las propiedades correspondientes. El

motivo es que, tanto en *teoría de la medida* como en *cálculo de probabilidades*, juega un papel importante la estructura de σ -álgebra:

Definición 2.1 Dado un conjunto arbitrario Ω , una familia $\mathcal{F}$ de subconjuntos de Ω tal que

1. para cualquier $A \in \mathcal{F}$ es $A^c \in \mathcal{F}$,
2. si $\{A_n\}$ es una colección numerable² de subconjuntos de Ω que verifican $A_n \in \mathcal{F}$ para cada n , se cumple $\bigcup_n A_n \in \mathcal{F}$,

recibe el nombre de **σ -álgebra** de subconjuntos de Ω .

La definición anterior tiene algunas consecuencias inmediatas.

Proposición 2.1 Si $\mathcal{F}$ es una σ -álgebra de subconjuntos de Ω , se cumple

1. $\Omega \in \mathcal{F}$ y $\emptyset \in \mathcal{F}$.
2. Si $\{A_n\}$ es una colección numerable de subconjuntos de Ω tal que $A_n \in \mathcal{F}$ para cada n , se cumple $\bigcap_n A_n \in \mathcal{F}$.

• En efecto, dado cualquier $A \in \mathcal{F}$ es $A^c \in \mathcal{F}$ y, por tanto, $\Omega = A \cup A^c \in \mathcal{F}$. Luego, también, $\emptyset = \Omega^c \in \mathcal{F}$. Además, es bien conocida la ley de De Morgan:

$$\bigcap_n A_n = \left(\bigcup_n A_n^c \right)^c.$$

En caso de que los conjuntos A_n pertenezcan a $\mathcal{F}$, lo mismo ocurre con sus complementarios, A_n^c , y con $\bigcup_n A_n^c$. En consecuencia, $\bigcap_n A_n \in \mathcal{F}$. ■

La definición y las propiedades anteriores muestran que las σ -álgebras son aquellas clases de subconjuntos de Ω , cerradas por cualquiera de las operaciones conjuntistas (unión, intersección y complementario) siempre que se realicen con un número numerable de términos; en el sentido de que cualquier operación llevada a cabo con conjuntos que pertenecen a $\mathcal{F}$, sigue siendo un elemento de $\mathcal{F}$.

Ejemplo 2.4

Dentro de un conjunto cualquiera Ω , existen muchas σ -álgebras de mayor o menor interés.

La más grande es, sin duda, la familia $\mathcal{P}(\Omega)$ de todos los subconjuntos de Ω que se denomina a veces la *σ -álgebra discreta*.

La más pequeña, es la colección $\{\emptyset, \Omega\}$, formada exclusivamente por el conjunto vacío y el conjunto total que suele denominarse la *σ -álgebra trivial*.

Para cada conjunto $A \subset \Omega$, existe una σ -álgebra mínima que lo contiene; es concretamente

$$\{\emptyset, A, A^c, \Omega\}.$$

De forma similar, dado otro subconjunto B de Ω , que no sea disjunto con A ni con A^c , la menor σ -álgebra que contiene a A y a B es

$$\left\{ \emptyset, A, B, A^c, B^c, A \cap B, A \cap B^c, A^c \cap B, A^c \cap B^c, A \cup B, A \cup B^c, A^c \cup B, A^c \cup B^c, (A \cap B) \cup (A^c \cap B^c), (A \cap B^c) \cup (A^c \cap B), \Omega \right\}.$$

² Una colección o familia de términos es *numerable* si puede ponerse en correspondencia biyectiva con un subconjunto de $\mathbb{N}$. En particular, por tanto, incluye el caso de una colección finita de términos.

(Los conjuntos A y B dividen el espacio en cuatro piezas: $A \cap B$, $A \cap B^c$, $A^c \cap B$ y $A^c \cap B^c$; entre los conjuntos de la σ -álgebra hay 6 formados por dos piezas: A , A^c , B , B^c , $(A \cap B) \cup (A^c \cap B^c)$ y $(A \cap B^c) \cup (A^c \cap B)$; más otros 4 formados por tres piezas: $A \cup B$, $A \cup B^c$, $A^c \cup B$ y $A^c \cup B^c$. En total 2^4 conjuntos.

Al añadir más conjuntos la complejidad de la σ -álgebra crece a un ritmo vertiginoso; con n conjuntos cualesquiera puede constar hasta de $2^{(2^n)}$ elementos.

Puesto que hay múltiples σ -álgebras en cualquier conjunto, tiene interés la siguiente propiedad.

Proposición 2.2 Sea $\{\mathcal{F}_i\}_{i \in I}$ una colección arbitraria de σ -álgebras en un conjunto Ω , entonces

$$\mathcal{F} = \bigcap_{i \in I} \mathcal{F}_i$$

es una σ -álgebra de subconjuntos de Ω .

• En efecto, si $A \in \mathcal{F}$, es $A \in \mathcal{F}_i$ para cada $i \in I$ y, por consiguiente, $A^c \in \mathcal{F}_i$ para cada $i \in I$. Luego, $A^c \in \mathcal{F}$.

Análogamente, si $\{A_n\} \subset \mathcal{F}$, será $\{A_n\} \subset \mathcal{F}_i$ para cada $i \in I$; o bien, puesto que cada $\mathcal{F}_i$ es σ -álgebra, $\bigcup_n A_n \in \mathcal{F}_i$ para cada $i \in I$. Con lo cual, $\bigcup_n A_n \in \mathcal{F}$. ■

El resultado anterior permite dar la siguiente definición.

Definición 2.2 Dada una familia cualquiera $\mathcal{C}$ de subconjuntos de un conjunto Ω , la intersección de todas las σ -álgebras que contienen a $\mathcal{C}$:

$$\sigma(\mathcal{C}) = \bigcap_{\mathcal{F} \supset \mathcal{C}} \mathcal{F}$$

es la mínima σ -álgebra que contiene a $\mathcal{C}$, ya que está contenida en cualquier σ -álgebra $\mathcal{F}$ que contenga a $\mathcal{C}$. Se denomina a $\sigma(\mathcal{C})$ la σ -álgebra engendrada por $\mathcal{C}$.

En el ejemplo 2.4 se ha construido directamente la σ -álgebra engendrada por una familia de subconjuntos $\mathcal{C}$ que esté formada por un único subconjunto, A , o por dos de ellos, A y B . Pero, en el caso $\Omega = \mathbb{R}$, la definición tiene sobre todo interés cuando la familia $\mathcal{C}$ está formada por todos los intervalos.

Definición 2.3 Sea $\mathcal{I}$ la familia constituida por todos los intervalos de $\mathbb{R}$ (abiertos, cerrados o semiabiertos, finitos o infinitos). La σ -álgebra $\sigma(\mathcal{I})$, engendrada por $\mathcal{I}$, se denomina la σ -álgebra de Borel en $\mathbb{R}$ y suele denotarse por $\mathbb{B}$. Los elementos de $\mathbb{B}$ se denominan conjuntos de Borel.

No existe una caracterización explícita de los conjuntos de Borel. La propia definición indica que es de Borel cualquier conjunto que pueda formarse mediante una secuencia de operaciones conjuntistas a partir de un número numerable de intervalos. Así que lo difícil es imaginar un conjunto que no pertenezca a la σ -álgebra de Borel. Por ejemplo, se cumple:

Proposición 2.3 Cualquier conjunto abierto de $\mathbb{R}$ es un conjunto de Borel. Lo mismo sucede con cualquier conjunto cerrado, con cualquier intersección numerable de abiertos, con cualquier unión numerable de cerrados, etc.

- En efecto, es conocido que cualquier abierto A de $\mathbb{R}$ se puede expresar como unión numerable de intervalos abiertos disjuntos³. Por consiguiente, cualquier abierto es un conjunto de Borel.

Así mismo, si A es cerrado, A^c es un abierto y, por consiguiente, de Borel. Lo mismo sucede entonces con su complementario, A . ■

En ocasiones, se plantean problemas relativos a la longitud que no involucran a toda la recta real sino a un subconjunto fijo, A , que será normalmente un intervalo o un conjunto de Borel. En tal caso, los elementos de $\mathbb{B}$ disjuntos con A son inútiles y por ello se considera la “parte” de la σ -álgebra contenida en A . En general, la idea responde a la siguiente definición.

Definición 2.4 Sea $\mathcal{F}$ una σ -álgebra de subconjuntos de Ω y A un subconjunto cualquiera de Ω , la familia

$$\mathcal{F}_A = \{ B \subset A \mid B = A \cap C \text{ para algún } C \in \mathcal{F} \}$$

recibe el nombre de **σ -álgebra restringida** de $\mathcal{F}$ a A . Si $\Omega = \mathbb{R}$ y $\mathcal{F}$ es la σ -álgebra de Borel, $\mathbb{B}$, se denomina a $\mathbb{B}_A$ la σ -álgebra de Borel en A .

- La comprobación de que $\mathcal{F}_A$ es una σ -álgebra de subconjuntos de A es muy simple; si $B \in \mathcal{F}_A$ es $B = A \cap C$ con $C \in \mathcal{F}$, luego

$$B^c = A - B = A - A \cap C = A \cap (\Omega - C)$$

que pertenece a $\mathcal{F}_A$ por ser $\Omega - C \in \mathcal{F}$. Análogamente, si es $B_n = A \cap C_n$ con $C_n \in \mathcal{F}$, para cada n , será

$$\bigcup_n B_n = A \cap (\bigcup_n C_n)$$

que pertenece a $\mathcal{F}_A$ por ser $\bigcup_n C_n \in \mathcal{F}$. ■

En el caso de la σ -álgebra de Borel, se puede caracterizar a $\mathbb{B}_A$ como la σ -álgebra en A engendrada por las intersecciones de A con los intervalos de $\mathbb{R}$. Así lo indica la proposición siguiente.

Proposición 2.4 $\mathbb{B}_A = \sigma(\{ A \cap I \mid I \text{ es un intervalo de } \mathbb{R} \})$.

- En efecto, sea $\mathcal{F}$ la σ -álgebra en A que figura en el segundo miembro. Puesto que $\mathbb{B}_A$ es una σ -álgebra que (por definición) contiene a las intersecciones $A \cap I$, para cualquier intervalo I de $\mathbb{R}$, será $\mathcal{F} \subset \mathbb{B}_A$ (puesto que $\mathcal{F}$ es la mínima de las σ -álgebras que tienen esta propiedad).

Recíprocamente, la familia

$$\mathcal{F}' = \{ B \in \mathbb{B} \mid A \cap B \in \mathcal{F} \}$$

es una σ -álgebra; porque, si $B \in \mathcal{F}'$, es $A \cap B^c = A - A \cap B \in \mathcal{F}$, luego $B^c \in \mathcal{F}'$; además, si $\{B_n\} \subset \mathcal{F}'$, también $\bigcup_n B_n \in \mathcal{F}'$ puesto que

$$A \cap (\bigcup_n B_n) = \bigcup_n (A \cap B_n) \in \mathcal{F}.$$

³ La idea consiste en definir, para cada $x \in A$, la componente $I_x = (a_x, b_x)$, donde $a_x = \inf\{y \mid (y, x) \subset A\}$ y $b_x = \sup\{y \mid (x, y) \subset A\}$; así A será unión de sus diversas componentes que no pueden ser más de un número numerable (pues cada una contiene racionales diferentes).

Por otra parte, cualquier intervalo I de $\mathbb{R}$ pertenece a $\mathcal{F}'$; luego es $\mathcal{F}' = \mathbb{B}$. Ello significa que $\mathcal{F}$ contiene a la intersección $A \cap B$ de A con cualquier conjunto $B \in \mathbb{B}$; es decir $\mathbb{B}_A \subset \mathcal{F}$. ■

Un teorema sobre clases de conjuntos

La cualidad de las σ -álgebras de ser cerradas por cualquier operación conjuntista, con un número numerable de términos, hace que sean adecuadas para múltiples propósitos. Sin embargo, como no siempre es fácil comprobar que una familia de conjuntos es una σ -álgebra, se hace preciso definir clases de conjuntos con propiedades parciales.

Definición 2.5 Una familia $\mathcal{C}$ de subconjuntos de un conjunto Ω se denomina una π -clase si es $A_1 \cap A_2 \in \mathcal{C}$ para cualquier par $A_1, A_2 \in \mathcal{C}$.

Una familia $\mathcal{G}$ de subconjuntos de Ω se denomina una λ -clase si cumple

1. $\Omega \in \mathcal{G}$.
2. $A_1 - A_2 \in \mathcal{G}$ si $A_1, A_2 \in \mathcal{G}$ y $A_2 \subset A_1$.
3. $\bigcup_n A_n \in \mathcal{G}$ si $\{A_n\} \subset \mathcal{G}$ es una sucesión creciente de conjuntos.

Así pues, las π -clases son sólo cerradas por intersecciones finitas. En cambio las λ -clases, además de contener a Ω , son cerradas por diferencias propias y límites de sucesiones crecientes; ello obliga a que sean cerradas por uniones de conjuntos disjuntos, pero no por uniones de conjuntos arbitrarios (cf. Ej. 2.7). En cierto sentido, las propiedades de las σ -álgebras están repartidas entre las π -clases y las λ -clases, ya que se verifica:

Proposición 2.5 Si una familia $\mathcal{G}$ de subconjuntos de Ω es a la vez π -clase y λ -clase es una σ -álgebra.

• En efecto $A^c = \Omega - A \in \mathcal{G}$ en virtud de (1) y (2). Además, si $A_1, A_2 \in \mathcal{G}$ es

$$(A_1 \cup A_2)^c = A_1^c \cap A_2^c$$

donde $A_1^c \cap A_2^c \in \mathcal{G}$ (por ser π -clase) y, por tanto, $A_1 \cup A_2 \in \mathcal{G}$. Así pues $\mathcal{G}$ es cerrada por uniones finitas.

Entonces, para cualquier sucesión $\{A_n\} \subset \mathcal{G}$, como $\bigcup_{n=1}^N A_n$ es una sucesión creciente (al crecer N) de conjuntos pertenecientes a $\mathcal{G}$, se sigue de (3) que $\bigcup_{n=1}^{\infty} A_n$ pertenece a $\mathcal{G}$. En definitiva, $\mathcal{G}$ es una σ -álgebra. ■

El mismo tipo de razonamiento utilizado para probar la proposición 2.2 establece la afirmación siguiente.

Proposición 2.6 Si $\{\mathcal{G}_i\}_{i \in I}$ es una colección arbitraria de λ -clases en un conjunto Ω , la intersección de todas ellas es también una λ -clase.

Según ello, la intersección de todas las λ -clases en Ω que contienen a una familia cualquiera $\mathcal{C}$ de subconjuntos de Ω es una λ -clase, la más pequeña que contiene a $\mathcal{C}$, que se denomina la λ -clase engendrada por $\mathcal{C}$.

El siguiente resultado es una versión del *teorema de clases monótonas* que constituye una herramienta importante de la teoría de la medida.

Proposición 2.7 Si $\mathcal{C}$ es una π -clase y $\mathcal{G}$ es una λ -clase que contiene a $\mathcal{C}$, entonces $\mathcal{G}$ contiene a la σ -álgebra engendrada por $\mathcal{C}$. Es decir $\sigma(\mathcal{C}) \subset \mathcal{G}$.

• En efecto, sea $\mathcal{H}$ la λ -clase engendrada por $\mathcal{C}$; con lo cual $\mathcal{H} \subset \mathcal{G}$. Según la proposición 2.5, si se prueba que $\mathcal{H}$ es una π -clase, será una σ -álgebra; como contiene a $\mathcal{C}$, se cumplirá $\sigma(\mathcal{C}) \subset \mathcal{H}$ y la demostración estará completa. Para ver que $\mathcal{H}$ es una π -clase, considérese

$$\mathcal{G}_1 = \{ A \subset \Omega \mid A \cap B \in \mathcal{H} \text{ para todo } B \in \mathcal{C} \}.$$

Es fácil comprobar que $\mathcal{G}_1$ es una λ -clase⁴ y, también, que $\mathcal{C} \subset \mathcal{G}_1$ (puesto que $\mathcal{C}$ es una π -clase). Luego $\mathcal{H} \subset \mathcal{G}_1$. Ello significa que

$$A \cap B \in \mathcal{H} \quad \text{cualquiera que sean } A \in \mathcal{H} \text{ y } B \in \mathcal{C}. \quad (2.3)$$

En segundo lugar, consideremos

$$\mathcal{G}_2 = \{ B \subset \Omega \mid B \cap A \in \mathcal{H} \text{ para todo } A \in \mathcal{H} \}.$$

De nuevo es sencilla la comprobación de que $\mathcal{G}_2$ es una λ -clase que, según (2.3), contiene a $\mathcal{C}$; será pues $\mathcal{H} \subset \mathcal{G}_2$. Es decir

$$A \cap B \in \mathcal{H} \quad \text{cualquiera que sean } A \in \mathcal{H} \text{ y } B \in \mathcal{H}.$$

Así pues, $\mathcal{H}$ es una π -clase y el razonamiento está completo. ■

Nótese que $\mathcal{H}$ coincide con $\sigma(\mathcal{C})$ pues, por un lado, $\sigma(\mathcal{C}) \subset \mathcal{H}$ y, por otra parte, como $\sigma(\mathcal{C})$ es λ -clase, se cumple $\mathcal{H} \subset \sigma(\mathcal{C})$.

2.4. Los conjuntos con longitud

No expondremos aquí el método empleado en teoría de la medida para extender la longitud a conjuntos más generales que los intervalos. Sin embargo, algunas de sus consecuencias más importantes tienen un enunciado simple y claro. Concretamente, permite establecer la siguiente conclusión:

Se puede atribuir longitud a los conjuntos de una clase $\mathcal{L}$ de conjuntos, de forma que se verifiquen las propiedades (a), (b), (c) y (d) de la sección 2.2.

Más exactamente, se cumple:

- (a) Cualquier intervalo I pertenece a $\mathcal{L}$ y $\lambda(I) = |b-a|$ si a y b son los extremos de I .
- (b) Si $A_1, A_2 \in \mathcal{L}$ son conjuntos disjuntos entonces $A_1 \cup A_2 \in \mathcal{L}$ y

$$\lambda(A_1 \cup A_2) = \lambda(A_1) + \lambda(A_2).$$

⁴ Por ejemplo, si A_n es una sucesión creciente de conjuntos tales que $A_n \cap B \in \mathcal{H}$ para cualquier $B \in \mathcal{C}$, será

$$\left(\bigcup_n A_n \right) \cap B = \bigcup_n (A_n \cap B) \in \mathcal{H}$$

puesto que $A_n \cap B$ es una sucesión creciente y $\mathcal{H}$ es una λ -clase. El resto de las propiedades se comprueba de forma análoga.

- (c) Si $\{A_n\}$ es una sucesión creciente de conjuntos pertenecientes a $\mathcal{L}$, entonces $A = \bigcup_n A_n \in \mathcal{L}$ y

$$\lambda(A) = \lim_{n \rightarrow \infty} \lambda(A_n).$$

- (d) Si $A_1, A_2 \in \mathcal{L}$ y $A_2 \subset A_1$, entonces $A_1 - A_2 \in \mathcal{L}$ y, supuesto que sea $\lambda(A_2) < \infty$, se cumple

$$\lambda(A_1 - A_2) = \lambda(A_1) - \lambda(A_2).$$

De esta forma, el logro de Lebesgue consistió en probar que el problema de la medida tiene solución: los conjuntos medibles pueden agruparse en una determinada clase de conjuntos $\mathcal{L}$, a cuyos miembros se les puede asignar longitud respetando los postulados (a), (b), (c) y (d). La técnica para definir la longitud no tiene mucho interés práctico pero, en cambio, sí es relevante disponer de información acerca de la clase $\mathcal{L}$ de los conjuntos medibles.

Para ello, cabe observar que las propiedades (a), (c) y (d) establecen que $\mathcal{L}$ es una λ -clase que contiene a la π -clase, $\mathcal{I}$, de los intervalos. La proposición 2.7, asegura entonces que $\mathcal{L}$ contiene a la σ -álgebra de Borel, $\mathbb{B}$.

Este es el sentido de la siguiente proposición que recoge, para el caso de la longitud, las principales conclusiones del *teorema de extensión*, uno de los resultados claves que se demuestra en la teoría de la medida.

Proposición 2.8 *La longitud puede extenderse de manera única a todos los conjuntos de Borel de la recta real.*

• La prueba de la unicidad es sencilla. Sea $K = [-k, k]$ e imaginemos que λ y μ son formas distintas de atribuir longitud a los subconjuntos de Borel de K de manera que se verifiquen, en ambos casos, los postulados (a), (b), (c) y (d). Entonces

$$\mathcal{E} = \{ B \in \mathbb{B}_K \mid \lambda(B) = \mu(B) \}$$

es una λ -clase de subconjuntos de K , como puede comprobarse de forma inmediata⁵. Como $\mathcal{E}$ contiene a la π -clase de los subintervalos de K , coincide con $\mathbb{B}_K$; es decir $\lambda(B)$ y $\mu(B)$ coinciden para cualquier $B \in \mathbb{B}_K$. En virtud de (c), lo mismo ocurre para cualquier $B \in \mathbb{B}$, puesto que

$$\lambda(B) = \lim_{k \rightarrow \infty} \lambda(B \cap K) = \lim_{k \rightarrow \infty} \mu(B \cap K) = \mu(B). \quad \blacksquare$$

Según el resultado anterior la σ -álgebra de Borel adquiere un papel decisivo en relación con el problema de la medida. A pesar de que no sea viable especificar un conjunto de Borel genérico, ni proporcionar un método operativo para calcular la longitud de uno cualquiera de ellos, a efectos teóricos no hay problema de imprecisión al referirse a la longitud de uno de tales conjuntos. A la vez, se trata de una familia de conjuntos lo suficientemente amplia para que nunca tenga

⁵ Por ejemplo, si $A_1, A_2 \in \mathcal{E}$ y $A_2 \subset A_1$ es

$$\lambda(A_1 - A_2) = \lambda(A_1) - \lambda(A_2) = \mu(A_1) - \mu(A_2) = \mu(A_1 - A_2)$$

donde todos los valores son finitos puesto que $A_2 \subset A_1 \subset K$. Las restantes propiedades se comprueban de forma análoga.

interés práctico preocuparse por conjuntos que no sean de Borel. Es más, el conjunto W del ejemplo 2.3 es uno de los pocos conjuntos conocidos que no es de Borel.

Como ejemplo de las ventajas que supone precisar la clase de los conjuntos a los que se atribuye longitud, puede observarse que, en $\mathbb{B}$, la propiedad (d) pasa a ser una consecuencia directa de (b). Ya que, si $A_1, A_2 \in \mathbb{B}$ y $A_2 \subset A_1$, entonces $A_1 - A_2 = A_1 \cap A_2^c \in \mathbb{B}$. Están definidas, por consiguiente, las longitudes $\lambda(A_1)$, $\lambda(A_2)$ y $\lambda(A_1 - A_2)$ y (b) obliga a que sea $\lambda(A_1 - A_2) = \lambda(A_1) - \lambda(A_2)$, supuesto que $\lambda(A_2) < \infty$.

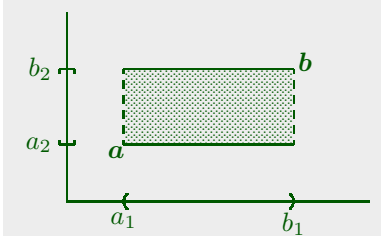
En el mismo sentido, no es difícil establecer ahora (véase Ej. 2.8) la afirmación previa al ejemplo 2.3. Concretamente:

Proposición 2.9 *Dados $A \in \mathbb{B}$ y $x \in \mathbb{R}$, si $x + A = \{x + y \mid y \in A\}$, entonces $x + A \in \mathbb{B}$ y $\lambda(x + A) = \lambda(A)$.*

2.5. El área en $\mathbb{R}^2$

La cuestión de asignar área a los subconjuntos del plano $\mathbb{R}^2$ tiene un planteamiento muy similar al de la longitud en $\mathbb{R}$, con la principal diferencia de que el papel que juegan los intervalos en $\mathbb{R}$, lo desempeñan en $\mathbb{R}^2$ los rectángulos de lados paralelos a los ejes. Para abreviar, sólo recibirán el nombre de rectángulos los conjuntos de la forma

$$(a_1, b_1) \times (a_2, b_2), \quad [a_1, b_1] \times [a_2, b_2], \quad (a_1, b_1] \times (a_2, b_2], \quad \text{etc.}$$



productos de un intervalo (abierto, cerrado o semiabierto, finito o infinito) sobre cada uno de los ejes, como se muestra en la figura adjunta.

Cada rectángulo de forma definida queda determinado por dos puntos: el extremo inferior $\mathbf{a} = (a_1, a_2)$ y el extremo superior $\mathbf{b} = (b_1, b_2)$.

La idea de área se rige por reglas casi idénticas a las empleadas en relación con la longitud:

- (a) Un rectángulo de extremos $\mathbf{a}$ y $\mathbf{b}$ tiene área igual a $(b_1 - a_1)(b_2 - a_2)$, independientemente de qué lados estén incluidos en él.
- (b) Si A_1 y A_2 son conjuntos disjuntos de $\mathbb{R}^2$ con área definida, su unión tiene como área la suma de las áreas de cada uno de ellos; es decir,

$$\lambda_2(A_1 \cup A_2) = \lambda_2(A_1) + \lambda_2(A_2)$$

donde $\lambda_2(A)$ designa el área de A .

- (c) Si $\{A_n\}$ es una sucesión creciente de subconjuntos de $\mathbb{R}^2$ con área definida y $A = \bigcup_n A_n$, se verifica

$$\lambda_2(A) = \lim_{n \rightarrow \infty} \lambda_2(A_n).$$

Tales principios estuvieron presentes desde el principio en el desarrollo de la integral como método de cálculo del área limitada por una función acotada, f , en un intervalo $[a, b]$. De hecho, es bien sabido que el concepto de integral de Riemann utiliza las sumas

$$s(f, P) = \sum_{i=0}^{n-1} (x_{i+1} - x_i) \inf_{x \in [x_i, x_{i+1}]} f(x)$$

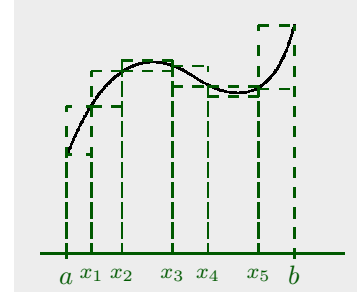
$$S(f, P) = \sum_{i=0}^{n-1} (x_{i+1} - x_i) \sup_{x \in [x_i, x_{i+1}]} f(x)$$

como aproximaciones a dicho área (por defecto y por exceso respectivamente) asociadas a una partición $P = \{x_1, \dots, x_n\}$ del intervalo $[a, b]$.

Ello corresponde a sumar las áreas de las dos series de rectángulos disjuntos que se muestran en la figura. Si, al variar la partición P , coinciden

$$\sup_P s(f, P) \quad \text{y} \quad \inf_P S(f, P)$$

el valor común define la integral de Riemann, $\int_a^b f(x) dx$.



No obstante, la integral de Riemann no es la mejor utilización posible de los principios que a los que obedece la idea de área, porque su existencia requiere de ciertas propiedades de continuidad del integrando. Así lo pone de relieve el siguiente ejemplo.

Ejemplo 2.5

Una función definida en el intervalo $[0, 1]$ que no es continua en ningún punto es la función de Dirichlet:

$$f(x) = \begin{cases} 1 & \text{si } x \in \mathcal{Q} \cap [0, 1] \\ 0 & \text{si } x \in [0, 1] - \mathcal{Q} \end{cases}$$

que vale 1 en los puntos racionales y 0 en los irracionales.

Desde luego, sea cual sea la partición P del intervalo $[0, 1]$ que se considere, en cada intervalo de la partición es $\inf f(x) = 0$ y $\sup f(x) = 1$; luego las sumas de Riemann valen $s(f, P) = 0$ y $S(f, P) = 1$, independientemente de la partición P . Por consiguiente, no existe la integral de Riemann de la función f .

Ahora bien, el conjunto de $\mathbb{R}^2$ “limitado” por la gráfica de la función de Dirichlet y el eje x está compuesto por segmentos verticales A_r , de longitud 1, situados sobre cada valor racional r del intervalo $[0, 1]$. Así que el área de dicho conjunto, A , es fácil de determinar: cada segmento A_r es un rectángulo de área nula y, como hay un número numerable de ellos, su unión tiene área nula. Esto es, escritos los racionales de $[0, 1]$ en forma de sucesión $\{r_n\}_{n \in \mathbb{N}}$, se tiene

$$\lambda_2(A) = \lim_{N \rightarrow \infty} \lambda_2\left(\bigcup_{n=1}^N A_{r_n}\right) = \lim_{N \rightarrow \infty} \sum_{n=1}^N \lambda_2(A_{r_n}) = 0.$$

De forma similar, la función

$$1 - f(x) = \begin{cases} 0 & \text{si } x \in \mathcal{Q} \cap [0, 1] \\ 1 & \text{si } x \in [0, 1] - \mathcal{Q} \end{cases}$$

tampoco es integrable en el sentido de Riemann, puesto que, como en el caso anterior, es $s(f, P) = 0$ y $S(f, P) = 1$ para cualquier partición P . Ahora, el subconjunto B de $\mathbb{R}^2$ comprendido entre su gráfica y el eje x se compone de un número no numerable de segmentos verticales, de longitud 1, situados sobre los irracionales del intervalo $[0, 1]$. No cabe sumar las áreas nulas de un número no numerable de segmentos para obtener el área de B , sino que hay que observar que A y B son conjuntos disjuntos cuya unión es todo el cuadrado $[0, 1]^2$. Si se puede asignar a B un área $\lambda_2(B)$ tendrá que ser

$$\lambda_2(A) + \lambda_2(B) = \lambda_2([0, 1]^2)$$

y, dado que el área del cuadrado $[0, 1]^2$ es 1 y $\lambda_2(A) = 0$, se obtiene $\lambda_2(B) = 1$.

Esta última manera de proceder pone de relieve un cuarto requisito de la noción de área que “casi” se deduce de (b):

- (d) Si $A_2 \subset A_1$ son conjuntos con área definida $\lambda_2(A_2)$ y $\lambda_2(A_1)$, supuesto que $\lambda_2(A_2) < \infty$, debe ser

$$\lambda_2(A_1 - A_2) = \lambda_2(A_1) - \lambda_2(A_2).$$

A la vista del ejemplo anterior pueden plantearse dos cuestiones de interés. Por un lado, lo mismo que en el caso de la longitud en $\mathbb{R}$, cabe preguntarse cuáles son los subconjuntos de $\mathbb{R}^2$ a los que se puede asignar un área, sin violar los principios (a), (b), (c) y (d). En segundo lugar, parece que la idea de integral puede modificarse para hacer integrables funciones para las que fracasa la definición de Riemann. Respecto al primer problema, una vez sabido lo que ocurre con la longitud, no cabe hacerse demasiadas ilusiones: hay conjuntos en $\mathbb{R}^2$ a los que no se puede atribuir área de forma coherente.

Ejemplo 2.6

Un conjunto sin área. Sea W el subconjunto de $\mathbb{R}$ sin longitud al que hacía referencia el ejemplo 2.3. El conjunto $V = W \times [0, 1]$ está constituido por segmentos verticales, de longitud 1, situados sobre los puntos de W .

Supuesto que pudiese atribuirse área a V , si se le somete a traslaciones horizontales de vector $(r, 0)$ para cada $r \in \mathcal{Q} \cap [-1, 1]$, se obtiene una colección numerable de conjuntos

$$V_r = \{(r + w, y) \mid w \in W, y \in [0, 1]\}$$

del mismo área y disjuntos entre sí, tales que

$$[0, 1]^2 \subset \bigcup_{r \in \mathcal{Q} \cap [-1, 1]} V_r \subset [-1, 2] \times [0, 1].$$

Por consiguiente, no puede atribuirse a V área cero, so pena de caer en la contradicción de que $[0, 1]^2$ está contenido en un subconjunto de área cero. Tampoco puede atribuirse a V área positiva, pues ello daría lugar a que el rectángulo $[-1, 2] \times [0, 1]$ contenga un conjunto de área infinita.

El proceso de extensión del área en $\mathbb{R}^2$ es idéntico al de la longitud en $\mathbb{R}$; establece que se puede atribuir área a los conjuntos de una familia $\mathcal{L}_2$ de subconjuntos de $\mathbb{R}^2$, de manera que se cumplan los postulados (a), (b), (c) y (d). Según ellos, $\mathcal{L}_2$ tiene estructura de λ -clase y, como contiene a la π -clase de los rectángulos, contendrá a la σ -álgebra engendrada por estos que, según la definición siguiente, se denomina la σ -álgebra de Borel en $\mathbb{R}^2$:

Definición 2.6 Sea $\mathcal{I}_2$ la familia constituida por todos los rectángulos de $\mathbb{R}^2$. La σ -álgebra $\sigma(\mathcal{I}_2)$, engendrada por $\mathcal{I}_2$, se denomina la **σ -álgebra de Borel en $\mathbb{R}^2$** y se representa habitualmente por $\mathbb{B}^2$. Los elementos de $\mathbb{B}^2$ reciben el nombre de **subconjuntos de Borel de $\mathbb{R}^2$** .

Así, en este caso, el teorema de extensión afirma:

Proposición 2.10 Todos los conjuntos de $\mathbb{B}^2$ tienen área definida y única.

La σ -álgebra $\mathbb{B}^2$ está constituida por aquellos conjuntos que pueden construirse mediante operaciones conjuntistas a partir de un número numerable de rectángulos y comprende a todos los subconjuntos de $\mathbb{R}^2$ que tienen un interés práctico. En particular, se cumple:

Proposición 2.11 Todo conjunto abierto de $\mathbb{R}^2$ es un conjunto de Borel. Lo mismo sucede con cualquier conjunto cerrado, con cualquier intersección numerable de abiertos, con cualquier unión numerable de cerrados, etc.

- De hecho, la familia de los cuadrados

$$C(\mathbf{r}, \varepsilon) = (r_1 - \varepsilon, r_1 + \varepsilon) \times (r_2 - \varepsilon, r_2 + \varepsilon)$$

donde r_1, r_2 y ε son racionales, es una colección numerable de rectángulos y cualquier conjunto abierto A de $\mathbb{R}^2$ puede expresarse como unión de tales cuadrados. En efecto, dado cualquier $\mathbf{x} \in A$, existe una bola abierta $B(\mathbf{x}, \delta)$ contenida en A , además en $B(\mathbf{x}, \delta/4)$ puede encontrarse un punto $\mathbf{r} = (r_1, r_2)$ de coordenadas racionales. Si se toma un racional ε con $\delta/4 < \varepsilon < \delta/2$, el cuadrado $C(\mathbf{r}, \varepsilon)$ contiene a $\mathbf{x}$ y está contenido en A (pues lo está en $B(\mathbf{x}, \delta)$). Luego

$$A = \bigcup C(\mathbf{r}, \varepsilon)$$

si la unión incluye todos los cuadrados $C(\mathbf{r}, \varepsilon)$ contenidos en A . En definitiva, cualquier abierto es un conjunto de Borel. Lo mismo ocurre con sus complementarios —los cerrados—, con la intersecciones numerables de abiertos o con las uniones numerables de cerrados. ■

También son conjuntos de Borel en el plano los productos de dos conjuntos de Borel arbitrarios de la recta real. Más exactamente:

Proposición 2.12 Si $\mathbb{B} \times \mathbb{B}$ designa la familia de todos los subconjuntos de $\mathbb{R}^2$ de la forma $B_1 \times B_2$, donde $B_1, B_2 \in \mathbb{B}$, la σ -álgebra engendrada por $\mathbb{B} \times \mathbb{B}$ coincide con la σ -álgebra de Borel en $\mathbb{R}^2$. Es decir, $\sigma(\mathbb{B} \times \mathbb{B}) = \mathbb{B}^2$.

• En efecto, como todo rectángulo es producto de dos conjuntos de Borel unidimensionales, se tiene $\mathcal{I}_2 \subset \mathbb{B} \times \mathbb{B} \subset \sigma(\mathbb{B} \times \mathbb{B})$. Por tanto, la mínima σ -álgebra que contiene a $\mathcal{I}_2$ es más pequeña que $\sigma(\mathbb{B} \times \mathbb{B})$; en símbolos:

$$\mathbb{B}^2 \subset \sigma(\mathbb{B} \times \mathbb{B}).$$

Dado que $\sigma(\mathbb{B} \times \mathbb{B})$ es la menor σ -álgebra que contiene a $\mathbb{B} \times \mathbb{B}$, establecer el contenido recíproco, se reduce a ver que $\mathbb{B} \times \mathbb{B}$ está contenido en $\mathbb{B}^2$. Para ello, consideremos

$$\mathcal{F} = \{ A \in \mathbb{B} \mid A \times I \in \mathbb{B}^2 \text{ para cualquier intervalo } I \}$$

y probemos que se trata de una σ -álgebra. En primer lugar, si $A \in \mathcal{F}$ e I es cualquier intervalo, será

$$A^c \times I = \mathbb{R} \times I - A \times I$$

siendo $\mathbb{R} \times I \in \mathcal{I}_2$ y $A \times I \in \mathbb{B}^2$, luego $A^c \times I \in \mathbb{B}^2$, o bien $A^c \in \mathcal{F}$. Por otra parte, si A_n es una sucesión de elementos de $\mathcal{F}$, se tiene

$$(\bigcup_n A_n) \times I = \bigcup_n (A_n \times I)$$

y, como $A_n \times I \in \mathbb{B}^2$, se cumple $(\bigcup_n A_n) \times I \in \mathbb{B}^2$; con lo cual $\bigcup_n A_n \in \mathcal{F}$. Obsérvese además que la σ -álgebra $\mathcal{F}$ contiene a cualquier intervalo de $\mathbb{R}$ y, a su vez, está contenida en $\mathbb{B}$. Tiene que ser pues $\mathcal{F} = \mathbb{B}$, lo cual asegura que $A \times I$ pertenece a $\mathbb{B}^2$ cualquiera que sean $A \in \mathbb{B}$ y el intervalo I .

Consideremos ahora

$$\mathcal{F}' = \{ B \in \mathbb{B} \mid A \times B \in \mathbb{B}^2 \text{ para cualquier } A \in \mathbb{B} \}$$

que, según puede probarse de manera análoga a la anterior, tiene estructura de σ -álgebra. Evidentemente $\mathcal{F}'$ está contenida en $\mathbb{B}$ y, según lo anterior, cualquier intervalo I pertenece a $\mathcal{F}'$; luego $\mathcal{F}' = \mathbb{B}$.

Tal igualdad establece, en definitiva, que $A \times B \in \mathbb{B}^2$ cualquiera que sean $A \in \mathbb{B}$ y $B \in \mathbb{B}$. ■

La σ -álgebra $\sigma(\mathbb{B} \times \mathbb{B})$ que aparece en el enunciado anterior responde a un concepto general:

Definición 2.7 Sean $\mathcal{F}_1$ y $\mathcal{F}_2$ σ -álgebras de subconjuntos de Ω_1 y Ω_2 respectivamente. La familia de subconjuntos de $\Omega_1 \times \Omega_2$:

$$\mathcal{F}_1 \times \mathcal{F}_2 = \{ A \times B \mid A \in \mathcal{F}_1, B \in \mathcal{F}_2 \}$$

engendra una σ -álgebra, $\sigma(\mathcal{F}_1 \times \mathcal{F}_2)$, llamada **σ -álgebra producto** de $\mathcal{F}_1$ y $\mathcal{F}_2$ que suele representarse por $\mathcal{F}_1 \otimes \mathcal{F}_2$. Es decir, $\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma(\mathcal{F}_1 \times \mathcal{F}_2)$.

En estos términos, la proposición anterior afirma: $\mathbb{B}^2 = \mathbb{B} \otimes \mathbb{B}$.

Hay una relación entre el área y la longitud que puede expresarse con facilidad en el caso de conjuntos de la forma $A \times B$ pertenecientes a $\mathbb{B} \times \mathbb{B}$.

Proposición 2.13 *Si $A, B \in \mathbb{B}$ se verifica*

$$\lambda_2(A \times B) = \lambda(A) \lambda(B). \quad (2.4)$$

• Para mostrarlo, es necesario restringirse primero al caso de conjuntos de Borel contenidos en un intervalo acotado cualquiera $K = [-k, k]$. Porque, entonces, es fácil ver que la familia

$$\mathcal{G}_1 = \{ A \in \mathbb{B}_K \mid \lambda_2(A \times I) = \lambda(A) \lambda(I) \text{ para cualquier intervalo } I \subset K \}$$

es una λ -clase⁶ que contiene, según (a), a la π -clase de los subintervalos de K . En consecuencia, $\mathcal{G}_1 = \mathbb{B}_K$, lo cual significa que (2.4) se cumple si $A \in \mathbb{B}_K$ y B es un subintervalo de K . A continuación, sea

$$\mathcal{G}_2 = \{ B \in \mathbb{B}_K \mid \lambda_2(A \times B) = \lambda(A) \lambda(B) \text{ para cualquier } A \in \mathbb{B}_K \}.$$

De nuevo, $\mathcal{G}_2$ es una λ -clase que, según lo anterior, contiene a la π -clase de los subintervalos de K . Por tanto, $\mathcal{G}_2 = \mathbb{B}_K$ y (2.4) se verifica cualquiera que sean $A, B \in \mathbb{B}_K$.

Por último, si A y B son conjuntos cualesquiera de $\mathbb{B}$, no necesariamente acotados, $A_k = A \cap K$, $B_k = B \cap K$ y $A_k \times B_k$ son sucesiones crecientes hacia A , B y $A \times B$ respectivamente. Se tiene pues

$$\lambda_2(A \times B) = \lim_{k \rightarrow \infty} \lambda_2(A_k \times B_k) = \lim_{k \rightarrow \infty} \lambda(A_k) \lambda(B_k) = \lambda(A) \lambda(B). \quad \blacksquare$$

Las dos últimas proposiciones indican que el concepto de área resultaría exactamente el mismo si se consideran como “rectángulos” todos los conjuntos de $\mathbb{B} \times \mathbb{B}$ y se sustituye el postulado (a) por la igualdad (2.4).

La integral de Lebesgue

Los resultados anteriores constituyen la base sobre la que Lebesgue elaboró un nuevo concepto de integral que puso fin a las dificultades teóricas con las que tropieza la integral de Riemann. Conseguía, en primer lugar, ampliar de forma considerable el conjunto de funciones integrables y, gracias a ello, mejorar de forma sustancial las propiedades de convergencia de la integral cuando el integrando converge.

La idea de Lebesgue es la siguiente: En vez de tomar una partición del eje de abscisas, se divide el eje de ordenadas, mediante los puntos de la forma $k/2^n$, donde n es un número natural fijo y k es entero.

⁶ Por ejemplo, si $A_1, A_2 \in \mathcal{G}_1$ y $A_1 \supset A_2$, se cumple

$$\begin{aligned} \lambda_2((A_1 - A_2) \times I) &= \lambda_2(A_1 \times I) - \lambda_2(A_2 \times I) = \lambda(A_1) \lambda(I) - \lambda(A_2) \lambda(I) \\ &= [\lambda(A_1) - \lambda(A_2)] \lambda(I) = \lambda(A_1 - A_2) \lambda(I) \end{aligned}$$

(la primera y la última igualdad requieren que las áreas y las longitudes sean finitas). Por consiguiente $A_1 - A_2 \in \mathcal{G}_1$. El resto de las propiedades se comprueba de forma similar.