CAPÍTULO 1

Espacios de Probabilidad

1.1. Introducción

El objetivo de la Teoría de Probabilidad es desarrollar y estudiar modelos matemáticos para experimentos cuyos resultados no pueden predecirse con exactitud. Aún cuando la historia de la teoría de probabilidades tiene ya varios siglos, y muchos autores consideran que se inició con la correspondencia entre Blaise Pascal y Pierre de Fermat sobre juegos de azar en el siglo XVII, se puede decir que no fue hasta el siglo XX cuando esta teoría alcanzó un desarrollo notable.

Uno de los principales problemas por los cuales este desarrollo no ocurrió antes fue la ausencia de una axiomatización adecuada de las probabilidades, que le diese una base sólida y le permitiese desarrollarse al igual que otras ramas de la Matemática. En 1933, A. N. Kolmogorov propone una axiomatización usando las ideas de la Teoría de Medida, desarrollada a principios del siglo XX por H. Lebesgue. Esta axiomatización propone modelar los experimentos que tienen comportamiento aleatorio usando un espacio de medida. Aún cuando el desarrollo pleno de estas ideas está más allá del alcance del material que presentamos, vamos a considerar este enfoque axiomático en las próximas secciones, presentando numerosas aplicaciones de estas ideas.

El objetivo de la Teoría de Probabilidades es desarrollar modelos para experimentos que están gobernados por el azar y estudiar sus propiedades y aplicaciones. El modelo fundamental para un experimento de este tipo, como el lanzamiento de un dado, es el Espacio de Probabilidad, que describimos en las siguientes secciones.

1.2. Espacios de Probabilidad.

Cada resultado posible de un experimento aleatorio será llamado *evento elemental* y el conjunto de los eventos elementales será el *espacio muestral*. Usualmente, denotaremos con Ω el espacio muestral, y mediante ω los eventos elementales (o puntos de Ω).

Veamos algunos ejemplos de experimentos aleatorios y sus espacios muestrales asociados.

1. En una fábrica se toma uno de los artículos producidos y se prueba para determinar si es defectuoso. En este caso podemos considerar $\Omega = \{B, D\}$, donde B indica bueno y D defectuoso. Si en cambio se extraen n artículos y se prueban, podemos considerar $\Omega = \{(\epsilon_1, \epsilon_2, \dots, \epsilon_n) : \epsilon_i = 0 \text{ ó } 1; i = 1, \dots, n\}$ donde $\epsilon_i = 0$ indica que el i-ésimo artículo es bueno y $\epsilon_i = 1$ indica que es defectuoso. Es decir, Ω

es el conjunto de n-uplas o vectores de dimensión n de ceros y unos. En este caso Ω consta de 2^n eventos elementales y, en particular, $\sum_{i=1}^{n} \epsilon_i$ representa el número de objetos defectuosos del evento elemental $(\epsilon_1, \epsilon_2, \dots, \epsilon_n)$.

2. En un punto de una carretera contamos el número de vehículos que pasan durante un cierto lapso de tiempo. En este caso podemos tomar $\Omega = \{0, 1, 2, \dots\}$, es decir el conjunto de los enteros no-negativos. Podemos, sin embargo, tomar otros conjuntos como espacio muestral en este caso. Por ejemplo, si sabemos que el número de vehículos considerados no supera los mil, podemos considerar $\Omega_1 = \{n : 0 \leq n \leq 1,000\}$, aunque no necesariamente del hecho de que Ω_1 sea subconjunto de Ω, se concluye que la descripción del experimento aleatorio mediante Ω_1 sea más simple que la que se obtiene usando Ω.

3. En una sucesión de cálculos realizados con una computadora, observamos los primeros k dígitos no tomados en cuenta al truncar los resultados de las operaciones en una cierta cifra decimal. En este caso podemos tomar como espacio muestral $\Omega = \{(a_1, \dots, a_k) : a_i \in \mathbb{Z},\ 0 \leq a_i \leq 9\}$

4. En una fábrica de componentes electrónicos se eligen varios al azar y se conecta cada uno de ellos hasta que se daña, observando en cada caso el tiempo de vida. Si se trata de un solo componente podemos tomar
$$\Omega = \{t : t \in \mathbb{R},\ t \geq 0\}$$
es decir, el conjunto de números reales no-negativos. Si se consideran n componentes, podemos tomar
$$\Omega = \{(t_1, t_2, \dots, t_n) : t_i \in \mathbb{R},\ t_i \geq 0\}.$$

5. Se lanza un dado repetidamente y se cuenta el número de lanzamientos hasta que salga el 6 por primera vez. En este caso el espacio muestral es el conjunto de los números naturales:
$$\Omega = \{1, 2, 3, \dots\}.$$

6. Se mide la presión y temperatura en una estación meteorológica. Aquí,
$$\Omega = \{(p, t) : p > 0,\ t \in \mathbb{R}\}.$$

7. Se escoge un punto al azar lanzando un dardo a un disco de radio un metro. En este caso el espacio muestral es el conjunto de puntos del plano que estan dentro de la circunferencia de radio 1:
$$\Omega = \{(x, y) : x^2 + y^2 \leq 1\}.$$

Definición 1.1 Si al realizar un experimento obtenemos como resultado el evento elemental ω, decimos que el evento $A \subset \Omega$ ha ocurrido si $\omega \in A$

En la práctica, al realizar un experimento nos interesa con frecuencia saber si algún subconjunto de Ω ha ocurrido. A estos subconjuntos los llamaremos *eventos* o *sucesos*. En el ejemplo 1 podemos estar interesados en el subconjunto: "entre los n artículos extraídos hay d defectuosos", es decir, en el subconjunto de Ω definido por
$$\{(\epsilon_1, \dots, \epsilon_n) : \epsilon_i = 0 \text{ ó } 1, \sum_{1}^{n} \epsilon_i = d\}.$$
En el ejemplo 3 nos interesará saber, por ejemplo, si la primera cifra no tomada en cuenta al truncar es mayor o igual que 5, o sea,
$$\{(a_1, \dots, a_k)\} : 0 \leq a_i \leq 9,\ a_1 \geq 5\}.$$

Análogamente, en la situación planteada en 6, nos interesarán eventos del tipo: "la presión está comprendida entre p_1 y p_2 y la temperatura entre t_1 y t_2", es decir

$$\{(p_i, t_i) : p_1 \leq p \leq p_2,\ t_1 \leq t \leq t_2\}.$$

Estamos interesados, por lo tanto, en considerar familias de subconjuntos de Ω, es decir, familias $\mathcal{A}$ de eventos. Estas familias son la segunda componente de nuestro modelo para un experimento que depende del azar, y deben satisfacer una serie de condiciones. Veamos qué condiciones debe cumplir la familia de eventos $\mathcal{A}$. En primer lugar

a. $\boxed{\Omega \in \mathcal{A}}$

   es decir que al realizar el experimento el resultado es un elemento de Ω. A Ω lo llamaremos *evento cierto*.

b. $\boxed{A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}}$

   donde $A^c = \Omega - A = \{\omega : \omega \in \Omega,\ \omega \notin A\}$ es el complemento de A.

   Es decir, si A es un evento, pediremos que "no ocurre A" también sea un evento.

   Finalmente, la familia $\mathcal{A}$ también debe satisfacer que si $A_1, A_2, \dots, A_n, \dots$ son eventos, "ocurre alguno de los A_n" también es un evento, o sea

c. $\boxed{A_n \in \mathcal{A}\ (n = 1, 2, \dots)\ \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}}$

Definición 1.2 Una familia $\mathcal{A}$ de subconjuntos de Ω que satisface las condiciones a, b y c se llama una *σ-álgebra* de subconjuntos o partes de Ω.

En adelante supondremos que las familias de eventos son σ-álgebras. Las siguientes propiedades se obtienen como consecuencia inmediata de la definición:

1. El conjunto vacío, $\emptyset$, es un evento, ya que $\emptyset = \Omega^c$.
2. $A_1, A_2, \dots A_k \in \mathcal{A} \Rightarrow \bigcup_{n=1}^{k} A_n \in \mathcal{A}$. Basta considerar $A_{n+1} = A_{n+2} = \cdots = \emptyset$ y aplicar la propiedad anterior y c.
3. $A_1, A_2, \dots, A_n, \dots \in \mathcal{A} \Rightarrow \bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$. En efecto, por las leyes de de Morgan,

$$\bigcap_{n=1}^{\infty} A_n = \left(\bigcup_{n=1}^{\infty} A_n^c \right)^c$$

   y basta ahora aplicar b y c.

Ejemplos.

8. Para cualquier conjunto Ω, la σ-álgebra más sencilla es la σ-álgebra trivial $\mathcal{T} = \{\Omega, \emptyset\}$. La mayor σ-álgebra de subconjuntos de Ω es $\mathcal{P}(\Omega)$, el conjunto de partes de Ω, es decir, la colección de todos los subconjuntos de Ω. Cualquier otra σ-álgebra debe contener a $\mathcal{T}$ y estar contenida en $\mathcal{P}(\Omega)$.

   En el caso de experimentos sencillos, por ejemplo experimentos con un conjunto finito de resultados, normalmente tomamos como σ-álgebra de eventos el conjunto de partes de Ω. En experimentos más complicados, con una cantidad no-numerable de resultados posibles, no siempre es posible tomar esta opción, y es necesario considerar σ-álgebras más pequeñas.

9. *Muestreo con reposición.* De la producción de una fábrica se extrae un artículo al azar y se determina si es bueno o defectuoso (B o D, respectivamente). Se devuelve este artículo al stock y se extrae de nuevo al azar un artículo, que puede ser el mismo. Esta operación se repite una vez más, de modo que en total se extraen tres.

   El espacio muestral es:

$$\Omega = \{BBB, BBD, BDB, DBB, BDD, DBD, DDB, DDD\}$$

   Observamos que hay 2^3 eventos elementales, ya que en cada una de las tres extracciones hay dos resultados posibles. Consideramos los siguientes eventos:

   A_1: "El segundo artículo resultó bueno"

   A_2: "Se obtuvo un solo defectuoso en las tres extracciones".

   A_3: "No hubo defectuosos".

   Los eventos definidos son:

$$A_1 = \{BBB, BBD, DBB, DBD\} \quad A_2 = \{BBD, BDB, DBB\} \quad A_3 = \{BBB\}$$

   El número de eventos elementales incluidos en A_1 es 2^2 ya que el resultado de la segunda extracción está fijo. El evento A_2 contiene 3 puntos muestrales, ya que hay tres lugares posibles para el objeto defectuoso en la muestra. Podemos ahora combinar estos eventos utilizando operaciones de conjuntos. Tenemos, por ejemplo,

$$\begin{aligned} A_1 \cap A_2 &= \{BBD, DBB\} \\ A_1^c \cup A_2^c &= \{BBB, BDB, BDD, DBD, DDB, DDD\} \\ A_1 \cap A_2^c &= \{BBB, DBD\} \end{aligned}$$

10. *Muestreo sin reposición.* De una población de N artículos entre los cuales hay n defectuosos, se extraen sucesivamente r sin reposición y se cuenta el número de defectuosos en la muestra. El espacio muestral contiene todos los subconjuntos de r elementos tomados entre los N dados.

La tercera componente del modelo es una (medida de) probabilidad, que definimos a continuación.

Definición 1.3 Sean Ω un espacio muestral y $\mathcal{A}$ una familia de eventos de Ω, es decir, una σ-álgebra de subconjuntos de Ω. Estamos interesados en asignar a cada evento $A \in \Omega$ un número real $P(A)$, que llamaremos la probabilidad de A, de modo tal que se cumplan las siguientes condiciones:

1. $\boxed{P(A) \geq 0 \text{ para todo } A \in \Omega}$

   La probabilidad de un evento cualquiera es un número real no negativo.

2. $\boxed{P(\Omega) = 1}$

   El evento cierto tiene probabilidad igual a 1.

3. Si $A_n \in \mathcal{A}$ para $n = 1, 2, \dots$ son eventos disjuntos dos a dos, es decir, tales que $A_i \cap A_j = \emptyset$ si $i \neq j$, entonces

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

Una terna $(\Omega, \mathcal{A}, P)$, formada por un espacio muestral Ω, una familia $\mathcal{A}$ de eventos y una probabilidad P se llama un *espacio de probabilidad*.

El problema de cómo definir la función P, o sea, de cómo asignar una probabilidad a cada evento, debe ser resuelto de acuerdo a las condiciones concretas de cada experimento aleatorio en consideración.

1.3. Algunas Consecuencias de la Definición.

Veamos a continuación algunas consecuencias de la definición anterior. Usaremos la notación $A + B$ para indicar la unión de los conjuntos A y B cuando ellos son disjuntos.

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

En efecto, consideremos $A_1 = \Omega$ y $A_i = \emptyset$, $i = 2, 3, \dots$ Entonces $A_i \in \mathcal{A}$ cualquiera que sea i y además si $i \neq j$ se tiene $A_i \cap A_j = \emptyset$. Resulta

$$P(\Omega) = P\left(\sum_{i=1}^{\infty} A_i\right) = P(\Omega) + \sum_{i=2}^{\infty} P(A_i).$$

Luego

$$\sum_{i=2}^{\infty} P(A_i) = 0$$

y como $P(A_i) \geq 0$ para todo i se tiene que $P(A_i) = 0$ para $i \geq 2$. En consecuencia $P(\emptyset) = 0$.

(2) $A_1 \cap A_2 = \emptyset \Rightarrow P(A_1 \cup A_2) = P(A_1) + P(A_2)$.

Basta considerar $A_i = \emptyset$, $i \geq 3$ y aplicar la condición 3 de la definición de espacio de probabilidad. De manera similar se puede demostrar que P es finitamente aditiva: Si $A_1, \dots, A_n$ son disjuntos dos a dos entonces

$$P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} P(A_k).$$

(3) $P(A^c) = 1 - P(A)$.

Como $A^c \cup A = \Omega$ y $A^c \cap A = \emptyset$ se tiene

$$P(A^c) + P(A) = 1.$$

(4) $A_1 \subset A_2 \Rightarrow P(A_1) \leq P(A_2)$.

Como $A_2 = A_1 + (A_2 \cap A_1^c)$ resulta

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

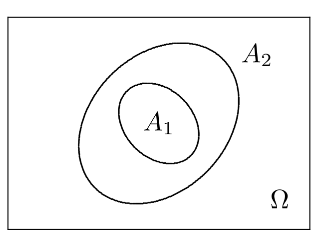

Figura 1.7

y en consecuencia

$$P(A_1) \leq P(A_2) \text{ ya que } P(A_2 \cap A_1^c) \geq 0.$$

(5) $P(A) \leq 1$ para todo $A \in \mathcal{A}$.

Esto es consecuencia inmediata del punto anterior al considerar que $A \subset \Omega$.

(6) $A_1 \subset A_2 \subset \cdots \subset A_n \subset \cdots \Rightarrow P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P(A_n)$.

Sean

$$B_1 = A_1 \quad \text{y} \quad B_n = A_n \cap A_{n-1}^c \;\text{ si } n > 1,$$

resulta

$$\bigcup_{i=1}^{\infty} A_i = \sum_{i=1}^{\infty} B_i$$

y entonces

$$\begin{aligned} P(\bigcup_{i=1}^{\infty} A_i) &= P(\sum_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} P(B_i) \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i) = \lim_{n\to\infty} P(\sum_{i=1}^{n} B_i) = \lim_{n\to\infty} P(A_n). \end{aligned}$$

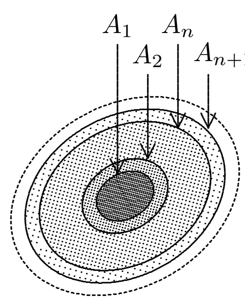

Figura 1.8

(7) $A_1 \supset A_2 \supset \cdots \supset A_n \supset \cdots \Rightarrow P\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P(A_n)$.

Como la sucesión $\{A_n^c\}$ es creciente, usando (7) obtenemos

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = P\left(\left(\bigcup_{n=1}^{\infty} A_n^c\right)^c\right) = 1 - P\left(\bigcup_{n=1}^{\infty} A_n^c\right)$$

$$= 1 - \lim_{n\to\infty} P(A_n^c) = 1 - \lim_{n\to\infty} (1 - P(A_n)) = \lim_{n\to\infty} P(A_n).$$

(8) $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$.

En efecto, considerando que

$$A_1 \cup A_2 = A_1 + (A_2 \cap A_1^c) \quad \text{y} \quad A_2 = (A_1 \cap A_2) + (A_1^c \cap A_2)$$

después de aplicar (2) a ambas igualdades y restar resulta la proposición (6).

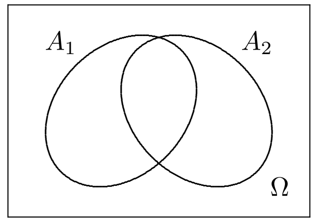

Figura 1.9

(9) $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$.

Para ver esto apliquemos (8) a los eventos $A \cup B$ y C, obteniendo

$$P(A \cup B \cup C) = P(A \cup B) + P(C) - P((A \cup B) \cap C)$$

y de manera similar

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ P((A \cup B) \cap C) &= P((A \cap C) \cup (B \cap C)) \\ &= P(A \cap C) + P(B \cap C) - P(A \cap B \cap C). \end{aligned}$$

Reemplazando las dos últimas expresiones en la primera obtenemos el resultado.

(10) $P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i) - \sum_{i<j=2}^n P(A_i \cap A_j) + \cdots + (-1)^{n+1} P(A_1 \cap \cdots \cap A_n)$.

Para demostrar esta proposición procedemos por inducción completa en n siguiendo la misma idea que en la anterior, que corresponde al caso $n = 3$. Para $n = 2$ es la propiedad (8).

Suponemos entonces que el resultado es cierto para n y queremos deducir que también lo es para $n + 1$. ¿Qué significa que el resultado es cierto para n? Significa que

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k+1} \sum_{1 \le i_1 < i_2 < \ldots < i_k \le n} P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) \tag{1.1}$$

Queremos deducir de (1.1) que también es válida una fórmula análoga para

$$P\left(\bigcup_{i=1}^{n+1} A_i\right).$$

Pongamos entonces $B = \bigcup_{i=1}^n A_i$ y apliquemos la propiedad (6) a

$$\begin{aligned} P\left(\bigcup_{i=1}^{n+1} A_i\right) &= P(B \cup A_{n+1}) = P(B) + P(A_{n+1}) - P(B \cap A_{n+1}) \\ &= P\left(\bigcup_{i=1}^n A_i\right) + P(A_{n+1}) - P(\bigcup_{i=1}^n (A_i \cap A_{n+1})). \end{aligned} \tag{1.2}$$

El primero de estos tres términos lo reemplazamos utilizando (1.1) y el último también sólo que, en lugar de cada A_i ponemos $A_i \cap A_{n+1}$. Veamos qué nos queda; en primer lugar,

$$P(A_1) + \cdots + P(A_n) + P(A_{n+1}),$$

los primeros n provenientes del primer sumando en (1.2) y el último del segundo sumando. En segundo lugar

$$-\sum_{1\leq i_1<i_2\leq n} P(A_{i_1}\cap A_{i_2}) - \sum_{i=1}^{n} P(A_i\cap A_{n+1}) = -\sum_{1\leq i_1<i_2\leq n+1} P(A_{i_1}\cap A_{i_2}).$$

Aquí el primer sumando viene del primero de (1.2) y el segundo, del tercero de (1.2). De la misma manera, para $k\leq n$, queda una suma de la forma

$$\begin{aligned}(-1)^{k+1}\sum_{1\leq i_1<i_2<\ldots<i_k\leq n} P(A_{i_1}\cap\cdots\cap A_{i_k}) - (-1)^k \sum_{1\leq i_1<i_2<\ldots<i_{k-1}\leq n} P(A_{i_1}\cap\cdots\cap A_{i_{k-1}}\cap A_{n+1})\\ = (-1)^{k+1}\sum_{1\leq i_1<i_2<\ldots<i_k\leq n+1} P(A_{i_1}\cap A_{i_2}\cap\cdots\cap A_{i_k}).\end{aligned}$$

Finalmente, para $k=n+1$, no tenemos ningún término en el primer sumando de (1.2) y tenemos uno sólo en el tercero que es:

$$(-1)^{n+2}P(A_1\cap A_2\cap\cdots\cap A_n\cap A_{n+1}).$$

Juntando todos los términos resulta

$$P\left(\bigcup_{i=1}^{n+1} A_i\right) = \sum_{k=1}^{n+1}(-1)^{k+1}\sum_{1\leq i_1<i_2<\ldots<i_k\leq n+1} P(A_{i_1}\cap A_{i_2}\cap\cdots\cap A_{i_k}).$$

1.4. Ejemplos y Aplicaciones.

1.4.1. Probabilidades en Espacios Finitos.

Sean $\Omega=\{\omega_1,\ldots,\omega_m\}$ un conjunto finito y $\mathcal{A}=\mathcal{P}(\Omega)$ la familia de todos los subconjuntos de Ω. Elijamos m números reales $p_i,\ i=1,2,\ldots m$, tales que

$$\begin{cases} p_i\geq 0, \quad \text{para todo } i,\\ \displaystyle\sum_{i=1}^{m} p_i = 1.\end{cases}$$

Poniendo $P(\omega_i)=p_i$ $(i=1,2,\ldots m)$, queda definida la probabilidad para todo evento $A\in\mathcal{A}$ mediante la asignación

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

Un caso particular de interés es aquel en el cual $p_i=1/m$ para todo i, y ahora si A tiene n elementos

$$P(A)=\frac{n}{m},$$

es decir que si todos los eventos elementales son igualmente probables, la probabilidad de un evento A es el cociente entre el número de elementos que pertenecen a A y el número total de elementos de Ω. Esta definición se conoce como la *definición clásica* y fue propuesta, entre otros, por Laplace.

En una situación como la descrita, en la cual todos los resultados posibles del experimento tienen la misma probabilidad de ocurrir, el problema de calcular la probabilidad de un evento se reduce a contar cuántos resultados posibles tiene el experimento y cuántos de estos pertenecen al evento que nos interesa. En este sentido, las técnicas combinatorias facilitan estos cálculos.

En un problema específico, podemos determinar si los resultados posibles tienen la misma probabilidad por consideraciones de simetría sobre el experimento que estamos analizando. Por ejemplo, si se trata del lanzamiento de un dado, en principio no hay razones para suponer que alguna cara tenga mayor o menor probabilidad de ocurrir que las demás, y por lo tanto asumimos como modelo que todos los resultados son equiprobables. Algo similar sucede con el lanzamiento de una moneda, el juego de ruleta o la extracción de una carta de un paquete que ha sido bien barajeado.

Por supuesto que en la práctica esto puede no ser cierto: el dado puede no ser perfectamente simétrico, o la ruleta puede estar desbalanceada y favorecer ciertos resultados. Para determinar si este es el caso existen procedimientos estadísticos que permiten contrastar la hipótesis de simetría, pero por los momentos no nos ocuparemos de este problema.

Veamos algunos ejemplos.

1. De los números del 1 al 10 escogemos tres al azar, en orden y sin reposición. ¿Cuál es la probabilidad de obtener 1, 2 y 3, en este orden?

   En este problema podemos describir el espacio muestral como el conjunto de todos los vectores de tres componentes tomadas de los enteros del 1 al 10, sin repetir ninguna componente.

   $$\Omega = \{(a, b, c) : 1 \leq a, b, c \leq 10, \text{ distintos}\}.$$

   Como estamos muestreando al azar, todos los vectores del espacio tienen la misma probabilidad.

   El evento que nos interesa corresponde a un resultado particular, el vector $(1, 2, 3)$. Por lo tanto tenemos que contar cuantos elementos hay en Ω para saber cuál es la probabilidad de cada uno de ellos. La primera componente del vector la podemos escoger de 10 maneras. Para la segunda sólo tenemos 9 posibilidades, porque no podemos repetir el número que ocupa la primera componente. Finalmente, para la tercera hay 8 posibilidades. Por lo tanto tenemos

   $$10 \times 9 \times 8 = 720$$

   puntos en el espacio muestral. Como todos tienen la misma probabilidad, la respuesta al ejemplo es $1/720$.

2. Si los números del ejemplo anterior se escogen con reposición ¿Cuál es la probabilidad de obtener 1, 2 y 3, en este orden?

   En este caso el espacio muestral incluye vectores con componentes repetidas:

   $$\Omega = \{(a, b, c) : 1 \leq a, b, c \leq 10\}.$$

   Para cada componente tenemos ahora 10 posibles valores, de modo que el espacio tiene $10^3 = 1,000$ puntos. Como todos tienen la misma probabilidad, la respuesta en este caso es $1/1,000 = 0.001$.

3. Si lanzamos dos dados, ¿Cuál es la probabilidad de que la suma sea 7?

   Vamos a suponer, para facilitar el razonamiento, que lanzamos un dado primero y luego el otro. Por lo tanto un espacio muestral adecuado para este experimento es el conjunto de pares ordenados formados con los enteros del 1 al 6, con reposición:

   $$\Omega = \{(a, b), 1 \leq a, b \leq 6\}.$$

   En este caso todos los eventos elementales de Ω tienen la misma probabilidad: $1/36$. Los resultados que tienen componentes cuya suma es 7 son

   $$(1, 6); \quad (2, 5); \quad (3, 4); \quad (4, 3); \quad (5, 2); \quad (6, 1).$$

Por lo tanto la probabilidad de que la suma de los dados sea 7 es

$$6 \times \frac{1}{36} = \frac{1}{6}.$$

En este ejemplo podemos considerar otro espacio muestral: el conjunto de las sumas posibles

$$\Omega' = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}.$$

El problema para usar este espacio como base para nuestro análisis es que sus elementos no son equiprobables. Por ejemplo, para tener una suma de 2, ambos dados tienen que salir 1, lo cual tiene probabilidad $1/36$, y acabamos de ver que la probabilidad de que la suma sea 7 es $1/6$.

4. Si lanzamos dos monedas, ¿cuál es la probabilidad de obtener un águila y un sol?

   Este problema lo hemos incluido para resaltar una dificultad importante que se ejemplifica con el razonamiento de D'Alembert, famoso matemático francés del siglo XVIII, quien argumentó que sólo hay tres casos posibles en esta situación:

   (1) dos aguilas, (2) dos soles, (3) un águila y un sol,

   y concluyó que la probabilidad de obtener un águila y un sol es $1/3$. Como hemos visto, el último caso en realidad debe separarse en dos:

   (3a) La primera moneda es águila y la segunda es sol.

   (3b) La primera moneda es sol y la segunda es águila.

   Esto es obvio si lanzamos una moneda tras otra y no simultáneamente, o si las monedas son distinguibles. Por lo tanto la respuesta correcta es $2/4 = 1/2$. Haremos una observación importante sobre este caso en la sección **??**.

5. Si lanzamos una moneda dos veces y una de las veces sale águila ¿Cuál es la probabilidad de que el otro lanzamiento haya sido sol?

   Para este ejemplo el espacio muestral es

   $$\Omega = \{SS, SA, AS, AA\}$$

   y todos los resultados tienen igual probabilidad de ocurrir. Si sabemos que uno de los lanzamientos fue A, nos quedan tres resultados posibles y de ellos en dos casos el otro lanzamiento es S. Por lo tanto la probabilidad es $2/3$.

   La situación sería distinta si nos dicen que el primer lanzamiento resultó A, pues en este caso el segundo tiene dos posibilidades A y S con igual probabilidad, y la respuesta en este caso sería que la probabilidad es $1/2$.

1.4.2. Probabilidades en Espacios Numerables.

Un caso similar al desarrollado en la sección anterior se presenta cuando Ω es un conjunto infinito numerable:

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_m, \dots\}, \quad \mathcal{A} = \mathcal{P}(\Omega) \quad \text{y} \quad P(A) = \sum_{i:\omega_i \in A} p_i,$$

donde los números p_i verifican

$$\begin{cases} p_i \geq 0, \quad \text{para todo } i, \\ \displaystyle\sum_{i=1}^{\infty} p_i = 1. \end{cases}$$

Claramente en este caso no es posible que los p_i sean todos iguales, ya que de ser así no pueden satisfacer las condiciones anteriores. En el capítulo 3 consideraremos en más detalle estos espacios y los del ejemplo anterior.

Veamos un ejemplo.

1. Lanzamos una moneda hasta que salga 'Aguila' por primera vez. Los resultados posibles de este experimento son los números naturales: $\Omega = \mathbb{N}$. La probabilidad de obtener 'Aguila' en el primer lanzamiento es $1/2$. La probabilidad de salga 'Sol' en el primer lanzamiento y 'Aguila' en el segundo es $(1/2) \times (1/2) = 1/4$. La probabilidad de tener 'Sol' dos veces y luego 'Aguila' es $1/8$ y así sucesivamente. Vemos que la probabilidad de obtener 'Aguila' por primera vez en el n-ésimo lanzamiento es $p_n = 1/2^n$. Tenemos que verificar que esta asignación define una probabilidad y para esto es necesario que
$$\sum_{n=1}^{\infty} p_n = 1.$$
Recordamos la fórmula para una serie geométrica:
$$1 + r + r^2 + r^3 + \cdots = \frac{1}{1-r} \tag{1.3}$$
y multiplicando ambos lados por r obtenemos
$$r + r^2 + r^3 + r^4 + \cdots = \frac{r}{1-r} \tag{1.4}$$
para $-1 < r < 1$.

   Si ponemos $r = 1/2$ en (1.4) obtenemos que la suma $\sum p_n$ vale 1. Además de comprobar que p_n define una probabilidad sobre Ω, este resultado muestra que con probabilidad 1 obtendremos un 'Aguila' en un número finito de lanzamientos, o equivalentemente, que la probabilidad de no obtener nunca 'Aguila' en una sucesión de lanzamientos de una moneda balanceada es 0.

   Sea ahora A el evento 'la primera Aguila se obtiene en un número par de lanzamientos'. Tenemos que $A = \{2, 4, 6, \dots\}$ y
$$P(A) = \frac{1}{4} + \frac{1}{16} + \frac{1}{64} + \cdots = \frac{1}{4} + \Big(\frac{1}{4}\Big)^2 + \Big(\frac{1}{4}\Big)^3 + \cdots$$
Poniendo $r = 1/4$ en la ecuación (1.4) obtenemos que
$$P(A) = \frac{1/4}{1 - 1/4} = \frac{1}{3},$$
de modo que la probabilidad de que la primera 'Aguila' salga en un número par de lanzamientos es $1/3$ y en un número impar, $2/3$.

1.4.3. Probabilidades en Espacios Continuos

Si el experimento consiste en seleccionar al azar un número en el intervalo $[0, 1]$, entonces la probabilidad de escoger un número en el intervalo $[c, d] \subset [0, 1]$ debe ser proporcional a la longitud del intervalo, pero como la probabilidad de que el número escogido caiga en el intervalo $[0, 1]$ es 1, vemos que no sólo es proporcional sino que es igual a la longitud del intervalo:
$$P([c, d]) = d - c, \qquad \text{para todo } [c, d] \subset [0, 1]. \tag{1.5}$$
Lamentablemente, no es posible definir una medida de probabilidad sobre todos los subconjuntos de $[0, 1]$ que satisfaga la propiedad (1.5). La demostración de este hecho está fuera de los objetivos de este curso,

pero esto implica que hay conjuntos que no son 'medibles', es decir, a los cuales no podemos asignarles una probabilidad.

Por lo tanto, es necesario restringirse a una clase más pequeña $\mathcal{F}$ de subconjuntos de $[0,1]$, que sea una σ-álgebra, es decir, que satisfaga las condiciones A1, A2 y A3. Una posibilidad es usar la clase de los conjuntos borelianos en $[0,1]$, que es la menor σ-álgebra generada por los subintervalos de $[0,1]$. Sin embargo, es importante observar que en este caso hay otras σ-álgebras que pueden considerarse.

Dada cualquier colección $\mathcal{C}$ de subconjuntos de Ω, es posible demostrar que existe una σ-álgebra, que denotaremos por $\sigma(\mathcal{C})$, que contiene a $\mathcal{C}$ y que es la menor de todas las σ-álgebras que contienen a $\mathcal{C}$ en el siguiente sentido: Si $\mathcal{D}$ es otra σ-álgebra que contiene a $\mathcal{C}$, entonces se cumple que $\sigma(\mathcal{C}) \subset \mathcal{D}$. $\sigma(\mathcal{C})$ se conoce como la σ-álgebra *generada* por $\mathcal{C}$ y es posible demostrar que siempre existe y es única.

En el ejemplo anterior mencionamos a la σ-álgebra de Borel $\mathcal{B}$ en $[0,1]$, que tiene gran importancia en el desarrollo de la teoría de la medida, y que introdujimos como la σ-álgebra generada por los subintervalos de $[0,1]$. De manera equivalente se puede definir como la σ-álgebra generada por la colección de los intervalos abiertos (a,b), $0 \leq a < b \leq 1$, o los intervalos cerrados $[a,b]$ o los de la forma $(a,b]$, o de la forma $[a,b)$. Es posible demostrar que todas estas definiciones son equivalentes.

También es posible definir la σ-álgebra de Borel como la σ-álgebra generada por los subconjuntos abiertos de $[0,1]$, y se puede demostrar que esta definición es equivalente a cualquiera de las anteriores. Esta definición tiene la ventaja de que podemos usarla en cualquier espacio que tenga una topología, por ejemplo, en cualquier espacio métrico.

1.4.4. Otros Ejemplos

(1) *Muestreo con Reposición.* Retomemos el ejemplo 1.2.9 sobre el muestreo con reposición, donde

$$\Omega = \{BBB, BBD, BDB, DBB, BDD, DBD, DDB, DDD\}$$

y sea $\mathcal{A} = \mathcal{P}(\Omega)$. Supongamos que la proporción de defectuosos en la población es $p = n/N$, donde n es el número de defectuosos en el total N de artículos en el stock. Por lo tanto, la proporción de buenos en la población es $1 - p = q$.

Consideremos el evento elemental $\{DDD\}$. Para asignarle la probabilidad correspondiente razonamos así: en cada una de las extracciones hay n formas posibles de elegir un defectuoso. En total resultan n^3 posibilidades de obtener los tres defectuosos y N^3 elecciones posibles de una terna cualquiera. Asignamos al evento $\{DDD\}$ la probabilidad

$$P(\{DDD\}) = \frac{n^3}{N^3} = p^3$$

y análogamente

$$\begin{gathered} P(\{BBB\}) = q^3, \\ P(\{BDD\}) = P(\{DDB\}) = P(\{DBD\}) = p^2q, \\ P(\{BBD\}) = P(\{BDB\}) = P(\{DBB\}) = pq^2. \end{gathered}$$

Se verifica que

$$P(\Omega) = p^3 + 3p^2q + 3pq^2 + q^3 = (p+q)^3 = 1.$$

Calculemos la probabilidad del evento A: "se obtiene al menos un defectuoso en la muestra". Como A es el complemento del evento A^c: "no se obtiene ningún defectuoso en la muestra", resulta

$$P(A) = 1 - P(A^c) = 1 - q^3.$$

Consideremos ahora la siguiente situación que se presenta en problemas vinculados a control de calidad. Supongamos que se ignora la proporción p de defectuosos en la población y estamos interesados en tener una estimación de ese valor. Extraemos una muestra de tres artículos entre los cuales hay uno solo defectuoso.

Analicemos la probabilidad del evento: "se obtiene un solo defectuoso en la muestra", según diversos valores de p, como indica el cuadro siguiente:

p	$3pq^2$
0.1	0.243
0.2	0.384
0.3	0.441
0.4	0.432
0.5	0.375
0.6	0.288
0.7	0.189
0.8	0.096
0.9	0.027

Si tuviéramos que seleccionar uno de estos valores para p, una opción posible sería admitir aquél que haga mayor la probabilidad del evento que ocurrió efectivamente, o sea 0.3.

Utilizando este criterio, y aceptando como posibles valores de p todos los números reales entre 0 y 1, adoptamos como estimación aquél que haga máxima la probabilidad $3pq^2 = 3p(1-p)^2$ del evento que efectivamente ocurrió. Este criterio de estimación se llama de *"máxima verosimilitud"*. Para maximizar esta función

$$L(p) = 3p(1-p)^2$$

calculamos su derivada

$$L'(p) = 3(1-p)(1-3p)$$

que se anula en $p = 1, \ p = 1/3$.

El gráfico de la función $L(p)$ es el que se indica en la figura 2.10, y el máximo para $p \in [0, 1]$ está en $p = 1/3$. Tomamos, por lo tanto, como estimación $\hat{p} = 1/3$, valor que obviamente se adecúa a lo que indica la intuición inmediata, dado que en la muestra de tres resultó uno defectuoso.

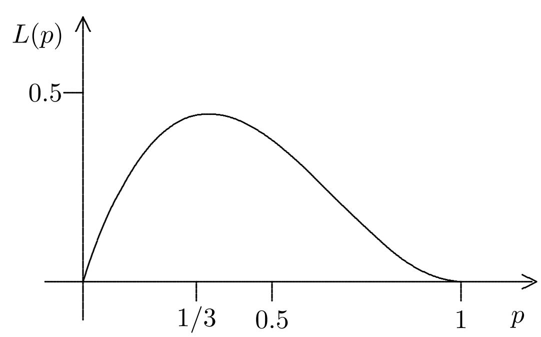

Figura 2.10

(2) *Error de Redondeo.* Consideremos nuevamente el caso del error de redondeo. Supongamos que se trunca el resultado de una operación aritmética en la parte entera, es decir, que en lugar del número real no negativo x se toma su parte entera $[x]$, esto es, el mayor entero que no supera x. El planteo es esencialmente el mismo si se trata del truncamiento en cualquier cifra decimal.

El error cometido al truncar es $x - [x]$, que podemos considerar como un evento elemental del intervalo $[0, 1) = \Omega$, tomado como espacio muestral.

Con frecuencia – como veremos al examinar este problema mas adelante – estaremos interesados en asignar al espacio muestral Ω una probabilidad uniforme en el siguiente sentido: intervalos de igual longitud deben tener igual probabilidad. No es difícil probar que una tal probabilidad P debe verificar

$$P([a,b)) = b - a \tag{1.6}$$

cualquiera que sea el intervalo $[a,b), 0 \leq a < b < 1$. Una manera de hacerlo es la siguiente: si P tiene esa propiedad, para n natural se tiene

$$P\left(\left[0, \frac{1}{n}\right)\right) = P\left(\left[\frac{1}{n}, \frac{2}{n}\right)\right) = \cdots = P\left(\left[\frac{n-1}{n}, 1\right)\right)$$

y como la suma de estas n probabilidades es $P(\Omega) = 1$ resulta que cada una de ellas es $1/n$.

Si m y n son enteros positivos, $m < n$, resulta que

$$P\left(\left[0, \frac{m}{n}\right)\right) = P\left(\left[0, \frac{1}{n}\right)\right) + \cdots + P\left(\left[\frac{m-1}{n}, \frac{m}{n}\right)\right) = \frac{m}{n}.$$

Si x es un número real cualquiera perteneciente al intervalo $(0,1)$, consideremos dos sucesiones de números racionales

$$\frac{m_k}{n_k} < x < \frac{m'_k}{n'_k}, \quad \frac{m_k}{n_k} \to x, \quad \frac{m'_k}{n'_k} \to x, \quad k \to \infty,$$

y se tiene

$$\frac{m_k}{n_k} = P\left(\left[0, \frac{m_k}{n_k}\right)\right) \leq P([0,x)) \leq P\left(\left[0, \frac{m'_k}{n'_k}\right)\right) = \frac{m'_k}{n'_k}.$$

Pasando al límite para $k \to \infty$ resulta

$$x \leq P([0,x)) \leq x \quad \Rightarrow \quad P([0,x)) = x,$$

o sea que la probabilidad de cada intervalo es su longitud.

Como familia de eventos podemos tomar la menor σ-álgebra que contiene a los intervalos, llamada σ-álgebra de Borel, que denotamos $\mathcal{B}$, y se puede probar que existe efectivamente una probabilidad P definida para la familia de eventos $\mathcal{B}$, que satisface (1.6), es decir, que asigna probabilidades iguales a intervalos de igual longitud.

Determinemos ahora la probabilidad del evento

$$A : \text{“ La primera cifra truncada es 9 ”}.$$

Resulta

$$P(A) = P([0.9,\ 1)) = 1 - 0.9 = 0.1.$$

¿Cuál es la probabilidad de que la segunda cifra truncada sea 9?

Este evento es

$$B = [0.09,\ 0.1) \cup [0.19,\ 0.2) \cup \cdots \cup [0.99,\ 1)$$

y su probabilidad es

$$P(B) = \overbrace{\frac{1}{100} + \cdots + \frac{1}{100}}^{10 \text{ veces}} = 0.1.$$

(3) Un dado está cargado de modo tal que la probabilidad de que salga la cara k es proporcional a k. Hallar la probabilidad de cada uno de los eventos:

a. El resultado de arrojar el dado es un número par.

b. El resultado es menor que 6.

Denotemos por p_k la probabilidad de que ocurra la cara k ($k = 1, 2, 3, 4, 5, 6$). Lo que establece el enunciado es que existe una constante C tal que $p_k = Ck$. Como $p_1 + p_2 + \cdots + p_6 = 1$, se deduce que

$$C(1 + 2 + \cdots + 6) = 1 \quad \Rightarrow \quad 21C = 1 \quad \Rightarrow \quad C = \frac{1}{21} \quad \Rightarrow \quad p_k = \frac{k}{21}$$

Resolvamos ahora a y b.

a. La probabilidad de obtener una cara par es

$$p_2 + p_4 + p_6 = \frac{12}{21} = \frac{4}{7}.$$

b. La probabilidad de obtener un resultado menor que 6 es

$$p_1 + p_2 + p_3 + p_4 + p_5 = \frac{15}{21} = \frac{5}{7}.$$

▲

(4) *El problema de los cumpleaños.* ¿Cuál es la probabilidad de que entre r personas al menos dos cumplan años el mismo día? (Se supone que la duración del año es de 365 días). ¿Cuál es el menor valor de r para el cual esta probabilidad es superior a $1/2$?

Tomamos como espacio muestral el conjunto de todas las r-uplas de fechas posibles:

$$\Omega = \{(f_1, f_2, \dots, f_r) \ : \ 1 \le f_i \le 365, \ i = 1, \dots, r\}$$

y la hipótesis natural es que todas las r-uplas son igualmente probables.

Llamemos A el evento de que entre los r individuos seleccionados, no hay dos que cumplan el mismo día, es decir que

$$A = \{(f_1, \dots, f_r) : 1 \le f_i \le 365, \text{ los } f_i \text{ son diferentes 2 a 2}\}.$$

La pregunta es ¿cuál es la probabilidad de que no ocurra A? Esto es

$$P(A^c) = 1 - P(A)$$

y como todos los eventos elementales de Ω son igualmente probables,

$$P(A) = \frac{\sharp A}{\sharp \Omega}$$

donde $\sharp A$ denota el cardinal del conjunto A. Para calcular $\sharp\Omega$ observamos que hay $N = 365$ fechas posibles para cada cumpleaños y queremos seleccionar un vector de longitud r de fechas con reposición. Cada componente del vector la podemos escoger de N maneras y por lo tanto hay N^r vectores posibles:

$$\sharp\Omega = N^k.$$

Por otro lado los vectores que pertenecen a A no pueden tener componentes repetidas. Por lo tanto, para escoger un vector que satisfaga esta condición, tenemos N valores posibles para la primera componente. Una vez que la hemos seleccionado, quedan $N - 1$ valores para escoger la segunda. Si ya tenemos las dos primeras quedan $N - 2$ posibilidades para la tercera y, continuando de esta manera, para la última componente quedan $N - r + 1$. En consecuencia

$$\sharp A = N(N-1)\cdots(N-r+1),$$

y por lo tanto

$$P(A^c) = 1 - \frac{N(N-1)\cdots(N-r+1)}{N^r} = 1 - \left(1 - \frac{1}{N}\right)\dots\left(1 - \frac{r-1}{N}\right).$$

Para acotar esta probabilidad utilizamos la desigualdad

$$1 - x \leq e^{-x}$$

válida para todo $x \in \mathbb{R}$, que puede ser demostrada usando un desarrollo de MacLaurin de orden 2 o verificando que la figura 2.11 es correcta.

Obtenemos

$$P(A^c) > 1 - e^{\frac{1}{N}+\frac{2}{N}+\cdots+\frac{r-1}{N}} = 1 - e^{-\frac{r(r-1)}{2N}}.$$

Para $r = 23$ y $N = 365$ obtenemos $P(A^c) > 0.50000175$. Así, en un grupo de 23 personas, con probabilidad mayor que 1/2, hay dos personas que cumplen años el mismo dia, lo cual es bastante sorprendente. Para $r = 30$ la probabilidad es superior a 0.696 y para $r = 50$, superior a 0.965.

n	$P(A_n)$
10	0.117
20	0.411
23	0.507
30	0.706
50	0.97
57	0.99
100	0.9999997

▲

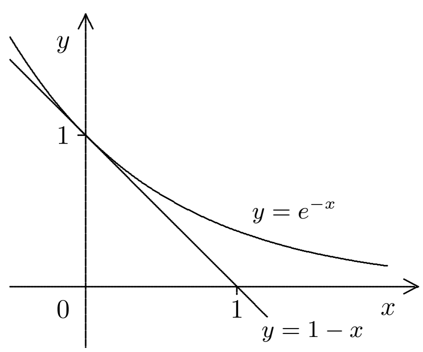

Figura 2.11

(5) Si la probabilidad de encontrar un artículo defectuoso en una población es $p = n/N$, donde N es el número de elementos de la población y n el de defectuosos, y realizamos muestreo con reposición extrayendo un artículo cada vez, calcular la probabilidad de encontrar el primer defectuoso en la m-ésima extracción. Si llamamos p_m a esta probabilidad, verificar que $\sum_{m=1}^{\infty} p_m = 1$.

Veremos ahora una solución al ejercicio con los elementos de que disponemos. Más adelante podremos tratarlo de manera más simple, utilizando conceptos que aún no hemos introducido. Comencemos por $m = 1$; p_1 es la probabilidad de extraer un defectuoso en la primera extracción, que es claramente

$$p_1 = \frac{n}{N} = p.$$

Sea ahora $m > 1$. El evento A_m: "el primer defectuoso es extraído en la m-ésima extracción", se escribe como

$$A_m = B_{m-1} \setminus B_m$$

donde B_m es el evento de que en las primeras m extracciones no hemos encontrado artículos defectuosos. La relación anterior expresa que "encontrar un defectuoso por primera vez en la m-ésima extracción" es lo mismo que "no extraer defectuosos en las $m-1$ primeras pero si en las m primeras". Como $B_m \subset B_{m-1}$ se tiene que $P(A_m) = P(B_{m-1}) - P(B_m)$. Por otra parte

$$P(B_m) = \frac{(N-n)^m}{N^m} = (1-p)^m$$

y, por lo tanto, deducimos que

$$\begin{aligned} p_m = P(A_m) &= (1-p)^{m-1} - (1-p)^m \\ &= (1-p)^{m-1}(1-(1-p)) = p(1-p)^{m-1}. \end{aligned}$$

En resumen, la fórmula

$$p_m = p(1-p)^{m-1}$$

vale para todo $m \geq 1$. Además, como $p > 0$,

$$\sum_{m=1}^{\infty} p(1-p)^{m-1} = p\sum_{m=0}^{\infty}(1-p)^m = \frac{p}{1-(1-p)} = 1.$$

Aquí hemos usado la suma de la serie geométrica:

$$\sum_{m=0}^{\infty} x^m = \frac{1}{1-x}, \qquad \text{válida para } |x| < 1.$$

◀

1.5. Ejercicios

1. Demuestre que la diferencia simétrica de dos conjuntos se puede escribir como $A \triangle B = (A \setminus B) \cup (B \setminus A)$.

2. Demuestre que para cualesquiera conjuntos $A,\ B,\ C$ y D, la siguiente proposición es cierta:

$$(A \cup B) \setminus (C \cup D) \subseteq (A \setminus C) \cup (B \setminus D).$$

3. Sea A, B, C, D subconjuntos de Ω. Demuestre las siguientes propiedades.

a. $((A \cap B) \cup (C \cap D))^c = (A^c \cup B^c) \cap (C^c \cup D^c)$; b. $A \triangle \Omega = A^c$.
c. $(A \cup B) \cap (A \cup B^c) \cap (A^c \cup B) \cap (A^c \cup B^c) = \emptyset$; d. $A \setminus B = A \cap (A \triangle B)$.
e. $A \cup B = (A \triangle B) \triangle (A \cap B)$; f. $A \triangle (B \triangle C) = (A \triangle B) \triangle C$.
g. $(A \cap B^c) \triangle (B \cap A^c) = A \triangle B$; h. $A \triangle B = C \triangle D \Rightarrow A \triangle C = B \triangle D$.
i. $A \cap (B \triangle C) = (A \cap B) \triangle (A \cap C)$; j. $A \triangle B = (A \triangle C) \triangle (C \triangle B)$.

4. ¿Cuándo son ciertas las siguientes relaciones?

a. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ b. $A \cap (B \cap C) = (A \cap B) \cap C$
c. $A \cup (B \cup C) = A \setminus (B \setminus C)$ d. $(A \setminus B) \setminus C = A \setminus (B \setminus C)$
e. $A \Delta (B \Delta C) = (A \Delta B) \Delta C$ f. $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$
g. $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$

5. Si $\{A_i,\ i \in I\}$ y $\{B_i,\ i \in I\}$ son dos colecciones de conjuntos, demuestre que

$$(\cup_{i \in I} A_i) \setminus (\cup_{i \in I} B_i) \subset \cup_{i \in I}(A_i \setminus B_i).$$

6. Sea A_n el conjunto de los enteros positivos divisibles por n. Halle los conjuntos a) $\cup_{n=2}^{\infty} A_n$, b) $\cap_{n=2}^{\infty} A_n$.

7. Halle los siguientes conjuntos a) $\cup_{n=1}^{\infty}\left[a+\frac{1}{n}, b-\frac{1}{n}\right]$; b) $\cap_{n=1}^{\infty}\left(a-\frac{1}{n}, b+\frac{1}{n}\right)$.

8. Sea A_α el conjunto de puntos sobre la curva $y = 1/x^\alpha$ para $0 < x < \infty$. Halle el conjunto $\cap_{\alpha \geq 1} A_\alpha$.

9. Sean A_1, A_2 y A_3 eventos de un espacio muestral. Expresar mediante uniones, intersecciones y complementos los siguientes eventos:

   a. Los tres eventos ocurren.
   b. Ocurre sólo A_1.
   c. Ocurren A_1 y A_2, pero no A_3.
   d. Ocurre al menos uno de los tres eventos.
   e. No ocurre ninguno.
   f. Ocurren al menos dos.
   g. Ocurren dos y no más.

10. Expresar como uniones disjuntas a) $A_1 \cup A_2$; b) $A_1 \cup A_2 \cup A_3$; c) $\bigcup_{i=1}^{n} A_i$.

11. Sea $\Omega = \{AAA, AAS, ASA, SAA, ASS, SAS, SSA, SSS\}$, donde A es águila y S es sol. Describa con palabras los siguientes eventos y calcule sus probabilidades:

    a. $B = \{AAA, AAS, ASA, ASS\}$;
    b. $C = \{AAA, SSS\}$.
    c. $D = \{AAS, ASA, SAA\}$;
    d. $E = \{AAS, ASA, SAA, ASS, SAS, SSA\}$.

12. El evento $A \setminus B$ quiere decir que A ocurre pero B no. Demuestre que las operaciones de unión, intersección y complemento se pueden expresar usando sólo esta operación.

13. Sean A, B y C tres eventos. Demuestre las siguientes propiedades:

    a. $P(A \cap B) + P((A \setminus B) \cup (B \setminus A)) + P(A^c \cap B^c) = 1$.

    b. $P(A^c \cap B^c) + P(A) + P(A^c \cap B) = 1$.

    c. $P(A^c \cap B^c \cap C^c) + P(A) + P(A^c \cap B) + P(A^c \cap B^c \cap C) = 1$.

    d. $P(A \cap B) - P(A)P(B) = P(A^c \cap B^c) - P(A^c)P(B^c)$.

14. Suponga que $P(A) \geq 0.9$, $P(B) \geq 0.8$ y $P(A \cap B \cap C) = 0$, demuestre que $P(C) \leq 0.3$.

15. Demuestre que si $A \cap B \cap C = \emptyset$, entonces

$$P((A \cup B) \cap (B \cup C) \cap (C \cup A)) = P(A \cap B) + P(B \cap C) + P(C \cap A).$$

16. Sea D el evento 'exactamente uno de los eventos A, B y C ocurre'. Exprese $P(D)$ en términos de $P(A), P(B), P(C), P(A \cap B), P(A \cap C), P(B \cap C)$ y $P(A \cap B \cap C)$.

17. Demuestre que

    a. $\min\{1, P(A) + P(B)\} \geq P(A \cup B) \geq \text{máx}\{P(A), P(B)\}$

    b. $\min\{P(A), P(B)\} \geq P(A \cap B) \geq \text{máx}\{0, P(A) + P(B) - 1\}$

    c. $P(\cap_1^n A_i) \geq \sum_1^n P(A_i) - (n-1)$.

18. La condición de σ-aditividad para una medida de probabilidad es equivalente a otras propiedades. Pruebe que es equivalente a las proposiones (a) y (b):

    a. Si $A_1 \subseteq A_2 \subseteq \cdots$ es una sucesión creciente de eventos y $A = A_1 \cup A_2 \cup \cdots$ entonces $P(A) = \lim_{n\to\infty} P(A_n)$.

    b. Si $A_1 \supseteq A_2 \supseteq \cdots$ es una sucesión decreciente de eventos y $A = A_1 \cap A_2 \cap \cdots$ entonces $P(A) = \lim_{n\to\infty} P(A_n)$.

19. Un experimento consiste de tomar al azar tres focos de la producción de una fábrica y probarlos, con resultados posibles defectuoso (1) o bueno (0). Sea A el evento 'El primer foco es defectuoso', B el evento 'el segundo foco es defectuoso' y C el evento 'el tercer foco es defectuoso'.

    a) Describa el espacio muestral Ω para este experimento.

    b) Liste los elementos de cada uno de los siguientes conjuntos: A, B, $A \cup C$, $B \cup C$, $A \cup B \cup C$, $A \cap B$, $A^c \cap B \cap C^c$, $A \cap B^c \cap C$, $(A \cup B^c) \cap C$, $(A^c \cap C) \cup (B \cap C)$.

20. Describa en detalle un espacio muestral para los siguientes experimentos: (a) Tres lanzamientos de un dado. (b) Calificaciones de una clase de 20 estudiantes en un examen. (c) Medición de la temperatura a mediodía en una estación meteorológica. (d) Medición de las velocidades de carros pasando por un punto dado. (e) Una sucesión infinita de lanzamientos de una moneda.

21. Una caja contiene n bolas rojas y n bolas blancas. Se extraen dos bolas al azar. ¿Cuál es el espacio muestral para este experimento? ¿Cuál es la probabilidad de que las dos bolas tengan colores distintos. Halle la probabilidad p_n de que las bolas sean del mismo color y evalúe $\lim_{n\to\infty} p_n$.

22. En una bolsa hay tres cartas numeradas 1, 2 y 3. Sacamos las cartas al azar sucesivamente y sin reposición. El resultado es una permutación de los números 1, 2 y 3. Describa el espacio muestral de este experimento. Para cada una de las siguientes descripciones, liste los elementos correspondientes y calcule la probabilidad del evento.

    (a) El 2 sale en primer lugar. (b) El 3 sale en segundo lugar. (c) El 2 sale en primer lugar y el 1 en tercer lugar. (d) O bien el 2 sale en primer lugar o bien el 1 sale en tercer lugar (o ambos). (e) Ningún número ocupa su lugar.

23. Repita el ejercicio anterior con 4 cartas numeradas de 1 a 4.

24. Una caja contiene 10 bolas negras y 5 bolas rojas. Se extraen 3 bolas al azar, con reposición. Calcular:

    a. La probabilidad de que sean 2 negras y una roja.

    b. La probabilidad de que sean las tres negras.

    c. Repetir el ejercicio suponiendo que la extracción es sin reposición.

25. Se extraen dos cartas sucesivamente de un juego de 52 cartas. Halle la probabilidad de que la segunda carta sea mayor que la primera.

26. Se lanzan al aire simultáneamente tres monedas balanceadas. Calcular:

    a. La probabilidad de obtener 3 caras.

    b. La probabilidad de obtener por lo menos 2 caras.

27. Lanzamos una moneda balanceada cuatro veces. Calcule la probabilidad de los siguientes eventos:

    a. Ocurren al menos tres Aguilas.

    b. Ocurren exactamente tres Aguilas.

    c. Ocurren al menos tres Aguilas consecutivas.

    d. Ocurren exactamente tres Aguilas consecutivas.

28. Se lanzan dos dados. Calcule la probabilidad de los siguientes eventos: a) obtenemos el mismo número en ambos dados; b) la suma es 7 u 11; c) los números son primos relativos, d) la suma es impar; e) el producto es impar; f) un número divide al otro.

29. Se realiza un test de conocimientos con 11 preguntas a contestar por *sí* o *no*. Se da por aprobada la prueba si se contestan correctamente al menos 6 de las 11 preguntas. ¿Cuál es la probabilidad de aprobar el examen contestando al azar?

30. Sean $P_1,\ P_2$ dos medidas de probabilidad definidas sobre la misma σ-álgebra $\mathcal{F}$ y sea $0 \leq \alpha \leq 1$. Demuestre que $\alpha P_1 + (1-\alpha)P_2$ también es una medida de probabilidad sobre $\mathcal{F}$. Generalice el resultado a n medidas de probabilidad.

31. a. Sea $p_i = a/i^2$ para $i \in \mathbb{N}$. Halle el valor de a para que p_i defina una probabilidad.

    b. Sea $p_i = b/i^2$ para $i = \pm 1, \pm 2, \dots$. Halle el valor de b para que p_i defina una probabilidad.

32. Para comenzar un cierto juego es necesario lanzar un 6 con un dado.

    a) ¿Cuál es la probabilidad de lanzar el primer 6 en el tercer intento?

    b) ¿Cuál es la probabilidad de necesitar más de tres intentos?

    c) ¿Cuántos lanzamientos hacen falta para que la probabilidad de haber lanzado un 6 sea al menos 0.95?

    d. ¿Cuál es la probabilidad de que el primer 6 ocurra en un número par de lanzamientos?

33. Sea $\mathcal{F}$ una σ-álgebra de eventos en un espacio finito. Demuestre que $\mathcal{F}$ no puede contener exactamente 6 eventos. ¿Qué enteros pueden ser el cardinal de $\mathcal{F}$?

34. Sean: $\Omega = [0,1]$, $\mathcal{B}$ la familia de conjuntos de Borel y P la probabilidad definida en el ejemplo 6 de la sección 2.4.

    a. Probar que $P(\{\omega\}) = 0$, donde $\{\omega\}$ es el subconjunto de Ω que consta sólo del punto ω. (Verificar previamente que $\{\omega\} \in \mathcal{B}$).

    b. Sean $Q = \{\omega : \omega \in [0,1] \text{ es racional}\}$ e $I = \{\omega : \omega \in [0,1] \text{ es irracional}\}$. Probar que $P(Q) = 0$ y $P(I) = 1$.

35. Se lanza reiteradamente una moneda balanceada. ¿Cuál es la probabilidad de que ocurran 4 caras antes que dos sellos?

36. Se lanzan cuatro dados y se multiplican los números que se obtienen. ¿Cuál es la probabilidad de que este producto sea divisible por 5? ¿Cuál es la probabilidad de que el último dígito en el producto sea 5?

37. Antonio y Bruno acuerdan una competencia de esgrima en una serie de mangas de modo que el primero en ganar dos mangas seguidas gana el combate. Antonio tiene probabilidad p de ganar una manga y Bruno probabilidad $q = 1 - p$. ¿Cuál es la probabilidad de que la competencia termine al cabo de k mangas?

38. En una caja tenemos n bolas con los números del 1 al n. Sea D_r el evento: 'se extrae una bola al azar y el número es divisible por r'. Halle $P(D_3),\ P(D_4),\ P(D_3 \cup D_4)$ y $P(D_3 \cap D_4$ y obtenga los límites de estas probabilidades cuando $n \to \infty$.

39. Definimos la función d sobre $\mathcal{F} \times \mathcal{F}$ por $d(A,B) = P(A \Delta B)$.

    a. Demuestre que para cualesquiera eventos $A,\ B$ y C

    $$d(A,B) + d(B,C) - d(A,C) = 2P(A \cap B^c \cap C) + P(A^c \cap B \cap C^c)$$

    b. ¿Cuándo vale $d(A,B) = 0$?

    c. Sea $A_1, A_2, \dots$ una sucesión no-decreciente de eventos: $A_i \subseteq A_j$ para $i \leq j$. Demuestre que para $i \leq j \leq k$,

    $$d(A_i, A_k) = d(A_i, A_j) + d(A_j, A_k).$$

40. En el juego de 'craps' el jugador lanza dos dados. Si el resultado es 7 u 11, el jugador gana. Si es 2, 3 ó 12, pierde. Si es cualquier otro resultado k, continua lanzando hasta obtener un 7, en cuyo caso pierde, o k, en cuyo caso gana. ¿Cuál es un espacio muestral adecuado para este juego? ¿Cuál es la probabilidad de ganar? ¿Cuál es la probabilidad de ganar en el primero o segundo lanzamiento? ¿Cuál es la probabilidad de ganar si el primer lanzamiento es 6?