



# INGENIERÍA ECONÓMICA

Probabilidad y teoría de juegos

Antonio M. Mendoza • Gabriel Robles C.



# INGENIERÍA ECONÓMICA

Gabriel Robles  
Contreras

Marco Antonio  
Molina Mendoza



Instituto Politécnico Nacional



# CONTENIDO

## Capítulo 1

### PROBABILIDAD 7

1.1	Espacio de sucesos . . . . .	8
1.2	Eventos . . . . .	9
1.3	Probabilidad de un evento . . . . .	10
1.4	Reglas aditivas . . . . .	12
1.5	Probabilidad condicional e independencia . . . . .	15
1.6	Regla de Bayes . . . . .	19
1.7	Variables aleatorias y distribución de probabilidad . . . . .	21
1.8	Esperanza matemática . . . . .	29
1.9	Loterías . . . . .	29
1.10	Racionalidad . . . . .	29
1.11	Teoría de la utilidad esperada . . . . .	29

## Capítulo 2

### JUEGOS CON INFORMACIÓN PERFECTA 31

## APÉNDICES

### Apéndice A

#### TEORÍA DE CONJUNTOS 35

A.1	Subconjuntos e igualdad de conjuntos . . . . .	37
A.2	El conjunto vacío . . . . .	39
A.3	La unión e intersección de conjuntos . . . . .	40
A.4	Diferencia de conjuntos . . . . .	43
A.5	Complemento de un conjunto . . . . .	44
A.6	Diagramas de Venn-Euler . . . . .	45
A.7	Producto cartesiano . . . . .	46
A.8	Representación geométrica del producto cartesiano . . . . .	49
A.9	Familias de conjuntos . . . . .	49
A.10	Conjunto potencia . . . . .	51

# 1

## Capítulo

# PROBABILIDAD

### Contenido

1.1	Espacio de sucesos . . . . .	8
1.2	Eventos . . . . .	9
1.3	Probabilidad de un evento . . .	10
1.4	Reglas aditivas . . . . .	12
1.5	Probabilidad condicional e independencia . . . . .	15
1.6	Regla de Bayes . . . . .	19
1.7	Variables aleatorias y distribución de probabilidad . . . . .	21
1.8	Esperanza matemática . . . . .	29
1.9	Loterías . . . . .	29
1.10	Racionalidad . . . . .	29
1.11	Teoría de la utilidad esperada . .	29

En el estudio de la estadística tratamos básicamente con la presentación e interpretación de resultados fortuitos que ocurren en un estudio planeado o en una investigación científica. Por ejemplo, en México, y con la finalidad de justificar la instalación de un semáforo, se podría registrar el número de accidentes que ocurren mensualmente en la intersección de Eje 1 Norte y Avenida Insurgentes. En una fábrica, se podrían clasificar los artículos que salen de la línea de ensamble como “defectuosos” o “no defectuosos”. En una reacción química, se podría revisar el volumen de gas que se libera cuando se varía la concentración de un ácido. Por ello, quienes se dedican a la estadística a menudo manejan datos numéricos que representan conteos o mediciones, o datos categóricos que se podrían clasificar de acuerdo con algún criterio.

En este capítulo, al referirnos a cualquier registro de información, ya sea numérico o categórico, utilizaremos el término *observación*. Por consiguiente, los números 2, 0, 1 y 2, que representan el número de accidentes que ocurrieron cada mes, de enero a abril, durante el año pasado en la intersección de Avenida Insurgentes y Eje 1 Norte, constituyen un conjunto de observaciones. Lo mismo ocurre con los datos categóricos N, D, N, N y D, que representan los artículos defectuosos o no defectuosos cuando se inspeccionan cinco artículos y se registran como observaciones.

Los estadísticos utilizan la palabra *experimento* para describir cualquier proceso que genere un conjunto de datos. Un ejemplo simple de experimento estadístico es el lanzamiento de una moneda al aire. En tal experimento, sólo hay dos resultados posibles: sol o águila. En estadística, nos interesan, en particular, las observaciones que se obtienen al repetir varias veces un experimento. En la mayoría de los casos, los resultados dependerán del azar, por lo tanto, no se pueden predecir con certeza. Si un químico realizara un análisis varias veces en las mismas condiciones, obtendría diferentes medidas, las cuales indicarían un elemento de probabilidad en el procedimiento experimental. Aun cuando lancemos una moneda al aire repetidas veces, no podemos tener la certeza de que en un lanzamiento determinado obtendremos

águila como resultado. Sin embargo, podemos conocer el conjunto completo de posibilidades para cada lanzamiento.

1.1 Espacio de sucesos

**Definición 1.1.1:** Al conjunto de todos los resultados posibles de un experimento se le llama *espacio de sucesos* y se representa con el símbolo  $\Omega$ .

A cada suceso en un espacio de sucesos se le llama *elemento* o *miembro* del espacio de sucesos. Si el espacio de sucesos tiene un número finito de elementos, podemos listar los miembros separados por comas y encerrarlos entre llaves. Por consiguiente, el espacio de sucesos  $\Omega$ , de los resultados posibles cuando se lanza una moneda al aire, se puede escribir como

$$\Omega = \{\text{sol}, \text{águila}\} .$$

**Ejemplo 1.1.1:** Considere el experimento de lanzar un dado. Si nos interesara el número que aparece en la cara superior, el espacio de sucesos sería

$$\Omega_1 = \{1, 2, 3, 4, 5, 6\} .$$

Si sólo estuviéramos interesados en si el número es par o impar, el espacio de sucesos sería simplemente

$$\Omega_2 = \{\text{par}, \text{impar}\} .$$

El ejemplo anterior ilustra el hecho de que se puede usar más de un espacio de sucesos para describir los resultados de un experimento. En este caso,  $\Omega_1$  brinda más información que  $\Omega_2$ . Si sabemos cuál elemento ocurre en  $\Omega_1$ , podremos indicar cuál resultado tiene lugar en  $\Omega_2$ ; no obstante, saber lo que pasa en  $\Omega_2$  no ayuda mucho a determinar qué elemento ocurre en  $\Omega_1$ . En general, lo deseable sería utilizar un espacio de sucesos que proporcione la mayor información acerca de los resultados del experimento. En algunos experimentos es útil listar los elementos del espacio de sucesos de forma sistemática utilizando un *diagrama de árbol*.

**Ejemplo 1.1.2:** Un experimento consiste en lanzar una moneda y después lanzarla una segunda vez si sale sol. Si en el primer lanzamiento sale águila, entonces se lanza un dado una vez. Para listar los elementos del espacio de sucesos que proporciona la mayor información construimos el diagrama de árbol de la figura 1.1.

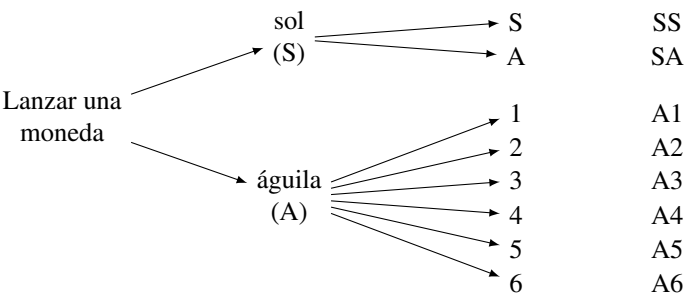


Figura 1.1 Diagrama de árbol

Si empezamos con la rama superior izquierda y nos movemos a la derecha a lo largo de la primera trayectoria, obtenemos el elemento SS, que indica la posibilidad de que ocurran caras en dos lanzamientos sucesivos de la moneda. De igual manera, el elemento A3 indica la posibilidad de que la moneda muestre un águila seguida por un 3 en el lanzamiento del dado. Al seguir todas las trayectorias, vemos que el espacio de sucesos es

$$\Omega = \{SS, SA, A1, A2, A3, A4, A5, A6\} .$$



Muchos de los conceptos de este capítulo se ilustran mejor con ejemplos que involucran el uso de dados y cartas. Es particularmente importante utilizar estas aplicaciones al comenzar el proceso de aprendizaje, ya que facilitan el flujo de esos conceptos nuevos en ejemplos científicos y de ingeniería.

## 1.2 Eventos

En cualquier experimento dado, podríamos estar interesados en la ocurrencia de ciertos *eventos*, más que en la ocurrencia de un elemento específico en el espacio de sucesos. Por ejemplo, quizás estemos interesados en el evento  $A$ , en el cual el resultado de lanzar un dado es divisible entre 3. Esto ocurrirá si el resultado es un elemento del subconjunto  $A = \{3, 6\}$  del espacio de sucesos  $\Omega_1$  del ejemplo 1.1.1.

Para cada evento asignamos un conjunto de elementos, que constituye un subconjunto del espacio de sucesos. Este subconjunto representa la totalidad de los elementos para los que el evento es cierto.

### Definición 1.2.1:

Un *evento* es un subconjunto de un espacio de sucesos.

### Ejemplo 1.2.1:

Dado el espacio de sucesos  $\Omega = \{t \mid t \geq 0\}$ , donde  $t$  es la vida en años de cierto componente electrónico, el evento  $A$  de que el componente falle antes de que finalice el quinto año es el subconjunto  $A = \{t \mid 0 \leq t < 5\}$ .

Es posible concebir que un evento puede ser un subconjunto que incluye todo el espacio de sucesos  $\Omega$ , o un subconjunto de  $\Omega$  que se denomina conjunto vacío y se denota con el símbolo  $\emptyset$ , que no contiene ningún elemento. Por ejemplo, si en un experimento biológico permitimos que  $A$  sea el evento de detectar un organismo microscópico a simple vista, entonces  $A = \emptyset$ . También, si

$$B = \{x \mid x \text{ es un factor par de } 7\},$$

entonces  $B$  debe ser el conjunto vacío, pues los únicos factores posibles de 7 son los números noes 1 y 7.

Considere un experimento en el que se registran los hábitos de tabaquismo de los empleados de una empresa industrial. Un posible espacio de suceso podría clasificar a un individuo como no fumador, fumador ocasional, fumador moderado o fumador empedernido. Si se determina que el subconjunto de los fumadores sea un evento, entonces la totalidad de los no fumadores corresponderá a un evento diferente, también subconjunto de  $\Omega$ , que se denomina complemento del conjunto de fumadores.

### Definición 1.2.2:

El complemento de un evento  $A$  respecto de  $\Omega$  es el subconjunto de todos los elementos de  $\Omega$  que no están en  $A$ . Denotamos el complemento de  $A$  mediante el símbolo  $A^C$ .

### Ejemplo 1.2.2:

Sea  $R$  el evento de que se seleccione una carta roja de una baraja ordinaria de 52 cartas, y sea  $\Omega$  toda la baraja. Entonces  $R^C$  es el evento de que la carta seleccionada de la baraja no sea una roja sino una negra.

### Definición 1.2.3:

La intersección de dos eventos  $A$  y  $B$ , que se denota con el símbolo  $A \cap B$ , es el evento que contiene todos los elementos que son comunes entre  $A$  y  $B$ .

### Ejemplo 1.2.3:

Sea  $E$  el evento de que una persona seleccionada al azar en un salón de clases sea estudiante de ingeniería, y sea  $F$  el evento de que la persona sea mujer. Entonces  $E \cap F$  es el evento de todas las estudiantes mujeres de ingeniería en el salón de clases.

**Ejemplo 1.2.4:** Sean  $V = \{a, e, i, o, u\}$  y  $C = \{l, r, s, t\}$ ; entonces, se deduce que  $V \cap C = \emptyset$ . Es decir,  $V$  y  $C$  no tienen elementos comunes, por lo tanto, no pueden ocurrir de forma simultánea.

Para ciertos experimentos estadísticos no es nada extraño definir dos eventos,  $A$  y  $B$ , que no pueden ocurrir de forma simultánea. Se dice entonces que los eventos  $A$  y  $B$  son mutuamente excluyentes. Expresado de manera más formal, tenemos la siguiente definición:

**Definición 1.2.4:** Dos eventos  $A$  y  $B$  son mutuamente excluyentes o disjuntos si  $A \cap B = \emptyset$ ; es decir, si  $A$  y  $B$  no tienen elementos en común.

A menudo nos interesamos en la ocurrencia de al menos uno de dos eventos asociados con un experimento. Por consiguiente, en el experimento del lanzamiento de un dado, si

$$A = \{2, 4, 6\} \quad \text{y} \quad B = \{4, 5, 6\},$$

podríamos estar interesados en que ocurran  $A$  o  $B$ , o en que ocurran tanto  $A$  como  $B$ . Tal evento, que se llama unión de  $A$  y  $B$ , ocurrirá si el resultado es un elemento del subconjunto  $\{2, 4, 5, 6\}$ .

**Definición 1.2.5:** La unión de dos eventos  $A$  y  $B$ , que se denota con el símbolo  $A \cup B$ , es el evento que contiene todos los elementos que pertenecen a  $A$  o a  $B$ , o a ambos.

**Ejemplo 1.2.5:** Sea  $P$  el evento de que un empleado de una empresa petrolera seleccionado al azar fume cigarrillos. Sea  $Q$  el evento de que el empleado seleccionado ingiera bebidas alcohólicas. Entonces, el evento  $P \cup Q$  es el conjunto de todos los empleados que beben o fuman, o que hacen ambas cosas.

### 1.3 Probabilidad de un evento

Quizá fue la insaciable sed del ser humano por el juego lo que condujo al desarrollo temprano de la teoría de la probabilidad. En un esfuerzo por aumentar sus triunfos, algunos pidieron a los matemáticos que les proporcionaran las estrategias óptimas para los diversos juegos de azar. Algunos de los matemáticos que brindaron tales estrategias fueron Pascal, Leibniz, Fermat y James Bernoulli. Como resultado de este desarrollo inicial de la teoría de la probabilidad, la inferencia estadística, con todas sus predicciones y generalizaciones, ha rebasado el ámbito de los juegos de azar para abarcar muchos otros campos asociados con los eventos aleatorios, como la política, los negocios, el pronóstico del clima y la investigación científica. Para que estas predicciones y generalizaciones sean razonablemente precisas, resulta esencial la comprensión de la teoría básica de la probabilidad.

En el resto de este capítulo consideraremos sólo aquellos experimentos para los cuales el espacio de sucesos contiene un número finito de elementos. La probabilidad de la ocurrencia de un evento que resulta de tal experimento estadístico se evalúa utilizando un conjunto de números reales denominados *probabilidades*, que van de 0 a 1. Para todo elemento en el espacio de sucesos asignamos una probabilidad tal que la suma de todas las probabilidades es 1. Si tenemos razón para creer que al llevar a cabo el experimento es bastante probable que ocurra cierto elemento, le tendríamos que asignar a éste una probabilidad cercana a 1. Por el contrario, si creemos que no hay probabilidades de que ocurra cierto elemento, le tendríamos que asignar a éste una probabilidad cercana a cero. En muchos experimentos, como lanzar una moneda o un dado, todos los elementos tienen la misma oportunidad de ocurrencia, por lo tanto, se les asignan probabilidades iguales. A los elementos fuera del espacio de sucesos, es decir, a los eventos simples que no tienen posibilidades de ocurrir, les asignamos una probabilidad de cero.

Para encontrar la probabilidad de un evento  $A$  sumamos todas las probabilidades que se asignan a los elementos en  $A$ . Esta suma se denomina probabilidad de  $A$  y se denota con  $P(A)$ .

Para medir la probabilidad de que ocurra un evento específico, necesitamos una herramienta matemática que nos permita cuantificar esta incertidumbre. Aquí es donde entra en juego la *función de probabilidad*. Esta función nos proporciona una manera sistemática de asignar un valor numérico a la probabilidad de cada posible resultado de un experimento aleatorio. En términos simples, la función de probabilidad nos ayuda a medir cuán probable es que ocurra un evento dado.

**Definición 1.3.1:**

Sea  $\Omega$  un espacio de sucesos y  $\mathcal{G}(\Omega)$  el conjunto potencia de  $\Omega$ . Definimos la función de probabilidad  $P : \mathcal{G}(\Omega) \rightarrow [0, 1]$ , tal que cumple los siguientes tres axiomas:

- i. Para cualquier evento  $A$ , se tiene que  $P(A) \geq 0$ .
- ii. La suma de las probabilidades de todos los sucesos posibles debe ser igual a 1, es decir,  $P(\Omega) = 1$ .

iii. Si  $A_1, A_2, A_3, \dots$  son sucesos disjuntos,  $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$ .

Observemos que la probabilidad del conjunto vacío  $\emptyset$  es igual a 0, es decir,  $P(\emptyset) = 0$ . Esto se debe a que el conjunto vacío no contiene ningún suceso posible, por lo tanto, la probabilidad de que ocurra un suceso en el conjunto vacío es nula. Si el lector aún no entiende lo anteriormente dicho, esto será mucho más claro con los ejemplos siguientes y cuando presentemos la definición 1.3.2.

**Ejemplo 1.3.1:**

Una moneda se lanza dos veces. ¿Cuál es la probabilidad de que salga al menos un sol (S)?

**Solución:** El espacio de sucesos para este experimento es

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

Si la moneda está balanceada, cada uno de estos resultados tendrá las mismas probabilidades de ocurrir. Por lo tanto, asignamos una probabilidad de  $\omega$  a cada uno de los elementos. Entonces,  $4\omega = 1$  o  $\omega = 1/4$ . Si  $A$  representa el evento de que salga al menos un sol (S), entonces

$$A = \{SS, SA, AS\} \quad \text{y} \quad P(A) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}.$$

**Ejemplo 1.3.2:**

Se carga un dado de forma que exista el doble de probabilidades de que salga un número par que uno impar. Si  $E$  es el evento de que ocurra un número menor que 4 en un solo lanzamiento del dado, calcule  $P(E)$ .

**Solución:** Es claro que el espacio de sucesos es  $\Omega = \{1, 2, 3, 4, 5, 6\}$ . Asignamos una probabilidad de  $\omega$  a cada número impar y una probabilidad de  $2\omega$  a cada número par. Como la suma de las probabilidades debe ser 1, tenemos

$$\omega + 2\omega + \omega + 2\omega + \omega + 2\omega = 9\omega = 1,$$

o bien,  $\omega = 1/9$ . Por lo tanto, asignamos probabilidades de  $1/9$  y  $2/9$  a cada número impar y par, respectivamente. Por consiguiente,

$$E = \{1, 2, 3\} \quad \text{y} \quad P(E) = \frac{1}{9} + \frac{2}{9} + \frac{1}{9} = \frac{4}{9}.$$

**Ejemplo 1.3.3:**

En el ejemplo anterior, sea  $A$  el evento de que resulte un número par y sea  $B$  el evento de que resulte un número divisible entre 3. Calcule  $P(A \cup B)$  y  $P(A \cap B)$ .

**Solución:** Es claro que los eventos son  $A = \{2, 4, 6\}$  y  $B = \{3, 6\}$ . Así que tenemos,

$$A \cup B = \{2, 3, 4, 6\} \quad \text{y} \quad A \cap B = \{6\}.$$

Al asignar una probabilidad de  $1/9$  a cada número impar y de  $2/9$  a cada número par, tenemos

$$P(A \cup B) = \frac{2}{9} + \frac{1}{9} + \frac{2}{9} + \frac{2}{9} = \frac{7}{9} \quad \text{y} \quad P(A \cap B) = \frac{2}{9}.$$

Si el espacio de sucesos para un experimento contiene  $N$  elementos, todos los cuales tienen las mismas probabilidades de ocurrir, asignamos una probabilidad igual a  $1/N$  a cada uno de los  $N$  elementos. La probabilidad de que cualquier evento  $A$  contenga  $n$  de estos  $N$  elementos es entonces el cociente del número de elementos en  $A$  y el número de elementos en  $\Omega$ .

**Definición 1.3.2:**

Si un experimento puede dar como resultado cualquiera de  $N$  diferentes resultados que tienen las mismas probabilidades de ocurrir, y si exactamente  $n$  de estos resultados corresponden al evento  $A$ , entonces la probabilidad del evento  $A$  es

$$P(A) = \frac{n}{N}.$$

**Ejemplo 1.3.4:**

A una clase de estadística para ingenieros asisten 25 estudiantes de ingeniería industrial, 10 de ingeniería mecánica, 10 de ingeniería eléctrica y 8 de ingeniería civil. Si el profesor elige al azar a un estudiante para que conteste una pregunta, ¿qué probabilidades hay de que el elegido sea

- a) estudiante de ingeniería industrial,
- b) estudiante de ingeniería civil o estudiante de ingeniería eléctrica?

**Solución:** Asignemos las especialidades de los estudiantes de ingeniería industrial, mecánica, eléctrica y civil por las letras  $I$ ,  $M$ ,  $E$  y  $C$ , respectivamente. El grupo está integrado por 53 estudiantes y todos tienen las mismas probabilidades de ser seleccionados.

- a) Como 25 de los 53 individuos estudian ingeniería industrial, la probabilidad del evento  $I$ , es decir, la de elegir al azar a alguien que estudia ingeniería industrial, es

$$P(I) = \frac{25}{53}.$$

- b) Como 18 de los 53 estudiantes son de las especialidades de ingeniería civil o eléctrica, se deduce que

$$P(C \cup E) = \frac{18}{53}.$$

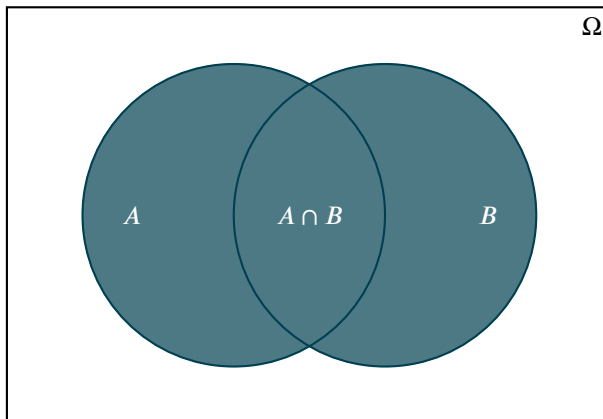
## 1.4 Reglas aditivas

A menudo resulta más sencillo calcular la probabilidad de algún evento a partir de las probabilidades conocidas de otros eventos. Esto puede ser cierto si el evento en cuestión se puede representar como la unión de otros dos eventos o como el complemento de algún evento. A continuación se presentan varias leyes importantes que con frecuencia simplifican el cálculo de las probabilidades. La primera, que se denomina *regla aditiva*, se aplica a uniones de eventos.

**Teorema 1.4.1:** Si  $A$  y  $B$  son dos eventos, entonces

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

**Demostración:** Consideremos el siguiente diagrama de Venn donde los círculos representan los eventos  $A$  y  $B$ :



**Figura 1.2** Regla aditiva de probabilidad

Notemos que  $P(A \cup B)$  es la suma de las probabilidades de los elementos en  $A \cup B$ . Así,  $P(A) + P(B)$  es la suma de todas las probabilidades en  $A$  más la suma de todas las probabilidades en  $B$ . Por lo tanto, sumamos dos veces las probabilidades en  $A \cap B$ . Como estas probabilidades se suman a  $P(A \cap B)$ , debemos restar esta probabilidad una vez para obtener la suma de las probabilidades en  $A \cup B$ .

**Corolario 1.4.1:** Si  $A$  y  $B$  son mutuamente excluyentes, entonces

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

**Demostración:** La demostración es inmediata, pues si  $A$  y  $B$  son mutuamente excluyentes,  $A \cap B = \emptyset$  y entonces  $P(A \cap B) = P(\emptyset) = 0$ .

En general, tenemos el siguiente corolario:

**Corolario 1.4.2:** Si  $A_1, A_2, \dots, A_n$  son mutuamente excluyentes, entonces

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n).$$

Un conjunto de eventos  $\{A_1, A_2, \dots, A_n\}$  de un espacio de sucesos  $\Omega$  se denomina partición de  $\Omega$  si  $A_1, A_2, \dots, A_n$  son mutuamente excluyentes y

$$A_1 \cup A_2 \cup \dots \cup A_n = \Omega.$$

Por lo tanto, tenemos

**Corolario 1.4.3:** Si  $A_1, A_2, \dots, A_n$  es una partición de un espacio de sucesos  $\Omega$ , entonces

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n) = P(\Omega) = 1.$$

Como se esperaba, el teorema 1.4.1 se extiende de forma análoga.

**Teorema 1.4.2:** Para tres eventos  $A, B$  y  $C$ ,

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

**Ejemplo 1.4.1:**

Al final del semestre John se va a graduar en la facultad de ingeniería industrial de una universidad. Después de tener entrevistas en dos empresas en donde quiere trabajar, determina que la probabilidad que tiene de lograr una oferta de empleo en la empresa  $A$  es 0.8, y que la probabilidad de obtenerla en la empresa  $B$  es 0.6. Si, por otro lado, considera que la probabilidad de recibir ofertas de ambas empresas es 0.5, ¿qué probabilidad tiene de obtener al menos una oferta de esas dos empresas?

**Solución:** Si usamos el teorema 1.4.1, entonces

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.8 + 0.6 - 0.5 = 0.9.$$

El teorema 1.4.1 y sus tres corolarios deberían ayudar al lector a comprender mejor la probabilidad y su interpretación. Los corolarios 1.4.1 y 1.4.2 sugieren el resultado muy intuitivo tratando con la probabilidad de que ocurra al menos uno de varios eventos, sin que puedan ocurrir dos de ellos simultáneamente. La probabilidad de que al menos ocurra uno es la suma de las probabilidades de ocurrencia de los eventos individuales. El tercer corolario simplemente establece que el valor mayor de una probabilidad (unidad) se asigna a todo el espacio de sucesos  $\Omega$ .

**Ejemplo 1.4.2:**

Las probabilidades de que un individuo que compra un automóvil nuevo elija uno de color verde, uno blanco, uno rojo o uno azul son 0.09, 0.15, 0.21 y 0.23, respectivamente, ¿cuál es la probabilidad de que un comprador dado adquiera un automóvil nuevo que tenga uno de esos colores?

**Solución:** Sean  $V$ ,  $B$ ,  $R$  y  $A$  los eventos de que un comprador seleccione, respectivamente, un automóvil verde, blanco, rojo o azul. Como estos cuatro eventos son mutuamente excluyentes, la probabilidad es

$$\begin{aligned} P(V \cup B \cup R \cup A) &= P(V) + P(B) + P(R) + P(A) \\ &= 0.09 + 0.15 + 0.21 + 0.23 \\ &= 0.68. \end{aligned}$$

A menudo es más difícil calcular la probabilidad de que ocurra un evento que calcular la probabilidad de que el evento no ocurra. Si éste es el caso para algún evento  $A$ , simplemente calculamos primero  $P(A^C)$  y, después, mediante el teorema 1.4.1, calculamos  $P(A)$  por sustracción.

**Teorema 1.4.3:**

Si  $A$  y  $A^C$  son eventos complementarios, entonces

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

**Demostración:** Puesto que los conjuntos  $A$  y  $A^C$  son disjuntos y  $A \cup A^C = \Omega$ , entonces

$$1 = P(\Omega) = P(A \cup A^C) = P(A) + P(A^C).$$

**Ejemplo 1.4.3:**

Si las probabilidades de que un mecánico automotriz dé servicio a 3, 4, 5, 6, 7, 8 o más vehículos en un día de trabajo dado son 0.12, 0.19, 0.28, 0.24, 0.10 y 0.07, respectivamente, ¿cuál es la probabilidad de que dé servicio al menos a 5 vehículos el siguiente día de trabajo?

**Solución:** Sea  $E$  el evento de que al menos 5 automóviles reciban servicio. Ahora bien,  $P(E) = 1 - P(E^C)$ , donde  $E^C$  es el evento de que menos de 5 automóviles reciban servicio. Como  $P(E^C) = 0.12 + 0.19 = 0.31$ , del teorema anterior se deduce que

$$P(E) = 1 - 0.31 = 0.69.$$

**Teorema 1.4.4:** Para cualquier evento  $A$ , se tiene que

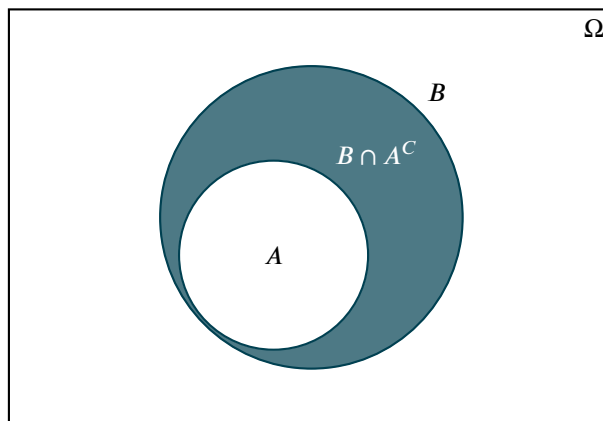
$$0 \leq P(A) \leq 1.$$

**Demostración:** Se sabe por el axioma 1 de la definición 1.3.1 que  $P(A) \geq 0$ . Si  $P(A) > 1$ , entonces por el teorema 1.4.3 resulta que  $P(A^C) < 0$ . Puesto que este resultado contradice el axioma 1 de la definición 1.3.1, que afirma que la probabilidad de todo evento debe ser no negativa, debe ser cierto que  $P(A) \leq 1$ .

**Teorema 1.4.5:** Sean  $A$  y  $B$  dos eventos. Si  $A \subset B$ , entonces

$$P(A) \leq P(B).$$

**Demostración:** Consideremos el siguiente diagrama de Venn donde los círculos representan los eventos  $A$  y  $B$ :



**Figura 1.3**  $B = A \cup (B \cap A^C)$

Como se ilustra en la figura anterior, el evento  $B$  puede ser tratado como la unión de los dos eventos disjuntos  $A$  y  $B \cap A^C$ . Por tanto,

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

Puesto que  $P(B \cap A^C) \geq 0$ , entonces  $P(B) \geq P(A)$ .

## 1.5 Probabilidad condicional e independencia

Un concepto muy importante en la teoría de probabilidad es la probabilidad condicional. En algunas aplicaciones el profesional se interesa por la estructura de probabilidad bajo ciertas restricciones. Por ejemplo, en epidemiología, en lugar de estudiar las probabilidades de que una persona de la población general tenga diabetes, podría ser más interesante conocer esta probabilidad en un grupo distinto, como el de las mujeres asiáticas cuya edad está en el rango de 35 a 50 años, o como el de los hombres hispanos cuya edad está entre los 40 y los 60 años. A este tipo de probabilidad se le conoce como probabilidad condicional.

### Probabilidad condicional

La probabilidad de que ocurra un evento  $A$  cuando se sabe que ya ocurrió algún evento  $B$  se llama probabilidad condicional y se denota con  $P(A | B)$ . El símbolo  $P(A | B)$  por lo general se lee como “la probabilidad de que ocurra  $A$ , dado que ocurrió  $B$ ”, o simplemente, “la probabilidad de  $A$ , dado  $B$ ”.

Considere el evento  $A$  de obtener un cuadrado perfecto cuando se lanza un dado. El dado se construye de modo que los números pares tengan el doble de probabilidad de ocurrencia que los números nones. Con base en el espacio de sucesos  $\Omega = \{1, 2, 3, 4, 5, 6\}$ , en el que a los números impares y a los pares se les asignaron probabilidades de  $1/9$  y  $2/9$ , respectivamente, la probabilidad de que ocurra  $A$  es de  $1/3$ . Suponga ahora que se sabe que el lanzamiento del dado tiene como resultado un número mayor que 3. Tenemos ahora un espacio de sucesos reducido,  $B = \{4, 5, 6\}$ , que es un subconjunto de  $\Omega$ . Para encontrar la probabilidad de que ocurra  $A$ , en relación con el espacio de sucesos  $B$ , debemos comenzar por asignar nuevas probabilidades a los elementos de  $B$ , que sean proporcionales a sus probabilidades originales de modo que su suma sea 1. Al asignar una probabilidad de  $\omega$  al número non en  $B$  y una probabilidad de  $2\omega$  a los dos números pares, tenemos  $5\omega = 1$  o  $\omega = 1/5$ . En relación con el espacio  $B$ , encontramos que  $A$  contiene sólo el elemento 4. Si denotamos este evento con el símbolo  $A | B$ , escribimos  $A | B = \{4\}$  y, en consecuencia,

$$P(A | B) = \frac{2}{5}.$$

Este ejemplo ilustra que los eventos pueden tener probabilidades diferentes cuando se consideran en relación con diferentes espacios de sucesos.

También podemos escribir

$$P(A | B) = \frac{2}{5} = \frac{\frac{2}{9}}{\frac{9}{9}} = \frac{P(A \cap B)}{P(B)},$$

donde  $P(A \cap B)$  y  $P(B)$  se calculan a partir del espacio de sucesos original  $\Omega$ . En otras palabras, una probabilidad condicional relativa a un subespacio  $B$  de  $\Omega$  se puede calcular en forma directa de las probabilidades que se asignan a los elementos del espacio de sucesos original  $\Omega$ .

#### Definición 1.5.1:

La probabilidad condicional de  $A$ , dado  $B$ , que se denota con  $P(A | B)$ , se define como

$$P(A | B) = \frac{P(B \cap A)}{P(B)} = \frac{P(A \cap B)}{P(B)}, \quad \text{siempre que } P(B) > 0.$$

Un ejemplo más: Suponga que tenemos un espacio de sucesos  $\Omega$  constituido por la población de adultos de una pequeña ciudad que cumplen con los requisitos para obtener un título universitario. Debemos clasificarlos de acuerdo con su género y situación laboral. Los datos se presentan en la siguiente tabla.

**Tabla 1.1** La clasificación de los adultos de una pequeña ciudad

	Empleado	Desempleado	Total
Hombre	460	40	500
Mujer	140	260	400
Total	600	300	900

Se seleccionará al azar a uno de estos individuos para que realice un viaje a través del país con el fin de promover las ventajas de establecer industrias nuevas en la ciudad. Nos interesaremos en los eventos siguientes:

$M$  : se elige a un hombre,

$E$  : el elegido tiene un empleo.

Al utilizar el espacio de sucesos reducido  $E$ , encontramos que

$$P(M | E) = \frac{460}{600} = \frac{23}{30}.$$



Sea  $|A|$  el número de elementos en cualquier conjunto  $A$ . Podemos utilizar esta notación, puesto que cada uno de los adultos tiene las mismas probabilidades de ser elegido, para escribir

$$P(M | E) = \frac{|E \cap M|}{|E|} = \frac{\frac{|E \cap M|}{|\Omega|}}{\frac{|E|}{|\Omega|}} = \frac{P(E \cap M)}{P(E)},$$

en donde  $P(E \cap M)$  y  $P(E)$  se calculan a partir del espacio de sucesos original  $\Omega$ . Para verificar este resultado observe que

$$P(E) = \frac{600}{900} = \frac{2}{3} \quad \text{y} \quad P(E \cap M) = \frac{460}{900} = \frac{23}{45}.$$

Por lo tanto,

$$P(M | E) = \frac{\frac{23}{45}}{\frac{2}{3}} = \frac{23}{30},$$

como antes.

### Ejemplo 1.5.1:

La probabilidad de que un vuelo programado normalmente salga a tiempo es  $P(D) = 0.83$ , la probabilidad de que llegue a tiempo es  $P(A) = 0.82$  y la probabilidad de que salga y llegue a tiempo es  $P(D \cap A) = 0.78$ . Calcule la probabilidad de que un avión

- llegue a tiempo, dado que salió a tiempo;
- salió a tiempo, dado que llegó a tiempo.

**Solución:** Al utilizar la definición 1.5.1 tenemos:

- La probabilidad de que un avión llegue a tiempo, dado que salió a tiempo es

$$P(A | D) = \frac{P(D \cap A)}{P(D)} = \frac{0.78}{0.83} = 0.94.$$

- La probabilidad de que un avión haya salido a tiempo, dado que llegó a tiempo es

$$P(D | A) = \frac{P(D \cap A)}{P(A)} = \frac{0.78}{0.82} = 0.95.$$

La noción de probabilidad condicional brinda la capacidad de reevaluar la idea de probabilidad de un evento a la luz de la información adicional; es decir, cuando se sabe que ocurrió otro evento. La probabilidad  $P(A | B)$  es una actualización de  $P(A)$  basada en el conocimiento de que ocurrió el evento  $B$ . En el ejemplo anterior es importante conocer la probabilidad de que el vuelo llegue a tiempo. Tenemos la información de que el vuelo no salió a tiempo. Con esta información adicional, la probabilidad más pertinente es  $P(A | D^C)$ , esto es, la probabilidad de que llegue a tiempo, dado que no salió a tiempo. A menudo las conclusiones que se obtienen a partir de observar la probabilidad condicional más importante cambian drásticamente la situación. En este ejemplo, el cálculo de  $P(A | D^C)$  es

$$P(A | D^C) = \frac{P(A \cap D^C)}{P(D^C)} = \frac{P(A) - P(A \cap D)}{P(D^C)} = \frac{0.82 - 0.78}{0.17} = 0.24.$$

Como resultado, la probabilidad de una llegada a tiempo disminuye significativamente ante la presencia de la información adicional.

## Eventos independientes

En el experimento del lanzamiento de un dado de la página 16 señalamos que  $P(A | B) = 2/5$ , mientras que  $P(A) = 1/3$ . Es decir,  $P(A | B) \neq P(A)$ , lo cual indica que  $A$  depende de  $B$ . Consideremos ahora un experimento en el que se sacan 2 cartas, una después de la otra, de una baraja ordinaria, con reemplazo. Los eventos se definen como

$A$  : la primera carta es un as,

$B$  : la segunda carta es una espada.

Como la primera carta se reemplaza, nuestro espacio de sucesos para la primera y segunda cartas consta de 52 cartas, que contienen 4 ases y 13 espadas. Entonces,

$$P(B | A) = \frac{13}{52} = \frac{1}{4} \quad \text{y} \quad P(B) = \frac{13}{52} = \frac{1}{4}.$$

Es decir,  $P(B | A) = P(B)$ . Cuando esto es cierto, se dice que los eventos  $A$  y  $B$  son *independientes*.

Aunque la probabilidad condicional permite alterar la probabilidad de un evento a la luz de material adicional, también nos permite entender mejor el muy importante concepto de independencia o, en el contexto actual, de eventos independientes. En el ejemplo 1.5.1 del aeropuerto,  $P(A | D)$  difiere de  $P(A)$ . Esto sugiere que la ocurrencia de  $D$  influye en  $A$  y esto es lo que, de hecho, se espera en este caso. Sin embargo, considere la situación en donde tenemos los eventos  $A$  y  $B$ , y

$$P(A | B) = P(A).$$

En otras palabras, la ocurrencia de  $B$  no influye en las probabilidades de ocurrencia de  $A$ . Aquí la ocurrencia de  $A$  es independiente de la ocurrencia de  $B$ . No podemos dejar de resaltar la importancia del concepto de independencia, ya que desempeña un papel vital en el material de casi todos los capítulos de este libro y en todas las áreas de la estadística aplicada.

### Definición 1.5.2:

Dos eventos  $A$  y  $B$  son *independientes* si y sólo si

$$P(B | A) = P(B) \quad \text{o} \quad P(A | B) = P(A),$$

si se asume la existencia de probabilidad condicional. De cualquier otra forma,  $A$  y  $B$  son dependientes.

Al multiplicar la fórmula de la definición 1.5.1 por  $P(B)$ , obtenemos la siguiente regla multiplicativa importante (o regla de producto), que nos permite calcular la probabilidad de que ocurran dos eventos

---

**Teorema 1.5.1:** Si en un experimento pueden ocurrir los eventos  $A$  y  $B$ , entonces

$$P(A \cap B) = P(B)P(A | B), \quad \text{siempre que } P(B) > 0.$$


---

Como los eventos  $A \cap B$  y  $B \cap A$  son equivalentes, del teorema anterior se deduce que también podemos escribir

$$P(A \cap B) = P(B \cap A) = P(A)P(B | A).$$

---

**Teorema 1.5.2:** Dos eventos  $A$  y  $B$  son independientes si y sólo si

$$P(A \cap B) = P(A)P(B).$$

Por lo tanto, para obtener la probabilidad de que ocurran dos eventos independientes simplemente calculamos el producto de sus probabilidades individuales.

---

**Ejemplo 1.5.2:**

Una pequeña ciudad dispone de un carro de bomberos y una ambulancia para emergencias. La probabilidad de que el carro de bomberos esté disponible cuando se necesite es 0.98 y la probabilidad de que la ambulancia esté disponible cuando se le requiera es 0.92. En el evento de un herido en un incendio, calcule la probabilidad de que tanto la ambulancia como el carro de bomberos estén disponibles, suponiendo que operan de forma independiente.

**Solución:** Sean  $A$  y  $B$  los respectivos eventos de que estén disponibles el carro de bomberos y la ambulancia. Entonces,

$$P(A \cap B) = P(A)P(B) = (0.98)(0.92) = 0.9016.$$

La regla multiplicativa se puede extender a situaciones con más de dos eventos.

**Teorema 1.5.3:**

Si, en un experimento, pueden ocurrir los eventos  $A_1, A_2, \dots, A_n$ , entonces

$$\begin{aligned} P(A_1 \cap A_2 \cap \dots \cap A_n) \\ = P(A_1)P(A_2 | A_1)P(A_3 | A_1 \cap A_2) \dots P(A_n | A_1 \cap A_2 \cap \dots \cap A_{n-1}). \end{aligned}$$

Si los eventos  $A_1, A_2, \dots, A_n$  son independientes, entonces

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)P(A_2) \dots P(A_n).$$

## 1.6 Regla de Bayes

La estadística bayesiana es un conjunto de herramientas que se utiliza en un tipo especial de inferencia estadística que se aplica en el análisis de datos experimentales en muchas situaciones prácticas de ciencia e ingeniería. La regla de Bayes es una de las normas más importantes de la teoría de probabilidad, ya que es el fundamento de la inferencia bayesiana.

### Probabilidad total

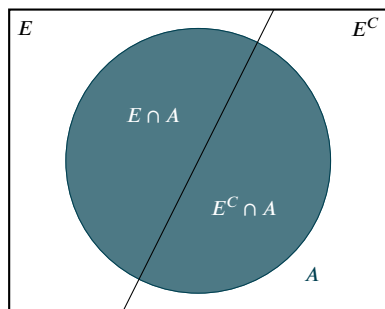


Figura 1.4

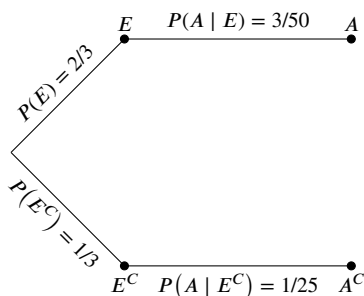


Figura 1.5

Regresemos al ejemplo de la sección anterior, en el que se selecciona un individuo al azar de entre los adultos de una pequeña ciudad para que viaje por el país promoviendo las ventajas de establecer industrias nuevas en la ciudad. Suponga que ahora se nos da la información adicional de que 36 de los empleados y 12 de los desempleados son miembros del Club Rotario. Deseamos encontrar la probabilidad del evento  $A$  de que el individuo seleccionado sea miembro del Club Rotario. Podemos remitirnos a la figura 1.4 y escribir  $A$  como la unión de los dos eventos mutuamente excluyentes  $E \cap A$  y  $E^C \cap A$ . Por lo tanto,  $A = (E \cap A) \cup (E^C \cap A)$ , y mediante el corolario 1.4.1 del teorema 1.4.1 y luego mediante el teorema 1.5.1, podemos escribir

$$\begin{aligned} P(A) &= P((E \cap A) \cup (E^C \cap A)) \\ &= P(E \cap A) + P(E^C \cap A) \\ &= P(E)P(A | E) + P(E^C)P(A | E^C). \end{aligned}$$

Los datos de la sección anterior, junto con los datos adicionales antes dados para el conjunto  $A$ , nos permiten calcular

$$P(E) = \frac{600}{900} = \frac{2}{3} \quad \text{y} \quad P(A | E) = \frac{36}{600} = \frac{3}{50},$$

y

$$P(E^C) = \frac{1}{3} \quad \text{y} \quad P(A | E^C) = \frac{12}{300} = \frac{1}{25}.$$

Si mostramos estas probabilidades mediante el diagrama de árbol de la figura 1.5, donde la primera rama da la probabilidad  $P(E)P(A | E)$  y la segunda rama da la probabilidad  $P(E^C)P(A | E^C)$ , deducimos que

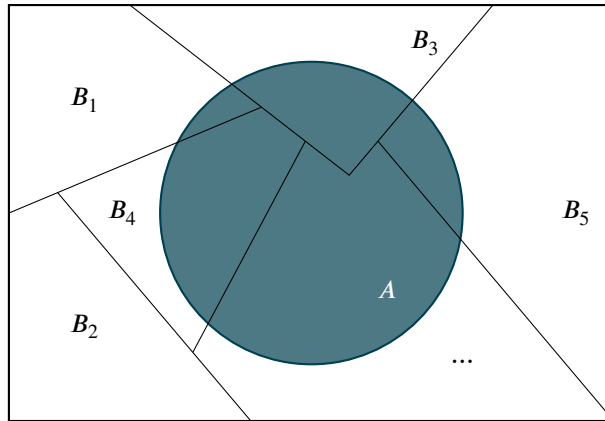
$$P(A) = \left(\frac{2}{3}\right)\left(\frac{3}{50}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{25}\right) = \frac{4}{75}.$$

Una generalización del ejemplo anterior para el caso en donde el espacio de sucesos se parte en  $k$  subconjuntos se cubre mediante el siguiente teorema, que algunas veces se denomina *teorema de probabilidad total* o *regla de eliminación*.

**Teorema 1.6.1:** Si los eventos  $B_1, B_2, \dots, B_k$  constituyen una partición del espacio de sucesos  $\Omega$ , tal que  $P(B_i) \neq 0$  para  $i = 1, 2, \dots, k$ , entonces, para cualquier evento  $A$  de  $\Omega$ ,

$$P(A) = \sum_{i=1}^k P(B_i \cap A) = \sum_{i=1}^k P(B_i)P(A | B_i).$$

**Solución:** Considere siguiente diagrama de Venn.



**Figura 1.6** Partición del espacio de sucesos  $\Omega$

Se observa que el evento  $A$  es la unión de los eventos mutuamente excluyentes

$$B_1 \cap A, B_2 \cap A, \dots, B_k \cap A;$$

es decir,

$$A = (B_1 \cap A) \cup (B_2 \cap A) \cup \dots \cup (B_k \cap A).$$

Por medio del corolario 1.4.2 del teorema 1.4.1 y el teorema 1.5.1 obtenemos

$$\begin{aligned} P(A) &= P((B_1 \cap A) \cup (B_2 \cap A) \cup \dots \cup (B_k \cap A)) \\ &= P(B_1 \cap A) + P(B_2 \cap A) + \dots + P(B_k \cap A) \\ &= \sum_{i=1}^k P(B_i \cap A) \\ &= \sum_{i=1}^k P(B_i)P(A | B_i). \end{aligned}$$

**Teorema 1.6.2:** Si los eventos  $B_1, B_2, \dots, B_k$  constituyen una partición del espacio de sucesos  $\Omega$ , donde  $P(B_i) \neq 0$  para  $i = 1, 2, \dots, k$ , entonces, para cualquier evento  $A$  en  $\Omega$ , tal que  $P(A) \neq 0$ ,

$$P(B_r | A) = \frac{P(B_r \cap A)}{\sum_{i=1}^k P(B_i \cap A)} = \frac{P(B_r)P(A | B_r)}{\sum_{i=1}^k P(B_i)P(A | B_i)}, \quad \text{para } r = 1, 2, \dots, k.$$

**Solución:** Es inmediato usando la definición de probabilidad condicional y luego usando teorema anterior.

**Ejemplo 1.6.1:**

Una empresa de manufactura emplea tres planos analíticos para el diseño y desarrollo de un producto específico. Por razones de costos los tres se utilizan en momentos diferentes. De hecho, los planos 1, 2 y 3 se utilizan para 30 %, 20 % y 50 % de los productos, respectivamente. La tasa de defectos difiere en los tres procedimientos de la siguiente manera,

$$P(D | P_1) = 0.01, \quad P(D | P_2) = 0.03, \quad P(D | P_3) = 0.02$$

en donde  $P(D | P_j)$  es la probabilidad de que un producto esté defectuoso, dado el plano  $j$ . Si se observa un producto al azar y se descubre que está defectuoso, ¿cuál de los planos tiene más probabilidades de haberse utilizado y, por lo tanto, de ser el responsable?

**Solución:** A partir del planteamiento del problema

$$P(P_1) = 0.30, \quad P(P_2) = 0.20, \quad P(P_3) = 0.50,$$

debemos calcular  $P(P_j | D)$  para  $j = 1, 2, 3$ . La regla de Bayes muestra

$$\begin{aligned} P(P_1 | D) &= \frac{P(P_1)P(D | P_1)}{P(P_1)P(D | P_1) + P(P_2)P(D | P_2) + P(P_3)P(D | P_3)} \\ &= \frac{(0.30)(0.01)}{(0.30)(0.01) + (0.20)(0.03) + (0.50)(0.02)} \\ &= 0.158. \end{aligned}$$

De manera análoga,

$$P(P_2 | D) = \frac{(0.03)(0.20)}{0.019} = 0.316 \quad \text{y} \quad P(P_3 | D) = \frac{(0.02)(0.50)}{0.019} = 0.526.$$

La probabilidad condicional de un defecto, dado el plano 3, es la mayor de las tres; por consiguiente, un defecto en un producto elegido al azar tiene más probabilidad de ser el resultado de haber usado el plano 3.

## 1.7 Variables aleatorias y distribución de probabilidad

La estadística realiza inferencias acerca de las poblaciones y sus características. Se llevan a cabo experimentos cuyos resultados se encuentran sujetos al azar. La prueba de un número de componentes electrónicos es un ejemplo de experimento estadístico, un concepto que se utiliza para describir cualquier proceso mediante el cual se generan varias observaciones al azar. A menudo es importante asignar una descripción numérica al resultado. Por ejemplo, cuando se prueban tres componentes electrónicos, el espacio de sucesos que ofrece una descripción detallada de cada posible resultado se escribe como

$$\Omega = \{NNN, NND, NDN, DNN, NDD, DND, DDN, DDD\}$$

donde  $N$  denota “no defectuoso”, y  $D$ , “defectuoso”. Es evidente que nos interesa el número de componentes defectuosos que se presenten. De esta forma, a cada punto en el espacio de sucesos se le asignará un valor numérico de 0, 1, 2 o 3. Estos valores son, por supuesto, cantidades aleatorias determinadas por el resultado del experimento. Se pueden ver como valores que toma la variable aleatoria  $X$ , es decir, el número de artículos defectuosos cuando se prueban tres componentes electrónicos.

**Definición 1.7.1:**

Una *variable aleatoria* es una función que asocia un número real con cada elemento del espacio de sucesos.

Utilizaremos una letra mayúscula, digamos  $X$ , para denotar una variable aleatoria, y su correspondiente letra minúscula,  $x$  en este caso, para uno de sus valores. En el ejemplo de la prueba de componentes electrónicos observamos que la variable aleatoria  $X$  toma el valor 2 para todos los elementos en el subconjunto

$$E = \{DDN, DND, NDD\}$$

del espacio de sucesos  $\Omega$ . Esto es, cada valor posible de  $X$  representa un evento que es un subconjunto del espacio de sucesos para el experimento dado.

**Ejemplo 1.7.1:**

De una urna que contiene 4 bolas rojas y 3 negras se sacan 2 bolas de manera sucesiva, sin reemplazo. Los posibles resultados y los valores  $y$  de la variable aleatoria  $Y$ , donde  $Y$  es el número de bolas rojas, son

Espacio de sucesos	$y$
$RR$	2
$RN$	1
$NR$	1
$NN$	0

**Tabla 1.2**

**Ejemplo 1.7.2:**

El empleado de un almacén regresa tres cascos de seguridad al azar a tres obreros de un taller siderúrgico que ya los habían probado. Si Smith, Jones y Brown, en ese orden, reciben uno de los tres cascos, liste los elementos para los posibles órdenes en que el empleado del almacén regresa los cascos, después calcule el valor  $m$  de la variable aleatoria  $M$  que representa el número de emparejamientos correctos.

**Solución:** Si  $S$ ,  $J$  y  $B$  representan, respectivamente, los cascos que recibieron Smith, Jones y Brown, entonces los posibles arreglos en los cuales se pueden regresar los cascos y el número de emparejamientos correctos son

Espacio de sucesos	$m$
$SJB$	3
$SBJ$	1
$BJS$	1
$JSB$	1
$JBS$	0
$BSJ$	0

**Tabla 1.3**

En cada uno de los dos ejemplos anteriores, el espacio de sucesos contiene un número finito de elementos. Por el contrario, cuando lanzamos un dado hasta que salga un 5, obtenemos un espacio de sucesos con una secuencia de elementos interminable,

$$\Omega = \{F, NF, NNF, NNNF, \dots\},$$

donde  $F$  y  $N$  representan, respectivamente, la ocurrencia y la no ocurrencia de un 5. Sin embargo, incluso en este experimento el número de elementos se puede igualar a la cantidad total de números enteros, de manera que hay un primer elemento, un segundo, un tercero y así sucesivamente, por lo que se pueden contar.

Hay casos en que la variable aleatoria es categórica por naturaleza en los cuales se utilizan las llamadas variables *ficticias*. Un buen ejemplo de ello es el caso en que la variable aleatoria es binaria por naturaleza, como se indica a continuación.

**Ejemplo 1.7.3:**

Considere la condición en que salen componentes de la línea de ensamble y se les clasifica como defectuosos o no defectuosos. Defina la variable aleatoria  $X$  mediante

$$X = \begin{cases} 1, & \text{si el componente está defectuoso,} \\ 0, & \text{si el componente no está defectuoso.} \end{cases}$$

Evidentemente la asignación de 1 o 0 es arbitraria, aunque bastante conveniente. La variable aleatoria en la que se eligen 0 y 1 para describir los dos posibles valores se denomina *variable aleatoria de Bernoulli*.

**Ejemplo 1.7.4:**

Los estadísticos utilizan planes de muestreo para aceptar o rechazar lotes de materiales. Suponga que uno de los planes de muestreo implica obtener una muestra independiente de 10 artículos de un lote de 100, en el que 12 están defectuosos. Si  $X$  representa a la variable aleatoria, definida como el número de artículos que están defectuosos en la muestra de 10, la variable aleatoria toma los valores 0, 1, 2, ..., 9, 10.

**Ejemplo 1.7.5:**

Existe interés por la proporción de personas que responden a cierta encuesta enviada por correo. Sea  $X$  tal proporción.  $X$  es una variable aleatoria que toma todos los valores de  $x$  para los cuales  $0 \leq x \leq 1$ .

**Ejemplo 1.7.6:**

Sea  $X$  la variable aleatoria definida como el tiempo que pasa, en horas, para que un radar detecte entre conductores sucesivos a los que exceden los límites de velocidad. La variable aleatoria  $X$  toma todos los valores de  $x$  para los que  $x \geq 0$ .

**Definición 1.7.2:**

Si un espacio de sucesos contiene un número finito de posibilidades, o una serie interminable con tantos elementos como números enteros existen, se llama *espacio de sucesos discreto*.

Los resultados de algunos experimentos estadísticos no pueden ser ni finitos ni contables. Éste es el caso, por ejemplo, en una investigación que se realiza para medir las distancias que recorre un automóvil de cierta marca, en una ruta de prueba preestablecida, con cinco litros de gasolina. Si se asume que la distancia es una variable que se mide con algún grado de precisión, entonces salta a la vista que tenemos un número infinito de distancias posibles en el espacio de suceso, que no se pueden igualar a la cantidad total de números enteros. Lo mismo ocurre en el caso de un experimento en que se registra el tiempo requerido para que ocurra una reacción química, en donde una vez más los posibles intervalos de tiempo que forman el espacio de sucesos serían un número infinito e incontable. Vemos ahora que no todos los espacios de sucesos necesitan ser discretos.

**Definición 1.7.3:**

Si un espacio de sucesos contiene un número infinito de posibilidades, igual al número de puntos en un segmento de recta, se le denomina *espacio de sucesos continuo*.

Una variable aleatoria se llama *variable aleatoria discreta* si se puede contar su conjunto de resultados posibles. En los ejemplos 1.7.1 a 1.7.3 las variables aleatorias son discretas. Sin embargo, una variable aleatoria cuyo conjunto de valores posibles es un intervalo completo de números no es discreta. Cuando una variable aleatoria puede tomar valor en una escala continua, se le denomina *variable aleatoria continua*. A menudo los posibles valores de una variable aleatoria continua son precisamente los mismos valores incluidos en el espacio de sucesos continuo. Es evidente que las variables aleatorias descritas en los ejemplos 1.7.5 y 1.7.6 son variables aleatorias continuas. En la mayoría de los problemas prácticos las variables aleatorias continuas representan datos medidos, como serían todos los posibles pesos, alturas, temperaturas, distancias o periodos de vida; en tanto que las variables aleatorias discretas representan datos por conteo, como el número de artículos defectuosos en una muestra de  $k$  artículos o el número de accidentes

de carretera por año en una entidad específica. Observe que tanto  $Y$  como  $M$ , las variables aleatorias de los ejemplos 1.7.1 y 1.7.2, representan datos por conteo:  $Y$  el número de bolas rojas y  $M$  el número de emparejamientos correctos de cascos.

## Distribuciones discretas de probabilidad

Una variable aleatoria discreta toma cada uno de sus valores con cierta probabilidad. Al lanzar una moneda tres veces, la variable  $X$ , que representa el número de caras, toma el valor 2 con  $3/8$  de probabilidad, pues 3 de los 8 elementos igualmente probables tienen como resultado dos caras y una cruz. Si se suponen pesos iguales para los eventos simples del ejemplo 1.7.2, la probabilidad de que ningún obrero reciba el casco correcto, es decir, la probabilidad de que  $M$  tome el valor cero, es  $1/3$ . Los valores posibles  $m$  de  $M$  y sus probabilidades son

$m$	0	1	3
$P(M = m)$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

Observe que los valores de  $m$  agotan todos los casos posibles, por lo tanto, las probabilidades suman 1.

Con frecuencia es conveniente representar todas las probabilidades de una variable aleatoria  $X$  usando una fórmula, la cual necesariamente sería una función de los valores numéricos  $x$  que denotaremos con  $f(x)$ ,  $g(x)$ ,  $r(x)$  y así sucesivamente. Por lo tanto, escribimos  $f(x) = P(X = x)$ ; es decir,  $f(3) = P(X = 3)$ . Al conjunto de pares ordenados  $(x, f(x))$  se le llama función de probabilidad, función de masa de probabilidad o distribución de probabilidad de la variable aleatoria discreta  $X$ .

### Definición 1.7.4:

El conjunto de pares ordenados  $(x, f(x))$  es una *función de probabilidad*, una *función de masa de probabilidad* o una *distribución de probabilidad* de la variable aleatoria discreta  $X$  si, para cada resultado  $x$  posible,

- i)  $f(x) \geq 0$ ,
- ii)  $\sum_x f(x) = 1$ ,
- iii)  $P(X = x) = f(x)$ .

### Ejemplo 1.7.7:

Un embarque de 20 computadoras portátiles similares para una tienda minorista contiene 3 que están defectuosas. Si una escuela compra al azar 2 de estas computadoras, calcule la distribución de probabilidad para el número de computadoras defectuosas.

**Solución:** Sea  $X$  una variable aleatoria cuyos valores  $x$  son los números posibles de computadoras defectuosas compradas por la escuela. Entonces  $x$  sólo puede asumir los números 0, 1 y 2. Así,

$$f(0) = P(X = 0) = \frac{\binom{3}{0}\binom{17}{2}}{\binom{20}{2}} = \frac{68}{95},$$

$$f(1) = P(X = 1) = \frac{\binom{3}{1}\binom{17}{1}}{\binom{20}{2}} = \frac{51}{190},$$

$$f(2) = P(X = 2) = \frac{\binom{3}{2}\binom{17}{0}}{\binom{20}{2}} = \frac{3}{190}.$$

Por consiguiente, la distribución de probabilidad de  $X$  es

$x$	0	1	2
$f(x)$	$\frac{68}{95}$	$\frac{51}{190}$	$\frac{3}{190}$

$$P(X = k) = \frac{\binom{3}{k}\binom{17}{2-k}}{\binom{20}{2}}, \quad k = 0, 1, 2.$$



Existen muchos problemas en los que desearíamos calcular la probabilidad de que el valor observado de una variable aleatoria  $X$  sea menor o igual que algún número real  $x$ . Al escribir  $F(x) = P(X \leq x)$  para cualquier número real  $x$ , definimos  $F(x)$  como la función de la distribución acumulativa de la variable aleatoria  $X$ .

**Definición 1.7.5:**

La función de la distribución acumulativa  $F(x)$  de una variable aleatoria discreta  $X$  con distribución de probabilidad  $f(x)$  es

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t), \quad \text{para } -\infty < x < \infty.$$

Para la variable aleatoria  $M$ , el número de emparejamientos correctos en el ejemplo 1.7.2, tenemos

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t) = f(0) + f(1) = \frac{1}{3} + \frac{1}{2} = \frac{5}{6}.$$

La función de la distribución acumulativa de  $M$  es

$$F(m) = \begin{cases} 0, & \text{para } m < 0, \\ \frac{1}{3}, & \text{para } 0 \leq m < 1, \\ \frac{5}{6}, & \text{para } 1 \leq m < 3, \\ 1, & \text{para } m \geq 1. \end{cases}$$

**i Nota**

Es necesario observar en particular el hecho de que la función de la distribución acumulativa es una función no decreciente monótona, la cual no sólo se define para los valores que toma la variable aleatoria dada sino para todos los números reales.

**Ejemplo 1.7.8:**

Calcule la función de la distribución acumulativa de la variable aleatoria  $X$  definida como se muestra a continuación:

$$f(x) = \frac{1}{16} \binom{4}{x}, \quad \text{para } x = 0, 1, 2, 3, 4.$$

Además, utilice  $F(x)$  para verificar que  $f(2) = 3/8$ .

**Solución:** El cálculo directo de la distribución de probabilidad da

$$f(0) = \frac{1}{16}, \quad f(1) = \frac{1}{4}, \quad f(2) = \frac{3}{8}, \quad f(3) = \frac{1}{4}, \quad f(4) = \frac{1}{16}.$$

Por lo tanto,

$$F(0) = f(0) = \frac{1}{16},$$

$$F(1) = f(0) + f(1) = \frac{5}{16},$$

$$F(2) = f(0) + f(1) + f(2) = \frac{11}{16},$$

$$F(3) = f(0) + f(1) + f(2) + f(3) = \frac{15}{16},$$

$$F(4) = f(0) + f(1) + f(2) + f(3) + f(4) = 1.$$

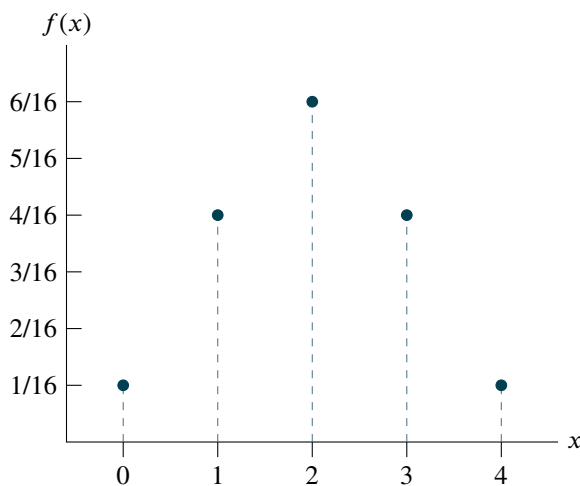
y por consiguiente,

$$F(x) = \begin{cases} 0, & \text{para } x < 0, \\ \frac{1}{16}, & \text{para } 0 \leq x < 1, \\ \frac{5}{16}, & \text{para } 1 \leq x < 2, \\ \frac{11}{16}, & \text{para } 2 \leq x < 3, \\ \frac{15}{16}, & \text{para } 3 \leq x < 4, \\ 1, & \text{para } x \geq 4. \end{cases} \quad \text{y} \quad \begin{aligned} f(2) &= F(2) - F(1) \\ &= \frac{11}{16} - \frac{5}{16} \\ &= \frac{3}{8}. \end{aligned}$$

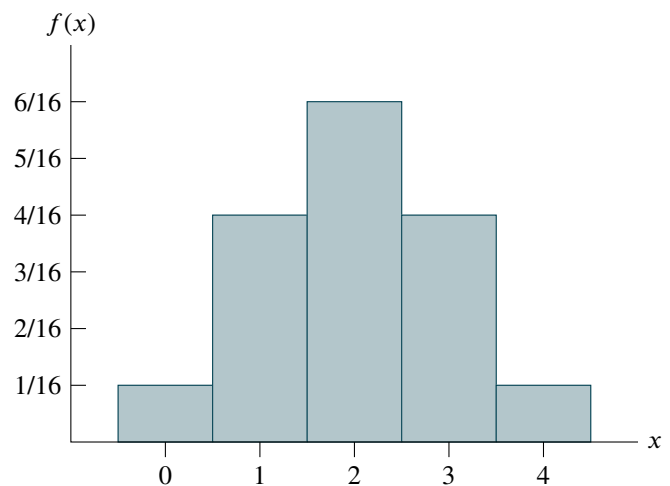
Muchas veces es útil ver una distribución de probabilidad en forma gráfica. Se pueden graficar los puntos  $(x, f(x))$  del ejemplo anterior para obtener la figura 1.7. Si unimos los puntos al eje  $x$ , ya sea con una línea punteada o con una línea sólida, obtenemos una gráfica de función de masa de probabilidad. La figura 1.7 permite ver fácilmente qué valores de  $X$  tienen más probabilidad de ocurrencia y, en este caso, también indica una situación perfectamente simétrica.

Sin embargo, en vez de graficar los puntos  $(x, f(x))$ , lo que hacemos más a menudo es construir rectángulos como en la figura 1.8. Aquí los rectángulos se construyen de manera que sus bases, con la misma anchura, se centren en cada valor  $x$ , y que sus alturas igualen a las probabilidades correspondientes dadas por  $f(x)$ . Las bases se construyen de forma tal que no dejen espacios entre los rectángulos. La figura 1.8 se denomina *histograma de probabilidad*.

Como cada base en la figura 1.8 tiene el ancho de una unidad,  $P(X = x)$  es igual al área del rectángulo centrado en  $x$ . Incluso si las bases no tuvieran el ancho de una unidad, podríamos ajustar las alturas de los rectángulos para que tengan áreas que igualen las probabilidades de  $X$  de tomar cualquiera de sus valores  $x$ . Este concepto de utilizar áreas para representar probabilidades es necesario para nuestro estudio de la distribución de probabilidad de una variable aleatoria continua.



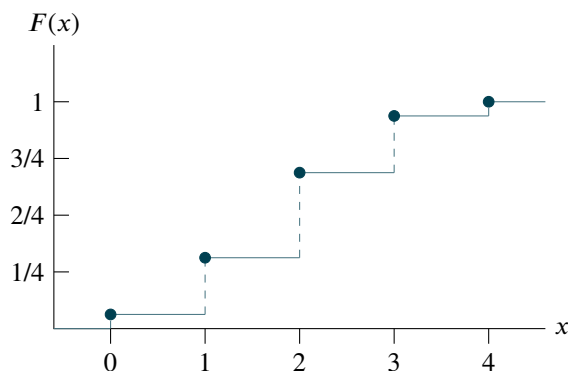
**Figura 1.7** Gráfica de función de masa de probabilidad



**Figura 1.8** Histograma de probabilidad

La gráfica de la función de la distribución acumulativa del último ejemplo, que aparece como una función escalonada en la figura 1.9, se obtiene graficando los puntos  $(x, F(x))$ .

**Figura 1.9** Gráfica de la función de distribución acumulativa discreta



Ciertas distribuciones de probabilidad se aplican a más de una situación física. La distribución de probabilidad del último ejemplo también se aplica a la variable aleatoria  $Y$ , donde  $Y$  es el número de soles que se obtienen cuando una moneda se

lanza 4 veces, o a la variable aleatoria  $W$ , donde  $W$  es el número de cartas rojas que resultan cuando se sacan 4 cartas al azar de una baraja de manera sucesiva, se reemplaza cada carta y se baraja antes de sacar la siguiente.

## Distribuciones continuas de probabilidad

Una variable aleatoria continua tiene una probabilidad 0 de adoptar exactamente cualquiera de sus valores. En consecuencia, su distribución de probabilidad no se puede presentar en forma tabular. En un principio esto parecería sorprendente, pero se vuelve más probable cuando consideramos un ejemplo específico. Consideremos una variable aleatoria cuyos valores son las estaturas de todas las personas mayores de 21 años de edad. Entre cualesquiera dos valores, digamos 163.5 y 164.5 centímetros, o incluso entre 163.99 y 164.01 centímetros, hay un número infinito de estaturas, una de las cuales es 164 centímetros. La probabilidad de seleccionar al azar a una persona que tenga exactamente 164 centímetros de estatura en lugar de una del conjunto infinitamente grande de estaturas tan cercanas a 164 centímetros que humanamente no sea posible medir la diferencia es remota, por consiguiente, asignamos una probabilidad 0 a tal evento. Sin embargo, esto no ocurre si nos referimos a la probabilidad de seleccionar a una persona que mida al menos 163 centímetros pero no más de 165 centímetros de estatura. Aquí nos referimos a un intervalo en vez de a un valor puntual de nuestra variable aleatoria.

Nos interesamos por el cálculo de probabilidades para varios intervalos de variables aleatorias continuas como  $P(a < X < b)$ ,  $P(W \geq c)$ , etc. Observe que cuando  $X$  es continua,

$$P(a < X \leq b) = P(a < X < b) + P(X = b) = P(a < X < b).$$

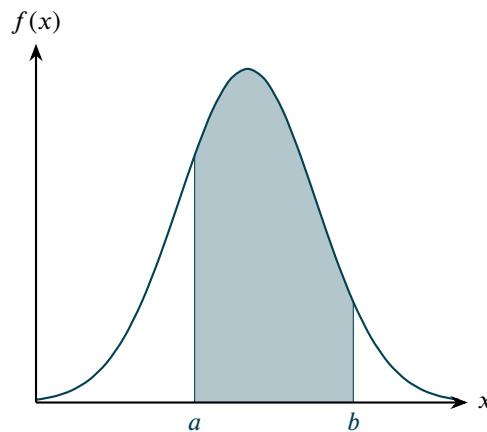
Es decir, no importa si incluimos o no un extremo del intervalo. Sin embargo, esto no es cierto cuando  $X$  es discreta.

Aunque la distribución de probabilidad de una variable aleatoria continua no se puede representar de forma tabular, sí es posible plantearla como una fórmula, la cual necesariamente será función de los valores numéricos de la variable aleatoria continua  $X$ , y como tal se representará mediante la notación funcional  $f(x)$ . Cuando se trata con variables continuas, a  $f(x)$  por lo general se le llama *función de densidad de probabilidad*, o simplemente *función de densidad* de  $X$ . Como  $X$  se define sobre un espacio de sucesos continuo, es posible que  $f(x)$  tenga un número finito de discontinuidades. Sin embargo, la mayoría de las funciones de densidad que tienen aplicaciones prácticas en el análisis de datos estadísticos son continuas y sus gráficas pueden tomar cualesquiera de varias formas, algunas de las cuales se presentan en la figura 1.10. Como se utilizarán áreas para representar probabilidades y éstas son valores numéricos positivos, la función de densidad debe caer completamente arriba del eje  $x$ .

Una función de densidad de probabilidad se construye de manera que el área bajo su curva limitada por el eje  $x$  sea igual a 1, cuando se calcula en el rango de  $X$  para el que se define  $f(x)$ . Esto asegura que la probabilidad total de todos los posibles valores de  $X$  sea 1, cumpliendo así con una de las propiedades fundamentales de las funciones de densidad de probabilidad.

Como este rango de  $X$  es un intervalo finito, siempre es posible extender el intervalo para que incluya a todo el conjunto de números reales. Esto se logra definiendo  $f(x)$  como cero en todos los puntos fuera del intervalo original. De esta manera, la función de densidad de probabilidad se puede considerar en todo el dominio de los números reales sin alterar su propiedad esencial de que el área bajo la curva es igual a 1.

**Figura 1.10**  $P(a < X < b)$



En la figura 1.10, la probabilidad de que  $X$  tome un valor entre  $a$  y  $b$  es igual al área sombreada bajo la función de densidad entre las ordenadas en  $x = a$  y  $x = b$ . Esta área representa la probabilidad acumulada de que la variable aleatoria  $X$  se encuentre dentro del intervalo  $[a, b]$ . Matemáticamente, esta probabilidad se calcula utilizando la integral definida de  $f(x)$  entre  $a$  y  $b$ , y está dada por la expresión:

$$P(a < X < b) = \int_a^b f(x)dx.$$

Esta integral suma todas las pequeñas probabilidades infinitesimales a lo largo del intervalo, proporcionando así la probabilidad total de que  $X$  esté entre  $a$  y  $b$ .

**Definición 1.7.6:**

La función  $f(x)$  es una *función de densidad de probabilidad* para la variable aleatoria continua  $X$ , definida en el conjunto de números reales, si

- i)  $f(x) \geq 0$ , para toda  $x \in \mathbb{R}$ .
- ii)  $\int_{-\infty}^{\infty} f(x)dx = 1$ .
- iii)  $P(a < X < b) = \int_a^b f(x)dx$ .

En el siguiente ejemplo, se demostrará que  $f(x)$  cumple con las dos primeras condiciones de la definición anterior. En cuanto a la tercera condición, esta es una propiedad inherente a las funciones de densidad de probabilidad y no requiere una verificación adicional específica. Esto se debe a que la integral representa la “acumulación” de probabilidad entre los límites  $a$  y  $b$ , una consecuencia directa del hecho de que la integral de  $f(x)$  sobre todo el dominio es igual a 1 (segundo inciso de la definición). Por lo tanto, dado que  $f(x)$  ya satisface las dos primeras condiciones, la tercera se deduce de manera natural, ya que la función está correctamente normalizada y no tiene valores negativos, lo que asegura que las probabilidades sean coherentes dentro de cualquier intervalo dado.

**Ejemplo 1.7.9:**

Suponga que el error en la temperatura de reacción, en °C, en un experimento de laboratorio controlado, es una variable aleatoria continua  $X$  que tiene la función de densidad de probabilidad

$$f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2, \\ 0, & \text{en otro caso.} \end{cases}$$

- a) Verifique que  $f(x)$  es una función de densidad.
- b) Calcule  $P(0 < X \leq 1)$ .

**Solución:** Usemos la definición anterior.

- a) Evidentemente,  $f(x) \geq 0$ . Para verificar la condición 2 de la definición anterior tenemos

$$\int_{-\infty}^{\infty} f(x) dx = \int_{-1}^2 \frac{x^2}{3} dx = \frac{x^3}{9} \Big|_{-1}^2 = \frac{8}{9} + \frac{1}{9} = 1.$$

- b) Si usamos la fórmula (iii) de la definición anterior, obtenemos

$$P(0 < X \leq 1) = \int_0^1 \frac{x^2}{3} dx = \frac{x^3}{9} \Big|_0^1 = \frac{1}{9}.$$

**Definición 1.7.7:**

La función de distribución acumulativa  $F(x)$ , de una variable aleatoria continua  $X$  con función de densidad  $f(x)$ , es

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt, \quad \text{para } -\infty < x < \infty.$$

Como una consecuencia inmediata de la definición anterior, se pueden escribir los dos resultados

$$P(a < X < b) = F(b) - F(a) \quad \text{y} \quad f(x) = \frac{dF(x)}{dx},$$

si existe la derivada.

**1.8 Esperanza matemática****1.9 Loterías****1.10 Racionalidad****1.11 Teoría de la utilidad esperada**



# 2

## Capítulo

# JUEGOS CON INFORMACIÓN PERFECTA

Contenido





# APÉNDICES



# A

## Apéndice

# TEORÍA DE CONJUNTOS

### Contenido

<b>A.1</b>	Subconjuntos e igualdad de conjuntos . . . . .	37
<b>A.2</b>	El conjunto vacío . . . . .	39
<b>A.3</b>	La unión e intersección de conjuntos . . . . .	40
<b>A.4</b>	Diferencia de conjuntos . . . .	43
<b>A.5</b>	Complemento de un conjunto .	44
<b>A.6</b>	Diagramas de Venn-Euler . . .	45
<b>A.7</b>	Producto cartesiano . . . . .	46
<b>A.8</b>	Representación geométrica del producto cartesiano . . . . .	49
<b>A.9</b>	Familias de conjuntos . . . . .	49
<b>A.10</b>	Conjunto potencia . . . . .	51

Partimos del supuesto de que la palabra conjunto es primitiva, y aceptamos que tenemos la idea de lo que es un conjunto. Para uniformizar dicha idea, diremos que los objetos que componen un conjunto son de cualquier especie y están bien determinados, esto último en el sentido de que dado un objeto arbitrario, se puede decidir si es o no del conjunto. Así por ejemplo, las mujeres bonitas no forman un conjunto; en cambio, los números pares sí. A los objetos que componen un conjunto se les llama elementos del conjunto.

Para nombrar conjuntos, generalmente se emplean letras mayúsculas  $A, B, C, D, \dots$ , y para nombrar a los elementos de un conjunto, si es el caso, se emplean usualmente letras minúsculas  $a, b, c, d, \dots$ . Para decir que un objeto  $x$  es elemento de un conjunto  $A$ , escribiremos  $x \in A$  y leemos:  $x$  es elemento de  $A$  o  $x$  está en  $A$  o  $x$  pertenece a  $A$ . Para decir que un objeto  $x$  no es elemento de un conjunto  $A$ , escribiremos  $x \notin A$ , y leemos:  $x$  no es elemento de  $A$  o  $x$  no está en  $A$  o  $x$  no pertenece a  $A$ .

Idealmente, nos gustaría que con cada enunciado o propiedad  $P$  (con solo una variable libre) se asociara un conjunto  $A$ , consistiendo de todos los objetos que satisfacen a  $P$ . Bajo esta situación escribiríamos

$$A = \{x \mid P(x)\}$$

(se lee:  $A$  es el conjunto de los objetos  $x$ , tal que  $x$  satisface a  $P$ ), donde  $P(x)$  es la condición o propiedad que debe satisfacer un objeto  $x$  para ser elemento de  $A$ . En este caso podemos construir los conjuntos

$$A = \{x \mid x \text{ es número real y } |x - 2| < 3\}$$

$$B = \{x \mid x = n^2, \text{ para algún número natural } n\}$$

$$C = \{x \mid x \text{ es número primo}\}$$

Pero también podríamos construir los conjuntos

$$D = \{x \mid x = x\}$$

$$E = \{x \mid x \text{ es humano}\}$$

$$F = \{x \mid x \text{ no es humano}\}$$

Claramente  $D \in D$ ,  $E \notin E$  y  $F \in F$ .

Análogamente podríamos construir los “conjuntos”:

$$F = \{x \mid x \text{ es conjunto}\} \quad \text{y} \quad G = \{x \mid x \text{ no es conjunto}\}.$$

Observemos que en este caso  $F \in F$  y  $G \notin G$ .

También podríamos construir el “conjunto”

$$X = \{x \mid x \text{ es conjunto y } x \notin x\}.$$

En este caso se tendría que

$$X \in X \implies X \notin X$$

y también que

$$X \notin X \implies X \in X;$$

que son contradicciones evidentes. A este ejemplo se le conoce como **Paradoja de Russell**, en honor al filósofo inglés Bertrand Russell quien la formuló.

De la exposición precedente, pudiera concluirse que la construcción de conjuntos en la forma  $\{x \mid P(x)\}$ , debe abandonarse, lo que no ha ocurrido; más bien se han hecho ciertas restricciones, a través del desarrollo de la teoría axiomática de conjuntos. En ese contexto, una propuesta es que hay dos tipos de colecciones, las clases que son conjuntos y las clases que no son conjuntos: cualquier colección de objetos especificados por alguna propiedad  $P$ , es una clase; mientras que un conjunto es una clase que es miembro de otra clase. Así que

$$A = \{x \mid x \text{ es conjunto y } x \notin x\}$$

es una clase que no es conjunto, con lo que se evita la paradoja de Russell.

Por otro lado, la expresión

$$A = \{2, \sqrt{2}, \pi\}$$

denota al conjunto

$$A = \{x \mid x = 2 \text{ o } x = \sqrt{2} \text{ o } x = \pi\}.$$

Análogamente, las expresiones

$$A = \{2, 4, 6, \dots, 100\}$$

y

$$B = \{2, 4, 6, \dots\}$$

son casos particulares de notación para los conjuntos

$$A = \{x \mid x \text{ es número par menor o igual que } 100\}$$

y

$$B = \{x \mid x \text{ es número par}\}$$

respectivamente.

Observemos que una persona de nombre Juan, no es elemento del conjunto

$$A = \{\text{Juan, Luis, silla, mesa, venus, tierra}\}$$

lo que es elemento del conjunto  $A$ , es la palabra (nombre) Juan. En este caso,  $A$  es un conjunto de palabras (nombres).

El otro modo de construir conjuntos es por medio del **axioma de elección**, el cual no enunciaremos pues no es el objetivo de esta obra. Pero, para tener una idea de esta construcción, veamos el ejemplo siguiente. Clasifiquemos a los números reales en diferentes conjuntos, de manera que: Dos números reales  $x$  y  $y$  están en un mismo conjunto, si y solo si,  $x - y$  es número racional. Digamos que  $\Omega$  es la colección de tales conjuntos. Por ejemplo, para cada número primo  $p$ , los conjuntos

$$A_p = \{x \mid x = \sqrt{p} + r, \text{ con } r \text{ número racional}\}$$

y

$$C = \{x \mid x \text{ es número racional}\}$$

son elementos de  $\Omega$ . Claramente  $\Omega$  es infinito y también cada  $S \in \Omega$  es infinito. Afirmamos que existe un conjunto  $T$  el cual posee solo un elemento en común con cada conjunto  $S \in \Omega$ . El conjunto  $T$  no puede construirse en la forma  $\{x \mid P(x)\}$ .

En resumen, hay dos maneras de construir conjuntos, una es en la forma  $\{x \mid P(x)\}$ , la cual requiere ciertas restricciones, y que se llama construcción por extensión; y la otra construcción es por medio del axioma de elección.

## A.1 Subconjuntos e igualdad de conjuntos

### Definición A.1.1:

Sean  $A$  y  $B$  conjuntos. Decimos que  $A$  es subconjunto de  $B$ , y escribimos  $A \subset B$ , si cada elemento de  $A$ , es también elemento de  $B$ .

Observemos que la negación de que  $A$  es subconjunto de  $B$ , es la negación de que cada elemento de  $A$  es también elemento de  $B$ . Por tanto,  $A$  no es subconjunto de  $B$  si existe al menos un elemento de  $A$  que no es elemento de  $B$ .

En lugar de  $A \subset B$  también se escribe  $B \supset A$ , y en cualquier caso se lee:  $A$  es subconjunto de  $B$  o  $A$  está contenido en  $B$  o  $B$  contiene a  $A$ . Para decir que  $A$  no es subconjunto de  $B$  escribimos  $A \not\subset B$  o  $B \not\supset A$ , y en cualquier caso se lee:  $A$  no es subconjunto de  $B$  o  $A$  no está contenido en  $B$  o  $B$  no contiene a  $A$ . Simbólicamente se tiene que:

$$(A \subset B) \iff (\forall x, x \in A \implies x \in B)$$

y

$$(A \not\subset B) \iff (\exists x \text{ tal que } x \in A \text{ y } x \notin B).$$

Es conveniente observar que si

$$A = \{x \mid P(x)\}$$

y

$$B = \{x \mid Q(x)\}$$

entonces

$$[A \subset B] \iff [P(x) \implies Q(x)].$$

### Ejemplo A.1.1:

Si

$$A = \{x \mid x \text{ es número real y } |x| \leq 3\},$$

$$B = \{x \mid x \text{ es número real y } |x| \leq 4\}$$

entonces  $A \subset B$  y  $B \not\subset A$ .

**Ejemplo A.1.2:** Si

$$A = \{x \mid x \text{ es número primo}\},$$

$$B = \{x \mid x \text{ es número impar}\}$$

entonces  $A \not\subset B$  y  $B \not\subset A$ .

**Ejemplo A.1.3:** Si

$$A = \{1, 2\},$$

$$B = \{1, 2, 2, 1\}$$

entonces  $A \subset B$  y  $B \subset A$ .

**Proposición A.1.1:** Si  $A$ ,  $B$  y  $C$  son conjuntos, entonces:

i.  $A \subset A$ .

ii.  $(A \subset B \text{ y } B \subset C) \implies A \subset C$ .

**Demostración:**

i. Es inmediato.

ii. Si  $A \subset B$  y  $B \subset C$ , entonces  $(x \in A \implies x \in B)$  y  $(x \in B \implies x \in C)$ , por tanto  $(x \in A \implies x \in C)$ , de donde se sigue que  $A \subset C$ .

**Definición A.1.2:**

Sean  $A$  y  $B$  conjuntos. Decimos que  $A$  es igual a  $B$ , y escribimos  $A = B$ , si  $A \subset B$  y  $B \subset A$ .

Simbólicamente se tiene que:

$$(A = B) \iff (A \subset B \text{ y } B \subset A).$$

Observemos que la negación de que los conjuntos  $A$  y  $B$  son iguales, lo que se expresa escribiendo  $A \neq B$ , y que se lee:  $A$  no es igual a  $B$  o  $A$  es distinto de  $B$ , es la negación de que  $[A \subset B \text{ y } B \subset A]$ . En consecuencia,

$$[A \neq B] \iff [A \not\subset B \text{ o } B \not\subset A].$$

Es decir,

$$[A \neq B] \iff [(\exists x \text{ tal que } x \in A \text{ y } x \notin B) \text{ o } (\exists x \text{ tal que } x \in B \text{ y } x \notin A)].$$

También conviene observar que, si  $A = \{x \mid P(x)\}$  y  $B = \{x \mid Q(x)\}$ , entonces

$$[A = B] \text{ si y solo si } [P(x) \iff Q(x)].$$

**Ejemplo A.1.4:** Si

$$A = \{x \mid x \text{ es número real y } x^2 - 1 = 0\}$$

y

$$B = \{-1, 1\},$$

entonces  $A = B$ .

**Ejemplo A.1.5:** Si

$$A = \{x \mid x \text{ es múltiplo de 6 y 9}\}$$

y

$$B = \{x \mid x \text{ es múltiplo de 18}\},$$

entonces  $A = B$ .

**Ejemplo A.1.6:**

Si

$$A = \{x \mid x \text{ es número primo y } x > 2\}$$

y

$$B = \{x \mid x \text{ es número impar}\},$$

entonces  $A \subset B$  y  $B \not\subset A$ , por tanto  $A \neq B$ .**Proposición A.1.2:**Si  $A$ ,  $B$  y  $C$  son conjuntos, entonces:

- i.  $A = A$ .
- ii.  $A = B \implies B = A$ .
- iii.  $(A = B \text{ y } B = C) \implies A = C$ .

**Demostración:**

- i.  $A \implies A \subset A$  y  $A \subset A \implies A = A$ .
- ii.  $A = B \implies A \subset B$  y  $B \subset A \implies B \subset A$  y  $A \subset B \implies B = A$ .
- iii.  $A = B$  y  $B = C \implies (A \subset B \text{ y } B \subset A)$  y  $(B \subset C \text{ y } C \subset B)$ . Entonces  $(A \subset B \text{ y } B \subset C)$  y  $(C \subset B \text{ y } B \subset A)$ . Por (ii) de la proposición A.1.1, se sigue  $A \subset C$  y  $C \subset A \implies A = C$ .

Hemos definido bajo qué condiciones un conjunto es *subconjunto* de otro. Así que, el término subconjunto sirve para relacionar conjuntos, pero no cualesquiera dos conjuntos pueden relacionarse bajo este término, es decir, es posible que dados dos conjuntos, ninguno sea subconjunto del otro. Naturalmente que la construcción de subconjuntos de un conjunto, puede hacerse por extensión o por medio del axioma de selección. Para construir por extensión un subconjunto  $A$  de un conjunto  $B$ , procedemos del modo siguiente:

$$A = \{x \mid x \in B \text{ y } P(x)\} \quad \text{o} \quad A = \{x \in B \mid P(x)\},$$

donde  $P(x)$  es una propiedad que puede satisfacer, o no, cada elemento  $x$  de  $B$ . Claro que

$$\{x \mid x \in B \text{ y } P(x)\} \subset \{x \mid x \in B \text{ y } Q(x)\},$$

si y solo si,  $\forall x \in B, P(x) \implies Q(x)$ . Finalmente, observemos que si  $A$  es conjunto, entonces

$$\begin{aligned} \{x \mid x \in A \text{ y } x = x\} &= \{x \mid x \in A\} \\ &= A. \end{aligned}$$

## A.2 El conjunto vacío

**Definición A.2.1:**Para cada conjunto  $A$ , definimos el conjunto

$$\emptyset_A = \{x \mid x \in A \text{ y } x \neq x\},$$

el cual es subconjunto de  $A$  y se le llama subconjunto vacío (o nulo) de  $A$ .

Claro que para cada conjunto  $A$ , el conjunto  $\emptyset_A$  no posee elementos, pues cada  $x \in A$  satisface  $x = x$ .

**Proposición A.2.1:**Si  $A$  y  $B$  son conjuntos, entonces

$$\emptyset_A = \emptyset_B.$$

**Demostración:** Se deja como ejercicio al lector.

**Definición A.2.2:**

Al único conjunto que no posee elementos, se le llama conjunto vacío y se le denota por  $\emptyset$ .

El conjunto  $\emptyset$  es subconjunto de cualquier conjunto, es decir, para cualquier conjunto  $A$ ,  $\emptyset \subset A$ . Así que, si  $A$  es conjunto no vacío, entonces al menos tiene como subconjuntos a  $\emptyset$  y a  $A$  mismo, los cuales se llaman *subconjuntos impropios* de  $A$ .

**Definición A.2.3:**

Decimos que un conjunto  $B$  es subconjunto propio de un conjunto  $A$ , si  $B \neq \emptyset$ ,  $B \subset A$  y  $B \neq A$ .

Observemos que

$$(B \subset A \text{ y } B \neq A) \iff (B \subset A \text{ y } A \not\subset B).$$

De la teoría axiomática de conjuntos, se deduce que ningún conjunto puede ser elemento de sí mismo, y que si  $A$  y  $B$  son conjuntos no vacíos, no es posible que se cumpla a la vez que  $A \in B$  y  $B \in A$ . También se deduce que

$$\{x \mid x \notin x\},$$

y

$$\{x \mid x = x\},$$

no son conjuntos.

### A.3 La unión e intersección de conjuntos

**Definición A.3.1:**

Sean  $A$  y  $B$  conjuntos. Definimos la unión de  $A$  y  $B$ , denotada por  $A \cup B$ , como el conjunto

$$A \cup B = \{x \mid x \in A \text{ o } x \in B\}.$$

La expresión  $A \cup B$ , se lee:  $A$  unión  $B$  o la unión de  $A$  y  $B$ .

Simbólicamente se tiene que:

$$(x \in A \cup B) \iff (x \in A \text{ o } x \in B)$$

y

$$(x \notin A \cup B) \iff (x \notin A \text{ y } x \notin B).$$

**Ejemplo A.3.1:**

Si

$$A = \{1, 3\}$$

y

$$B = \{0, 1, 2\},$$

entonces  $A \cup B = \{0, 1, 2, 3\}$ .

**Ejemplo A.3.2:**

Si

$$A = \{x \mid x \text{ es número par}\}$$

y

$$B = \{x \mid x \text{ es número impar}\},$$

entonces  $A \cup B = \{x \mid x \text{ es número entero positivo}\}$ .

**Ejemplo A.3.3:**

Si  $A = \emptyset$  y  $B = \emptyset$ , entendemos que

$$A \cup B = \emptyset.$$



**Proposición A.3.1:** Si  $A$ ,  $B$  y  $C$  son conjuntos, entonces:

- i.  $A \cup A = A$ .
- ii.  $A \cup B = B \cup A$ .
- iii.  $\emptyset \cup A = A$ .
- iv.  $(A \cup B) \cup C = A \cup (B \cup C)$ .
- v.  $A \subset A \cup B$  y  $B \subset A \cup B$ .
- vi.  $(A \subset C \text{ y } B \subset C) \iff A \cup B \subset C$ .
- vii.  $A \subset B \iff A \cup B = B$ .
- viii.  $A \subset B \implies A \cup C \subset B \cup C$ .
- ix.  $A \subset B \text{ y } C \subset D \implies A \cup C \subset B \cup D$ .

**Demostración:** Solo se demostrará (viii) y (ix), los demás dejan como ejercicio al lector.

viii.  $x \in A \cup C \implies x \in A \text{ o } x \in C \implies x \in B \text{ o } x \in C \implies x \in B \cup C$ .

ix. Por hipótesis,  $A \subset B$  y  $C \subset D$ , entonces, aplicando (viii), se tiene

$$\begin{aligned} A \cup C &\subset B \cup C \\ &= C \cup B \\ &\subset D \cup B \\ &= B \cup D. \end{aligned}$$

**Definición A.3.2:**

Sean  $A$  y  $B$  conjuntos. Definimos la intersección de  $A$  y  $B$ , denotada por  $A \cap B$ , como el conjunto

$$A \cap B = \{x \mid x \in A \text{ y } x \in B\}.$$

La expresión  $A \cap B$ , se lee:  $A$  intersección  $B$  o la intersección de  $A$  y  $B$ .

Simbólicamente se tiene que:

$$(x \in A \cap B) \iff (x \in A \text{ y } x \in B)$$

y

$$(x \notin A \cap B) \iff (x \notin A \text{ o } x \notin B).$$

**Ejemplo A.3.4:** Si

$$A = \{1, 2, 3\}$$

y

$$B = \{0, 1, 2\},$$

entonces  $A \cap B = \{1, 2\}$ .

**Ejemplo A.3.5:** Si

$$A = \{x \mid x \text{ es número par}\}$$

y

$$B = \{x \mid x \text{ es número impar}\},$$

entonces  $A \cap B = \emptyset$ .

**Ejemplo A.3.6:** Si

$$A = \{x \mid x \text{ es número par}\}$$

y

$$B = \{x \mid x \text{ es número primo}\},$$

entonces  $A \cap B = \{2\}$ .

**Proposición A.3.2:** Si  $A$ ,  $B$  y  $C$  son conjuntos, entonces:

- i.  $A \cap A = A$ .
- ii.  $A \cap B = B \cap A$ .
- iii.  $\emptyset \cap A = \emptyset$ .
- iv.  $(A \cap B) \cap C = A \cap (B \cap C)$ .
- v.  $A \cap B \subset A$  y  $A \cap B \subset B$ .
- vi.  $(C \subset A \text{ y } C \subset B) \iff C \subset A \cap B$ .
- vii.  $(A \subset B) \iff (A \cap B = A)$ .
- viii.  $A \subset B \implies A \cap C \subset A \cap B$ .
- ix.  $A \subset B \text{ y } C \subset D \implies A \cap C \subset B \cap D$ .

**Demostración:** Se proponen como ejercicio al lector.

**Definición A.3.3:** Sean  $A$  y  $B$  conjuntos. Decimos que  $A$  y  $B$  son ajenos, si  $A \cap B = \emptyset$ .

Observemos que el conjunto vacío es ajeno con cualquier conjunto.

La siguiente proposición enuncia las propiedades distributivas de la unión respecto de la intersección, y viceversa.

**Proposición A.3.3:** Si  $A$ ,  $B$  y  $C$  son conjuntos, entonces:

- i.  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ .
- ii.  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ .

**Demostración:**

- i. Por definición de igualdad, debemos probar que

$$A \cap (B \cup C) \subset (A \cap B) \cup (A \cap C)$$

y que

$$(A \cap B) \cup (A \cap C) \subset A \cap (B \cup C).$$

En este caso podemos probar simultáneamente ambas contenciones. De esta forma, tenemos que

$$\begin{aligned} x \in A \cap (B \cup C) &\iff x \in A \text{ y } x \in (B \cup C) \\ &\iff x \in A \text{ y } (x \in B \text{ o } x \in C) \\ &\iff (x \in A \text{ y } x \in B) \text{ o } (x \in A \text{ y } x \in C) \\ &\iff x \in A \cap B \text{ o } x \in A \cap C \\ &\iff x \in (A \cap B) \cup (A \cap C) \end{aligned}$$

- ii. Para demostrar la igualdad, debemos verificar ambas contenciones. Primero, debemos probar que

$$A \cup (B \cap C) \subset (A \cup B) \cap (A \cup C).$$

Luego, debemos mostrar que

$$(A \cup B) \cap (A \cup C) \subset A \cup (B \cap C).$$

Pero también podemos proceder de la siguiente manera:

$$\begin{aligned} (A \cup B) \cap (A \cup C) &= [(A \cup B) \cap A] \cup [(A \cup B) \cap C] \\ &= [A] \cup [(A \cap C) \cup (B \cap C)] \\ &= [A \cup (A \cap C)] \cup (B \cap C) \\ &= A \cup (B \cap C) \end{aligned}$$

## A.4 Diferencia de conjuntos

### Definición A.4.1:

Sean  $A$  y  $B$  conjuntos. Definimos la diferencia de  $A$  y  $B$ , denotada por  $A - B$ , como el conjunto

$$A - B = \{x \mid x \in A \text{ y } x \notin B\}.$$

La expresión  $A - B$  se lee:  $A$  menos  $B$  o la diferencia de  $A$  y  $B$ .

Simbólicamente se tiene que:

$$(x \in A - B) \iff (x \in A \text{ y } x \notin B)$$

y

$$(x \notin A - B) \iff (x \notin A \text{ o } x \in B).$$

### Ejemplo A.4.1:

Si

$$A = \{1, 2, 3, 4, 5, 6\}$$

y

$$B = \{2, 4, 6, 8, 10\},$$

entonces

$$A - B = \{1, 3, 5\}$$

y

$$B - A = \{8, 10\}.$$

### Ejemplo A.4.2:

Si

$$A = \{x \mid x \text{ es número natural}\}$$

y

$$B = \{x \mid x \text{ es número par}\},$$

entonces

$$A - B = \{x \mid x \text{ es número impar}\}$$

y

$$B - A = \emptyset.$$

De los ejemplos anteriores se deduce que, en general,

$$A - B \neq B - A;$$

es decir, que la diferencia de conjuntos no es conmutativa. Esto significa que el orden en el que se restan los conjuntos importa, ya que restar  $B$  de  $A$  no produce el mismo resultado que restar  $A$  de  $B$ . Además, la diferencia de conjuntos no es asociativa, es decir, en general,

$$(A - B) - C \neq A - (B - C).$$

Esto implica que la forma en que agrupamos los conjuntos al restarlos también afecta el resultado.

### Proposición A.4.1:

Si  $A$  y  $B$  son conjuntos, entonces:

- i.  $A - A = \emptyset$ .
- ii.  $A - \emptyset = A$ .
- iii.  $A - B \subset A$ .
- iv.  $A \subset B \iff A - B = \emptyset$ .

**Demostración:** Se dejan como ejercicio al lector.

## A.5 Complemento de un conjunto

### Definición A.5.1:

Sean  $A$  y  $B$  conjuntos. Definimos el complemento de  $A$  relativo a  $B$ , denotado por  $\mathcal{C}_B(A)$ , como el conjunto

$$\mathcal{C}_B(A) = B - A.$$

Si  $A$  es un conjunto, no es posible hablar de  $\{x \mid x \notin A\}$ , pues nos llevaría a conclusiones paradójicas. Por ejemplo, si  $A = \{1, 2\}$  y  $B = \{x \mid x \notin A\}$ , entonces  $A \in B$  y también  $B \in B$ . Sin embargo, siempre que trabajamos con conjuntos, podemos suponer que estos son subconjuntos de otro conjunto “más grande” denominado **conjunto universal**. Tal conjunto universal depende del discurso, esto es, de los conjuntos con que se trabaje en una situación dada. En general llamaremos  $X$  al conjunto universal de un discurso, y esto significa que los conjuntos que se mencionen en el discurso, serán subconjuntos de  $X$ . Lo anterior permitirá simplificar la escritura cuando hablemos del complemento de un conjunto  $A$ , relativo al conjunto universal  $X$ ; pues en lugar de  $\mathcal{C}_X(A)$ , escribiremos simplemente  $\mathcal{C}(A)$  o  $A^C$ . Por tanto

$$A^C = X - A.$$

Vamos a convenir en escribir

$$x \in A^C \iff x \notin A,$$

en lugar de

$$x \in A^C \iff x \in X \text{ y } x \notin A,$$

y por tanto

$$x \notin A^C \iff x \in A.$$

### Proposición A.5.1:

Si  $A$  es un conjunto y  $X$  es el conjunto universal, entonces:

- i.  $A \cup A^C = X$ .
- ii.  $A \cap A^C = \emptyset$ .
- iii.  $A - A^C = A$ .
- iv.  $(A^C)^C = A$ .
- v.  $\emptyset^C = X$ .
- vi.  $X^C = \emptyset$ .
- vii.  $A \cup X = X$ .
- viii.  $A \cap X = A$ .
- ix.  $A - X = \emptyset$ .
- x.  $A - B = A \cap B^C$ .
- xi.  $A \subset B \iff B^C \subset A^C$ .
- xii.  $A \subset B^C \iff B \subset A^C$ .

**Demostración:** Solo se demostrará (i), los demás incisos se dejan como ejercicio al lector.

i. Por definición,

$$A^C = \{x \in X \mid x \notin A\}.$$

Ahora, consideremos

$$A \cup A^C = \{x \in X \mid x \in A \text{ o } x \in A^C\}.$$

Dado que  $A$  y  $A^C$  cubren todos los elementos del conjunto universal  $X$ , cualquier elemento  $x \in X$  estará en  $A$  o en  $A^C$ . Por lo tanto

$$A \cup A^C = X.$$

El siguiente teorema es conocido como el Teorema de De Morgan. Este teorema establece las relaciones entre las operaciones de unión e intersección de conjuntos y sus complementos, proporcionando una herramienta fundamental en la teoría de conjuntos y la lógica matemática.

---

**Teorema A.5.1:** Si  $A$  y  $B$  son conjuntos y  $X$  es el conjunto universal, entonces:

- i.  $(A \cup B)^C = A^C \cap B^C$ .
- ii.  $(A \cap B)^C = A^C \cup B^C$ .

**Demostración:**

- i. Para demostrar que  $(A \cup B)^C = A^C \cap B^C$ , consideremos un elemento  $x$  y analicemos su pertenencia a los conjuntos involucrados:

$$\begin{aligned} x \in (A \cup B)^C &\Leftrightarrow x \notin A \cup B \\ &\Leftrightarrow x \notin A \text{ y } x \notin B \\ &\Leftrightarrow x \in A^C \text{ y } x \in B^C \\ &\Leftrightarrow x \in A^C \cap B^C \end{aligned}$$

- ii. Para demostrar que  $(A \cap B)^C = A^C \cup B^C$ , consideremos un elemento  $x$  y analicemos su pertenencia a los conjuntos involucrados:

$$\begin{aligned} x \in (A \cap B)^C &\Leftrightarrow x \notin A \cap B \\ &\Leftrightarrow x \notin A \text{ y } x \notin B \\ &\Leftrightarrow x \in A^C \text{ o } x \in B^C \\ &\Leftrightarrow x \in A^C \cup B^C \end{aligned}$$


---

**Proposición A.5.2:** Si  $A$ ,  $B$ ,  $C$  y  $D$  son conjuntos y  $X$  es el conjunto universal, entonces:

- i.  $A - B = A - (A \cap B)$ .
- ii.  $(A \cap B = \emptyset \text{ y } A \cup B = X) \Rightarrow B = A^C$ .
- iii.  $A - B = A \Leftrightarrow B - A = B$ .
- iv.  $(A \cup B) - (A \cap B) = (A - B) \cup (B - A)$ .

**Demostración:** Se dejan como ejercicio al lector.

---

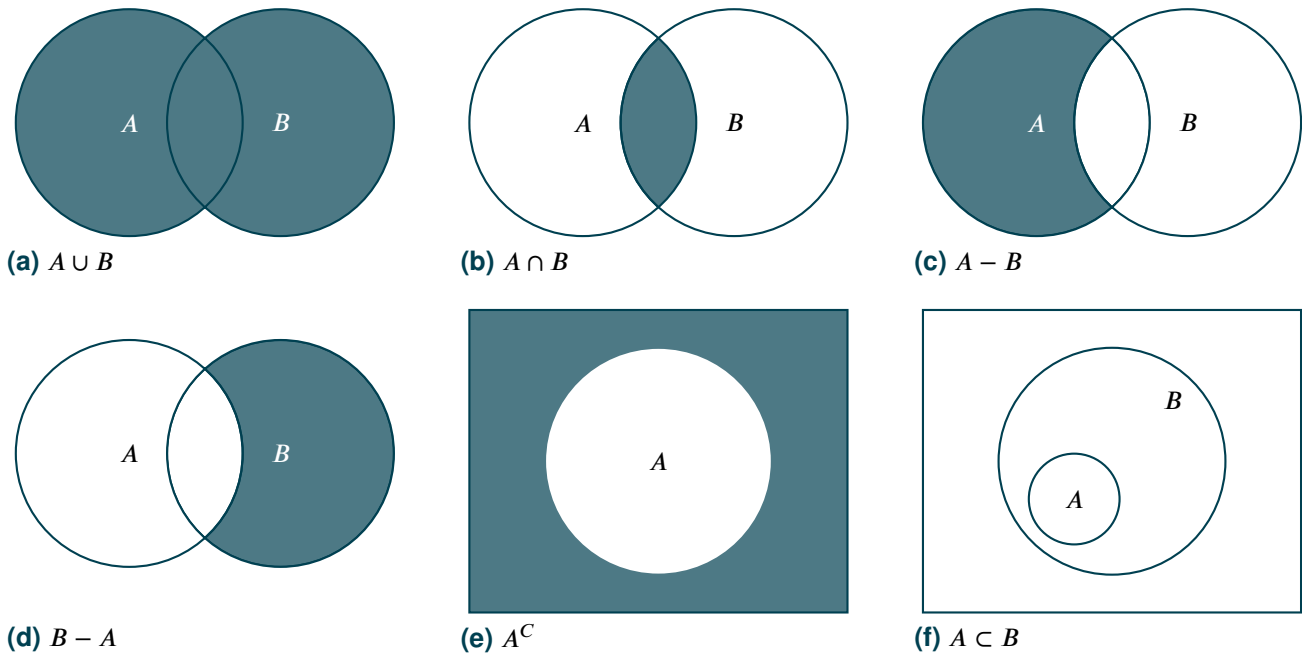
## A.6 Diagramas de Venn-Euler

En los diagramas de Venn-Euler, a un conjunto no vacío se le representa con una figura cerrada, dentro de un rectángulo que representa el conjunto universal y se utilizan para comparar, contrastar y analizar la intersección, la unión y la diferencia de dichos conjuntos.

**Figura A.1** Ejemplo de diagrama de Venn-Euler que muestra tres conjuntos  $A$ ,  $B$  y  $C$  dentro de un conjunto universal  $X$ . El conjunto  $A$  está representado por el círculo más grande a la izquierda, el conjunto  $B$  por el círculo mediano a la derecha, y el conjunto  $C$  por el círculo más pequeño en el centro.



Las intersecciones entre las figuras se utilizan para mostrar la relación entre estos conjuntos, ya sea para señalar la presencia de elementos comunes o para evidenciar su exclusividad. Además, podemos hacer diagramas de las operaciones antes vistas:



**Figura A.2** Diagramas de Venn que ilustran varias operaciones de conjuntos. (a) La unión de  $A$  y  $B$ , representada por el área sombreada que cubre ambos conjuntos. (b) La intersección de  $A$  y  $B$ , mostrada como el área común sombreada entre ambos conjuntos. (c) La diferencia  $A - B$ , que es el área de  $A$  excluyendo la intersección con  $B$ . (d) La diferencia  $B - A$ , que es el área de  $B$  excluyendo la intersección con  $A$ . (e) El complemento de  $A$ ,  $A^C$ , que es el área fuera del conjunto  $A$  dentro del conjunto universal. (f) La relación de subconjunto  $A \subset B$ , donde  $A$  está completamente contenido dentro de  $B$ .

## A.7 Producto cartesiano

### Definición A.7.1:

Sean  $A$  y  $B$  conjuntos no vacíos, y sean  $a \in A$  y  $b \in B$ . Definimos la pareja ordenada  $(a, b)$ , donde  $a$  es la primera entrada (o primera coordenada) y  $b$  es la segunda entrada (o segunda coordenada), como

$$(a, b) = \{\{a\}, \{a, b\}\}.$$

### Proposición A.7.1:

Sean  $A$  y  $B$  conjuntos no vacíos y sean  $a \in A$  y  $b \in B$ . Entonces

$$(a, b) = (b, a) \iff a = b.$$

**Demostración:** Se deja como ejercicio al lector.

En consecuencia de la proposición anterior, tenemos que

$$(a, b) \neq (b, a) \iff a \neq b.$$

### Teorema A.7.1:

Sean  $A$  y  $B$  conjuntos no vacíos, sean  $a, c \in A$  y sean  $b, d \in B$ . Entonces

$$(a, b) = (c, d) \iff a = c \text{ y } b = d.$$

**Demostración:** Se deja como ejercicio al lector.

En consecuencia del teorema anterior que

$$(a, b) \neq (c, d) \iff a \neq c \text{ o } b \neq d.$$

**Definición A.7.2:**

Sean  $A$  y  $B$  conjuntos. Si  $A \neq \emptyset$  y  $B \neq \emptyset$ , se define el producto cartesiano de  $A$  y  $B$ , denotado por  $A \times B$ , como el conjunto

$$A \times B = \{(a, b) \mid a \in A \text{ y } b \in B\};$$

y si  $A = \emptyset$  o  $B = \emptyset$ , se define  $A \times B = \emptyset$ . La expresión  $A \times B$  se lee:  $A$  cruz  $B$  o el producto cartesiano de  $A$  y  $B$ .

Observemos que:

1.  $(a, b) \in A \times B \iff a \in A \text{ y } b \in B$ .
2.  $x \in A \times B \iff \exists a \in A \text{ y } \exists b \in B \text{ tales que } x = (a, b)$ .
3.  $(a, b) \notin A \times B \iff a \notin A \text{ o } b \notin B$ .

**Ejemplo A.7.1:**

Si  $A = \{0, 1, 2\}$  y  $B = \{2, 3\}$ , entonces

$$A \times B = \{(0, 2), (0, 3), (1, 2), (1, 3), (2, 2), (2, 3)\}$$

y

$$B \times A = \{(2, 0), (2, 1), (2, 2), (3, 0), (3, 1), (3, 2)\}.$$

Del ejemplo anterior se sigue que, en general,

$$A \times B \neq B \times A.$$

Esto significa que el producto cartesiano de dos conjuntos no es conmutativo. En otras palabras, el orden en el que se toman los conjuntos importa.

**Proposición A.7.2:**

Si  $A, B, C$  y  $D$  son conjuntos no vacíos, entonces:

- i.  $A \times B = B \times A \iff A = B$ .
- ii.  $A \subset C \text{ y } B \subset D \iff A \times B \subset C \times D$ .
- iii.  $A \times (B \cup C) = (A \times B) \cup (A \times C)$ .
- iv.  $(A \cup B) \times C = (A \times C) \cup (B \times C)$ .
- v.  $A \times (B \cap C) = (A \times B) \cap (A \times C)$ .
- vi.  $(A \cap B) \times C = (A \times C) \cap (B \times C)$ .
- vii.  $A \times (B - C) = (A \times B) - (A \times C)$ .
- viii.  $(A - B) \times C = (A \times C) - (B \times C)$ .

**Demostración:** Probaremos los incisos (ii) y (iii), los otros se dejan como ejercicio al lector.

- ii. a) Supongamos que  $A \subset C$  y  $B \subset D$ :  $(a, b) \in A \times B \implies a \in A \text{ y } b \in B \implies$  (por hipótesis)  $a \in C \text{ y } b \in D \implies (a, b) \in C \times D$ . En consecuencia,  $A \times B \subset C \times D$ .
- b) Supongamos ahora que  $A \times B \subset C \times D$ :  $a \in A \text{ y } b \in B \implies (a, b) \in A \times B \implies$  (por hipótesis)  $(a, b) \in C \times D \implies a \in C \text{ y } b \in D$ . En consecuencia,  $A \subset C$  y  $B \subset D$ .
- iii.  $(a, b) \in A \times (B \cup C) \iff a \in A \text{ y } b \in (B \cup C) \iff a \in A \text{ y } (b \in B \text{ o } b \in C) \iff (a \in A \text{ y } b \in B) \text{ o } (a \in A \text{ y } b \in C) \iff (a, b) \in A \times B \text{ o } (a, b) \in A \times C \iff (a, b) \in (A \times B) \cup (A \times C)$ .

**Definición A.7.3:**

Sean  $A_1, A_2, \dots, A_n$  conjuntos no vacíos. Sean  $a_1 \in A_1, a_2 \in A_2, \dots, a_n \in A_n$ . Definimos la  $n$ -ada ordenada, con  $n \geq 3$ , como

$$(a_1, a_2, \dots, a_n) = ((a_1, a_2, \dots, a_{n-1}), a_n).$$

Al elemento  $a_k$ , donde  $k \in \{1, 2, \dots, n\}$ , se le llama la  $k$ -ésima entrada, o  $k$ -ésima coordenada, de la  $n$ -ada  $(a_1, a_2, \dots, a_n)$ .

Observemos que por definición  $(a_1, a_2, a_3) = ((a_1, a_2), a_3)$ .

**Proposición A.7.3:**

Sean  $A_0, A_1, \dots, A_n$  conjuntos no vacíos; y sean  $a_i, b_i \in A_i, \forall i = 0, 1, \dots, n$ . Entonces

$$(a_0, a_1, \dots, a_n) = (b_0, b_1, \dots, b_n) \iff a_i = b_i, \forall i = 0, 1, \dots, n.$$

**Demostración:** Procederemos por inducción sobre  $n$ .

i. Si  $n = 1$ , por el teorema A.7.1, se sabe que

$$(a_0, a_1) = (b_0, b_1) \iff a_0 = b_0 \text{ y } a_1 = b_1,$$

con lo que se tiene lo deseado.

ii. Suponemos el resultado cierto para  $n = k$ , es decir, suponemos que

$$(a_0, a_1, \dots, a_k) = (b_0, b_1, \dots, b_k) \iff a_i = b_i, \forall i = 0, 1, \dots, k.$$

iii. Probaremos, a partir de (ii), que el resultado es cierto para  $n = k + 1$ , entonces

$$(a_0, a_1, \dots, a_k, a_{k+1}) = (b_0, b_1, \dots, b_k, b_{k+1}),$$

si y solo si

$$((a_0, a_1, \dots, a_k), a_{k+1}) = ((b_0, b_1, \dots, b_k), b_{k+1}),$$

si y solo si

$$(a_0, a_1, \dots, a_k) = (b_0, b_1, \dots, b_k) \text{ y } a_{k+1} = b_{k+1},$$

si y solo si  $a_i = b_i, \forall i = 0, 1, \dots, k$  y  $a_{k+1} = b_{k+1}$ .

**Definición A.7.4:**

Sean  $A_1, A_2, \dots, A_n$  conjuntos.

i. Si  $A_i \neq \emptyset, \forall i = 1, 2, \dots, n$ , entonces se define el producto cartesiano de  $A_1, A_2, \dots, A_n$ , denotado por  $A_1 \times A_2 \times \dots \times A_n$ , como el conjunto

$$A_1 \times A_2 \times \dots \times A_n = \{(a_1, a_2, \dots, a_n) \mid a_i \in A_i, \forall i = 1, \dots, n\}.$$

ii. Si  $A_i = \emptyset$ , para algún  $i = 1, 2, \dots, n$ , se define

$$A_1 \times A_2 \times \dots \times A_n = \emptyset.$$

**Definición A.7.5:**

Si  $A$  es un conjunto y  $n \in \mathbb{N}$ , definimos  $A^n$  como

$$A^n = \underbrace{A \times A \times \dots \times A}_{n \text{ veces}}.$$

**Ejemplo A.7.2:**

$$A^2 = A \times A \text{ y } A^3 = A \times A \times A.$$



## A.8 Representación geométrica del producto cartesiano

Sean  $X$  y  $Y$  conjuntos, digamos que

$$X = \{a, b, c, \dots\}$$

y

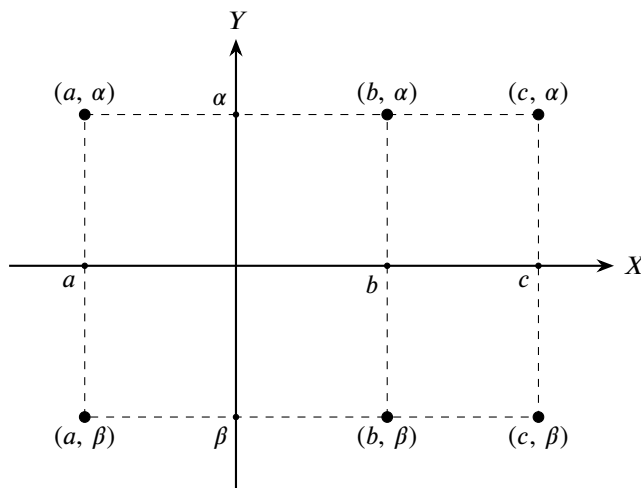
$$Y = \{\alpha, \beta, \dots\}.$$

Geométricamente

$$X \times Y = \{(a, \alpha), (a, \beta), (b, \alpha), (b, \beta), (c, \alpha), (c, \beta), \dots\}$$

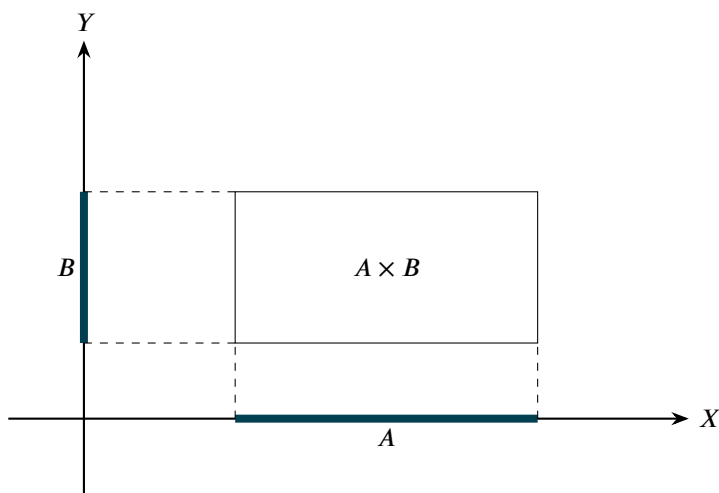
se representa como sigue:

**Figura A.3** Diagrama que representa el producto cartesiano  $X \times Y$ . Los ejes  $X$  e  $Y$  están etiquetados, y las líneas punteadas delimitan las regiones donde se encuentran los puntos. Los puntos  $(a, \alpha)$ ,  $(b, \alpha)$ ,  $(c, \alpha)$ ,  $(a, \beta)$ ,  $(b, \beta)$  y  $(c, \beta)$  están marcados, mostrando las combinaciones posibles de elementos de  $X$  y  $Y$ .



Si  $A \subset X$  y  $B \subset Y$ , geométricamente  $A \times B$  se representa como sigue:

**Figura A.4** Diagrama que representa el producto cartesiano  $A \times B$ . Los ejes  $X$  e  $Y$  están etiquetados, y el rectángulo sombreado representa el conjunto  $A \times B$ . Las líneas gruesas en los ejes indican los intervalos correspondientes a los conjuntos  $A$  y  $B$ . Las líneas punteadas muestran las proyecciones de los límites del rectángulo sobre los ejes.



## A.9 Familias de conjuntos

Si  $I$  es un conjunto no vacío y cada  $\alpha \in I$  tiene asignado un conjunto  $A_\alpha$ , entonces la colección de conjuntos  $\{A_\alpha \mid \alpha \in I\}$  se llama familia de conjuntos, e  $I$  se llama conjunto de índices para la familia. No se requiere que conjuntos con índices diferentes, sean diferentes. Observemos que cualquier conjunto  $\mathcal{F}$  cuyos elementos son conjuntos, puede ser convertido en una familia de conjuntos, por autoindización, esto es, se usa a  $\mathcal{F}$  mismo como un conjunto de índices, y se asigna a cada miembro de  $\mathcal{F}$  el conjunto que representa.

**Definición A.9.1:**

Sea

$$\mathcal{F} = \{A_\alpha \mid \alpha \in I\}$$

una familia de conjuntos.

- i. Definimos la unión de la familia  $\mathcal{F}$ , misma que denotamos por

$$\cup \mathcal{F} = \cup \{A_\alpha \mid \alpha \in I\}$$

o por  $\bigcup_{\alpha \in I} A_\alpha$ , como el conjunto

$$\bigcup_{\alpha \in I} A_\alpha = \{x \mid x \in A_\alpha, \text{ para algún } \alpha \in I\}$$

- ii. Definimos la intersección de la familia  $\mathcal{F}$ , que se denota por

$$\cap \mathcal{F} = \cap \{A_\alpha \mid \alpha \in I\}$$

o por  $\bigcap_{\alpha \in I} A_\alpha$ , como el conjunto

$$\bigcap_{\alpha \in I} A_\alpha = \{x \mid x \in A_\alpha, \forall \alpha \in I\}.$$

En lugar de  $\bigcup_{\alpha \in I} A_\alpha$ , también se escribe  $\bigcup_{\alpha} A_\alpha$ . Análogamente, en lugar de  $\bigcap_{\alpha \in I} A_\alpha$ , también se escribe  $\bigcap_{\alpha} A_\alpha$ .

**Ejemplo A.9.1:**

- i. Para cualquier conjunto  $X$ ,  $X = \bigcup_{x \in X} \{x\}$ .

- ii. Para cada  $k \in \mathbb{N}$ , sea  $A_k = \{n \mid n \in \mathbb{N} \text{ y } n \geq k\}$ . Nótese que tenemos  $A_1 \supset A_2 \supset A_3 \dots$ . Claramente  $\bigcup_{k \in \mathbb{N}} A_k = A_1$  y  $\bigcap_{k \in \mathbb{N}} A_k = \emptyset$ .

- iii. Para cada  $n \in \mathbb{N}$ , sean

$$A_n = \left[0, 1, -\frac{1}{2^n}\right] \quad \text{y} \quad B_n = \left[0, 1, -\frac{1}{3^n}\right].$$

Obsérvese que  $A_n \subset B_n$ ,  $\forall n \in \mathbb{N}$ , y que  $A_n \neq B_m$ ,  $\forall n \in \mathbb{N}$  y  $\forall m \in \mathbb{N}$ . Sin embargo,

$$\bigcup_{n \in \mathbb{N}} A_n = [0, 1) \quad \text{y} \quad \bigcap_{n \in \mathbb{N}} B_n = [0, 1).$$

**Teorema A.9.1:** Si  $\mathcal{F} = \{A_\alpha \mid \alpha \in I\}$  y  $\mathcal{G} = \{B_\beta \mid \beta \in J\}$  son familias de conjuntos, entonces:

- i.  $\left(\bigcup_{\alpha \in I} A_\alpha\right) \cap \left(\bigcup_{\beta \in J} B_\beta\right) = \bigcup_{(\alpha, \beta) \in I \times J} (A_\alpha \cap B_\beta).$
- ii.  $\left(\bigcap_{\alpha \in I} A_\alpha\right) \cup \left(\bigcap_{\beta \in J} B_\beta\right) = \bigcap_{(\alpha, \beta) \in I \times J} (A_\alpha \cup B_\beta).$
- iii.  $\left(\bigcup_{\alpha \in I} A_\alpha\right) \times \left(\bigcup_{\beta \in J} B_\beta\right) = \bigcup_{(\alpha, \beta) \in I \times J} (A_\alpha \times B_\beta).$
- iv.  $\left(\bigcap_{\alpha \in I} A_\alpha\right) \times \left(\bigcap_{\beta \in J} B_\beta\right) = \bigcap_{(\alpha, \beta) \in I \times J} (A_\alpha \times B_\beta).$

**Demostración:** Se deja como ejercicio al lector.

**Teorema A.9.2:** Si  $\mathcal{F} = \{A_\alpha \mid \alpha \in I\}$  es una familia de subconjuntos de un conjunto universal  $X$ , entonces:

$$\text{i. } \left( \bigcup_{\alpha \in I} A_\alpha \right)^C = \bigcap_{\alpha \in I} A_\alpha^C.$$

$$\text{ii. } \left( \bigcap_{\alpha \in I} A_\alpha \right)^C = \bigcup_{\alpha \in I} A_\alpha^C.$$

**Demostración:** Se deja como ejercicio al lector.

## A.10 Conjunto potencia

**Definición A.10.1:**

Sea  $A$  un conjunto. Se define el conjunto potencia de  $A$ , denotado por  $\wp(A)$  o por  $2^A$ , como el conjunto

$$\wp(A) = \{B \mid B \subset A\}.$$

Observemos que

1. Por definición,  $B \subset A \iff B \in \wp(A)$ .
2. Para cualquier conjunto  $A$ , se cumple que  $\emptyset \in \wp(A)$  y  $A \in \wp(A)$ .
3.  $a \in A \iff \{a\} \in \wp(A)$ .

**Ejemplo A.10.1:**

Si  $A = \{0, 1\}$ , entonces

$$\wp(A) = \{\emptyset, \{0\}, \{1\}, A\}.$$

**Teorema A.10.1:**

Si  $X$  y  $Y$  son conjuntos, entonces

$$X \subset Y \iff \wp(X) \subset \wp(Y).$$

**Demostración:** Procedamos por casos.

- i. Probemos primero que  $X \subset Y \implies \wp(X) \subset \wp(Y)$ : Si  $A \in \wp(X)$ , entonces  $A \subset X$ , y como por hipótesis  $X \subset Y$ , entonces  $A \subset Y$ , entonces  $A \in \wp(Y)$ .
- ii. Ahora probaremos que  $\wp(X) \subset \wp(Y) \implies X \subset Y$ : Si  $x \in X$ , entonces  $\{x\} \in \wp(X)$ , y como por hipótesis  $\wp(X) \subset \wp(Y)$ , entonces  $\{x\} \in \wp(Y)$ , entonces  $x \in Y$ .