
Probabilidades y Estadística

Ramiro Dibur

2025

Índice general

1. Nociones básicas	2
1.1. Modelo probabilístico	2
1.2. Eventos	3
1.3. Eventos aleatorios	4
1.4. Definición de probabilidad	6
1.5. Probabilidad condicional	9
1.6. Probabilidad total	11

Prefacio

Estas son mis notas de Probabilidades y Estadística del segundo cuatrimestre de 2025. Las escribo más que nada para estudiar yo, pero las publico por si le llegan a ser útil a alguien.

Nociones básicas

Formalizamos algunos conceptos de probabilidad que vienen de la intuición.

1.1 Modelo probabilístico

Consideremos un experimento con distintos posibles resultados.

Definición 1.1. El *espacio muestral* de un experimento es el conjunto de posibles resultados del experimento.

Usualmente denotamos un espacio muestral con Ω .

Observación 1.2. Todo resultado corresponde con un único elemento $\omega \in \Omega$.

Veamos algunos ejemplos.

Ejemplo 1.3. Consideremos el siguiente experimento:

1. Se tira un dado balanceado de 6 caras.
2. Se graba el resultado.

En este caso, el espacio muestral es el conjunto

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

Cabe aclarar que no importa de qué manera escribimos los resultados siempre y cuando la correspondencia con el resultado sea clara. Por ejemplo, podríamos haber definido el espacio muestral como

$$\Omega = \{\square, \square, \square, \square, \square, \square\}.$$

Ejemplo 1.4. Consideremos el siguiente experimento:

1. Se tira una moneda 3 veces.
2. Se graba el resultado.

El espacio muestral es

$$\Omega = \{CCC, CCS, \dots, SSC, SSS\}.$$

Nótese que Ω se puede escribir como $\{C, S\}^3$.

Ejemplo 1.5. Consideremos el siguiente experimento:

1. Se elije un habitante de Buenos Aires al azar.
2. Se mide su altura en metros.

El espacio muestral podría ser

$$\Omega = \mathbb{R}.$$

Uno podría argumentar que el espacio muestral debería ser

$$\Omega = [0, 3],$$

ya que es imposible que alguien mida -1 m o 100 m. Sin embargo, lo único que nos interesa es que, al medir a alguien, caiga dentro de Ω .

1.2 Eventos

Definición 1.6. Sea Ω un espacio muestral. Un *evento* es un subconjunto de Ω .

Veámoslo en algunos ejemplos.

Ejemplo 1.7. Consideramos el experimento del ejemplo 1.3. El conjunto

$$A = \{\text{el resultado es un número par}\} = \{2, 4, 6\}$$

es un evento dado que $A \subseteq \Omega$.

Por ahora, usemos la noción intuitiva de probabilidades.

La probabilidad se le asigna a un evento, no a un resultado. Por ejemplo, cuando decimos

$$P(\{\varpi\}) = \frac{1}{6}$$

en realidad queremos decir

$$P(\{\varpi\}) = \frac{1}{6}.$$

No obstante, por practicidad acudiremos a la primera notación.

Usualmente calculamos la probabilidad de un evento de la siguiente manera:

$$P(A) = \frac{\# \text{ casos donde sucede } A}{\# \text{ casos totales}}.$$

Veamos por qué esto no es generalizable.

Ejemplo 1.8. Consideremos el siguiente experimento:

1. Se tiran 2 dados balanceados de 6 caras.
2. Se suman los números de las caras.
3. Se graba el resultado.

Un espacio muestral podría ser

$$\Omega = \{2, 3, \dots, 12\}.$$

Sin embargo, $P(2) \neq \frac{1}{10}$. Esto se puede resolver tomando el espacio muestral

$$\Omega = \{\square\square, \square\square, \dots, \text{12}\text{12}, \text{12}\text{12}\}.$$

Por lo tanto, para todo resultado $\omega \in \Omega$,

$$P(\omega) = \frac{\# \text{ casos donde sucede } A}{\# \text{ casos totales}} = \frac{1}{36}.$$

A partir de este ejemplo surge una definición.

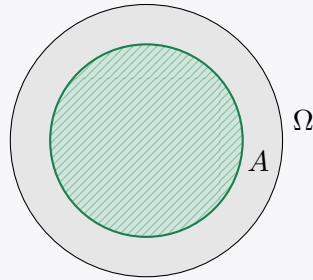
Definición 1.9. Sea Ω un espacio muestral. Si Ω es finito y todos sus elementos tienen la misma probabilidad, decimos que Ω es un *espacio muestral equiprobable*.

Sin embargo, hasta ahora únicamente tratamos con espacios muestrales finitos. Veamos qué pasa con los infinitos.

Ejemplo 1.10. Consideremos el siguiente experimento:

1. Se elije un punto del disco unitario $D = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ al azar.
2. Se graba el resultado.

Sea $A = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq \frac{1}{2}\}$.



Por lo que

$$P(A) = \frac{\text{área}(A)}{\text{área}(\Omega)}.$$

1.3 Eventos aleatorios

No podemos definirle una probabilidad a todos los eventos de un espacio muestral.

Definición 1.11. Un evento al cual le podemos definir una probabilidad es llamado un *evento aleatorios*.

Agregamos algunas reglas adicionales.

Definición 1.12. Llamamos \mathcal{F} a una familia de eventos a los cuales podemos calcularles su probabilidad si cumple los siguientes axiomas:

(F1) $\Omega \in \mathcal{F}$.

(F2) Si $A \in \mathcal{F}$, entonces $A^c \in \mathcal{F}$.

(F3) Si $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{F}$, entonces $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{F}$.

Observación 1.13. Si una familia cumple los axiomas (F1), (F2) y (F3), entonces se llama una σ -álgebra de conjuntos.

Ejemplo 1.14. En el ejemplo 1.3, el espacio muestral

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

tiene a la familia de eventos $\mathcal{F} = \mathcal{P}(\Omega)$ que se les puede asignar una probabilidad.

Veamos algunas propiedades que podemos deducir.

Proposición 1.15. Sea Ω un espacio muestral y $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ una σ -álgebra. Las siguientes proposiciones son verdaderas:

1. $\emptyset \in \mathcal{F}$.
2. Si $\{A_n\}_{1 \leq n \leq N} \subseteq \mathcal{F}$, entonces $\bigcup_{1 \leq n \leq N} A_n \in \mathcal{F}$.
3. Si $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{F}$, entonces $\bigcap_{n \in \mathbb{N}} A_n \in \mathcal{F}$. (También la versión finita.)
4. Si $A, B \in \mathcal{F}$, entonces $A \setminus B \in \mathcal{F}$.

Demostración. (1.) Dado que $\Omega \in \mathcal{F}$ por (F1), obtenemos que $\emptyset = \Omega^c \in \mathcal{F}$ por (F2).

(2.) Sea $\{A_n\}_{1 \leq n \leq N} \subseteq \mathcal{F}$. Consideremos la familia $\{B_n\}_{n \in \mathbb{N}}$ tal que

$$B_n = \begin{cases} A_n & \text{si } 1 \leq n \leq N, \\ \emptyset & \text{si } n \geq N. \end{cases}$$

Entonces, por (F3),

$$\bigcup_{1 \leq n \leq N} A_n = \bigcup_{n \in \mathbb{N}} B_n \in \mathcal{F}.$$

(3.) Sea $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{F}$. Por (F3),

$$\bigcap_{n \in \mathbb{N}} A_n = \left(\bigcup_{n \in \mathbb{N}} A_n^c \right)^c \in \mathcal{F}.$$

(4.) Sean $A, B \in \mathcal{F}$. Dado que $B^c \in \mathcal{F}$,

$$A \setminus B = A \cap B^c \in \mathcal{F}.$$

□

Ejemplo 1.16. Consideremos el siguiente experimento:

1. Se elije un número real del intervalo $[0, 1]$ al azar.
2. Se graba el resultado.

Sea $\Omega = [0, 1]$ el espacio muestral y sea \mathcal{F} la familia de eventos a los cuales les podemos asignar una probabilidad. Para un intervalo $[a, b]$ la probabilidad se puede calcular como

$$P([a, b]) = b - a.$$

Aplicando las propiedades de la proposición 1.15, podemos deducir que en \mathcal{F} están los eventos:

- Los intervalos abiertos y cerrados.
- Uniones e intersecciones numerables de cerrados y/o abiertos.
- Los puntos $\{x\}$ con $x \in [0, 1]$.
- Los números racionales \mathbb{Q} .

¿Cuál es la probabilidad de \mathbb{Q} ? Basta con tomar $\{B_m\} = \{B(q_m, \frac{\varepsilon}{2^{m+1}})\}_{m \in \mathbb{N}}$ y ver que

$$\begin{aligned} P\left(\bigcup_{m \in \mathbb{N}} B_m\right) &\leq \sum_{m \in \mathbb{N}} P(B_m) \\ &\leq \sum_{m \in \mathbb{N}} \frac{\varepsilon}{2^m} \\ &\leq \varepsilon. \end{aligned}$$

Tomando $\varepsilon \rightarrow 0$, obtenemos que $P(\mathbb{Q}) = 0$.

1.4 Definición de probabilidad

La *idea de Laplace* de probabilidad consta en que la probabilidad de un evento es el límite de la frecuencia con la que sucede cuando la cantidad de ensayos tiende a infinito.

Por otro lado, está la *axiomatización de Kolmogorov*:

Definición 1.17. Sea Ω un espacio muestra y $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ una σ -álgebra. Una función probabilidad es una función $P : \mathcal{F} \rightarrow [0, 1]$ que cumple los siguientes axiomas:

- (P1) $P(\Omega) = 1$.
- (P2) $P(A) \geq 0$ para todo $A \in \mathcal{F}$.
- (P3) Si $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{F}$ una familia de eventos disjuntos, entonces $P(\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n=1}^{\infty} P(A_n)$.

Con esto podemos definir un espacio de probabilidad.

Definición 1.18. Sea Ω un espacio muestral, $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ una familia de eventos y $P : \mathcal{F} \rightarrow [0, 1]$ una función probabilidad. Entonces, la terna (Ω, \mathcal{F}, P) es un **espacio de probabilidad**.

Probamos algunos resultados inmediatos.

Proposición 1.19. Sea (Ω, \mathcal{F}, P) un espacio de probabilidad. Entonces, las siguientes proposiciones son verdaderas:

1. $P(\emptyset) = 0$.
2. Si A y B son eventos disjuntos, entonces $P(A \cup B) = P(A) + P(B)$.
3. Si $\{A_n\}_{1 \leq n \leq N}$ es una familia de eventos disjuntos, entonces $P(\bigcup_{1 \leq n \leq N} A_n) = \sum_{n=1}^N P(A_n)$.
4. $P(A^c) = 1 - P(A)$.
5. Si $A \subseteq B$, entonces $P(A) \leq P(B)$.
6. $P(\bigcup_{n \in \mathbb{N}} A_n) \leq \sum_{n=1}^{\infty} P(A_n)$.

Demostración. (1.) Consideremos la familia $\{\Omega, \emptyset, \dots\}$. Por (P3),

$$\begin{aligned} P(\Omega) &= P(\Omega \cup \emptyset \cup \dots) \\ &= P(\Omega) + \underbrace{P(\emptyset) + \dots}_{=0} \end{aligned}$$

(2.) La propiedad sale utilizando (P3) y tomando la familia $\{A, B, \emptyset, \dots\}$.

(3.) Se prueba por inducción y usando la proposición anterior.

(4.) Podemos escribir como unión disjunta $\Omega = A \cup A^c$. Y con la propiedad 2. obtenemos que

$$1 = P(\Omega) = P(A) + P(A^c),$$

entonces

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

(5.) Como $B = (B \setminus A) \cup A$,

$$P(A) \leq P(B \setminus A) + P(A) = P(B).$$

(6.) Sea $\{A_n\}_{n \in \mathbb{N}}$ una familia de eventos. Consideremos $\{B_n\}_{n \in \mathbb{N}}$ tal que

$$B_n = A_n \setminus \bigcup_{1 \leq k \leq n-1} A_k.$$

Como los eventos B_n son disjuntos dos a dos,

$$P\left(\bigcup_{n \in \mathbb{N}} A_n\right) = P\left(\bigcup_{n \in \mathbb{N}} B_n\right) = \sum_{n=1}^{\infty} P(B_n) \leq \sum_{n=1}^{\infty} P(A_n).$$

□

Ejemplo 1.20. Consideremos el siguiente experimento:

1. Se tira una moneda balanceada hasta que salga cara.
2. Se anota la cantidad total de lanzamientos.

El espacio muestral es

$$\Omega = \mathbb{N},$$

con

$$P(1) = \frac{1}{2}, \quad P(i) = \frac{1}{2^i}, \quad i \geq 1.$$

Ahora pensémoslo con infinitas tiradas de moneda.

Ejemplo 1.21. Consideremos el experimento del ejemplo 1.20. Entonces, el espacio muestral es

$$\Omega = \{0, 1\}^{\mathbb{N}}.$$

Sea $C_k = \{\text{la } k\text{-ésima moneda es cara}\}$ y definimos

$$A_n = \{\text{las primeras } n \text{ monedas son cara}\} = \bigcap_{k=1}^n C_k.$$

Entonces, $P(A_n) = \frac{1}{2^n}$. Consideremos

$$A = \{\text{sale siempre cara}\} = \bigcap_{n \in \mathbb{N}} A_n.$$

Por lo tanto $A \subseteq A_n$ para todo $n \in \mathbb{N}$. Así, $P(A) \leq P(A_n) = \frac{1}{2^n} \rightarrow 0$ y $P(A) = 0$.

Veamos otro evento:

$$B = \{\text{a partir de algún momento salen sólo caras}\}.$$

Es decir, existe $n_0 \in \mathbb{N}$ tal que para todo $k \geq n_0$ la moneda k es cara. En notación,

$$B = \bigcup_{n_0 \in \mathbb{N}} \bigcap_{k=n_0}^{\infty} C_k.$$

Para cada n_0 fijo,

$$\begin{aligned} P\left(\bigcap_{k=n_0}^{\infty} C_k\right) &= \lim_{m \rightarrow \infty} P\left(\bigcap_{k=n_0}^m C_k\right) \\ &= \lim_{m \rightarrow \infty} \frac{1}{2^{m-n_0+1}} \\ &= 0. \end{aligned}$$

Entonces, por subaditividad numerable,

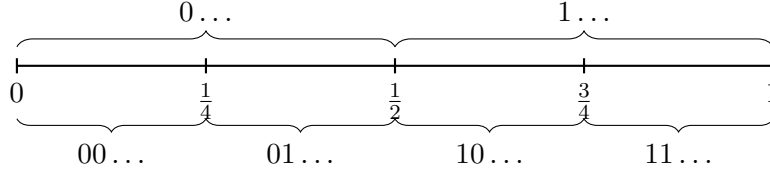
$$P(B) \leq \sum_{n_0 \in \mathbb{N}} 0 = 0.$$

Podemos mirar a $\omega = (\omega_k)_{k \in \mathbb{N}} \in \{0, 1\}^{\mathbb{N}}$ como la expansión binaria de un número real

$x \in [0, 1]$:

$$x = \sum_{k=1}^{\infty} \frac{\omega_k}{2^k}.$$

De esta forma, elegir una secuencia infinita de lanzamientos es equivalente a elegir un número real en $[0, 1]$.



Proposición 1.22. Sea $\{A_n\}_{n \in \mathbb{N}}$ una familia de eventos con $A_n \subseteq A_{n+1}$ para todo n . Entonces

$$P\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \lim_{n \rightarrow \infty} P(A_n).$$

Demostración. Definimos $B_n = A_n \setminus A_{n-1}$ (con $A_0 = \emptyset$). Entonces

$$\bigcup_{n \in \mathbb{N}} B_n = \bigcup_{n \in \mathbb{N}} A_n \quad \text{y} \quad B_i \cap B_j = \emptyset \text{ si } i \neq j.$$

Como $P(B_n) = P(A_n) - P(A_{n-1})$, tenemos

$$\begin{aligned} P\left(\bigcup_{n \in \mathbb{N}} A_n\right) &= \sum_{n=1}^{\infty} P(B_n) \\ &= \lim_{N \rightarrow \infty} \sum_{n=1}^N (P(A_n) - P(A_{n-1})) \\ &= \lim_{N \rightarrow \infty} (P(A_N) - P(A_0)). \end{aligned}$$

Como $A_0 = \emptyset$, queda

$$P\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \lim_{N \rightarrow \infty} P(A_N).$$

□

Corolario 1.23. Tomando complementos se obtiene que, si $A_n \supseteq A_{n+1}$, entonces

$$P\left(\bigcap_{n \in \mathbb{N}} A_n\right) = 1 - P\left(\bigcup_{n \in \mathbb{N}} A_n^c\right) = 1 - \lim_{n \rightarrow \infty} P(A_n^c) = \lim_{n \rightarrow \infty} P(A_n).$$

1.5 Probabilidad condicional

Ya vimos cómo calcular probabilidades de eventos. Ahora queremos formalizar qué significa calcular probabilidades *condicionadas* a que cierto evento ya ocurrió.

Ejemplo 1.24. Tiro dos dados y sumo los resultados. Consideremos los eventos

$$A = \{\text{la suma es } 5\},$$

$$B = \{\text{el segundo dado es par}\}.$$

El espacio muestral es

$$\Omega = \{(a, b) \mid a, b \in \{1, \dots, 6\}\}.$$

En particular,

$$A = \{(1, 4), (4, 1), (2, 3), (3, 2)\}.$$

Luego,

$$P(A \mid B) = \frac{\#(A \cap B)}{\#B} = \frac{2}{18} = \frac{1}{9}.$$

El cálculo anterior nos motiva a introducir la definición general de probabilidad condicional.

Definición 1.25. Dados A y B eventos, con $P(B) \neq 0$, la *probabilidad condicional* de A dado B se define como

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

Veamos cómo aplica la definición en un ejemplo.

Ejemplo 1.26. Consideremos una urna con 9 bolitas de la siguiente forma:

- 5 naranjas,
- 4 violetas, de las cuales 3 tienen una cruz.

Denotemos $N = \{\text{naranja}\}$, $V = \{\text{violeta}\}$, $C = \{\text{tiene cruz}\}$. Entonces,

$$P(N) = \frac{5}{9}, \quad P(N \mid V) = \frac{1}{3}, \quad P(C \mid V) = \frac{1}{2}.$$

A partir de la definición, podemos verificar fácilmente que $P(\cdot \mid B)$ cumple los axiomas de probabilidad.

Proposición 1.27. Sea B un evento con $P(B) \neq 0$. Entonces la aplicación

$$A \mapsto P(A \mid B)$$

define una probabilidad sobre el espacio muestral condicionado.

Demostración. Para todo B con $P(B) > 0$:

$$(P1) \quad P(\Omega \mid B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1.$$

$$(P2) \quad \text{Como } P \text{ es una probabilidad, } P(A \mid B) \geq 0.$$

(P3) Si $\{A_n\}_{n \in \mathbb{N}}$ son eventos disjuntos dos a dos,

$$\begin{aligned} P\left(\bigcup_{n \in \mathbb{N}} A_n \mid B\right) &= \frac{P(\bigcup_{n \in \mathbb{N}} (A_n \cap B))}{P(B)} \\ &= \frac{\sum_{n \in \mathbb{N}} P(A_n \cap B)}{P(B)} \\ &= \sum_{n \in \mathbb{N}} \frac{P(A_n \cap B)}{P(B)} \\ &= \sum_{n \in \mathbb{N}} P(A_n \mid B). \end{aligned}$$

□

Observación 1.28. Es importante aclarar que $A \mid B$ no es un evento. Por lo tanto, expresiones como $P(A \mid B \mid C)$ no tienen sentido.

Proposición 1.29. Para todo par de eventos A, B con $P(B) \neq 0$, se cumple

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

Este resultado es útil para calcular probabilidades conjuntas a partir de probabilidades condicionales.

Ejemplo 1.30. Sacamos dos cartas de un mazo de 52 cartas sin reemplazo. Sea

$$\begin{aligned} A_1 &= \{\text{la primera carta es trébol}\}, \\ A_2 &= \{\text{la segunda carta es trébol}\}. \end{aligned}$$

Entonces,

$$\begin{aligned} P(A_1) &= \frac{13}{52}, \\ P(A_2 \mid A_1) &= \frac{12}{51}. \end{aligned}$$

Aplicando la proposición anterior,

$$P(A_1 \cap A_2) = P(A_1) P(A_2 \mid A_1) = \frac{13}{52} \cdot \frac{12}{51}.$$

Proposición 1.31 (Regla del producto). Para eventos A_1, A_2, \dots, A_n se cumple

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) P(A_2 \mid A_1) P(A_3 \mid A_1 \cap A_2) \dots P(A_n \mid A_1 \cap \dots \cap A_{n-1}).$$

Idea de demostración. La demostración se hace fácilmente por inducción en n . □

1.6 Probabilidad total

La noción de probabilidad condicional nos permite descomponer probabilidades en términos de eventos disjuntos.

Proposición 1.32 (Probabilidad total en dos partes). Sea $B \subseteq \Omega$ un evento. Entonces, para todo $A \subseteq \Omega$,

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap B^c) \\ &= P(A | B) P(B) + P(A | B^c) P(B^c). \end{aligned}$$

Más en general, si $\{A_i\}_{i \in I}$ es una partición numerable de Ω , se cumple la *ley de la probabilidad total*:

$$P(B) = \sum_{i \in I} P(B | A_i) P(A_i).$$

Ejemplo 1.33. Consideremos el siguiente experimento:

1. En una caja hay 3 monedas: una con dos caras y dos equilibradas.
2. Se elige una moneda al azar.
3. Se lanza la moneda.

Definamos los eventos

$$\begin{aligned} C &= \{\text{sale cara}\}, \\ A &= \{\text{la moneda elegida es la de dos caras}\}. \end{aligned}$$

Entonces,

$$\begin{aligned} P(C) &= P(C | A) P(A) + P(C | A^c) P(A^c) \\ &= 1 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{2}{3} \\ &= \frac{2}{3}. \end{aligned}$$

Este cálculo es un primer ejemplo de la *fórmula de la probabilidad total*, que nos permite descomponer la probabilidad de un evento en función de una partición del espacio.

Ejemplo 1.34. Ahora consideremos otro experimento. Lanzamos un dado. Sea

$$\begin{aligned} A &= \{\text{saco un 5}\}, \\ B &= \{\text{saco un 7}\}, \\ C &= \{\text{ni 5 ni 7}\} = (A \cup B)^c. \end{aligned}$$

Supongamos que ganar depende de estos eventos: si ocurre A se gana siempre, si ocurre B nunca se gana, y si ocurre C se gana con cierta probabilidad.

Denotemos $G = \{\text{ganar}\}$. Entonces, por probabilidad total,

$$P(G) = P(G | A) P(A) + P(G | B) P(B) + P(G | C) P(C).$$

Como $P(G | A) = 1$, $P(G | B) = 0$, y $P(A) = P(B) = \frac{1}{6}$, $P(C) = \frac{4}{6}$, se obtiene

$$P(G) = 1 \cdot \frac{1}{6} + 0 \cdot \frac{1}{6} + P(G | C) \frac{4}{6}.$$

Si además sabemos que $P(G | C) = \frac{1}{2}$, queda

$$P(G) = \frac{1}{6} + \frac{1}{2} \cdot \frac{4}{6} = \frac{6}{15}.$$

Estos ejemplos muestran cómo las probabilidades condicionales y la probabilidad total se combinan para calcular probabilidades en situaciones más complejas.