Capítulo 3

Probabilidad Condicional e Independencia

3.1. Introducción.

Consideremos una población de 20 estudiantes entre los cuales hay 14 que estudian Medicina y 6 que estudian Ingeniería. De esta población se escogen sin reposición dos estudiantes al azar y se consideran los eventos:

E_1: "El primer estudiante seleccionado estudia Medicina".

E_2: "El segundo estudiante seleccionado estudia Medicina".

El espacio muestral que consideramos consiste de la colección de todos los pares ordenados

$$(a_i, a_j); \quad (a_i, b_k); \quad (b_k, a_i); \quad (b_k, b_h)$$

donde los a_i son estudiantes de Medicina y los b_j son de Ingeniería, $i \neq j; k \neq h$; $i, j \leq 14; h, k \leq 6$. El número de eventos elementales es 20×19.

La siguiente tabla de doble entrada indica el número de puntos muestrales correspondientes a la partición de Ω según los eventos E_1, E_2 y sus complementos. En la última fila aparecen los totales correspondientes a cada columna y en la última columna los correspondientes a cada fila.

	E_2	E_2^c	
E_1	14×13	14×6	14×19
E_1^c	6×14	6×5	6×19
	14×19	6×19	20×19

Tabla 3.1

Utilizando este cuadro podemos calcular fácilmente las probabilidades de eventos tales como

$$P(E_1 \cap E_2) = \frac{14 \times 13}{20 \times 19}; \qquad P(E_1) = \frac{14 \times 19}{20 \times 19}; \qquad P(E_1^c \cap E_2) = \frac{6 \times 14}{20 \times 19}.$$

Si suponemos que el primer estudiante estudia Medicina, ¿cuál es la probabilidad de que el segundo también?

En este caso vemos, a partir de la tabla, que hay 14×19 resultados posibles, de los cuales 14×13 son favorables al evento E_2 y por lo tanto la probabilidad que deseamos calcular es

$$\frac{14 \times 13}{14 \times 19} = \frac{(14 \times 13)/(20 \times 19)}{(14 \times 19)/(20 \times 19)} = \frac{P(E_1 \cap E_2)}{P(E_1)}.$$

La probabilidad que hemos calculado se llama "*probabilidad condicional de* E_2 *dado* E_1" y se denota $P(E_2|E_1)$.

Observamos que $P(E_2) = \frac{14\times 19}{20\times 19} = \frac{7}{10}$ no coincide con $P(E_2|E_1) = \frac{13}{19}$. Al saber que ocurrió E_1 disponemos de cierta información adicional que modifica nuestro espacio muestral: la nueva población, para la segunda extracción, no coincide con la original, ya que sólo quedan 13 estudiantes de Medicina de un total de 19 estudiantes posibles.

Notemos además que si las extracciones se realizan con reposición esto no ocurre, ya que el resultado de la primera extracción no nos da ninguna información sobre la segunda. En este caso se tiene:

$$P(E_2|E_1) = P(E_2) = \frac{7}{10}.$$

Definición 3.1 Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad y sea $B \in \mathcal{A}$ tal que $P(B) > 0$. Definimos una nueva función $P(\cdot|B)$ de $\mathcal{A}$ en $\mathbb{R}$ mediante

$$P(A|B) = \frac{P(A\cap B)}{P(B)} \qquad \text{para todo } A \in \mathcal{A}.$$

Esta función $P(\cdot|B)$ que hemos definido es una probabilidad; en efecto:

1. $P(A|B) \geq 0$ para todo $A \in \mathcal{A}$.

2. $P(\Omega|B) = \frac{P(\Omega\cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$.

3. Sean $A_1, A_2, \dots$ conjuntos disjuntos en $\mathcal{A}$, entonces

$$\begin{aligned} P(\bigcup_{i=1}^{\infty} A_i|B) &= \frac{P(B\cap \bigcup_{i=1}^{\infty} A_i)}{P(B)} = \frac{P(\bigcup_{i=1}^{\infty} B\cap A_i)}{P(B)} \\ &= \frac{\sum_{i=1}^{\infty} P(B\cap A_i)}{P(B)} = \sum_{i=1}^{\infty} P(A_i|B). \end{aligned}$$

$P(\cdot|B)$ se llama probabilidad condicional dado B.

Dos propiedades elementales de la probabilidad condicional son las siguientes:

- Si A y B son disjuntos, entonces $P(A|B) = 0$. En efecto,

$$A\cap B = \emptyset \Rightarrow P(A\cap B) = 0 \quad \text{y} \quad P(A|B) = \frac{P(A\cap B)}{P(B)} = 0.$$

- Si $B \subset A$ entonces $P(A|B) = 1$ ya que en este caso $P(A\cap B) = P(B)$.

Ejemplos

1. Se lanzan dos dados hasta que la suma de los puntos sea 7 u 8. Si sale 7 gana el jugador A, si sale 8 gana B. ¿Cuál es la probabilidad de que gane A?

► Vamos a resolver este problema de dos maneras. Supongamos primero que al cabo de n lanzamientos gana A. Esto quiere decir que en los $n-1$ primeros lanzamientos no salió ni 7 ni 8, y que en el n-ésimo se obtuvo 7. Este evento tiene probabilidad

$$p_n = \frac{1}{6}\left(\frac{25}{36}\right)^{n-1}.$$

De esta manera, la probabilidad de que A gane es:

$$\begin{aligned}P(A) = \sum_{n=1}^{\infty} p_n &= \sum_{n=1}^{\infty} \frac{1}{6}\left(\frac{25}{36}\right)^{n-1} \\ &= \frac{1}{6}\sum_{n=0}^{\infty}\left(\frac{25}{36}\right)^{n} \\ &= \frac{1}{6}\,\frac{1}{1-\frac{25}{36}} = \frac{6}{11}.\end{aligned}$$

Otra forma de resolver el problema es observar que la probabilidad de que gane A es la probabilidad condicional de que la suma de puntos sea 7, dado que el juego termina en un cierto lanzamiento, es decir, dado que la suma es 7 u 8 (hay que observar que la probabilidad de que el juego no termine es cero) o sea

$$P(A) = P(\{7\}|\{7,8\}) = \frac{P(\{7\})}{P(\{7,8\})} = \frac{\frac{6}{36}}{\frac{11}{36}} = \frac{6}{11}.$$

▲

2. Se lanza un dado dos veces.

a) Si la suma de los resultados es 8, ¿cuál es la probabilidad de que el primer lanzamiento haya resultado en $k,\ 1 \le k \le 6$?

b) Si el primer lanzamiento resultó en 3, ¿cuál es la probabilidad de que el segundo sea $k,\ 1 \le k \le 6$?

c) Si el primer lanzamiento resultó en 3, ¿cuál es la probabilidad de que la suma de ambos lanzamientos sea 7?

► Sea X el resultado del primer lanzamiento e Y el del segundo. Sabemos que $P(X = k) = P(Y = k) = 1/6,\ 1 \le k \le 6$.

a) Queremos calcular

$$P(X = k|X + Y = 8) = \frac{P((X = k) \cap (X + Y = 8))}{P(X + Y = 8)}.$$

Veamos primero la probabilidad en el denominador. Observamos que hay 5 resultados cuya suma es ocho de un total de 36 resultados posibles, que corresponden a los pares ordenados $(2,6)$; $(3,5)$; $(4,4)$; $(5,3)$; $(6,2)$ y por lo tanto la probabilidad en el denominador vale $5/36$. Por otro lado, la probabilidad en el numerador vale 0 si tomamos $k = 1$. Para $2 \le k \le 6$ hay un solo resultado del segundo lanzamiento para el cual la suma es ocho: $Y = 8 - k$ y en consecuencia la probabilidad en el numerador es $1/36$. Finalmente tenemos

$$P(X = k|X + Y = 8) = \begin{cases} 1/5 & \text{si } 2 \le k \le 6, \\ 0 & \text{si } k = 1. \end{cases}$$

b) Ahora nos interesa calcular

$$P(Y = k|X = 3) = \frac{P((Y = k) \cap (X = 3))}{P(X = 3)}.$$

Sabemos que $P(X = 3) = 1/6$. Para calcular la probabilidad del numerador observamos que de un total de 36 resultados posibles, sólo uno corresponde al evento $(Y = k) \cap (X = 3)$ y por lo tanto esta probabilidad es $1/36$. En consecuencia

$$P(Y = k|X = 3) = \frac{1/36}{1/6} = \frac{1}{6}.$$

Este resultado es igual a $P(Y = k)$ y concuerda con lo que uno esperaría intuitivamente, ya que saber el resultado del primer lanzamiento no debe afectar en manera alguna el resultado del segundo.

c) Nos interesa ahora

$$P(X + Y = 7|X = 3) = \frac{P((Y + Y = 7) \cap (X = 3))}{P(X = 3)}$$

pero

$$(X + Y = 7) \cap (X = 3) = (3 + Y = 7) \cap (X = 3) = (Y = 4) \cap (X = 3),$$

por lo tanto

$$P(X + Y = 7|X = 3) = \frac{P((Y = 4) \cap (X = 3))}{P(X = 3)}$$

y por el resultado de la parte b de este ejercicio sabemos que esta probabilidad es $1/6$. ▲

3. Consideremos ahora una situación que se presenta con frecuencia en casos de controles masivos aplicados en prevención médica y control de calidad de productos.

   Para controlar una cierta enfermedad en una población donde la proporción de enfermos es p se usa un determinado examen médico para detectar enfermos. Se sabe que la probabilidad de que al aplicar el examen a un enfermo lo muestre como tal es de 0.90, y que la probabilidad de que el examen aplicado a una persona sana la señale como enferma es 0.01. Calcular la probabilidad de que una persona esté realmente enferma si el examen médico lo mostró como tal.

► Para responder esta pregunta elegimos al azar una persona en la población y consideramos los eventos

E: "la persona está enferma".

R: "el examen la detecta como enferma".

Queremos calcular

$$P(E|R) = \frac{P(E \cap R)}{P(R)}.$$

Sabemos que

$$\begin{aligned} P(E) &= p, \\ P(R|E) &= \frac{P(E \cap R)}{P(E)} = 0.90, \\ P(R|E^c) &= \frac{P(E^c \cap R)}{P(E^c)} = 0.01. \end{aligned}$$

A partir de las dos primeras igualdades obtenemos

$$P(E \cap R) = 0.90p$$

y a partir de la tercera

$$P(E^c \cap R) = 0.01(1 - p).$$

Por lo tanto

$$\begin{aligned} P(R) &= P(R \cap E) + P(R \cap E^c) \\ &= 0.90p + 0.01(1 - p), \end{aligned}$$

y en consecuencia

$$P(E|R) = \frac{0.90p}{0.90p + 0.01(1 - p)} = \frac{90p}{89p + 1}.$$

En particular, si $p = \frac{1}{30}$ resulta

$$P(E|R) \simeq 0.76.$$

Llama la atención que esta probabilidad no sea muy próxima a 1, como uno esperaría. Analizando el comportamiento de $P(E|R)$ como función de p (ver figura 3.1) observamos que si la proporción p de enfermos en la población es pequeña, el método de control masivo es insuficiente, dado que $P(E|R)$ está lejos de 1. Por ejemplo, si $p = 0.001$, resulta $P(E|R) \simeq 0.083$.

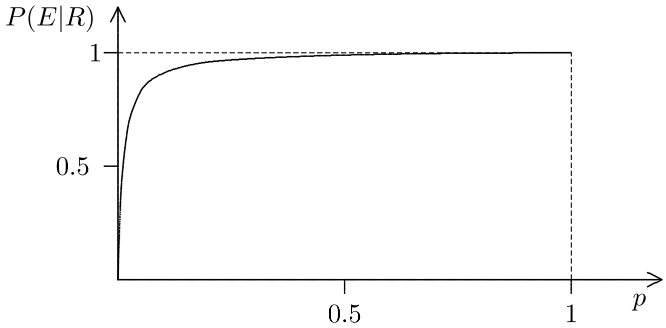

Figura 3.1

Los datos de este ejemplo han sido elegidos arbitrariamente, pero observamos, sin embargo, que cualquier método de control que pueda implicar errores en la clasificación considerada, está sujeto a dificultades como la que hemos encontrado anteriormente, que requieren un análisis previo a su aplicación. ▲

4. Estudiemos la distribución del tiempo de vida de un aparato eléctrico bajo la hipótesis de que el aparato no envejece, sino que se destruye por una perturbación aleatoria. Lo que queremos decir es que el objeto en cuestión cumple con la siguiente condición: dado que el aparato está en funcionamiento en el instante s, se comporta respecto a su tiempo de vida futuro a partir de ese instante como si hubiera comenzado a funcionar en s, es decir que

$$P(T > s + h | T > s) = P(T > h), \qquad h \geq 0,$$

o sea

$$\frac{P(T > s + h)}{P(T > s)} = P(T > h).$$

Si llamamos $\varphi(t) = P(T > t)$, la igualdad anterior se puede escribir

$$\varphi(s + h) = \varphi(s)\varphi(h) \tag{3.1}$$

para cualquier par de números positivos s y h. Si φ no es identicamente nula veamos que

$$\varphi(t) = e^{-\lambda t} \qquad \text{donde } t \geq 0, \lambda \geq 0.$$

► De (3.1), si $s = h = 0$, resulta

$$\varphi(0) = (\varphi(0))^2$$

de donde $\varphi(0) = 0$ ó $\varphi(0) = 1$. Ahora bien,

$$\varphi(0) = 0 \Rightarrow \varphi(s) = \varphi(0)\varphi(s) = 0, \text{ para todo } s \geq 0$$

y tenemos una solución trivial.

Consideremos ahora el caso $\varphi(0) = 1$. A partir de (3.1) deducimos

$$\varphi(ns) = \varphi(s + s + \cdots + s) = (\varphi(s))^n, \text{ para todo } s \geq 0.$$

Si $s = 1/n$ resulta

$$\varphi(1/n) = (\varphi(1))^{\frac{1}{n}}$$

poniendo ahora $s = n/m$, se tiene

$$\varphi(n/m) = (\varphi(1/m))^n = (\varphi(1))^{\frac{n}{m}},$$

es decir que

$$\varphi(t) = (\varphi(1))^t \qquad \text{si } t \in \mathbb{Q},\ t \geq 0.$$

Falta mostrar que la relación anterior es cierta para cualquier número real positivo t. Sean (t_k), (t'_k) dos sucesiones de racionales positivos tales que

$$t_k < t < t'_k, \quad t_k \to t, \quad t'_k \to t,$$

entonces, dado que la función φ es monótona decreciente, resulta

$$(\varphi(1))^{t_k} = \varphi(t_k) \leq \varphi(t) \leq \varphi(t'_k) = (\varphi(1))^{t'_k}$$

haciendo $k \to \infty$ tenemos

$$(\varphi(1))^t \leq \varphi(t) \leq (\varphi(1))^t$$

y por lo tanto $\varphi(t) = (\varphi(1))^t$. Poniendo $\varphi(1) = e^{-\lambda}$ obtenemos

$$\varphi(t) = e^{-\lambda t}$$

donde λ debe ser positivo para que φ sea decreciente. Concluimos entonces que $P(T > t) = e^{-\lambda t}$ y decimos que en este caso que T tiene distribución exponencial.

Como ejemplo numérico de esta situación supongamos que el tiempo de vida T medido en horas de una cierta marca de focos tiene una distribución exponencial con parámetro $\lambda = 0.001$.

a) Calcule la probabilidad de que un foco de esta marca dure 2000 h.

b) Calcule la probabilidad de que dure 2000 h. dado que ha durado 1000 h.

a) $P(T \geq 2000) = e^{-0.001 \times 2000} = e^{-2} \approx 0.13533$.

b) $P(T \geq 2000 | T \geq 1000) = \frac{P((T \geq 2000) \cap (T \geq 1000))}{P(T \geq 1000)} = \frac{P(T \geq 2000)}{P(T \geq 1000)} = \frac{e^{-2}}{e^{-1}} = e^{-1} \approx 0.367879$.

▲

5. Consideremos las familias que tienen dos niños y supongamos que varones y hembras son igualmente probables. Si escogemos al azar una familia y en ella hay un hijo varón, ¿cuál es la probabilidad de que el otro hijo sea también varón?

► En este problema el espacio muestral es

$$\Omega = \{(v, v); (v, h); (h, v); (h, h)\}$$

donde v es varón, h es hembra y el orden del par es el orden de los nacimientos. Cada uno de los puntos tiene probabilidad $1/4$.

Al considerar este problema uno podría llegar a la conclusión de que, como ambos sexos son igualmente probables, la probabilidad que buscamos es $1/2$. Veremos, sin embargo, que esto no es cierto. Sea

$$A : \text{“uno de los hijos es varón”}$$
$$B : \text{“ambos hijos son varones”}$$

entonces $B \subset A$ y $A \cap B = B$. Por lo tanto

$$P(B|A) = \frac{P(B)}{P(A)} = \frac{1/4}{3/4} = \frac{1}{3},$$

que es la respuesta correcta al problema. ▲

6. Veamos un problema parecido al anterior. Consideremos de nuevo las familias que tienen dos hijos. Si escogemos un niño al azar entre estas familias y resulta ser varón, ¿cuál es la probabilidad de que el otro niño en la familia también sea varón?

▶ En este ejercicio, una representación apropiada del espacio muestral es la siguiente:

$$\Omega' = \{h_v, h_h, v_h, v_v\}$$

donde los puntos del espacio muestral no son ahora las familias sino los hijos en las familias y h_v : "hembra que tiene un hermano", h_h : "hembra que tiene una hermana", v_h : "varón que tiene una hermana" y v_v : "varón que tiene un hermano". Sea

$$A' : \text{“ el niño escogido es varón''}$$
$$B' : \text{“ el niño escogido tiene un hermano''}$$

entonces

$$A' \cap B' = \{v_v\}$$

y por lo tanto

$$P(B'|A') = \frac{P(A' \cap B')}{P(A')} = \frac{1/4}{1/2} = \frac{1}{2}.$$

▲

7. **El Problema de Selección Optima.** Vamos a considerar la siguiente situación: tenemos n objetos del mismo tipo (por ejemplo, televisores, teléfonos celulares, etc.) y suponemos que si nos dan cualquier par de objetos, siempre podemos decidir cuál es mejor, sin que hayan empates. Queremos seleccionar el mejor de ellos con las siguientes restricciones: nos ofrecen los objetos uno a uno al azar y en cada caso debemos decidir si lo aceptamos o no. Si lo aceptamos tenemos que quedarnos con ese objetos. Si no lo aceptamos se nos muestra el siguiente, pero los objetos rechazados no pueden ser escogidos posteriormente.

   Adoptamos la siguiente regla general: 'nunca escogemos un objeto que sea de calidad inferior a los que hemos rechazado previamente'. Entonces podemos escoger el primer objeto y detener la búsqueda, o rechazamos el primer objeto y examinamos los posteriores con la esperanza de obtener alguno mejor a lo largo de la sucesión. Por supuesto, es posible que con esta estrategia nunca seleccionemos el mejor objeto, o que lo rechacemos inicialmente y nunca podamos hacer una selección. Por otro lado, si el número de objetos es grande, es razonable rechazar el primero con la esperanza de encontrar uno mejor posteriormente.

   Supongamos que, siguiendo la regla descrita anteriormente, seleccionamos el i-ésimo objeto, de modo que este objeto debe ser mejor que todos los anteriores. ¿Cuál es la probabilidad de que el objeto que escogemos sea realmente el mejor de todos?

▶ Sea B el evento que el i-ésimo objeto sea el mejor de los observados hasta ese momento y sea A el evento que ocurre si el i-ésimo objeto es el mejor de los n objetos. Queremos calcular $P(A|B)$, para lo cual necesitamos $P(B)$ y $P(A \cap B)$. Tenemos que $A \subset B$, de modo que $A \cap B = A$ y $P(A \cap B) = P(A)$. Por hipótesis los objetos se presentan al azar, de modo que todos los posibles órdenes tienen la misma probabilidad. Por lo tanto $P(B)$ es la probabilidad de que un objeto dado, el mejor de todos, ocupe el lugar i. Como hay $i!$ permutaciones de i objetos y hay $(i-1)!$ permutaciones si fijamos el lugar de un objeto, la probabilidad es

$$P(B) = \frac{(i-1)!}{i!} = \frac{1}{i}.$$

De manera similar $P(A)$ es la probabilidad de que en una permutación hecha al azar de n objetos distinguibles, un objeto dado, el mejor, ocupe el i-ésimo lugar, de modo que

$$P(A) = \frac{(m-1)!}{m!} = \frac{1}{m}.$$

Por lo tanto la probabilidad condicional es

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} = \frac{i}{m}.$$

▲

3.2. Resultados Básicos sobre Probabilidad Condicional.

Estudiaremos en esta sección algunas propiedades sencillas pero fundamentales de la probabilidad condicional.

Proposición 3.1 *Para cualquier colección finita de eventos $A_1, \dots, A_n$ se tiene*

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)\cdots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$$

siempre que $P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) > 0$.

Demostración. Como

$$P(A_1) \geq P(A_1 \cap A_2) \geq \cdots \geq P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) > 0$$

todas las probabilidades condicionales que aparecen están bien definidas. Si escribimos explícitamente el segundo miembro de la igualdad obtenemos

$$P(A_1)\frac{P(A_1 \cap A_2)}{P(A_1)}\frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)}\cdots\frac{P(A_1 \cap \cdots \cap A_n)}{P(A_1 \cap \cdots \cap A_{n-1})}$$

y simplificando se obtiene el primer miembro de la igualdad. ■

Como ejemplo consideremos el siguiente problema.

1. Se extraen sin reposición tres cartas de un juego completo de barajas. Calcule la probabilidad de que ninguna sea trébol.

► Sea A_i: "la i-ésima carta no es trébol". Queremos calcular

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) = \frac{39}{52}\frac{38}{51}\frac{37}{50} = 0.4135$$

▲

Una de las propiedades más útiles de la probabilidad condicional es la Ley de la Probabilidad Total, que demostramos a continuación.

Proposición 3.2 (Ley de la Probabilidad Total) *Sea $B_1, B_2, \dots$ una familia finita o numerable de conjuntos disjuntos dos a dos cuya unión es Ω. Entonces, para cualquier evento A,*

$$P(A) = \sum_i P(B_i)P(A|B_i)$$

donde la suma se toma sobre todos los índices i para los cuales $P(B_i) > 0$.

Demostración.

$$\begin{aligned} P(A) &= P(A \cap \Omega) = P(A \cap (\cup_i B_i)) \\ &= P(\cup_i (A \cap B_i)) = \sum_i P(A \cap B_i) \\ &= \sum_i P(B_i)P(A|B_i). \end{aligned}$$

■

Veamos ahora algunos ejemplos.

2 En una bolsa se colocan n tarjetas con nombres escritos en ellas y se extraen dos, sucesivamente y sin reposición. Si $m < n$ de los nombres son de mujer, calcule la probabilidad de que el segundo nombre extraído sea de mujer.

▶ Sea A el evento cuya probabilidad deseamos calcular y F: “el primer nombre extraído es femenino”. Los eventos F y F^c son disjuntos y forman una partición del espacio muestral. En consecuencia

$$P(A) = P(A|F)P(F) + P(A|F^c)P(F^c)$$

y

$$P(F) = \frac{m}{n}; \quad P(F^c) = \frac{n-m}{n}; \quad P(A|F) = \frac{m-1}{n-1}; \quad P(A|F^c) = \frac{m}{n-1}.$$

Por lo tanto

$$\begin{aligned} P(A) &= \frac{(m-1)}{(n-1)}\frac{m}{n} + \frac{m}{(n-1)}\frac{(n-m)}{n} \\ &= \frac{m}{(n-1)n}(m-1+n-m) = \frac{m}{n}. \end{aligned}$$

▲

3 Cada vez que el repartidor de una pizzeria regresa a buscar los pedidos para repartir, se encuentra que pueden haber entre 1 y 5 encargos esperando, y cada una de estas posibilidades tiene la misma probabilidad. Si, en promedio, la mitad de los clientes le da propina, calcule la probabilidad de que obtenga al menos dos propinas en un viaje.

▶ Sea A: “obtiene al menos dos propinas en el viaje” y sean B_i para $i = 1, \dots, 5$ los eventos “hay i encargos por repartir al iniciar el viaje”, respectivamente. Estos conjuntos B_i forman una partición del espacio muestral y por lo tanto

$$P(A) = \sum_{i=1}^{5} P(A|B_i)P(B_i).$$

Sabemos que $P(B_i) = 1/5$ y además, si comienza con i encargos, el número de propinas tiene distribución binomial con parámetros $n = i$ y $p = 1/2$ (ver ejemplo 2.4.4). Las probabilidades condicionales son:

$$\begin{aligned}
&P(A|B_1) = 0, \qquad P(A|B_2) = \left(\frac{1}{2}\right)^2, \\
&P(A|B_3) = P(2 \text{ éxitos en 3 intentos}) + P(3 \text{ éxitos en 3 intentos}) \\
&\qquad\qquad = \binom{3}{2}\frac{1}{2^3} + \binom{3}{3}\frac{1}{2^3} = \frac{1}{2}, \\
&P(A|B_4) = \left[\binom{4}{2} + \binom{4}{3} + \binom{4}{4}\right]\frac{1}{2^4}, \\
&P(A|B_5) = \left[\binom{5}{2} + \binom{5}{3} + \binom{5}{4} + \binom{5}{5}\right]\frac{1}{2^5},
\end{aligned}$$

y finalmente obtenemos $P(A) = \frac{9}{20}$. ▲

4 **La Ruina del Jugador.** Un jugador con capital inicial k juega en un casino una sucesión de juegos hasta lograr un capital m o hasta que se arruine. Si la probabilidad de ganar en cada juego es p ¿Cuál es la probabilidad de que se arruine?

► La probabilidad de ruina depende de k y de m. Supongamos que m está fijo y sea u_k la probabilidad de ruina del jugador si al comienzo tiene un capital de k unidades. Si el jugador gana en el primer juego, la probabilidad de que se arruine es ahora u_{k+1}, mientras que si pierde, la probabilidad de ruina es u_{k-1}.

Sea G_1 el evento "el jugador gana el primer juego", mientras que R es el evento "el jugador se arruina". Entonces

$$P(R|G_1) = u_{k+1}, \qquad P(R|G_1^c) = u_{k-1}.$$

Por lo tanto,

$$\begin{aligned}
P(R) &= P(R|G_1)P(G_1) + P(R|G_i^c)P(G_1^c) \\
&= pP(R|G_1) + qP(R|G_1^c),
\end{aligned}$$

es decir

$$u_k = pu_{k+1} + qu_{k-1}, \qquad 1 \leq k \leq m-1,$$

con las condiciones iniciales

$$u_0 = 1, \qquad u_m = 0.$$

Para resolver esta ecuación introducimos las diferencias $\delta_k = u_k - u_{k-1}$ para $k = 1, 2, \ldots, m$. Como $p + q = 1$ tenemos

$$u_k = (p+q)u_k = pu_{k+1} + qu_{k-1}$$

de donde obtenemos

$$p(u_{k+1} - u_k) = q(u_k - u_{k-1})$$

es decir

$$p\delta_{k+1} = q\delta_k \quad \Rightarrow \quad \delta_{k+1} = \frac{q}{p}\delta_k$$

Iterando esta relación obtenemos que

$$\delta_{k+1} = \left(\frac{q}{p}\right)^{k-1}\delta_1$$

Para regresar a las u_k observamos que $\delta_1 = u_1 - u_0 = u_1 - 1$ y $\delta_1 + \delta_2 + \cdots + \delta_k = u_k - 1$. Por lo tanto,

$$\begin{aligned} u_k &= 1 + \delta_1 + \delta_2 + \cdots + \delta_k \\ &= 1 + \delta_1 + \big(\frac{q}{p}\big)\delta_1 + \cdots + \big(\frac{q}{p}\big)^{k-1}\delta_1 \\ &= 1 + \Big[1 + \big(\frac{q}{p}\big) + \cdots + \big(\frac{q}{p}\big)^{k-1}\Big]\delta_1 \end{aligned}$$

Falta determinar δ_1. Como $u_m = 0$,

$$0 = u_m = 1 + \Big[1 + \big(\frac{q}{p}\big) + \cdots + \big(\frac{q}{p}\big)^{m-1}\Big]\delta_1$$

y obtenemos

$$\delta_1 = \frac{-1}{1 + \big(\frac{q}{p}\big) + \cdots + \big(\frac{q}{p}\big)^{m-1}}$$

de modo que

$$\begin{aligned} u_k &= 1 - \frac{1 + \big(\frac{q}{p}\big) + \cdots + \big(\frac{q}{p}\big)^{k-1}}{1 + \big(\frac{q}{p}\big) + \cdots + \big(\frac{q}{p}\big)^{m-1}} \\ &= \begin{cases} 1 - \frac{k}{m} = \frac{m-k}{m} & \text{si } p = q = 1/2 \\ 1 - \frac{1-(q/p)^k}{1-(q/p)^m} = \frac{(q/p)^k - (q/p)^m}{1-(q/p)^m} & \text{si } p \neq q \end{cases} \end{aligned}$$

▲

5 Un caminante sale del punto o y escoge uno de los caminos $ob_1,\ ob_2,\ ob_3,\ ob_4$ al azar (ver figura 3.2). En cada uno de los cruces siguientes de nuevo escoge al azar. ¿cuál es la probabilidad de que el caminante llegue al punto a?

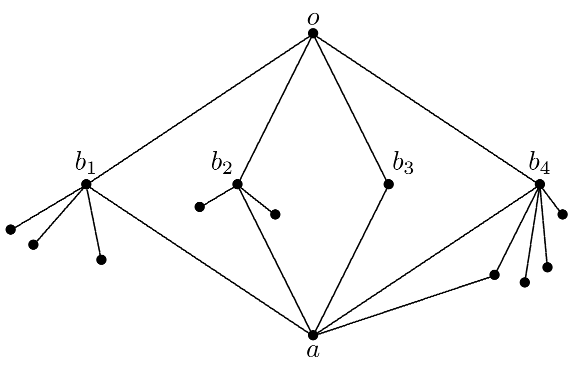

Figura 3.2

► Sea $B_k,\ k = 1, 2, 3, 4$, el evento "el caminante pasa por el punto b_k", estos eventos forman una partición de Ω y además son equiprobables. Por lo tanto $P(B_k) = 1/4,\ k = 1, 2, 3, 4$.

Si el caminante llega a b_1, hay cuatro caminos que puede escoger y sólo uno de ellos lo lleva a a, de donde la probabilidad condicional de llegar a a pasando por b_1 es $1/4$. Si A es el evento "el caminante llega a a", tenemos

$$P(A|B_1) = \frac{1}{4}.$$

De manera similar obtenemos

$$P(A|B_2) = \frac{1}{3}; \quad P(A|B_3) = 1; \quad P(A|B_4) = \frac{2}{5}.$$

Por lo tanto, la probabilidad de llegar a a es

$$P(A) = \sum_{i=1}^{4} P(A|B_i)P(B_i) = \frac{1}{4}\left(\frac{1}{4} + \frac{1}{3} + 1 + \frac{2}{5}\right) = \frac{119}{240}.$$

▲

6 ¿Cuál es la probabilidad de obtener 6 números distintos al lanzar 6 dados simétricos?

► Consideremos los siguientes eventos:

E_1: “se obtiene cualquier resultado en el primer lanzamiento”.

E_2: “el segundo dado es distinto al primero”.

E_3: “el tercer dado es distinto a los dos primeros”.

y así sucesivamente. Entonces tenemos:

$$\begin{aligned} P(E_1) &= 1 \\ P(E_2|E_1) &= 5/6 \\ P(E_3|E_1 \cap E_2) &= 4/6 \\ &\vdots \\ P(E_6|E_1 \cap E_2 \cap \cdots \cap E_5) &= 1/6 \end{aligned}$$

y usando la proposición 3.1

$$P(E_1 \cap E_2 \cap \cdots \cap E_6) = 1 \times \frac{5}{6} \times \frac{4}{6} \times \frac{3}{6} \times \frac{2}{6} \times \frac{1}{6} = \frac{6!}{6^6}.$$

▲

7 Se extraen dos cartas al azar de un juego de barajas. Calcule la probabilidad de que las cartas sean un As y un diez.

► Consideremos los siguientes eventos:

A_1 : “la primera carta es un As” $\quad B_1$: “la primera carta es un diez”

A_2 : “la segunda carta es un As” $\quad B_2$: “la segunda carta es un diez”

C : “se extraen un As y un diez”.

Tenemos

$$\begin{aligned} P(C) &= P(B_2|A_1)P(A_1) + P(A_2|B_1)P(B_1) \\ &= \frac{4}{51} \times \frac{4}{52} + \frac{4}{51} \times \frac{4}{52} \\ &= \frac{8}{663}. \end{aligned}$$

▲

3.3. El Teorema de Bayes

El otro resultado fundamental sobre probabilidad condicional se conoce como el Teorema de Bayes y es útil en situaciones en las cuales se conocen probabilidades de la forma $P(A|B_i)$ y $P(B_i)$ y se desea determinar $P(B_i|A)$.

Proposición 3.3 (Teorema de Bayes) *Sea $B_1, B_2, \ldots$ una partición finita o numerable de Ω y sea A un evento cualquiera con $P(A) > 0$. Entonces*

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_j P(A|B_j)P(B_j)}.$$

Demostración. Por la definición de probabilidad condicional tenemos:

$$P(B_i|A) = \frac{(A \cap B_i)}{P(A)} \tag{3.2}$$

y

$$P(A \cap B_i) = P(A|B_i)P(B_i). \tag{3.3}$$

Por la proposición 3.2 tenemos

$$P(A) = \sum_j P(A|B_j)P(B_j). \tag{3.4}$$

Usando (3.3) y (3.4) en (3.2) obtenemos el resultado. ■

Ejemplos.

1. De cien pacientes en un hospital con una cierta enfermedad se escogen diez para usar un tratamiento que aumenta la probabilidad de sanar de 0.50 a 0.75. Si posteriormente un doctor encuentra a un paciente curado, ¿cuál es la probabilidad de que haya recibido el tratamiento?

► Sean

C: "el paciente está curado".

T: "el paciente recibió el tratamiento".

A partir de la información dada obtenemos,

$$P(T) = \frac{10}{100} = 0.1; \qquad P(T^c) = \frac{90}{100} = 0.9;$$
$$P(C|T) = 0.75; \qquad P(C|T^c) = 0.50.$$

Usando el Teorema de Bayes

$$\begin{aligned} P(T|C) &= \frac{P(C|T)P(T)}{P(C|T)P(T) + P(C|T^c)P(T^c)} \\ &= \frac{(0.75)(0.1)}{(0.75)(0.1) + (0.5)(0.9)} = \frac{1}{7}. \end{aligned}$$

▲

2. Tres cajas contienen dos monedas cada una. En la primera, C_1, ambas son de oro; en la segunda, C_2, ambas son de plata y en la tercera, C_3, una es de oro y otra es de plata. Se escoge una caja al azar y luego una moneda también al azar. Si la moneda es de oro, ¿cuál es la probabilidad de que venga de la caja que contiene dos monedas de oro?

▶ Sabemos que $P(C_i) = 1/3$. Sea O: "se escoge una moneda de oro". Usando el Teorema de Bayes,

$$\begin{aligned} P(C_1|O) &= \frac{P(O|C_1)P(C_1)}{P(O|C_1)P(C_1) + P(O|C_2)P(C_2) + P(O|C_3)P(C_3)} \\ &= \frac{1 \times \frac{1}{3}}{1 \times \frac{1}{3} + 0 \times \frac{1}{3} + \frac{1}{2} \times \frac{1}{3}} = \frac{2}{3} \end{aligned}$$

Si uno considera el problema sin mucho cuidado podría razonar de la siguiente manera: luego de sacar una moneda de oro sólo hay dos cajas de las cuales pudo provenir, y como estamos escogiendo las cajas al azar, la probabilidad de que venga de C_1 es $1/2$. El problema con este razonamiento es que es más probable sacar una moneda de oro de C_1 que de C_3, y el argumento no toma esto en cuenta. ▲

3. Tres enfermedades distintas y excluyentes A, B y C producen el mismo conjunto de síntomas H. Un estudio clínico muestra que las probabilidades de contraer las enfermedades son 0.01; 0.005 y 0.02 respectivamente. Además, la probabilidad de que el paciente desarrolle los síntomas H para cada enfermedad son 0.90; 0.95 y 0.75, respectivamente. Si una persona enferma tiene los síntomas H, ¿cuál es la probabilidad de que tenga la enfermedad A?

▶ Sabemos lo siguiente:

$$\begin{aligned} P(A) &= 0.01, \qquad P(B) = 0.005, \qquad P(C) = 0.02 \\ P(H|A) &= 0.9, \qquad P(H|B) = 0.95, \quad P(H|C) = 0.75 \end{aligned}$$

Usando el Teorema de Bayes tenemos

$$\begin{aligned} P(A|H) &= \frac{P(H|A)P(A)}{P(H|A)P(A) + P(H|B)P(B) + P(H|C)P(C)} \\ &= \frac{0.9 \times 0.01}{0.9 \times 0.01 + 0.95 \times 0.005 + 0.75 \times 0.02} = \frac{9}{28.75} = 0.313. \end{aligned}$$

▲

4. Un estudiante responde una pregunta de un examen de múltiple selección que tiene cuatro respuestas posibles. Suponga que la probabilidad de que el estudiante conozca la respuesta a la pregunta es 0.8 y la probabilidad de que adivine es 0.2. Si el estudiante adivina, la probabilidad de que acierte es 0.25. Si el estudiante responde acertadamente la pregunta, ¿cuál es la probabilidad de que el estudiante realmente supiera la respuesta?

▶ Definamos los siguientes eventos:

C: "el estudiante conoce la respuesta".

A: "el estudiante responde acertadamente".

Queremos calcular $P(C|A)$ y sabemos que

$$P(C) = 0.8; \qquad P(A|C^c) = 0.25; \qquad P(A|C) = 1.$$

Usando el Teorema de Bayes

$$\begin{aligned} P(C|A) &= \frac{P(A|C)P(C)}{P(A|C)P(C) + P(A|C^c)P(C^c)} \\ &= \frac{0.8}{0.8 + (0.25)(0.2)} = \frac{0.8}{0.85} = 0.941. \end{aligned}$$

▲

3.4. Eventos Independientes

En los ejemplos que hemos considerado en relación al concepto de probabilidad condicional, hemos visto que en ciertos casos, cuando la ocurrencia del evento A no aporta información respecto a la ocurrencia o no del evento B, se tiene

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

En este caso decimos que B es independiente de A. Ahora bien, como

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

podemos expresar la relación anterior en la forma siguiente, que es simétrica en ambos sucesos:

$$P(A \cap B) = P(A)P(B). \tag{3.5}$$

Observamos además que (3.5) es válida aún cuando A o B sean sucesos de probabilidad nula. Usaremos esta relación como definición.

Definición 3.2 Dos eventos A y B son independientes si vale (3.5).

Es fácil verificar que Ω y $\emptyset$ son independientes de cualquier evento.

Proposición 3.4 *Si A y B son independientes, también lo son: a) A y B^c, b) A^c y B^c.*

Demostración. Veamos b), la demostración de a) es similar.

$$\begin{aligned} P(A^c \cap B^c) &= P((A \cup B)^c) = 1 - P(A \cup B) \\ &= 1 - P(A) - P(B) + P(A \cap B) \\ &= 1 - P(A) - P(B) + P(A)P(B) \\ &= [1 - P(A)][1 - P(B)] \\ &= P(A^c)P(B^c). \end{aligned}$$

■

Ejemplo

1. Un lote de diez objetos contiene cuatro defectuosos y seis en buen estado. Se extraen dos objetos sucesivamente y sin reposición. Sean los eventos D_1: "el primer objeto es defectuoso" y D_2: "el segundo objeto es defectuoso". ¿Son independientes estos eventos? ¿Qué sucede si los objetos se extraen con reposición?

▶ En el primer caso podemos calcular $P(D_2)$ de la siguiente manera:

$$\begin{aligned} P(D_2) &= P(D_2|D_1)P(D_1) + P(D_2|D_1^c)P(D_1^c) \\ &= \frac{3}{9}\frac{4}{10} + \frac{4}{9}\frac{6}{10} = \frac{36}{90} = \frac{2}{5}. \end{aligned}$$

Por otro lado,

$$P(D_2|D_1) = \frac{3}{9} \neq \frac{2}{5} = P(D_2)$$

de modo que D_1 y D_2 no son independientes.

En cambio, si los objetos se extraen con reposición tenemos $P(D_1) = P(D_2) = 4/10$, mientras que $P(D_1 \cap D_2) = (4/10)^2$ y los eventos son independientes. ▲

La definición de independencia se puede generalizar a una familia cualquiera de eventos:

Definición 3.3 Sea $\mathcal{C} = \{A_i, i \in I\}$ una familia de eventos. Diremos que los eventos A_i son independientes si para cualquier colección finita de eventos $A_{i_1}, A_{i_2}, \dots, A_{i_n}$ de $\mathcal{C}$, se cumple:

$$P(\bigcap_{j=1}^{n} A_{i_j}) = \prod_{j=1}^{n} P(A_{i_j}). \tag{3.6}$$

En este caso decimos también que $\mathcal{C}$ es una familia de eventos independientes.

Observemos que en la definición *sólo* intervienen colecciones finitas de eventos de $\mathcal{C}$, pero intervienen *todas* las colecciones finitas. Por ejemplo, si la familia consta de tres eventos, no es suficiente verificar (3.6) para las parejas de eventos. En efecto, sea el experimento de lanzar dos dados y $\mathcal{C} = \{A_1, A_2, A_3\}$ con

$$\begin{aligned} A_1 &: \text{“se obtiene un 6 en el primer dado''}. \\ A_2 &: \text{“se obtiene un 1 en el segundo dado''}. \\ A_3 &: \text{“la suma de los dos dados es 7''}. \end{aligned}$$

Claramente,

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{6}$$

y

$$P(A_1 \cap A_2) = P(A_1 \cap A_3) = P(A_2 \cap A_3) = \frac{1}{6} \times \frac{1}{6}.$$

Pero

$$P(A_1 \cap A_2 \cap A_3) = \frac{1}{36} \neq \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6},$$

de modo que los tres eventos no son independientes.

Ejemplos.

2 Si A y B son eventos independientes y la probabilidad de que ambos ocurran es 0.16, mientras que la probabilidad de que ninguno ocurra es 0.36, calcule $P(A)$ y $P(B)$.

▶ Sabemos que $P(A \cap B) = P(A)P(B) = 0.16$ y $P((A \cup B)^c) = 0.36$, de donde obtenemos

$$\begin{aligned} P(A \cup B) &= 0.64 \\ &= P(A) + P(B) - P(A \cap B) \\ &= P(A) + P(B) - 0.16. \end{aligned}$$

Por lo tanto tenemos el siguiente par de ecuaciones

$$\begin{aligned} P(A) + P(B) &= 0.8 \\ P(A)P(B) &= 0.16 \end{aligned}$$

de donde se obtiene que $P(A) = P(B) = 0.4$.

▲

3 ¿Cuál es la probabilidad de obtener tres 6 al lanzar ocho dados?

▶ El problema es equivalente a calcular la probabilidad de obtener tres 6 al lanzar ocho veces el mismo dado. Sea E_i el evento “obtenemos un 6 en el i-ésimo lanzamiento”. Calculemos primero la probabilidad de que los primeros tres sean 6 y el resto no, es decir,

$$P(E_1 \cap E_2 \cap E_3 \cap E_4^c \cap E_5^c \cap E_6^c \cap E_7^c \cap E_8^c)$$

y por independencia esto es

$$\left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^5.$$

Es fácil ver que esta es la probabilidad de cualquier conjunto de ocho lanzamientos de un dado en el cual haya exactamente tres 6. Como hay $\binom{8}{3}$ conjuntos de este tipo, tenemos que la probabilidad deseada es

$$\binom{8}{3}\left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^5.$$

▲

4 Una muestra de tamaño 4 se extrae con reposición de una bolsa que contiene 6 bolas, de las cuales 4 son blancas. Sea A el evento "la primera bola es blanca" y B: "la tercera bola es blanca". ¿Son independientes estos eventos? ¿Qué sucede si el muestreo se hace sin reposición?

▶ En el primer caso (muestreo con reposición) tenemos

$$P(A) = \frac{4}{6}, \quad P(B) = \frac{4}{6} \;\text{ y }\; P(A \cap B) = \left(\frac{4}{6}\right)^2,$$

de modo que los eventos son independientes.

En el segundo caso, de nuevo $P(A) = 4/6$ y para calcular $P(B)$ podemos proceder como en el ejemplo 3.2.2 obteniendo $P(B) = 4/6$. Para calcular $P(A \cap B)$ consideremos el evento C definido por "la segunda bola es blanca". Entonces

$$\begin{aligned}
P(A \cap B) &= P(A \cap B \cap C) + P(A \cap B \cap C^c)\\
&= P(A)P(C|A)P(B|A \cap C) + P(A)P(C^c|A)P(B|A \cap C^c)\\
&= \frac{4}{6}\frac{3}{5}\frac{2}{4} + \frac{4}{6}\frac{2}{5}\frac{3}{4} = \frac{2}{5}
\end{aligned}$$

y en este caso los eventos no son independientes. ▲

5 Demuestre: si $P(B|A) = P(B|A^c)$ entonces A y B son independientes.

▶ La relación se puede escribir

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

Multiplicando por $P(A^c)$ ambos lados obtenemos

$$\begin{aligned}
P(B \cap A^c) &= \frac{(1 - P(A))P(B \cap A)}{P(A)}\\
&= \frac{P(B \cap A)}{P(A)} - P(B \cap A),
\end{aligned}$$

de donde

$$\frac{P(B \cap A)}{P(A)} = P(B \cap A) + P(B \cap A^c) = P(B)$$

es decir

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

▲

6 El primer hijo que tiene una mujer es un niño hemofílico. La mujer, en cuya historia familiar no aparecen casos de hemofilia, tiene dudas sobre tener un segundo hijo, pero piensa que su hijo no heredó la hemofilia de ella y que su enfermedad es debida a una mutación. Por lo tanto, la probabilidad de que un segundo hijo tenga hemofilia es la de que nuevamente la enfermedad venga de una mutación, y esto es un número pequeño, m (digamos $m = 10^{-5}$). Calcule cuál es en realidad la probabilidad de que el segundo hijo tenga hemofilia si el primero nació hemofílico.

► Definamos los siguientes eventos:

$$\begin{aligned}
&A_1 : \text{“la madre es portadora de la enfermedad''}.\\
&A_2 : \text{“el primer hijo es hemofílico''}.\\
&A_3 : \text{“el segundo hijo es hemofílico''}.
\end{aligned}$$

Un varón normal tiene un par de cromosomas XY y tiene hemofilia si y sólo si, en lugar del cromosoma X tiene un cromosoma X', que lleva un gen que produce la enfermedad. Sea m la probabilidad de que un cromosoma X mute en un cromosoma X'.

La madre tiene dos cromosomas X y el evento A_1 ocurre sólo si al menos uno de estos dos cromosomas X es mutante, lo cual ocurre con probabilidad

$$P(A_1) = 1 - (1-m)^2 = 2m - m^2 \approx 2m$$

donde hemos supuesto que las mutaciones ocurren en forma independiente, y hemos desechado el término m^2 por ser mucho menor que $2m$.

Si la madre es portadora de la enfermedad y uno de sus cromosomas es X', su hijo tendrá probabilidad $1/2$ de heredar el cromosoma X, es decir

$$P(A_2|A_1) = P(A_2^c|A_1) = \frac{1}{2}$$

Si en cambio la madre no lleva la enfermedad, su hijo será hemofílico si su cromosoma X sufre una mutación:

$$P(A_2|A_1^c) = m.$$

Además, por independencia, tenemos que

$$\begin{aligned}
P(A_3|A_1\cap A_2) &= P(A_3^c|A_1\cap A_2) = P(A_3|A_1\cap A_2^c)\\
&= P(A_3^c|A_1\cap A_2^c) = \frac{1}{2},\\
P(A_3|A_1^c\cap A_2) &= P(A_3|A_1^c\cap A_2^c) = m.
\end{aligned}$$

Queremos calcular

$$P(A_3|A_2) = \frac{P(A_2\cap A_3)}{P(A_2)}$$

pero

$$\begin{aligned}
P(A_2\cap A_3) &= P(A_1\cap A_2\cap A_3) + P(A_1^c\cap A_2\cap A_3)\\
&= P(A_1)P(A_2|A_1)P(A_3|A_2\cap A_1) + P(A_1^c)P(A_2|A_1^c)P(A_3|A_2\cap A_1^c)\\
&= 2m\left(\frac{1}{2}\right)^2 + (1-2m)m^2\\
&= \frac{m}{2} + m^2 - 2m^3 \approx \frac{m}{2}
\end{aligned}$$

y además

$$\begin{aligned}P(A_2) &= P(A_2|A_1)P(A_1) + P(A_2|A_1^c)P(A_1^c) \\ &= \frac{1}{2}2m + m(1-2m) \\ &\approx 2m.\end{aligned}$$

Por lo tanto

$$P(A_3|A_2) \approx \frac{m/2}{2m} = \frac{1}{4}.$$

▲

7 *Muestreo con Reposición.* Analicemos de nuevo, desde el punto de vista de la independencia, el muestreo con reposición de una población con dos tipos de individuos A y B en proporciones p y $1-p$ respectivamente. Deseamos calcular la probabilidad de extraer r del tipo A cuando tomamos una muestra de tamaño n, $(r \leq n)$ extrayendo uno a la vez.

► Supongamos que la extracción dió como resultado que los primeros r fueron del tipo A y los $n-r$ restantes del tipo B. ¿Cómo podemos calcular la probabilidad de que esto ocurra? Para ello, llamemos A_i al evento "en la i-ésima extracción obtuvimos un individuo de tipo A". En virtud de la independencia,

$$\begin{aligned}P(A_1 \cap A_2 \cap \cdots \cap A_r \cap A_{r+1}^c \cap \cdots \cap A_n^c) &= \prod_{i=1}^{r} P(A_i) \prod_{i=r+1}^{n} P(A_i^c) \\ &= p^r(1-p)^{n-r}\end{aligned}$$

Si observamos ahora que cualquiera sea el orden en que extraigamos los r individuos de tipo A y los $n-r$ de tipo B, la probabilidad es la misma, tendremos que la probabilidad buscada es

$$\binom{n}{r} p^r(1-p)^{n-r}, \qquad 0 \leq r \leq n,$$

es decir, la distribución binomial que habíamos encontrado anteriormente. ▲

8 Consideremos de nuevo la población del ejemplo anterior y calculemos la probabilidad de obtener por primera vez un individuo de tipo A en la n-ésima extracción.

► Sea A_i el evento "en la i-ésima extracción se obtiene un individuo de tipo A" y sea $\mathcal{C} = \{A_i, i \geq 1\}$. $\mathcal{C}$ es una familia de eventos independientes. También son familias de eventos independientes las obtenidas a partir de $\mathcal{C}$ al sustituir cualesquiera conjuntos A_i por sus complementos (ver la proposición 3.4). Entonces

$$\begin{aligned}p_n &= P(\text{ el primero de tipo } A \text{ es el } n\text{-esimo}) \\ &= P(A_1^c \cap \cdots \cap A_{n-1}^c \cap A_n) = (1-p)^{n-1}p,\end{aligned}$$

que es la distribución geométrica que vimos en el capítulo anterior.

Las probabilidades p_n que hemos calculado determinan la distribución de un "tiempo de vida" discreto, es decir, del número de repeticiones de un experimento hasta que ocurra un evento determinado. Es de interés observar que las probabilidades p_n pueden ser utilizadas como aproximación discreta de la situación que hemos estudiado anteriormente en relación a la distribución del tiempo de vida de aparatos que "no envejecen" (ver ejemplo 3.1.4).

Sea T el tiempo de vida de un aparato de esa clase. Para calcular $P(T > t)$ dividimos el intervalo $[0, t]$ en n intervalos iguales $I_1,\ I_2, \ldots, I_n$ (ver figura 3.3).

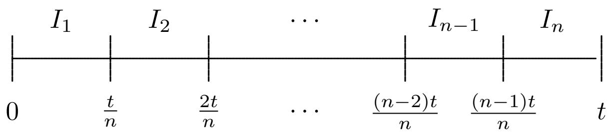

Figura 3.3

Si A_k: "el aparato no se daña en el intervalo I_k", es claro que

$$\{T > t\} = \bigcap_{k=1}^{n} A_k.$$

Ahora bien, la hipótesis de "no envejecimiento" se traduce en el hecho de que los eventos $A_1, \dots, A_n$ son independientes, es decir que el hecho de que no ocurran desperfectos en los intervalos $I_1, \dots, I_k$ no suministra información sobre la probabilidad de daños en el intervalo I_{k+1}.

Si suponemos que cuando la longitud h de un intervalo es pequeña, la probabilidad es aproximadamente proporcional a h, o más precisamente, que

$$P(\text{se daña en } [0, h]) = \lambda h + o(h)$$

donde λ es la constante de proporcionalidad y $o(h)$ es una función de h que satisface $\frac{o(h)}{h} \to 0$ $(h \to 0)$, resulta que

$$\begin{aligned} P(T > t) &= P(\cap_{k=1}^{n} A_k) = \prod_{k=1}^{n} P(A_k) \\ &= \left(1 - \frac{\lambda t}{n} + o(\frac{t}{n})\right)^n \\ &\to e^{-\lambda t} \qquad (n \to \infty) \end{aligned}$$

y volvemos a encontrar la distribución exponencial. ▲

9 Al probar independientemente dos aparatos electrónicos iguales hasta su inutilización, se obtuvieron los siguientes valores para la duración de vida: 21 horas y 30.2 horas. Suponiendo que el aparato no sufre envejecimiento, esto es, deterioro progresivo, tenemos dos observaciones correspondientes a la distribución

$$P(T > x) = e^{-\lambda x} \qquad \text{para cierto } \lambda.$$

¿Cómo obtener un valor razonable de λ, esto es, *una estimación de* λ?

► Naturalmente, el valor "razonable" de λ depende de qué consideremos "razonable", o sea, del criterio de estimación que se adopte. Como ya hemos visto el criterio de máxima verosimilitud, trataremos de utilizarlo.

En una primera etapa simplificaremos el problema. Supongamos que sólo observamos el primer aparato hasta su destrucción, es decir, 21 horas. Además, consideremos que en realidad no medímos exactamente las 21 horas, sino que sólo observamos que el aparato quedó inútil entre las 21 y 21+h horas de vida ($h > 0$). Tenemos

$$P(21 < T < 21 + h) = \int_{21}^{21+h} \lambda e^{-\lambda t} dt.$$

De acuerdo al criterio adoptado, debemos maximizar esta probabilidad. Como h está fijo, esto es equivalente a maximizar lo siguiente:

$$\frac{1}{h} P(21 < T < 21 + h) = \frac{1}{h} \int_{21}^{21+h} \lambda e^{-\lambda t} dt$$

pero si h es pequeño, esto es aproximadamente igual a

$$\lambda e^{-21\lambda}$$

debido al Teorema del Valor Medio del Cálculo Integral. Por lo tanto, el problema se reduce a maximizar la función $\lambda e^{-\lambda t}$ en el punto $t = 21$. Obtenemos

$$\hat{\lambda} = \frac{1}{21}.$$

Después de estas consideraciones resulta fácil considerar el caso original. Usando la independencia, debemos maximizar el producto

$$\frac{1}{h}\ P(21 < T < 21 + h)\ \frac{1}{k}\ P(30.2 < T < 30.2 + k) \simeq \lambda e^{-21\lambda}\ \lambda e^{-30.2\lambda}$$

si h y k son pequeños. Se obtiene

$$\hat{\lambda} = \frac{2}{21 + 30.2}.$$

Resulta sencillo ahora generalizar la aplicación anterior al caso de n observaciones independientes $t_1, t_2, \ldots, t_n$. Queremos obtener el valor de $\hat{\lambda}$ (si es único) que maximice

$$L(\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda t_i} = \lambda^n e^{-\lambda \sum_{i=1}^{n} t_i}$$

como

$$L'(\lambda) = \left(\frac{n}{\lambda} - \sum_{i=1}^{n} t_i \right) L(\lambda)$$

resulta que un valor posible para el máximo en el intervalo $(0,\infty)$ es

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} t_i} = \frac{1}{\bar{t}}$$

donde $\bar{t}$ es el promedio de las n observaciones. Por otra parte, la función $L(\lambda)$ es positiva para $\lambda \in (0, \infty)$, pero $L(\lambda) \to 0$ cuando $\lambda \to 0^+$ y $L(\lambda) \to 0$ cuando $\lambda \to \infty$, de modo que si consideramos un intervalo cerrado $[a, A]$ con $a > 0$ suficientemente pequeño y A suficientemente grande, $L(\lambda)$ tiene necesariamente un máximo en este intervalo y lo alcanza en $\hat{\lambda}$ ya que

$$L(\hat{\lambda}) = \left(\frac{n}{\sum_{i=1}^{n} t_i} \right)^n e^{-n}$$

es estrictamente positivo y finito a menos que $t_1 = t_2 = \cdots = t_n = 0$ y esto sólo puede suceder con probabilidad 0. ▲

10 Decimos que A atrae a B si $P(B|A) > P(B)$ y A repele a B si $P(B|A) < P(B)$. (a) Demuestre que si B atrae a A entonces A atrae a B y B^c repele a A. (b) Un documento que buscamos está en alguno de los n folders de un cajón y B_j respresenta el evento de que esté en el j-ésimo folder, donde $P(B_j) = b_j > 0$. Por otro lado F_j representa el evento de que en una búsqueda rápida en el folder j no encontremos el documento, y $P(F_j|B_j) = a_j < 1$. Demuestre que B_j y F_j son mutuamente repelentes, pero que F_j atrae a B_i para $i \neq j$.

► a) Como B atrae a A, usando la definición tenemos que $P(A \cap B) > P(A)P(B)$. Dividiendo ahora por $P(A)$ obtenemos $P(B|A) > P(B)$. Por otro lado

$$P(A|B^c)P(B^c) = P(A) - P(A|B)P(B) < P(A)(1 - P(B)) = P(A)P(B^c),$$

y al dividir por $P(B^c) > 0$ obtenemos que B^c repele a A.

b) Por el teorema de Bayes,

$$P(B_j|F_j) = \frac{P(F_j|B_j)P(B_j)}{\sum_i P(F_i|B_i)P(B_i)} = \frac{a_j b_j}{1 - b_j + a_j b_j}$$

ya que, para $i \neq j$, $P(F_j|B_i) = 1$. Por lo tanto

$$P(B_j) - P(B_j|F_j) = \frac{b_j(1-b_j)(1-a_j)}{1 - b_j + a_j b_j} > 0,$$

de modo que F_j repele a B_j. Además, para $i \neq j$

$$P(B_i|F_j) - P(B_i) = \frac{b_i}{1 - b_j + a_j b_j} - b_i = \frac{b_i b_j(1-a_j)}{1 - b_j + a_j b_j} > 0,$$

▲

11 **El Problema de los Puntos** Dos jugadores que llamaremos A y B juegan una serie de juegos independientes y en cada uno ambos tienen igual probabilidad de ganar. Ambos apuestan una cierta cantidad C y el acuerdo es que el primero en ganar n juegos gana la apuesta. Sin embargo, la serie de juegos se interrumpe antes de que alguno de los jugadores haya ganado n juegos. La pregunta es ¿cómo debe repartirse el dinero de la apuesta?

► Este problema, conocido como el problema de los puntos, tiene una larga historia, fue considerado por Pascal y Fermat en su correspondencia y resuelto por ambos. La idea de Pascal es que el monto de la apuesta debe repartirse en proporción a la probabilidad que cada jugador tiene de ganar si la serie de juegos continuara.

Veamos un caso extremo para comenzar. Supongamos que $n = 5$ y la serie de juegos se detiene cuando A ha ganado 4 juegos y B no ha ganado ninguno. En esta situación la única manera en que B puede ganar es si gana 5 partidas seguidas. Como los juegos son independientes y la probabilidad de ganar en cada juego es 1/2, la probabilidad de que B gane es $(1/2)^5 = 1/32$. Por lo tanto, de acuerdo a Pascal B debería recibir 1/32 del total de la apuesta y A debería recibir el resto.

Veamos otro caso particular. Supongamos que A necesita un juego para ganar mientras que B necesita 3. Para decidir quién gana A y B deben jugar a lo sumo tres juegos más. Veamos cuáles son los resultados posibles de tres juegos:

$$\begin{array}{cccc} \underline{\text{AAA}} & \underline{\text{AAB}} & \underline{\text{ABB}} & \text{BBB} \\ & \underline{\text{ABA}} & \underline{\text{BAB}} & \\ & \underline{\text{BAA}} & \underline{\text{BBA}} & \end{array}$$

En siete de ellos, los que están subrayados, A gana al menos un juego antes de que B gane tres. Como todas las series de tres juegos tienen la misma probabilidad 1/8, A gana con probabilidad 7/8 y A debe llevarse 7/8 partes de la apuesta.

Se puede argumentar que si A gana la primera partida, la serie termina allí pues ya se sabe que A es el ganador, de modo que en realidad sólo tendríamos las siguientes posibilidades:

$$\underline{\text{A}} \qquad \underline{\text{BA}} \qquad \underline{\text{BBA}} \qquad \text{BBB}$$

Sin embargo, estos eventos no son igualmente probables: las probabilidades respectivas son 1/2, 1/4, 1/8 y 1/8. Como en los tres primeros casos A gana la serie, la probabilidad de que esto ocurra es

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8}$$

que es el resultado que obtuvimos anteriormente. La ventaja del primer enfoque es que facilita el cálculo pues como los eventos son equiprobables, sólo hay que contar en cuantos casos gana A. Este es el enfoque que vamos a tomar para resolver el caso general, que formulamos a continuación.

Si A ha ganado $n-a$ juegos y B ha ganado $n-b$, el problema se puede se puede plantear de manera equivalente como sigue: ¿Cuál es la probabilidad de que A gane a juegos antes de que B gane b juegos?

Llamemos $u(a,b)$ la probabilidad que queremos calcular. Para que el evento que nos interesa ocurra, A y B deben jugar a lo sumo $a+b-1$ juegos. Por lo tanto la probabilidad de que A gane la serie es la probabilidad de que A gane al menos a juegos en una serie de $a+b-1$. La probabilidad de que A gane exactamente k juegos en esta serie tiene una distribución binomial con $a+b-1$ ensayos y probabilidad de éxito en cada ensayo $1/2$, esto es

$$\binom{a+b-1}{k}\Big(\frac{1}{2}\Big)^{a+b-1}$$

Como $u(a,b)$ es la probabilidad de que A gane al menos a juegos en una serie de $a+b-1$ tenemos

$$u(a,b) = \sum_{k=a}^{a+b-1}\binom{a+b-1}{k}\Big(\frac{1}{2}\Big)^{a+b-1} = \sum_{k=0}^{b-1}\binom{a+b-1}{k}\Big(\frac{1}{2}\Big)^{a+b-1}.$$

▲

12 **Cadenas de Markov.** En un proceso de fabricación se sabe que una máquina produce siempre la misma cantidad C de artículos, que se dividen en dos categorías, A y B. La máquina puede operar en dos estados, E_1 y E_2, y las probabilidades respectivas de producir un artículo de tipo A son p_1 y p_2, donde $p_2 > p_1$. Las condiciones del proceso determinan que la máquina debe operar en un mismo estado toda una jornada.

Si la máquina trabaja un día en el estado E_i, $i = 1, 2$, al día siguiente hay que determinar si va a operar en el mismo estado o en el otro. El problema que vamos a enfocar es el de regular el proceso. Tomamos las decisiones del siguiente modo:

- Si la máquina opera un día en el estado E_1, observamos la cantidad de artículos del tipo A producidos. Sea r esta cantidad. Si $r \le s$, donde s es un número fijo, al día siguiente se utilizará la máquina en el estado E_2. Si en cambio $r > s$, se la mantiene en el estado E_1 al día siguiente. Es decir, si el número de artículos de categoría A es muy pequeño, cambiamos al estado E_2, donde tenemos mayor probabilidad ($p_2 > p_1$) de producir artículos de categoría A. Pero si el número de artículos de categoría A es grande, decidimos que la máquina opere en el mismo estado E_1.
- Análogamente, si una jornada la máquina fue utilizada en el estado E_2, para decidir el estado en el cual trabajará al día siguiente, observamos el número r de artículos de tipo A producidos. Si $r \le t$, donde t ha sido fijado previamente, al día siguiente se mantiene en el estado E_2, en el cual tenemos mayor probabilidad de producir artículos de categoría A. Si en cambio $r > t$, cambiamos al estado E_1.

Este esquema de decisiones nos dice que si estamos en el estado E_1, al día siguiente nos mantenemos en este estado o lo cambiamos al estado E_2, según ciertas probabilidades a y $1-a$, que podemos calcular a partir de la distribución binomial, conocidos C, p_1 y s. De manera similar, si estamos en el estado E_2, lo mantenemos o no en la jornada siguiente según ciertas probabilidades b y $1-b$, calculadas también con la distribución binomial conocidos C, p_2 y t.

Llamando E_i^n, $i = 1, 2;\ n = 1, 2, \dots$ el suceso "en el n-ésimo día la máquina funciona en el estado E_i" tenemos

$$P(E_1^{n+1}|E_1^n) = a \qquad P(E_2^{n+1}|E_1^n) = 1 - a \tag{3.7}$$
$$P(E_2^{n+1}|E_2^n) = b \qquad P(E_1^{n+1}|E_2^n) = 1 - b \tag{3.8}$$

para cualquier $n \geq 1$.

Observamos que las probabilidades de pasar de un estado a otro dependen sólamente del estado en el cual la máquina operó el día previo, y no de los restantes días pasados, cualquiera sea el número $n + 1$.

Representemos las probabilidades de cambio de estado por una tabla:

		Dia $n+1$	
		E_1	E_2
Dia n	E_1	a	$1-a$
	E_2	$1-b$	b

Tabla 3.2

de donde obtenemos la matriz

$$M = \begin{pmatrix} a & 1-a \\ 1-b & b \end{pmatrix}.$$

Observamos que las sumas de las filas de M tienen el valor constante 1. Una matriz con esta propiedad, y para la cual $a_{ij} \geq 0$, se llama *estocástica* o *de Markov*. Podemos, ahora, calcular probabilidades de interés, como por ejemplo la probabilidad de pasar del estado E_i al estado E_j en m pasos. Sea

$$p_{ij}^{(m)} = P(\text{pasar del estado } E_i \text{ al } E_j \text{ en } m \text{ pasos})$$

donde $i, j = 1, 2$.

Ahora bien, considerando el estado ocupado en el primer paso, esta probabilidad se puede expresar:

$$\begin{aligned} p_{ij}^{(m)} &= p_{i1}^{(1)} p_{1j}^{(m-1)} + p_{i2}^{(1)} p_{2j}^{(m-1)} \\ &= \sum_{k=1}^{2} p_{ik}^{(1)} p_{kj}^{(m-1)} \end{aligned}$$

Los términos $p_{ij}^{(1)}$ para i fijo, son los términos de la i-ésima fila de M, y los términos $p_{ij}^{(1)}$ para j fijo, son los términos de la columna j de M.

Si $m = 2$, de modo que $p_{kj}^{(m-1)} = p_{kj}^{(1)}$, obtenemos

$$p_{11}^{(2)} = p_{11}^{(1)} p_{11}^{(1)} + p_{12}^{(1)} p_{21}^{(1)} \qquad p_{12}^{(2)} = p_{11}^{(1)} p_{12}^{(1)} + p_{12}^{(1)} p_{22}^{(1)}$$

$$p_{21}^{(2)} = p_{21}^{(1)} p_{11}^{(1)} + p_{22}^{(1)} p_{21}^{(1)} \qquad p_{22}^{(2)} = p_{21}^{(1)} p_{12}^{(1)} + p_{22}^{(1)} p_{22}^{(1)}$$

de donde se observa que los términos $p_{ij}^{(2)}$ son los términos de la matriz M^2, esto es

$$M^2 = \begin{pmatrix} a^2 + (1-a)(1-b) & a(1-a)(1-b)b \\ a(1-b) + b(1-b) & (1-b)(1-a) + b^2 \end{pmatrix}$$

donde M^2 resulta también estocástica.

Procediendo por inducción obtenemos que M^m es estocástica y tiene como términos las probabilidades $p_{ij}^{(m)}$ de pasar del estado i al j en m pasos.

Si al comenzar el proceso, el primer día del año, partimos del estado E_i, al cabo de $m+1$ días estaremos en el estado E_j con cierta probabilidad. Es fácil calcular esta probabilidad porque, para $m=1$ (esto es, en el segundo día), las probabilidades de estar en E_1 y E_2 están dadas por

$$(1,0)\begin{pmatrix} a & 1-b \\ 1-b & b \end{pmatrix} = (a,\ 1-b)$$

es decir, la primera fila de M, y en general, procediendo por inducción, tendremos para el $(m-1)$-ésimo día:

$$(1,0)M^m.$$

Si en el primer día del proceso estamos en E_1 o E_2 con probabilidades p y $1-p$, determinadas por algún procedimiento azaroso, al cabo de m pasos, es decir el día $m+1$, tendremos para cada estado las probabilidades determinadas por el vector

$$(p,\ 1-p)M^m$$

Se nos presenta ahora el problema de calcular las potencias de M, y observar si cuando $m \to \infty$ resulta algún valor límite para M^m. Ahora bien, como M^m es estocástica, es suficiente calcular los elementos de la diagonal de la matriz M^m. Sea $a_m = p_{11}^{(m)},\ a_1 = a,\ b_m = p_{22}^{(m)},\ b_1 = b$. Tenemos

$$\begin{aligned} a_{m+1} &= a_m a + (1-a_m)(1-b) \\ b_{m+1} &= b_m b + (1-b_m)(1-a) \end{aligned}$$

Estudiemos la primera de estas igualdades recurrentes, ya que la segunda se obtiene cambiando a por b y a_m por b_m. Resulta

$$\begin{aligned} a_{m+1} - a_m &= a_m a + (1-a_m)(1-b) - a_m \\ &= a_m a + (1-a_m)(1-b) - a_{m-1}a - (1-a_{m-1})(1-b) \\ &= (a_m - a_{m-1})a - (1-b)(a_m - a_{m-1}) \end{aligned}$$

y si llamamos $\Delta_m = a_{m+1} - a_m$ obtenemos

$$\begin{aligned} \Delta_m &= (-1+b+a)\Delta_{m-1} \\ &= (-1+b+a)^2\Delta_{m-2} \\ &= (-1+b+a)^{m-1}\Delta_1 \end{aligned}$$

Supongamos por el momento que $0 < b+a < 2$, de modo que $|-1+b+a| < 1$. Obtenemos entonces

$$\begin{aligned} a_{m+1} &= \Delta_m + \Delta_{m-1} + \cdots + \Delta_1 + a \\ &= a + \Delta_1 \sum_{j=0}^{m-1}(-1+b+a)^j \\ &= a + \frac{\Delta_1(-1+b+a)^m - 1}{-2+b+a} \end{aligned}$$

y como

$$\begin{aligned} \Delta_1 = a_2 - a_1 &= a^2 + (1-a)(1-b) - a \\ &= (a-1)(-1+b+a) \end{aligned}$$

entonces

$$a_{m-1} = a + (a-1)(-1+b+a)\frac{(-1+b+a)^m - 1}{-2+b+a}$$

y de manera similar

$$b_{m+1} = b + (b-1)(-1+b+a)\frac{(-1+b+a)^m - 1}{-2+b+a}$$

y los demás términos de la matriz se obtienen como $1 - a_{m+1}$ y $1 - b_{m+1}$.

En particular, haciendo $m \to \infty$ obtenemos

$$\begin{aligned} a_m \to a_\infty &= a + \frac{(a-1)(-1+b+a)}{2-a-b} \\ &= \frac{1-b}{2-a-b} \end{aligned}$$

y la matriz M^m converge a M^∞ donde

$$M^\infty = \begin{pmatrix} \frac{1-b}{2-a-b} & \frac{1-a}{2-a-b} \\ \frac{1-b}{2-a-b} & \frac{1-a}{2-a-b} \end{pmatrix}$$

como se verifica fácilmente luego de algunos cálculos sencillos.

Por consiguiente, si las probabilidades de partir de E_1 o E_2 el primer día eran p y $1-p$, tenemos en el límite

$$(p, 1-p)\begin{pmatrix} \frac{1-b}{2-a-b} & \frac{1-a}{2-a-b} \\ \frac{1-b}{2-a-b} & \frac{1-a}{2-a-b} \end{pmatrix} = \left(\frac{1-b}{2-a-b}, \frac{1-a}{2-a-b}\right)$$

Observamos que el resultado obtenido, que es la probabilidad límite de estar en el primer o segundo estado, es el mismo cualquiera haya sido el vector inicial de probabilidades $(p, 1-p)$.

La distribución límite

$$\left(\frac{1-b}{2-a-b}, \frac{1-a}{2-a-b}\right)$$

se llama *estacionaria* porque tiene la siguiente propiedad:

$$\left(\frac{1-b}{2-a-b}, \frac{1-a}{2-a-b}\right) M = \left(\frac{1-b}{2-a-b}, \frac{1-a}{2-a-b}\right)$$

y por lo tanto es invariante en cualquier número de pasos, o sea que si el día inicial la distribución estacionaria nos da las probabilidades de partir del estado E_1 o del estado E_2, cualquier otro día tendremos la misma distribución de probabilidades.

Resta por estudiar los casos $a+b=2$ y $a+b=0$. Si $a+b=2$, necesariamente $a=b=1$ y resulta

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{de modo que} \quad M^m = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{y} \quad M^\infty = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

pero este es un caso sin interés porque la máquina permanece siempre en el estado del cual parte. Por otro lado, si $a+b=0$, es decir $a=b=0$ resulta

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

y no existe distribución límite porque saltamos con probabilidad 1 de un estado al otro en cada paso.

Señalamos también, que resulta intuitivo que obtenemos de este modo una regulación efectiva del número de artículos de categoría A producidos, al menos a la larga, o sea, después de un número grande de días de aplicar el sistema de decisiones. ▲

3.5. Ejercicios.

1. Sea A, B y C eventos con probabilidad estrictamente positiva. Demuestre las siguientes relaciones:

   a. $P(A^c \cup B^c) = 1 - P(B)P(A|B)$
   b. $P(A \cap B | B \cup C) = P(A \cap B|B)P(B|B \cup C)$
   c. $P(B \cap C|A) = P(C|A)P(B|A \cap C)$ si $P(A \cap C \neq 0$
   d. $P(A|B)P(B|C)P(C|A) = P(B|A)P(C|B)P(A|C)$
   e. $\frac{P(A|A\cup B)}{P(B|A \cup B)} = \frac{P(A)}{P(B)}$

2. Demuestre: $\frac{P(B^c|A)}{P(B)} + \frac{P(A^c)}{P(A)} = \frac{P(A^c|B)}{P(A)} + \frac{P(B^c}{P(B)}$.

3. Un detector de mentiras muestra una señal positiva (indicando una mentira) 10 % de las veces que el sujeto dice la verdad y 94 % de las veces que miente. Si dos personas son sospechosas de un crimen que se sabe ha cometido uno solo de ellos, y ambos dicen ser inocentes, ¿cuál es la probabilidad de que una señal positiva del detector corresponda al culpable?

4. Se obtiene una muestra de cuatro bolas a partir de una bolsa que contiene doce, de las cuales ocho son blancas. Si el muestreo es sin reposición, halle la probabilidad de que la tercera bola sea blanca, si sabemos que la muestra tiene tres bolas blancas. ¿Que sucede si el muestreo se hace con reposición?

5. Se lanza un par de dados simétricos. Calcule la probabilidad de que la suma sea 7 dado que:

   a. La suma es impar,
   b. La suma es mayor que 6,
   c. El resultado del primer dado fue impar,
   d. El resultado del segundo dado fue par,
   e. El resultado de al menos un dado fue impar,
   f. Los dos dados tuvieron el mismo resultado,
   g. Los dos dados tuvieron distintos resultados,
   h. La suma de los dos dados fue 13.

6. Una bolsa contiene cuatro bolas blancas y dos negras y otra contiene tres de cada color. Se escoge una bolsa al azar y luego se selecciona una bola, también al azar. ¿Cuál es la probabilidad de que sea blanca?

7. Una caja contiene 10 focos, cuatro malos y seis buenos. Los focos se prueban de la siguiente manera: se extraen al azar y se prueban sin reemplazarlos. Este proceso se repite hasta localizar los cuatro en mal estado. ¿Cuál es la probabilidad de que el último en mal estado se identifique en la quinta prueba? ¿y en la décima?

8. Cierta vacuna brinda protección parcial contra una enfermedad, de modo que una persona vacunada tiene probabilidad 0.4 de contraer la enfermedad, mientras que para una persona no vacunada esta probabilidad es de 0.8. Si 75 % de la población está vacunada, ¿Cuál es la probabilidad de que una persona que tiene la enfermedad haya sido vacunada?

9. Luego de una serie de pruebas para evaluar un nuevo tipo de examen para detectar cáncer, se ha determinado que 97 % de los pacientes cancerosos de un hospital reaccionan positivamente, mientras que sólo 5 % de aquellos que no tienen cáncer muestran un resultado positivo. Si 2 % de los pacientes del hospital tienen cáncer, ¿cuál es la probabilidad de que un paciente elegido al azar que reacciona positivamente al examen realmente tenga cáncer?

10. Suponga que 5 % de los hombres y 25 de cada 10.000 mujeres son daltónicos. Se escoge un daltónico al azar, ¿cuál es la probabilidad de que sea hombre?

11. Una ferretería tiene tres cajas con igual cantidad de tornillos. Dos de las cajas contienen 5 % de tornillos defectuosos y la otra contiene 10 %. Se escogen dos cajas al azar y se mezclan y de ella se extraen cinco tornillos, uno de los cuales es defectuoso, ¿cuál es la probabilidad de que una de las cajas usadas en la mezcla haya sido la que contenía 10 % de defectuosos?

12. Tres sucursales de una tienda tienen 8, 12, y 14 empleados de los cuales 4, 7 y 10 son mujeres, respectivamente.

    a. Se escoge una sucursal al azar y de ella se escoge un empleado. Si éste es una mujer, ¿cuál es la probabilidad de que ella trabaje en la sucursal con 12 empleados?

    b. Si se escoge un segundo empleado de la misma sucursal, ¿cuál es la probabilidad de que se escoja una mujer?

13. Las señales telegráficas "*punto*" y "*raya*" se envían en proporción 3:4. Debido a ciertas condiciones que causan una transmisión muy errática, un punto se transforma en una raya con probabilidad 1/4, y una raya en un punto con probabilidad 1/3. Si se recibe un punto, ¿cuál es la probabilidad de que haya sido enviado un punto?

14. En una bolsa hay cinco bolas blancas y tres negras y en otra hay tres blancas y siete negras. Se escoge una bolsa al azar y se selecciona una bola, ¿cuál es la probabilidad de que sea blanca?

15. La caja I contiene 50 tornillos y 70 clavos. La caja II contiene 40 tornillos y 20 clavos.

    a. Calcule la probabilidad de extraer un tornillo si se selecciona una caja al azar y luego se extrae un objeto.

    b. Calcule la probabilidad de extraer un tornillo si se mezclan los contenidos de ambas cajas y luego se extrae un objeto.

    c. Si se selecciona una caja al azar y luego se extrae un objeto que resulta ser un tornillo, ¿cuál es la probabilidad de que provenga de la caja I?

16. Lanzamos una moneda repetidamente hasta que obtener sol diez veces. a) ¿Cuál es la probabilidad de no haber obtenido dos águilas en sucesión para ese momento? b) ¿Cuál es la probabilidad de no haber obtenido dos soles en sucesión para ese momento?

17. Se extrae una bola de una caja que contiene cuatro blancas y dos negras. Si la bola es blanca se la deja fuera de la bolsa, mientras que si es negra se la vuelve a colocar dentro. Extraemos luego otra bola. Sea A el evento "la primera bola es blanca" y $B =$ "la segunda bola es blanca". Diga si las siguientes proposiciones son ciertas o falsas:

    a. $P(A) = 2/3$ b. $P(B) = 3/5$
    c. $P(B|A) = 3/5$ d. $P(A|B) = 9/4$
    e. Los eventos A y B son disjuntos.

18. Lanzamos una moneda tres veces y consideramos los siguientes eventos: A: el primer lanzamiento es águila, B: el segundo lanzamiento es sol, C: el tercer lanzamiento es águila, D: los tres lanzamientos son iguales, E: hay exactamente un águila en los tres lanzamientos.

    a. ¿Cuáles de los siguientes eventos son independientes?

    (i) A, B (ii) A, D (iii) A, E (iv) D, E.

    b. ¿Cuáles de los siguientes tríos de eventos son independientes?

    (i) A, B, C (ii) A, B, D (iii) C, D, E.

19. Sean A y B dos eventos disjuntos. Demuestre que si A y B son independientes, alguno de los dos tiene probabilidad 0.

20. De un ejemplo de tres eventos A, B, C tales que $P(A \cap B \cap C) = P(A)P(B)P(C)$ pero $P(A^c \cap B^c \cap C^c) \neq P(A^c)P(B^c)P(C^c)$.

21. Sea A, B y C eventos independientes y $P(C) \neq 0$. Demuestre:

    a. $P(A \cap B|C) = P(A|C)P(B|C)$. b. $P(A \cup B|C) = P(A|C) + P(B|C) - P(A \cap B|C)$.

    c. $P(A|B \cap B) = P(A)$ siempre que $P(B \cap C) \neq 0$.

22. Sean $G = \{1, 2, 3\}$, $H = \{4, 5, 6\}$ Lanzamos dos dados y sean los eventos A : 'el primer dado cae en H, B :'El segundo dado cae en H, C : un dado cae en G y el otro en H, D : el total es cuatro, E : el total es cinco y F : el total es siete. ¿Cuáles de las siguientes proposiciones son ciertas?

    a. A y F son independientes.
    b. A y D son independientes.
    c. A y E son independientes.
    d. $P(A \cap B \cap C) = P(A)P(B)P(C)$.
    e. A y C son independientes.
    f. C y E son independientes.
    g. $P(A \cap C \cap E) = P(A)P(C)P(E)$.
    h. A, C y E son independientes.

23. a. De un ejemplo de tres eventos A, B y C tales que $P(A \cap B \cap C) = P(A)P(B)P(C)$ pero $P(A^c \cap B^c \cap C^c) \neq P(A^c)P(B^c)P(C^c)$.

    b. Demuestre que si $P(A \cap B \cap C) = P(A)P(B)P(C)$; $P(A^c \cap B \cap C) = P(A^c)P(B)P(C)$; $P(A \cap B^c \cap C) = P(A)P(B^c)P(C)$ y $P(A \cap B \cap C^c) = P(A)P(B)P(C^c)$ entonces A, B y C son independientes.

24. Demuestre que si
$$\frac{P(A)}{P(A \cap B)} + \frac{P(B)}{P(A \cap B)} = \frac{1}{P(A)} + \frac{1}{P(B)},$$
entonces A y B son independientes.

25. Lanzamos un dado cinco veces, a. Si el dado sale 1 al menos una vez, ¿cuál es la probabilidad de que salga 1 al menos dos veces? b. Si el primer lanzamiento due 1, ¿cuál es la probabilidad de que salga 1 al menos dos veces?

26. Lanzamos una moneda repetidamente hasta que sol haya ocurrido diez veces. a) ¿Cuál es la probabilidad de que para ese momento no hayan ocurrido dos águilas seguidas? b) ¿Cuál es la probabilidad de que para ese momento no hayan ocurrido dos soles seguidas?

27. Sea $\Omega = \{a, b, c, d, e\}$ y $P(a) = P(b) = 1/8$, $P(c) = P(d) = P(e) = 3/16$. Sean $A = \{a, d, e\}$, $B = \{a, c, e\}$ y $C = \{a, c, d\}$. Demuestre que $P(A \cap B \cap C) = P(A)P(B)P(C)$ pero ningún par de eventos son independientes.

28. Se toma una muestra de cinco transistores producidos por una máquina que en promedio produce 20 % de transistores defectuosos.

    a. Calcule la probabilidad de que si seleccionamos un transistor de la muestra, éste resulte defectuoso.

    b. Suponga que seleccionamos al azar un transistor de la muestra y resulta defectuoso, ¿cuál es la probabilidad de que un segundo transistor seleccionado al azar resulte defectuoso?

29. Si $A_1, \dots, A_n$ son eventos independientes, muestre que
$$P(A_1 \cup A_2 \cup \dots \cup A_n) = 1 - (1 - P(A_1))(1 - P(A_2)) \cdots (1 - P(A_n)).$$

30. a. Si sabemos que una mano de poker tiene al menos tres aces ¿cuál es la probabilidad de que tenga los cuatro? b. Si sabemos que una mano de poker tiene los aces de corazón, trébol y diamante ¿cuál es la probabilidad de que también tenga el as de pica? c. Halle la probabilidad de que una mano de poker tenga los dos aces negros dado que tiene al menos tres aces.

31. En un closet hay tres pares de medias blancas y dos de media negras. Las medias con todas nuevas y del mismo estilo y tamaño. Si seleccionamos dos medias al azar, ¿cuál es la probabilidad de que formen un par? Resuelva el mismo problema cambiando media por zapatos.

32. En un closet tenemos seis pares diferentes de zapatos. Si sacamos cinco zapatos al azar ¿Cuál es la probabilidad de obtener al menos un par de zapatos?

33. Una muestra de tamaño 4 se extrae con reposición de una bolsa que contiene 6 bolas, de las cuales 4 son blancas. Sea A el evento "exactamente una de las dos primeras bolas extraídas es blanca" y sea $B =$ "la cuarta bola es blanca". ¿Son A y B independientes? ¿Qué sucede si el muestreo se realiza sin reposición?

34. Considere de nuevo el ejercicio anterior y definamos C como el evento "exactamente dos de las bolas extraídas son blancas" ¿Son A, B y C independientes? ¿Son B y C independientes?

35. Un fabricante de motos inscribe tres corredores en una carrera. Sea $A_i, i = 1, 2, 3$ el evento definido por "el i-ésimo corredor finaliza la carrera en alguno de los tres primeros lugares". Si los eventos A_1, A_2 y A_3 son independientes y $P(A_1) = P(A_2) = P(A_3) = 0,1$ calcule las probabilidades de los siguientes eventos:

    a. Ninguno de los corredores termina entre los tres primeros.

    b. Al menos uno termina entre los tres primeros.

    c. Al menos dos terminan entre los tres primeros.

    d. Todos terminan entre los tres primeros.

36. Sean A y B eventos independientes tales que con probabilidad $1/6$ ocurren simultáneamente, y con probabilidad $1/3$ ninguno de ellos ocurre. Halle $P(A)$ y $P(B)$. ¿Están determinadas de forma única estas probabilidades?

37. ¿Cuál es el menor valor de n para el cual la probabilidad de obtener al menos un 6 en una serie de n lanzamientos de un dado es mayor que $3/4$?

38. Los eventos $A_1, A_2, \dots$ son independientes y $P(A_j) = p$, $j = 1, 2, \dots$ Halle el menor valor de n para el cual $P(\cup_1^n A_k) \geq p_0$ donde p_0 es un número fijo.

39. Una máquina consiste de 4 componentes conectados en paralelo, de modo que la máquina falla sólo si los cuatro componentes fallan. Supongamos que los componentes son independientes entre sí. Si los componentes tienen probabilidades 0.1; 0.2; 0.3 y 0.4 de fallar cuando la máquina es encendida, ¿cuál es la probabilidad de que funcione al encenderla?

40. En una fábrica de computadoras un inspector revisa un lote de 20 máquinas y encuentra que 3 de ellas necesitan ajustes antes de empacarlas. Otro empleado descuidado mezcla las computadoras que habían sido revisadas, de modo que el inspector tiene que hacerlo otra vez. a. ¿Cuál es la probabilidad de que no haga falta probar más de 17 computadoras? b. ¿Cuál es la probabilidad de que haga falta probar exactamente 17 computadoras?

41. Supongamos que A y B son independientes, y B y C son independientes.

    a. ¿Son A y C independientes en general?

    b. ¿Es B independiente de $A \cup C$?

    c. ¿Es B independiente de $A \cap C$?

42. Se lanza una moneda balanceada n veces. Muestre que la probabilidad condicional de águila en cualquier lanzamiento específico dado que hay k águilas en n lanzamientos es k/n $(k > 0)$.

43. Suponga que A y B son eventos tales que $P(A|B) = P(B|A)$, $P(A \cup B) = 1$ y $P(A \cap B) > 0$. Demuestre que $P(A) > 1/2$.

44. Un arquero acierta al centro de la diana con probabilidad 0.9. a) ¿Cuál es la probabilidad de que logre acertar 8 veces si lanza 10 flechas? b) ¿Si lanza flechas hasta acertar 8 veces al centro de la diana, ¿Cuál es la probabilidad de que necesite a lo sumo lanzar 10 flechas?

45. Considere estos dos sistemas de ensayos de Bernoulli: 1) Lanzamos una moneda y águila es éxito; 2) Lanzamos un dado y 6 es éxito. Para cada uno de estos casos calcule $P(A)/P(B)$, donde A: 'el tercer éxito ocurre en el quinto ensayo', B: 'tres de los primeros cinco ensayos resultan en éxito'. Generalice reemplazando 3 por i y 5 por j.

46. Lanzamos cuatro dados, ¿Cuál es la probabilidad de que el mayor número sea 5 y el menor sea 3?

47. La caja A tiene dos bolas rojas y tres negras. La caja B tiene cinco rojas y una blanca. Se selecciona una bola al azar de la caja A y se coloca en la caja B y luego se escoge una bola al azar de la caja B. a) ¿Cuál es la probabilidad de que ambas bolas sean rojas? b) ¿Cuál es la probabilidad de que la segunda bola sea roja? c) Dado que la segunda bola es roja ¿Cuál es la probabilidad de que la primera también haya sido roja? d) Dado que la segunda bola es blanca ¿cuál es la probabilidad de que la primera haya sido roja?

48. Una fábrica produce 300 automóviles al día. La fábrica compra baterías de dos proveedores. La compañía A le vende 100 baterías al día, de las cuales 99 % funcionan correctamente. Las otras 200 baterías son producidas por la compañía B, de las cuales 5 % son defectuosas. Si seleccionamos un auto al azar de la producción de un día y la batería es defectuosa ¿cuál es la probabilidad de que haya sido producida por la empresa B?

49. Un empleado debe verificar el funcionamiento de una máquina que produce tornillos al inicio del día. Esta máquina necesita repararse una vez cada 10 días, en promedio y cuando necesita repararse, todos los tornillos que produce son defectuosos. Cuando la máquina trabaja adecuadamente, 5 % de los tornillos producidos son defectuosos y aparecen al azar a lo largo de la producción del día. ¿Cuál es la probabilidad de que la máquina esté funcionando bien si a. el primer tornillo que el inspector revisa es defectuoso? b. los dos primeros tornillos que el inspector revisa son defectuosos? c. los tres primeros son defectuosos?

50. Una carta de un juego de naipes se ha perdido. Trece cartas se extraen de las 51 restantes y resultan ser tres diamantes, dos picas, cuatro corazones y cuatro tréboles. Halle la probabilidad de que la carta perdida sea de cada una de las pintas.

51. Resuelva el problema de los puntos cuando los dos jugadores tienen probabilidades p y $q = 1 - p$ de ganar.

52. **Las cajas de cerillos de Banach.** Una persona tiene dos cajas de n cerillos, una en el bolsillo derecho y otra en el izquierdo. Cuando necesita un cerillo escoge una caja al azar hasta que se encuentra una caja vacía. ¿Cuál es la probabilidad de que la otra caja tenga k cerillos? (El celebre matemático polaco Stefan Banach solía reunirse con otros matemáticos en el Café Escocés en Lwów, Polonia, en donde había un cuaderno en el cual se anotaban los problemas planteados y sus soluciones. Esta libreta se conoce como el Libro Escocés. El problema anterior es el último problema incluido en este libro).

53. **La paradoja de Galton.** Si lanzamos tres monedas al menos dos de ellas son iguales, y la tercera tiene probabilidad 1/2 de caer águila o sol, de modo que la probabilidad de que las tres sean iguales es 1/2. En realidad la probabilidad de que las tres monedas sean iguales es 1/4. ¿Qué está mal en el razonamiento anterior?

54. **El modelo de Pólya.** El matemático polaco G. Pólya propuso el siguiente modelo para el proceso de contagio. Comenzamos con una caja que contiene una bola blanca y otra negra. Cada segundo escogemos una bola al azar de la caja, la reemplazamos y añadimos otra del mismo color. Haga una simulación d este proceso con una computadora y trata de hacer una predicción sobre la proporción de bolas blancas luego de un tiempo largo. ¿Es cierto que las proporciones de bolas de un color dado tienen una tendencia a estabilizarse?

55. En una colección de 65 monedas una de ellas tiene dos águilas y el resto son monedas normales. Si seleccionamos una moneda al azar y al lanzarla sale águila seis veces seguidas, ¿Cuál es la probabilidad de que hayamos escogido la moneda con dos águilas?

56. Te dan dos cajas y cincuenta bolas, la mitad blancas y la otra mitad negras. Debs distribuir las bolas en las cajas sin restricciones pero quieres maximizar la probabilidad de obtener una bola blanca si escoges una caja al azar y luego una bola al azar. ¿Cómo debes distribuir las bolas? Justifica tu respuesta.

57. Lanzamos una moneda n veces y obtenemos k soles. Demuestra que la probabilidad condicional de obtener un sol en cualquier lanzamiento específico, dado que hay k soles en total es k/n.

58. **La paradoja de Simpson.** Un fabricante de focos tiene dos plantas. La planta A vende lotes de focos que consisten de 1000 focos regulares y 2000 focos ahorradores. A través de pruebas de control de calidad se sabe que, en promedio, hay 2 focos regulares y 11 ahorradores defectuosos por lote. En la planta B se venden lotes de 2000 focos regulares y 1000 ahorradores, y en promedio hay 5 regulares y 6 ahorradores defectuosos por lote.

    El gerente de la planta A afirma que ellos son más eficientes pues sus tasas de focos defectuosos son 0.2 % y 0.55 % mientras que para la otra planta son 0.25 % y 0.6 %. Por su parte el gerente de la planta B responde diciendo 'cada lote de 3000 focos que producimos contiene 11 focos defectuosos, comparado con 13 defectuosos para los focos producidos por A, de modo que nuestra tasa de 0.37 % de focos defectuosos es inferior a la de ellos, que es 0.43 %. ¿Quién tiene la razón?

59. El evento A atrae al evento B si $P(B|A) > P(B)$ (ver el ejercicio 3.4.1). a) ¿Es transitiva esta relación? b) Demuestre que si A atrae a B y a C pero repele a $B \cap C$, entonces A atrae a $B \cup C$. c) ¿Es posible que A atraiga a B y a C pero repela a $B \cup C$? d) Demuestre que si $B_1, \dots B_n$ es una partición dl espacio muestral y A atrae algún B_j entonces debe repeler a algún B_i.

60. Demuestre que si $A_1, \dots A_n$ son eventos independientes definidos en un espacio muestral Ω y $0 < P(A_i) < 1$ para todo i, entonces Ω debe tener al menos 2^n puntos.