Probabilidad

Àngel J. Gil Estallo

P08/75057/02304

Universitat Oberta
de Catalunya
www.uoc.edu

Índice

Introducción a la probabilidad

1. Introducción

Comencemos por introducir algunos ejemplos y preguntas que nos puedan guiar en el desarrollo posterior de la teoría; para hacerlo, nos basaremos en un ejemplo muy sencillo: el estudio del comportamiento de un dado.

Lanzar un dado y observar cuál es el número de puntos que aparece en la cara superior es lo que se denomina un **experimento aleatorio**, ya que, si bien sabemos cuáles son los posibles resultados (que salgan 1, 2, 3, 4, 5 ó 6 puntos), no podemos saber cuál será el número de puntos que saldrá en cada tirada particular. El conjunto de todos los resultados posibles se denomina **espacio muestral** y se suele designar con la letra Ω.

Experimento aleatorio

Seguro que cuando lancemos el dado siempre saldrá un número de puntos menor o igual que 6, pero no podemos decir exactamente cuántos puntos conseguiremos en cada tirada.

> **Experimento aleatorio** es aquel que tiene diferentes resultados posibles, de los que no tenemos certeza sobre cuál se producirá realmente. Además, es preciso que el experimento se pueda repetir en condiciones idénticas tantas veces como sea necesario.
>
> **Espacio muestral** es el conjunto de resultados posibles que podemos obtener al realizar un experimento aleatorio. Se designa por la letra Ω.

Ω = espacio muestral.
ω_i = resultado posible.

A partir de este momento, y si no se dice lo contrario, consideraremos que el espacio muestral es finito y que tenemos k resultados posibles: $\omega_1, \omega_2, \ldots, \omega_k$.

Es importante fijar desde el primer momento cuál es el espacio muestral asociado al experimento aleatorio.

Ejemplos de espacio muestral

En el caso del dado, el espacio muestral es:

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

Ver qué sale si lanzamos una moneda al aire es un experimento aleatorio (los resultados posibles son cara o cruz, pero en cada tirada no podemos decir cuál de los dos saldrá) con espacio muestral:

$$\Omega = \{\text{cara, cruz}\}$$

2. Acontecimientos o sucesos aleatorios

> Un **acontecimiento** o **suceso aleatorio** es cualquier subconjunto del espacio muestral. Es decir, cualquier subconjunto del conjunto de resultados de un experimento aleatorio.

Denotaremos los acontecimientos por letras mayúsculas A, B…

Podemos preguntarnos ahora cuáles son los acontecimientos más simples de todos: evidentemente, serán aquellos que constan de un único resultado; son los llamados **acontecimientos elementales**.

Identificaremos los resultados del experimento aleatorio con los **acontecimientos elementales**, que son los acontecimientos formados por un único resultado.

En general, los acontecimientos contendrán más de un resultado, y muchas veces nos interesará conocer el número de resultados que contienen.

Notación

Card( ) es la función **cardinal** de un conjunto.

En otros textos Card(A) se denota por #A o por $n(A)$.

Card(A) denotará el número de resultados que contiene el acontecimiento A.

Así pues, después de realizar un experimento aleatorio, tenemos acontecimientos de dos tipos:

a) Acontecimientos elementales: son estrictamente los que podemos obtener como resultado del experimento.

b) Acontecimientos: son agrupaciones de uno o más resultados. A menudo podremos dar una descripción del acontecimiento a partir de una característica común a todos los resultados del acontecimiento (del tipo "ser par", "ser primero", etc.).

El acontecimiento "ser par"

Los dados no saben si los números que salen son pares o no. Tampoco llevan escrita la palabra *par*. El hecho de que el número que aparece en un dado sea par se corresponde a la descripción que hacemos de cierta situación que no se corresponde a un único valor del dado. Por tanto, ¡ser par no es un resultado!

Así pues, si el acontecimiento contiene un único resultado, diremos que se trata de un acontecimiento *elemental*; si contiene más de uno, diremos simplemente que es un acontecimiento.

Ejemplos de resultados y acontecimientos

a) En el caso del dado, el conjunto de posibles resultados (espacio muestral) es $\Omega = \{1, 2, 3, 4, 5, 6\}$ y podemos tener, entre otros, los sucesos siguientes:

- B = "el número de puntos es par", que corresponde a la agrupación de los resultados $\{2, 4, 6\}$; es decir, en el fondo $B = \{2, 4, 6\}$. Evidentemente, Card(B) = 3.
- $C = \{3, 4, 5, 6\}$. En este caso el suceso C corresponde a "sacar un valor superior o igual a 3"; Card(C) = 4.
- $D = \{1, 6\}$. D es un suceso formado por los resultados 1 y 6.
- $F = \{3\}$. F es el suceso que corresponde a "sacar un 3".

b) "Sacar un número mayor que 5" es un acontecimiento elemental o resultado, ya que se corresponde a sacar un 6. "Sacar un número par" no es un suceso elemental, ya que no se corresponde con un único valor concreto del dado.

Los **resultados favorables** a un suceso son los resultados que contiene. Card(A) es el número de resultados favorables a A.

A continuación definiremos dos sucesos muy especiales: el suceso seguro y el imposible.

2.1. El suceso seguro

El **suceso seguro** es el que está formado por todos los resultados posibles; es decir, es el mismo espacio muestral Ω.

El suceso seguro contiene todos los resultados; así, pase lo que pase, "seguro" que cualquier resultado pertenece al suceso "seguro".

Ejemplo de un suceso seguro

Lanzamos un dado y consideramos el suceso "sacar un número menor que 10". ¡Esto seguro que siempre pasa!

2.2. El suceso imposible

El **suceso imposible**, denotado por $\emptyset$ (**conjunto vacío**), es el suceso que no ocurre nunca. Evidentemente, $\text{Card}(\emptyset) = 0$.

Ejemple de suceso imposible

Lanzamos un dado y consideramos el suceso "sacar un 26". ¡Seguro que eso no pasa nunca! Otro ejemplo de suceso imposible es "sacar un número par y múltiplo de 5".

3. Operaciones con sucesos

Acabamos de ver que podemos determinar un suceso por medio del conjunto de resultados posibles de un experimento aleatorio. Evidentemente, la descripción puede ser muy compleja y a veces interesa mezclar u operar ciertos sucesos para describir situaciones más complicadas. También hemos visto que un acontecimiento es, de hecho, el conjunto de los resultados que contiene: por tanto, todas las propiedades de los conjuntos y de las operaciones con conjuntos son válidas para sucesos.

A continuación describiremos varias operaciones conjuntivas aplicadas a los sucesos aleatorios; en concreto, hablaremos de la unión, de la intersección y del complementario de sucesos.

3.1. Intersección de sucesos y sucesos incompatibles

Supongamos que tenemos dos sucesos A y B:

El suceso $A \cap B$ (leído *A intersección B*) está formado por aquellos resultados favorables a A y a B simultáneamente.

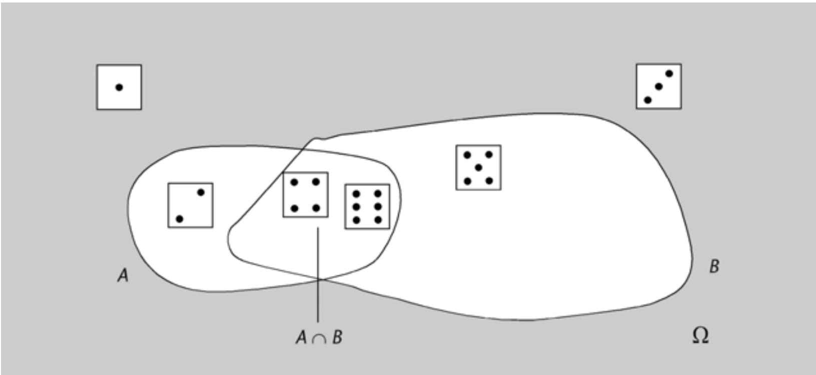
Es decir, la intersección de los sucesos A y B es un suceso menor que A y que B, en el que sólo aparecen los resultados que están en A y también en B. Para ilustrar gráficamente estas operaciones, nos será útil utilizar los conocidos diagramas de Venn.

Número de elementos de $A \cap B$

No existe una fórmula que nos diga el número de elementos de $A \cap B$ a partir de los elementos de A y de B. En la mayoría de los casos deberemos mirar directamente cuántos resultados hay en $A \cap B$. Lo seguro es que $\text{Card}(A \cap B) \leq \text{Card}(A)$ y $\text{Card}(A \cap B) \leq \text{Card}(B)$.

Ejemplo de intersección de sucesos

Si el suceso A es "sacar un número par" y el suceso B es "sacar un número mayor que 3", entonces $A = \{2, 4, 6\}$ y $B = \{4, 5, 6\}$ y, evidentemente, $A \cap B = \{$"sacar un número par mayor que 3"$\} = \{4, 6\}$, que son los únicos resultados que están al mismo tiempo en A y en B. Gráficamente:



3.1.1. Sucesos incompatibles

Más adelante será muy importante saber si dos sucesos tienen resultados en común o no; esto conduce a la definición de sucesos incompatibles.

Dos acontecimientos son incompatibles si no tienen ningún resultado en común, es decir, si $A \cap B$ es el suceso imposible. Dicho de otra manera, A y B son incompatibles si son conjuntos disyuntos, es decir, si $A \cap B = \emptyset$. En este caso $\text{Card}(A \cap B) = 0$.

Ejemplo de sucesos incompatibles

Los sucesos "sacar un número menor que 2" y "sacar un número par mayor o igual que 4" son sucesos incompatibles.

3.2. Unión de sucesos

Supongamos que tenemos dos sucesos A y B. El suceso $A \cup B$ (leído *A unión B*) está formado por aquellos resultados favorables a A, a B o a ambos a la vez.

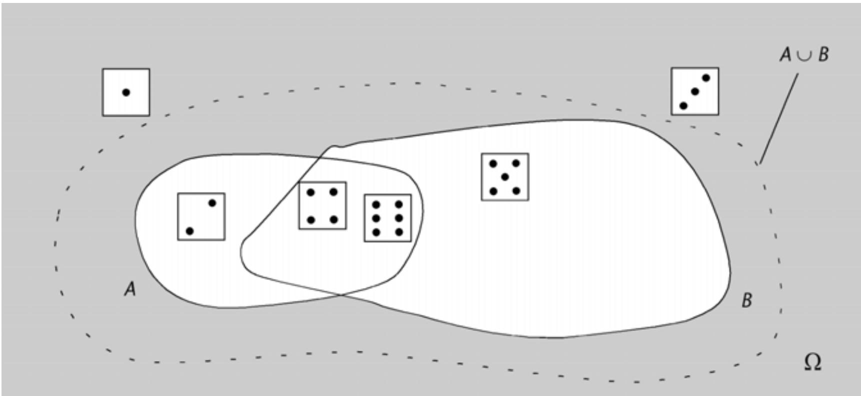
Es decir, unir los sucesos A y B sirve para crear un suceso mayor que contiene los resultados de A más los resultados de B. Es fácil ver que:

$$\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B) - \text{Card}(A \cap B)$$

Observad que en Card(A) + Card(B) cuentan dos veces los elementos de la intersección (una vez por pertenecer a A y una vez por pertenecer a B); en Card(A) + + Card(B) – Card($A \cap B$), cada elemento de la unión está contado una sola vez.

Ejemplo de unión de sucesos

Si el suceso A es "sacar un número par" y el suceso B es "sacar un número mayor que 3", entonces $A = \{2, 4, 6\}$ y $B = \{4, 5, 6\}$ y, $A \cup B$ = "sacar un número par o mayor que 3" $= \{2, 4, 5, 6\}$ y $A \cap B = \{4, 6\}$. Así, Card($A \cup B$) = 4, que coincide con Card(A) + + Card(B) – Card($A \cap B$), ya que Card(A) = 3, Card(B) = 3 y Card($A \cap B$) = 2. Gráficamente:



3.3. Complementario de un suceso

Supongamos que estudiamos un suceso A.

El suceso A^C (leído **complementario** de A) está formado por aquellos resultados que no son favorables a A.

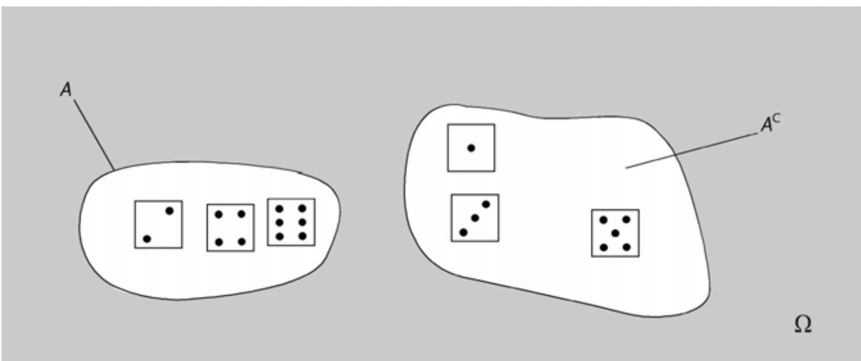
Notación

A veces se escribe $\bar{A}$ o $C(A)$ en lugar de A^C.

Es decir, el complementario del suceso A es el suceso que únicamente contiene los resultados que no están en A.

Ejemplo de suceso complementario

Si $B = \{1, 2, 5\}$, entonces $B^c = \{3, 4, 6\}$, precisamente los resultados que no están en B. Si el suceso es A = "sacar un número par", entonces $A = \{2, 4, 6\}$ y $A^C = \{1, 3, 5\}$, que se corresponde, como cabía esperar, con los números impares. Gráficamente:



Evidentemente, el número de elementos de A^C es el número de posibles resultados menos el número de elementos de A, es decir:

$$\text{Card}(A^C) = \text{Card}(\Omega) - \text{Card}(A)$$

Complementario e incompatibilidad

A y A^c siempre son sucesos disyuntos, ya que no puede ser que un resultado esté en A y fuera de A simultáneamente. Por ejemplo, un número no puede ser par e impar simultáneamente.

3.4. Tablas de sucesos

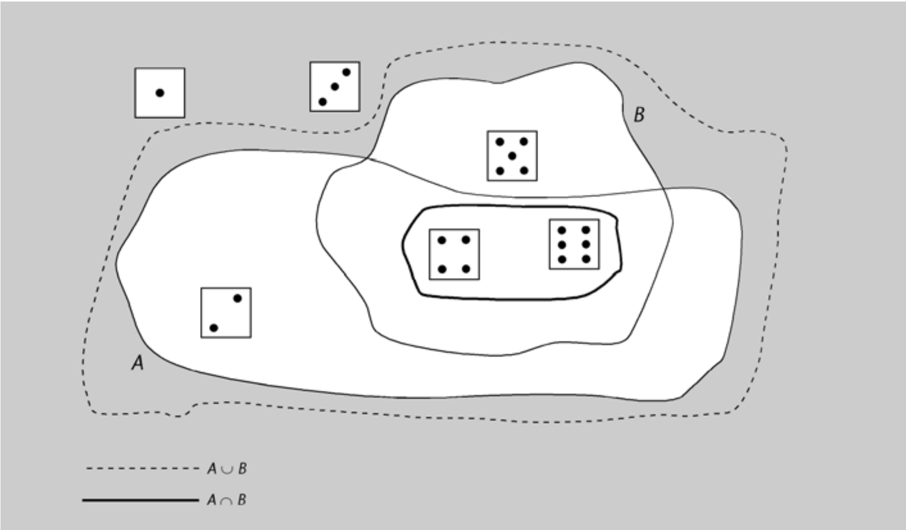
En ocasiones nos puede resultar útil confeccionar una tabla en la que se vean claramente los resultados que pertenecen a cada suceso y a ciertas operaciones con los sucesos. Por ejemplo, en el caso del dado y si el suceso A es "sacar un número par" y el suceso B es "sacar un número mayor que 3", entonces $A = \{2, 4, 6\}$ y $B = \{4, 5, 6\}$, podemos construir una tabla como ésta, en la que vamos marcando qué resultado pertenece a cada suceso:

Posible resultado: Valor del dado	¿Es de A?	¿Es de B?	¿Es de $A \cup B$?	¿Es de $A \cap B$?	¿Es de A^C?
1	No	No	No	No	Sí
2	Sí	No	Sí	No	No
3	No	No	No	No	Sí
4	Sí	Sí	Sí	Sí	No
5	No	Sí	Si	No	Sí
6	Sí	Sí	Sí	Si	No

En dicha tabla, como podéis observar, para tener...

- "Sí" a la pregunta "¿Es de $A \cup B$?", debéis tener como mínimo un "Sí" (o bien a "¿Es de A?" o bien a "¿Es de B?");
- "Sí" a la pregunta "¿Es de $A \cap B$?", debemos tener un "Sí" a "¿Es de A?" y también un "Sí" a "¿Es de B?";
- "Sí" a la pregunta "¿Es de A^C?", debemos tener un "No" a "¿Es de A?".

Podemos visualizar el contenido de la tabla mediante un gráfico, en el que podemos ver que el suceso $A \cup B$ engloba los resultados de A y los de B, mientras que el suceso $A \cap B$ contiene exclusivamente los resultados que están en A y también en B.



A es "sacar un número par" y el suceso B es "sacar un número mayor que 3".

4. Resumen

En esta sesión hemos introducido el concepto de experimento aleatorio, a partir del cual se desarrollará la teoría de la probabilidad. Todo experimento aleatorio tiene asociado el conjunto de sus resultados, llamado *espacio muestral*. Los resultados se pueden agrupar en sucesos (también llamados *acontecimientos*), que se denominan *elementales* si contienen un único resultado. Se insiste en la noción de resultado favorable a un suceso (de hecho, son los resultados que forman parte del mismo).

Después se recuerdan las operaciones de unión, intersección y complementario de conjuntos, aplicadas al caso de los experimentos aleatorios. Se tratan dos tipos particulares de sucesos: el **suceso seguro** y el **suceso imposible**. A partir de este último se definen los **sucesos incompatibles**, que son aquéllos en los que su intersección está vacía (es decir, su intersección es el suceso imposible). Para favorecer la visualización de los sucesos y las operaciones que se pueden realizar sobre éstos, se utilizan los **diagramas de Venn** y las **tablas de sucesos**.

Ejercicios

1. Supongamos que hemos examinado el sistema operativo y el procesador de diez ordenadores de nuestra empresa. Los resultados han sido los siguientes:

Ordenador	Sistema operativo	Procesador
1	Doors95	Lenteron
2	Doors98	FortiumII
3	Doors95	Lenteron
4	Doors2000	FortiumII
5	Doors95	Lenteron
6	Doors98	Lenteron
7	Doors95	Lenteron
8	Doors98	FortiumII
9	Doors95	FortiumII
10	Doors2000	FortiumII

Consideremos los sucesos A y B, que son:

- A = "el sistema operativo es Doors98"
- B = "el procesador es FortiumII"

a) Interpretad la situación en términos de experimentos aleatorios y determinad el espacio muestral correspondiente.
b) Determinad qué resultados forman los sucesos $A \cup B$, $A \cap B$, A^c, B^c, $A^c \cap B$, $A^c \cup B$ y dad una descripción de cada uno.

2. Lanzamos un dado dos veces:
a) Calculad el espacio muestral del experimento y el número de resultados posibles.
b) Determinad cuáles son los sucesos siguientes y el número de resultados favorables a cada uno de éstos:

- A = "las dos tiradas sale el mismo número"
- B = "la suma de los dos número es mayor que 7"
- $A \cap B$, $A \cup B$, B^C

Solucionario

1.
El experimento aleatorio consiste en escoger un ordenador al azar y, por tanto, cada uno de los diez ordenadores es un resultado posible. Es decir, $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, si suponemos que tenemos los ordenadores numerados del 1 al 10.

Podemos construir la tabla de sucesos siguiente:

Ordenador	¿Es de A?	¿Es de B?	¿Es de $A \cup B$	¿Es de $A \cap B$	¿Es de A^c?	¿Es de B^c?	¿Es de $A^c \cap B$?	¿Es de $A^c \cup B$?
1	No	No	No	No	Sí	Sí	No	Sí
2	Sí	Sí	Sí	Sí	No	No	No	Sí
3	No	No	No	No	Sí	Sí	No	Sí
4	No	SÍ	Sí	No	Sí	No	Sí	Sí
5	No	No	No	No	Sí	Sí	No	Sí
6	Sí	No	Sí	No	No	Sí	No	No
7	No	No	No	No	Sí	Sí	No	Sí
8	Sí	Sí	Sí	Sí	No	No	No	Sí
9	No	SÍ	Sí	No	Sí	No	Sí	Sí
10	No	SÍ	Sí	No	Sí	No	Sí	Sí

Y también podemos dar una descripción de cada uno de los sucesos:

a) $A \cup B$ = "el sistema es Doors98 o tiene FortiumII" = {2, 4, 6, 8, 9, 10}

b) $A \cap B$ = "el sistema es Doors98 y tiene FortiumII" = {2, 8}

c) A^c = "el sistema no es Doors98" = {1, 3, 4, 5, 7, 9, 10} (en este caso el sistema será o Doors95 o Doors2000).

d) B^c = "no tener FortiumII" = {1, 3, 5, 6, 7} (en este caso será tener Lenteron)

e) $A^c \cap B$ = "el sistema no es Doors98 y tiene FortiumII" = {4, 9, 10}; por tanto, será Doors95 o Doors2000 con FortiumII.

f) $A^c \cup B$ = "el sistema no es Doors98 o tiene FortiumII" = {1, 2, 3, 4, 5, 7, 8, 9, 10}; por tanto, será Doors95 o Doors2000 o tendrá FortiumII .

2.

El espacio muestral se puede distribuir en forma de tabla en la que la fila representa el valor de la primera tirada y la columna, el valor de la segunda.

	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Por tanto, hay 36 resultados posibles, es decir, $\text{Card}(\Omega) = 36$.

Los sucesos favorables a A son:

$$\{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)\}$$

y $\text{Card}(A) = 6$.

Los sucesos favorables a B son:

$$\{(2,6), (3,5), (3,6), (4,4), (4,5), (4,6), (5,3), (5,4),$$
$$(5,5), (5,6), (6,2), (6,3), (6,4), (6,5), (6,6)\}$$

y, por tanto, $\text{Card}(B) = 15$.

Inspeccionando los sucesos A y B, tenemos que $A \cap B = \{(4,4), (5,5), (6,6)\}$ y $\text{Card}(A \cap B) = 3$. Ahora tenemos que:

$$A \cup B = \{(1,1), (2,2), (3,3), (2,6), (3,5), (3,6), (4,4), (4,5),$$
$$(4,6), (5,3), (5,4), (5,5), (5,6), (6,2), (6,3), (6,4), (6,5), (6,6)\}$$

y $\text{Card}(A \cup B) = 18$, también podemos calcular de forma indirecta:

$$\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B) - \text{Card}(A \cap B) = 6 + 15 - 3 = 18$$

$$B^C = \{(1,1), (1,2),(1,3), (1,4), (1,5), (1,6), (2,1),(2,2), (2,3), (2,4),$$
$$(2,5), (3,1), (3,2), (3,3), (3,4), (4,1), (4,2), (4,3), (5,1), (5,2), (6,1)\}$$

que, evidentemente, corresponden al caso de que la suma de los dados sea menor o igual que 7; en total, $\text{Card}(B^C) = 21$, número que también se puede calcular haciendo:

$$\text{Card}\ (B^C) = \text{Card}(\Omega) - \text{Card}(B) = 36 - 15 = 21$$

Combinatoria y técnicas de recuento

En el estudio de un experimento aleatorio es importante contar los resultados que forman parte de cierto suceso, así como contar los posibles resultados que se pueden obtener al llevar a cabo el experimento. En esta sesión recordaremos brevemente algunas técnicas básicas de combinatoria para poder tratar las situaciones más comunes. En concreto trabajaremos:

1) La regla del producto

2) Las variaciones

3) Las permutaciones

4) Las combinaciones y los números combinatorios

En general, se tratará de, dados unos cuantos objetos, construir agrupaciones utilizando sólo dichos objetos y teniendo en cuenta los aspectos siguientes:

1) ¿Se tienen que agrupar todos los objetos disponibles o sólo unos cuántos?

2) ¿Podemos repetir los objetos en las agrupaciones o bien en éstas los objetos deben ser todos diferentes?

3) ¿Importa el orden? Es decir, ¿interesa considerar las diferentes posibilidades de ordenación dentro de cada agrupación o son indiferentes?

Según la respuesta a estas preguntas, deberemos utilizar alguna de las técnicas que comentamos a continuación.

1. La regla del producto

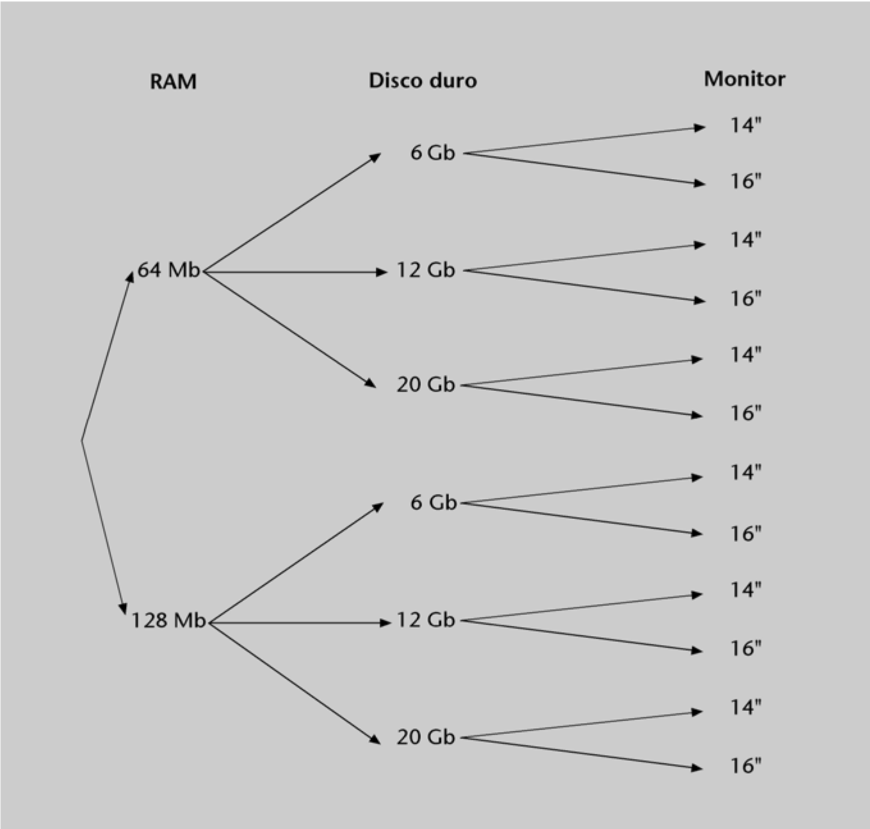
En este apartado enunciaremos una regla general de mucha utilidad en numerosos casos.

> Para determinar el total de posibilidades que tenemos si debemos efectuar una secuencia de elecciones sucesivas, hay que multiplicar el número de posibilidades de la primera elección por el número de posibilidades de la segunda elección, y así sucesivamente.

Estos procesos en los que hacemos elecciones sucesivas se pueden representar fácilmente en forma de árboles, en los que de cada nodo salen tantas ramas como posibilidades tenemos, tal como veremos en el ejemplo siguiente.

Ejemplo de aplicación de la regla del producto

Nuestra empresa monta ordenadores y los clientes pueden escoger algunas de las características del equipo: concretamente, pueden escoger entre 64 ó 128 MB de memoria RAM, disco duro de 6, 12 ó 20 GB y pantalla de 14 ó de 16 pulgadas, opciones que se pueden mezclar de todas las maneras posibles. ¿Cuántos modelos diferentes de ordenador podemos ofrecer a nuestros clientes? Mediante un diagrama de árbol podemos disponer fácilmente de todas las configuraciones posibles.



Se ve claramente que en total podemos ofrecer $2 \cdot 3 \cdot 2 = 12$ configuraciones diferentes en nuestros equipos.

2. Variaciones

Supongamos que tenemos N objetos y nos piden que escogamos k de éstos, de manera que no haya dos repetidos y que cualquier cambio en la ordenación dé lugar a otro grupo diferente: ¿de cuántas maneras podemos hacerlo?

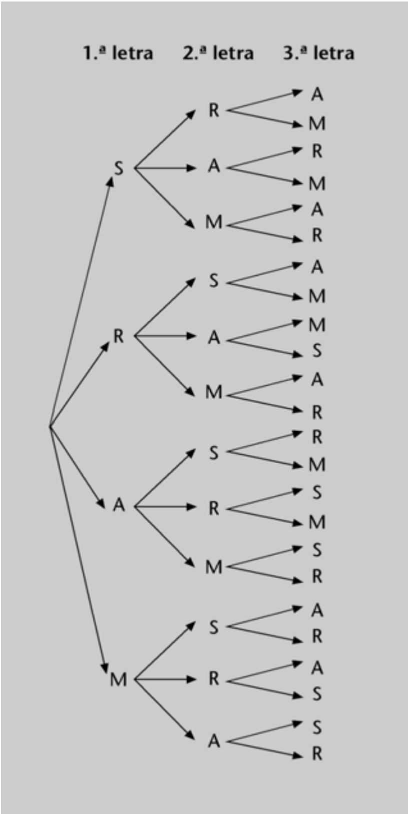
Características de las variaciones

En las variaciones no es preciso que agrupemos todos los objetos de golpe, no podemos repetir los objetos y, además, importa el orden.

Cada una de las maneras en las que podemos escoger y ordenar k elementos de entre N dados –sin repetir ninguno– es una de las **variaciones** de los N elementos cogidos de k en k.

Variaciones de las letras de la palabra *SRAM*

Supongamos que queremos considerar todas las posibles palabras (en el sentido de secuencias de letras, aunque no tengan sentido) de tres letras que se pueden formar usando sólo letras de la palabra *SRAM* (por tanto, $N = 4$). Lo haremos mediante un árbol en el que se vea que a cada paso tenemos una posibilidad de elección menos y que necesitamos tres pasos para construir los subconjuntos ordenados de tres elementos. Observad que, de hecho, es como si escogiésemos tres letras de la palabra *SRAM* y después las ordenásemos de todas las formas posibles, ya que ordenaciones diferentes dan lugar a "palabras" diferentes.



Observamos que en el primer paso tenemos $N = 4$ posibilidades de elección, en el segundo $N - 1 = 3$ y en el tercero $N - 2 = 2$; en total, tenemos $N(N - 1)(N - 2) = 4 \cdot 3 \cdot 2 = 24$ palabras diferentes.

Observamos que tenemos que hacer k elecciones y que en cada una tenemos una posibilidad menos que en la anterior. Aplicando la regla del producto, podemos saber fácilmente el número total de posibles agrupaciones.

El **número de las variaciones** de N elementos cogidos de k en k se obtiene haciendo el producto de k factores a partir de N, restando cada vez una unidad:

$$N(N-1)\ldots(N-k+1) = \frac{N!}{(N-k)!}$$

(donde hay k factores decrecientes a partir de N).

Ejemplo

La revista *PC Universe* nos envía una lista de veintitrés portátiles y nos pide que devolvamos la lista de los que creamos que son los cinco mejores, ordenada de la forma: primero, segundo, tercero, cuarto y quinto. En este caso tenemos que ordenar 5 (= k) portátiles como mejores portátiles entre 23 (= N) posibles y, por tanto, tenemos:

$N(N-1)\ldots(N-5+1) = 23 \cdot 22 \cdot 21 \cdot 20 \cdot 19 = 4.037.880$ posibles respuestas diferentes.

2.1. Variaciones con repetición

Si nos encontramos en la situación descrita por las variaciones, pero con la particularidad de que se pueden repetir los objetos tantas veces como sea necesario, nos encontraremos con un problema de **variaciones con repetición**. La diferencia con las variaciones normales es que en cada elección podemos volver a escoger entre todos los objetos iniciales (se pueden repetir).

El **número de las variaciones con repetición** de N elementos cogidos de k en k se obtiene haciendo el producto de k factores iguales a N, es decir, es igual a:

$$N^k$$

Características de las variaciones con repetición

En las variaciones con repetición no es preciso que agrupemos todos los objetos de golpe, podemos repetir los objetos y, además, importa el orden.

Ejemplo de variaciones con repetición

¿Cuántos caracteres diferentes se pueden codificar con cadenas de ocho ceros y ocho unos? Tenemos dos objetos (0 y 1), tenemos que cogerlos ocho veces y podemos repetir el 0 y el 1 tantas veces como convenga; dado que tenemos ocho elecciones y cada una de éstas con dos posibilidades, el número total es 2^8. Observad que las cadenas 01000000 y 1000000 son diferentes (codifican caracteres diferentes) y, por tanto, importa el orden.

3. Permutaciones

A continuación nos planteamos de cuántas maneras diferentes se puede ordenar un conjunto de objetos.

Una **permutación** de un conjunto de objetos es cualquier posible ordenación de estos objetos.

Características de las permutaciones

En las permutaciones debemos agrupar todos los objetos de golpe, no podemos repetir los objetos y, además, evidentemente, importa el orden.

Calcular el número total de permutaciones de un conjunto de objetos resulta muy fácil utilizando la regla del producto: supongamos que tenemos N objetos; para decidir cuál irá en primer lugar, tenemos N posibles elecciones. Una vez que hemos escogido cuál irá en primer lugar, tenemos $N-1$ posibilidades de elección para escoger cuál irá en segundo lugar, y así sucesivamente. Es decir, en cada paso tenemos una posibilidad menos de elección, ya que hemos ido fijando los objetos y, por tanto, cada vez disponemos de un objeto menos que hay que situar; aplicando la regla del producto, obtenemos el número total de posibles ordenaciones.

El **número de las permutaciones** de N objetos es:

$$N! = N \cdot (N-1) \cdot (N-2) \cdot \ldots \cdot 3 \cdot 2 \cdot 1$$

Otra manera de verlo

También se puede calcular pensando que el número de permutaciones es el número de las variaciones de N elementos cogidos de N en N, ya que, por definición, se trata de ordenarlos todos.

Ejemplo de aplicación de las permutaciones

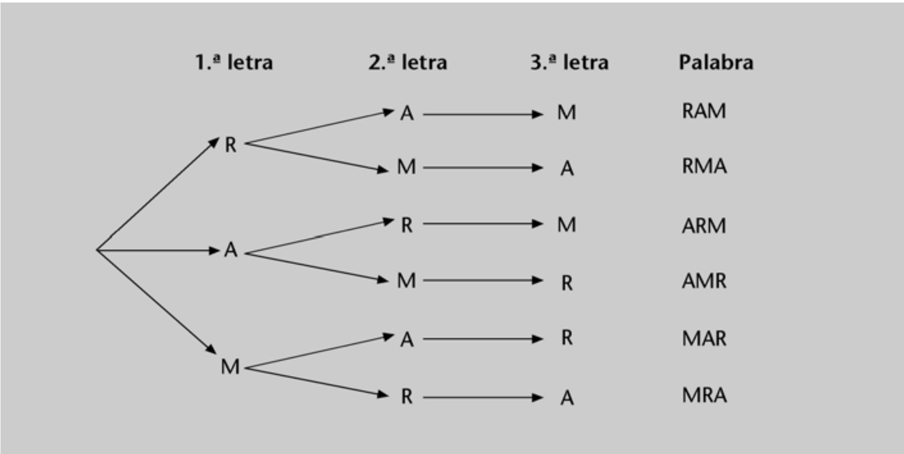
¿De cuántas formas se pueden ordenar las letras de la palabra *RAM*? Si hacemos un árbol, observamos que en el primer paso tenemos tres opciones; en el segundo sólo tenemos dos, ya que una letra está fijada en el primer lugar, y en el tercer paso tenemos una única opción, que es la letra que todavía no hemos situado. Por tanto, el número de permutaciones será $3! = 3 \cdot 2 \cdot 1 = 6$.

Las ordenaciones de veintitrés portátiles

Si la revista *PCUniverse* nos pide que ordenemos de todas las formas posibles los veintitrés portátiles, tendremos que escribir 23! ordenaciones diferentes.

23! es un número enorme, concretamente es:

25.852.016.738.884.976.640.000.



4. Combinaciones

En otras situaciones interesa considerar agrupaciones en las que no hay que tener en cuenta el orden, bien porque no es relevante, bien porque no afecta al resultado final. En este caso trabajaremos directamente con subconjuntos.

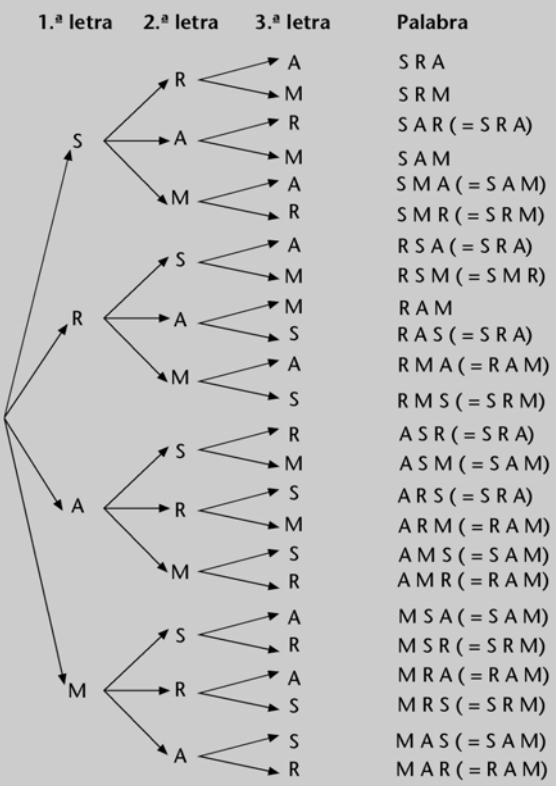
Dados N objetos diferentes, una **combinación** de estos N objetos tomados de k en k es cualquier subconjunto de k elementos que se pueda formar usando sólo los N objetos iniciales.

Características de las combinaciones

En las combinaciones no es preciso que agrupemos todos los objetos, no podemos repetir los objetos y no importa el orden.

Ejemplo de aplicación de combinaciones

Supongamos que tenemos que escribir todos los posibles subconjuntos de tres elementos que se pueden formar utilizando únicamente letras de la palabra *SRAM*. Lo haremos mediante un árbol en el que se ve que en cada paso tenemos una posibilidad de elección menos y que necesitamos tres pasos para construir los subconjuntos de tres elementos. El paréntesis indica que el subconjunto ya ha sido construido anteriormente y con qué subconjunto se corresponde.



Comentario

SAR es el mismo subconjunto que *SRA* y, por tanto, no lo consideramos como un grupo diferente.

También observamos que hay muchas repeticiones porque, por lo que respecta a los conjuntos, es lo mismo considerar $\{S, R, A\}$ como $\{R, A, S\}$. En el árbol hemos indicado las repeticiones y, por tanto, sólo tenemos que considerar los subconjuntos siguientes: $\{S, R, A\}$, $\{S, R, M\}$, $\{S, A, M]$, $\{R, A, M\}$. Así pues, tenemos cuatro combinaciones posibles; este número se puede deducir a partir del árbol: el árbol tiene en total $4 \cdot 3 \cdot 2 = 24$ ramas (que, de hecho, se corresponden a las variaciones de cuatro elementos tomados de tres en tres), pero cada una de éstas aparece 3! veces, ya que las tres letras aparecen en todos los órdenes posibles; por tanto, el número de combinaciones es igual a $4 \cdot 3 \cdot 2 / 3!$, que, escrito utilizando factoriales, se puede expresar como $4! / (3! \cdot 1!)$.

Notación

El número combinatorio *N* sobre *k* también se puede escribir $C(N, k)$ o bien C_N^k.

El **número de combinaciones** de *N* objetos cogidos de *k* en *k* se calcula haciendo:

$$\binom{N}{k} = \frac{\text{Número de variaciones de } N \text{ elementos cogidos de } k \text{ en } k}{\text{Permutaciones de } k \text{ elementos}} =$$

$$= \frac{N(N-1)\ldots(N-k+1)}{k!} = \frac{N!}{k!(N-k)!}$$

$\binom{N}{k}$ es el llamado **número combinatorio** "*N* sobre *k*".

Los cinco mejores portátiles

PCUniverse nos envía una lista de veintitrés portátiles y nos pide que devolvamos la lista de los que creemos que son los cinco mejores de la lista. En este caso, debemos escoger 5(= *k*) entre 23 (= *N*) posibles y, por tanto, tenemos:

$$C(N, K) = \frac{23!}{(5!\ (23-5)!)} =$$

$$= \frac{23!}{(5!\ 18!)} = 33.649$$

posibilidades de escoger cinco portátiles (la revista no pide que los ordenemos, sólo quiere saber los que creemos que son los cinco mejores, sin orden de preferencia entre sí).

5. Resumen

En esta sesión se hace un recorrido por algunas técnicas de recuento muy utilizadas en el contexto del estudio de la probabilidad. En el cuadro siguiente resumimos los casos estudiados, sus características y las fórmulas de combinatoria más importantes que aparecen, suponiendo que dispongamos de N individuos y tengamos que cogerlos de k en k:

	¿Hay que cogerlos todos?	¿Se pueden repetir?	¿Importa el orden?	Número total
Variaciones	No	No	Sí	$N(N-1)\ldots(N-k+1)$
Variaciones con repetición	No	Si	Sí	N^k
Permutaciones	Sí	No	Sí	$N!$
Combinaciones	No	No	No	$\binom{N}{k} = \frac{N!}{k!(N-k)!}$

Las fórmulas se deducen a partir de la llamada *regla del producto*. También se insiste en cómo distinguir un caso de otro y en qué casos hay que aplicar cada una de las definiciones.

Ejercicios

1. El póquer
La baraja de póquer consta de cincuenta y dos cartas con cuatro palos (corazones, diamantes, tréboles y picas) y trece cartas de cada palo, numeradas así: as, 2 ..., 10, *J*, *Q* y *K*. Cada vez se reparten cinco cartas; esto se llama *una mano de cinco cartas*.

a) ¿Cuántas manos de cinco cartas diferentes pueden aparecer al jugar al póquer?
b) ¿Cuántas de estas manos tienen exactamente un as?
c) ¿Cuántas manos no tienen ningún as?
d) ¿Cuántos póquers se pueden formar? Un póquer son cuatro cartas del mismo número y otra de un número diferente.
e) ¿En cuántas manos hay al menos una *K*?

2. Las sillas de Guillermo Puertas
Se hace una reunión de los diez empresarios más importantes del sector informático del mundo; a la reunión asisten Guillermo Puertas y su feroz competidor, director de Oracel. Para evitar problemas de protocolo, se sientan en una hilera de diez sillas:

a) ¿De cuántas maneras diferentes se pueden sentar los diez empresarios?
b) ¿De cuántas maneras diferentes se pueden sentar los diez empresarios, pero de manera que Guillermo Puertas y su feroz competidor no estén uno al lado del otro?

Solucionario

1. El póquer:

a) Se trata de hacer subconjuntos de cinco cartas en los que no importa el orden y en los que, evidentemente, no se pueden repetir las cartas.

En total, pues, $\binom{52}{5} = 2.598.960$ manos diferentes.

b) Miremos las posibilidades de escoger cuatro cartas que no sean ases, que son:

$$\binom{48}{4} = 194.580$$

Ahora, para completar una mano, tenemos que escoger entre cuatro ases; por tanto, las posibilidades totales son:

$$\binom{48}{4} \cdot 4 = 778.320$$

c) Hay que escoger cinco cartas entre las cuarenta y ocho que no son un as. Por tanto:

$$\binom{48}{5} = 1.712.304$$

d) Para formar un póquer, debemos seguir un proceso como éste:

- Escoger un número (del as a la K): trece posibilidades.
- Escoger cuatro cartas de este número: una posibilidad, ya que sólo hay cuatro cartas de cada número.
- Escoger una carta de entre las que quedan: cuarenta y ocho posibilidades.

En total tenemos 13 · 48 = 624 póquers diferentes.

e) Primero miramos en cuántas manos no hay ninguna K:

$$\binom{52-4}{5} = 1.712.300$$

Quitamos las cuatro K de la baraja y nos quedan cuarenta y ocho cartas, de las que tengo que coger cinco. Por tanto, hay al menos una K en:

$$\binom{52}{5} - \binom{52-4}{5} = 886.656 \text{ manos}$$

2. Las sillas de Guillermo Puertas:

a) Son las permutaciones de diez personas: 10! = 3.628.800.

b) Primero calculamos todas las maneras posibles de que se sienten juntos. Puesto que tienen que sentarse juntos, es como si tuviéramos nueve objetos por ordenar; teniendo en cuenta que un objeto (el par Guillermo Puertas-director de Oracel) vale por dos, ya que Guillermo Puertas tanto puede estar a la derecha como a la izquierda, en total tenemos, pues, 2 · 9! = 725.760 maneras de que se sienten juntos. Por tanto, tenemos:

Haremos los cálculos...

... con ordenador o usando la función factorial de algunas calculadoras.

10! – 2 · 9! = 10 · 9! – 2 · 9! = 8 · 9! = 2.903.040 posibilidades de maneras de sentarse separados.

Probabilidad

Aquí comenzamos el estudio de la probabilidad. A grandes rasgos, la probabilidad de un suceso es una medida de la tendencia que tiene a darse dicho suceso. Esta medida será un número situado entre dos valores: el 0, que será la probabilidad de un suceso que no se pueda dar nunca (el suceso imposible) y el 1, que se corresponderá con un suceso que se da siempre (el suceso seguro).

1. Introducción y frecuencia relativa

Para ilustrar las ideas sobre probabilidad que se introducirán más adelante de manera formal, comenzaremos por reflexionar sobre la frecuencia relativa de un resultado de un experimento aleatorio. Así pues, consideramos un experimento aleatorio y su espacio muestral Ω y vemos, en primer lugar, un ejemplo que nos guiará en esta introducción.

La frecuencia relativa en un dado trucado (I)

Lanzamos un dado $R = 100$ veces y anotamos cuántas veces aparece cada resultado en la tabla siguiente:

Resultado	1	2	3	4	5	6
Apariciones	12	28	20	20	5	15

La frecuencia relativa del resultado 2 es 28 / 100; el resultado 5 tiene frecuencia relativa 5/100. Como podéis observar, hay números que tienen más tendencia a salir que otros, lo que nos puede hacer sospechar que el dado no es perfectamente “neutral”. La frecuencia relativa es un indicador numérico de la tendencia a darse que tiene cada resultado.

Repetimos el experimento aleatorio un número R de veces; si dividimos el número de veces que se da un resultado por R, obtenemos la **frecuencia relativa** del resultado. Evidentemente, la frecuencia relativa de cualquier resultado es un número entre 0 y 1. La suma de las frecuencias relativas de todos los resultados debe ser igual a 1.

Partiendo del hecho de que un suceso es un conjunto de resultados, también es posible determinar su frecuencia relativa.

La frecuencia relativa de un suceso se obtiene dividiendo el número de veces que el resultado que se obtiene al realizar el experimento es favorable al suceso por el número de repeticiones del experimento.

Para calcular la frecuencia relativa de un suceso también podemos utilizar la frecuencia relativa de los resultados que contiene.

La frecuencia relativa de cualquier suceso es igual a la suma de las frecuencias relativas de sus resultados favorables. Es evidente que la frecuencia relativa de un suceso es un número entre cero y uno.

Significado de la frecuencia relativa = 0

Frecuencia relativa = 0 significa que no ha pasado nunca; frecuencia relativa = 1 significa que todas las veces hemos obtenido un resultado favorable al suceso.

La frecuencia relativa en un dado trucado (II)

Siguiendo con el ejemplo del dado trucado: ¿cuál es la frecuencia relativa del suceso P = "sacar un número par"? De las cien veces que hemos lanzado el dado, 28 + 20 + 15 veces hemos obtenido un número par; por tanto, la frecuencia relativa de P es:

$$(28 + 20 + 15)/100 = 63/100$$

que es precisamente igual a:

$$28/100 + 20/100 + 15/100$$

es decir, la suma de las frecuencias relativas del 2, el 4 y el 6 (los resultados pares). Así pues, ¡el 63% de las veces ha salido un número par!

Es muy importante que os deis cuenta de que a partir de las frecuencias relativas de ciertos sucesos podemos deducir la frecuencia relativa de otros.

Si dos sucesos son incompatibles, la frecuencia relativa de su unión es la suma de las respectivas frecuencias relativas.

Recordad que...

... dos sucesos son incompatibles si son disyuntos.

Frecuencia relativa de la unión de sucesos (I)

Siguiendo con el ejemplo del dado trucado, la frecuencia relativa del suceso P = "sacar un número par" es 63% y la del suceso Q = "salir un 3 o un 5" es 25%; la frecuencia relativa de $P \cup Q$ es 88%, que es precisamente igual a 63% + 25%, ya que P y Q son disyuntos.

Fijaos en que si los sucesos no son incompatibles, la frecuencia relativa de su unión no es la suma de sus frecuencias relativas.

Frecuencia relativa de la unión de sucesos (II)

Si consideramos P = "sacar un número par" y Q = "sacar un número mayor que 3"; la frecuencia relativa de P es 63%, la de Q es 40%, pero la de $P \cup Q$ no es 63% + 40% = 103%, ya que la frecuencia relativa no puede ser mayor que 100% = 1. De hecho, la frecuencia relativa de $P \cup Q$ corresponde a la suma de las frecuencias relativas de los resultados 2, 4, 5, 6 y es igual al 68%.

En realidad, la frecuencia relativa de la unión de dos sucesos es la suma de sus frecuencias relativas menos la frecuencia relativa de su intersección.

Frecuencia relativa de la unión de sucesos (III)

Podemos comprobar la propiedad citada en el caso del ejemplo del dado trucado, ya que la frecuencia relativa de la intersección de $P \cap Q$ es 35% y, por tanto, la frecuencia relativa de $P \cup Q$ es igual a 63% + 40% – 35% = 68%.

¿Qué otras propiedades de la frecuencia relativa podemos destacar? De entre las muchas que tiene, comentaremos dos:

1) La frecuencia relativa del suceso seguro es 1, ya que se da siempre que efectuemos el experimento. La frecuencia relativa del suceso imposible es cero, ya que no se da nunca.

2) La frecuencia relativa del complementario de *A* es 1 menos la frecuencia relativa de *A*.

Frecuencia relativa del suceso seguro

¿Cuál es la frecuencia relativa del suceso "sacar un número menor que 10" cuando lanzamos un dado mil veces? Evidentemente 1, ya que las mil veces hemos obtenido un número menor que diez. Y esta frecuencia será siempre 1, independientemente del número de tiradas que efectuemos.

Frecuencia relativa de la A^c

Si un suceso se da el 23% de las veces, su complementario se tiene que dar el 77% (100% – 23%) de las veces.

2. La teoría de la probabilidad

La teoría de la probabilidad es una teoría matemática que establece cómo podemos asignar una probabilidad a sucesos complejos a partir de la probabilidad de sucesos más simples. Existen varias formas de definir estas probabilidades, pero todas tienen que cumplir unos requisitos de coherencia que se corresponden con algunas de las propiedades mencionadas para la frecuencia relativa de los sucesos que hemos comentado en la sección anterior.

La cuestión de decidir qué propiedades básicas se tienen que exigir a lo que llamamos *probabilidad* fue estudiada por Kolmogorov hacia los años treinta; en sus trabajos llegó a la conclusión de que era suficiente con tres propiedades (axiomas) para establecer el concepto de probabilidad. Estas propiedades son las siguientes:

1) La probabilidad de cualquier suceso debe ser un número entre 0 y 1.

2) Si dos sucesos son disyuntos, la probabilidad de su unión es la suma de las probabilidades de ambos sucesos.

3) La probabilidad del conjunto de todos los posibles resultados (suceso seguro) debe ser 1.

Expresiones para la probabilidad

Puesto que la probabilidad es un número entre 0 y 1, a veces la expresaremos en porcentaje, tal como se suele hacer con la frecuencia relativa. Así, diremos que cierta probabilidad es 0,3 o bien del 30%.

Si comenzamos por definir la probabilidad como una función *P* que hace corresponder a cada suceso su probabilidad, podemos reescribir las propiedades anteriores de la manera siguiente:

P1.) $0 \leq P(A) \leq 1$ para todo suceso *A*.

P2.) Si *A* y *B* son sucesos disyuntos, entonces $P(A \cup B) = P(A) + P(B)$.

P3.) $P(\Omega) = 1$ (en el que Ω es el espacio muestral, que se corresponde con el suceso que contiene todos los posibles resultados).

Probabilidad y frecuencia relativa

Observad que estas propiedades son algunas de las mencionadas por la frecuencia relativa, sólo que cambiando la expresión "frecuencia relativa" por "probabilidad".

3. Propiedades que se derivan de la definición de probabilidad

Probabilidad y frecuencia relativa

Observad que en la definición de probabilidad no existe ninguna regla para calcular la $P(A \cap B)$ en función de $P(A)$ y $P(B)$.

A continuación veremos que, con las propiedades P1, P2 y P3, podemos llegar a conclusiones muy interesantes sobre la probabilidad de algunos sucesos determinados.

3.1. La probabilidad del suceso imposible

Con respecto al suceso imposible $\varnothing$, tenemos que:

$$P(\varnothing) = 0$$

Esto se debe a los hechos siguientes:

a) Puesto que $\varnothing$ y Ω son disyuntos, por aplicación de la propiedad P2:

$$P(\varnothing \cup \Omega) = = P(\varnothing) + P(\Omega)$$

b) Pero dado que $\varnothing \cup \Omega = \Omega$, por la propiedad P3, $P(\varnothing \cup \Omega) = P(\Omega) = 1$.

c) En consecuencia, $1 = P(\varnothing) + 1$, de lo que podemos concluir que $P(\varnothing) = 0$.

3.2. La probabilidad del complementario

Un razonamiento análogo nos permite tratar la probabilidad del complementario de un suceso. Fijémosnos en que, para todo suceso A, $A \cup A^c = \Omega$; por las propiedades P2 y P3, tenemos que $P(A \cup A^c) = P(A) + P(A^c) = P(\Omega) = 1$; entonces:

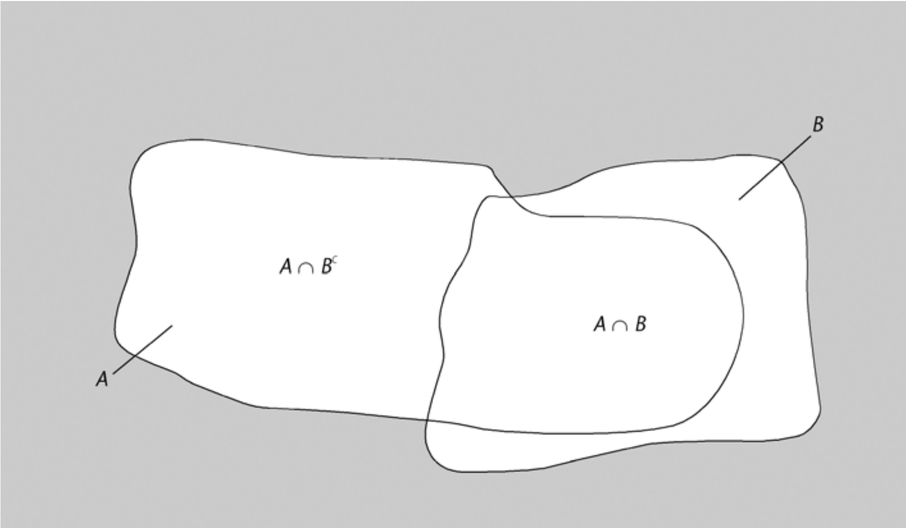
$$P(A^c) = 1 - P(A)$$

3.3. La probabilidad de la unión

Otra propiedad muy interesante que se deriva de $A = (A \cap B) \cup (A \cap B^c)$ y del hecho de que $(A \cap B)$ y $(A \cap B^c)$ son disyuntos es la siguiente:

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

Gráficamente:



Una propiedad de la unión

$A = (A \cap B) \cup (A \cap B^c)$, ya que los resultados de A son "los de A que están en B" unión "los de A que no están en B".

La propiedad **2** permite calcular la probabilidad de la unión de sucesos que sean incompatibles, pero ¿qué pasa si los sucesos no lo son? En este caso aplicaremos la regla siguiente:

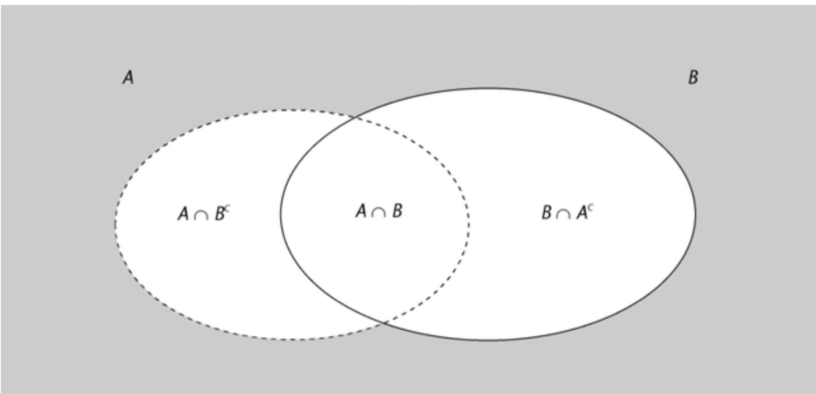
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Según ésta, para calcular la probabilidad de la unión hay que calcular la suma de probabilidades y después restar la probabilidad de la intersección. Por tanto, siempre tenemos que $P(A \cup B) \leq P(A) + P(B)$.

Demostración de la regla de la probabilidad de la unión

La regla según la cual $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ se puede demostrar muy fácilmente a partir de los hechos siguientes:

1) Tal como se ve en el gráfico siguiente:



tenemos que $A \cup B = B \cup (A \cap B^c)$. Puesto que estos conjuntos son evidentemente disyuntos, tenemos que $P(A \cup B) = P(B) + P(A \cap B^c)$;

2) Puesto que $P(A) = P(A \cap B) + P(A \cap B^c)$, podemos deducir que $P(A \cap B^c) = P(A) - P(A \cap B)$.

3) Finalmente, $P(A \cup B) = P(B) + P(A \cap B^c) = P(B) + P(A) - P(A \cap B)$.

4. Asignación de probabilidad cuando los resultados son equiprobables. Regla de Laplace

Hasta ahora hemos visto cómo permite la teoría de la probabilidad asignar probabilidades a ciertos sucesos a partir de la probabilidad de otros sucesos más simples. Ahora falta ver cómo podemos asignar probabilidades a los sucesos más simples de todos (los resultados de un experimento aleatorio), de manera que podamos obtener, gracias a la probabilidad, conclusiones interesantes sobre los sucesos y sobre las situaciones descritas por los sucesos que interesan.

Primero comenzaremos por definir lo que se entiende por resultados equiprobables.

Dos resultados de un experimento aleatorio son **equiprobables** si tienen la misma probabilidad de suceder.

Un espacio muestral es **uniforme** si todos los resultados son equiprobables.

Probabilidad y equiprobabilidad

Puede parecer paradójico definir la probabilidad a partir del concepto de sucesos equiprobables. Lo que pasa es que todo el mundo entiende perfectamente que, si el dado es perfecto, tantas posibilidades hay de sacar un uno como un dos, etc. Es decir, el concepto de equiprobabilidad no depende de cómo midamos efectivamente la probabilidad de cada uno de los números del dado.

En este caso, si $\Omega = \{\omega_1, \omega_2 \dots \omega_k\}$ tiene k posibles resultados, tendremos que $P(\omega_1) = = P(\omega_2) = \dots = P(\omega_k)$. Por aplicación de la propiedad **2** y partiendo del hecho de que los resultados de un experimento aleatorio son siempre disyuntos –ya que si se da un resultado, no se puede dar otro diferente–, tenemos que:

$$1 = P(\Omega) = P(\omega_1 \cup \omega_2 \cup \dots \cup \omega_k) = P(\omega_1) + P(\omega_2) + \dots + P(\omega_k) = kP(\omega_1)$$

y, por tanto, la probabilidad de cada resultado es:

$$P(\omega_i) = \frac{1}{k} = \frac{1}{\text{Card}(\Omega)}$$

¿Cuál será la probabilidad de un suceso cualquiera? Supongamos que el suceso A resulta estar formado por s resultados, con lo que Card$(A) = s$; supongamos que, por ejemplo, $A = \{\omega_1, \omega_2, \dots, \omega_s\}$. En este caso:

$$P(A) = P(\omega_1) + P(\omega_2) + \dots + p(\omega_s) = (1/k) + \dots + (1 / k) = s/k$$

Probabilidad de un dado perfecto

Si lanzamos un dado perfecto, todos los números tienen la misma probabilidad de salir. ¿Cuál es ésta probabilidad? Evidentemente, será 1/6.

Es decir, para calcular la probabilidad de A, deberemos determinar cuántos resultados forman parte del suceso A –los resultados favorables a A– y dividir este número por el total de resultados posibles; ésta es la llamada regla de Laplace.

La **regla de Laplace** afirma que en un espacio uniforme la probabilidad de un suceso A se calcula mediante el cociente siguiente:

$$P(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)} = \frac{\text{Número de resultados favorables a } A}{\text{Número total de posibles resultados}}$$

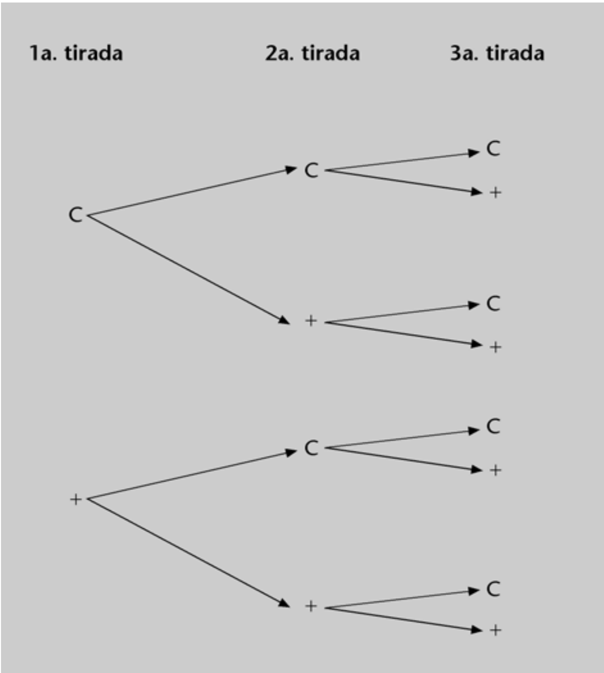
El uso de esta regla permite interpretar la probabilidad como porcentaje, ya que podemos considerar $P(A)$ como el porcentaje de veces que se da un resultado de A con respecto al total de posibles resultados.

Ejemplo de cálculo de probabilidades

Lanzamos un dado perfecto (sin trucar) y consideramos los sucesos A = "sale un número mayor que 4", B = "sale un número par", C = "sale un número mayor que 8". En este caso, el espacio muestral tiene 6 posibles resultados equiprobables y $P(A) = 2/6$, $P(B) = 3/6 = 1/2$, $P(C) = 0/6 = 0$.

Ejemplo de aplicación de la regla de Laplace

Lanzamos una moneda perfecta tres veces y consideramos los sucesos siguientes: A = "salen 3 caras", B = "sale una única cruz", D = "no salen ni dos caras ni dos cruces consecutivas". En primer lugar determinamos el espacio muestral del experimento, espacio que se puede obtener fácilmente mediante el árbol de los posibles resultados (en los que C representa cara y + representa cruz):



por tanto:

$$\Omega = \{(CCC), (CC+), (C+C), (C++), (+CC), (+C+), (++C), (+++)\}$$

es decir, tenemos ocho posibles resultados, y todos ellos son equiprobables (ya que a cada tirada la probabilidad de que salga cara es la misma probabilidad de que salga cruz). A continuación podemos calcular cada uno de los sucesos:

$$A = \{(CCC)\},\ B = \{(+CC), (C+C), (CC+)\},\ D = \{(C+C), (+C+)\}$$

Las probabilidades correspondientes son $P(A) = 1/8$, $P(B) = 3/8$, $P(D) = 2/8$.

5. Probabilidades en espacios muestrales no uniformes y frecuencia relativa

No siempre podemos asegurar que los resultados de un experimento aleatorio sean equiprobables. Pensemos en el caso de un dado del que sospechamos que está trucado; ahora nos enfrentamos a dos nuevos problemas: cómo podemos confirmar esta sospecha y cómo podemos calcular la probabilidad de cada resultado. En los dos casos la respuesta pasa por calcular la frecuencia relativa de cada uno de los resultados después de repetir el experimento una y otra vez.

Así, si después de repetir el experimento (lanzar el dado) mil veces, obtenemos cien unos, cien doses, cien treses, cien cuatros, cien cincos y quinientos seises, podemos pensar que la probabilidad de obtener un seis es aproximadamente 1/2, mientras que la probabilidad de obtener cualquier otro número será de aproximadamente 1/10.

Claro que estas aproximaciones a la probabilidad de cada resultado no serán definitivas, ya que al ir aumentando el número de repeticiones del experimento, irán cambiando las frecuencias relativas. Para evitar esta ambigüedad, se va repitiendo el experimento y se observa si la frecuencia relativa de cada resultado tiende hacia algún valor a medida que se van repitiendo los experimentos.

Estas probabilidades, obtenidas como el valor al que tienden las frecuencias relativas después de repetir muchas veces un experimento, se denominan **probabilidades empíricas**.

Necesidad de la repetibilidad

Recordad que una de las características que pedimos a los experimentos aleatorios es que se tienen que poder repetir tantas veces como sea necesario.

A partir de estas probabilidades empíricas, y aplicando las propiedades que definen la probabilidad, podemos asignar probabilidad a todos los otros sucesos.

Un dado curioso y bastante trucado

Lanzamos un dado 150.000 veces y obtenemos los resultados siguientes:

Resultado	1	2	3	4	5	6
Frecuencia	1	100.000	40.000	9.999	0	0

Para calcular las probabilidades empíricas de cada uno de los resultados, debemos dividir su frecuencia por 150.000 (el número de veces que hemos repetido el experimento). La probabilidad de que el número sea par es (100.000 / 150.000) + (9.999 / 150.000). La probabilidad de que el número sea mayor que 5 es 0.

6. Probabilidad condicionada

En este apartado nos concentraremos en el concepto de probabilidad condicionada. Comenzaremos por un ejemplo en el que se ilustra la influencia que unos sucesos pueden tener sobre otros.

Probabilidad condicionada en un dado perfecto

Si lanzamos un dado perfecto, la probabilidad de que el resultado sea un número par es 0,5. A continuación lanzamos el mismo dado y cae debajo de la mesa, y antes de mirar el número que ha salido, alguien nos dice que es mayor que 3; en este caso, ¿cuál es la probabilidad de que sea par? Lo primero que debemos preguntarnos es cuáles son los posibles resultados: son {4, 5, 6}, ya que la información previa de la que disponemos descarta que haya salido un uno, un dos o un tres. Así, de los posibles resultados mayores que 3, dos son pares. Por tanto, dado que tenemos tres resultados en los que el número que sale es mayor que 3 –y de éstos, dos son pares–, la probabilidad de que salga par sabiendo que el número es mayor que 3 es 2/3.

A continuación definimos el concepto de probabilidad condicionada, que recoge la influencia que pueden tener unos sucesos sobre otros.

La probabilidad del suceso A **condicionado** a B, probabilidad que denotaremos por $P(A \mid B)$, es la probabilidad de que al realizar el experimento aleatorio, el resultado obtenido sea de A, sabiendo que el resultado obtenido es de B.

Observación sobre la probabilidad condicionada

Observamos que si el resultado es de B, puede ser de A (con lo que será de $A \cap B$) o no.

En el ejemplo del dado perfecto, si B = "ser mayor que 3" y A = "ser par", tendremos que si el resultado es de B ("ser mayor que 3"), puede ser de A ("ser par") o no ("no ser par").

Comenzaremos calculando de forma intuitiva probabilidades condicionadas en una situación relacionada con los ordenadores de nuestra empresa.

Consideremos la tabla siguiente, que registra las diferentes configuraciones de los cincuenta ordenadores de nuestra empresa, considerando la RAM de la que disponen y el tipo de procesador (el tipo de procesador aparece en una etiqueta pegada al lado del botón de puesta en marcha, mientras que para saber la RAM, hay que poner en funcionamiento el ordenador):

	64 MB	128 MB
Patinum1	10	20
Patinum22	10	10

De la tabla obtenemos que tenemos diez ordenadores con 64 MB de RAM y Patinum1, veinte ordenadores con 128 MB de RAM y Patinum1, diez ordenadores con 64 MB de RAM y Patinum22 y, finalmente, diez ordenadores de 128 MB de RAM y Patinum22. Si escogemos un ordenador al azar, podemos calcular fácilmente muchas probabilidades; por ejemplo, P ("estar equipado con Patinum1") = 30/50, mientras que P ("tener 64 MB de RAM") = 20/50.

A continuación introducimos la probabilidad condicionada: supongamos que escogemos un ordenador al azar y leemos en la etiqueta que su procesador es un Patinum22; nos preguntamos cuál es la probabilidad de que tenga 64 Mb de RAM. ¿Cuál es esta probabilidad? Sabemos que el ordenador lleva un Patinum22 y, por tanto, debe ser uno de los veinte ordenadores que tienen este procesador. De entre estos veinte, sabemos que diez tienen 64 MB de RAM, por tanto:

$$P(\text{“tiene 64 MB de RAM”} \mid \text{“lleva Patinum22”}) = 10 / 20 = 1 / 2$$

Otro caso: hemos puesto en marcha un ordenador sin mirar el adhesivo del procesador y resulta que tiene 128 MB de RAM. ¿Cuál es la probabilidad de que tenga Patinum1? De la misma manera que antes, hay treinta ordenadores con 128 MB de RAM, de los cuales veinte tienen Patinum1; por tanto:

$$P(\text{“lleva Patinum1”} \mid \text{“tiene 128 MB de RAM”}) = 20/30$$

Si en esta última fórmula dividimos numerador y denominador por el total de individuos, obtenemos:

$$P(\text{“lleva Patinum1”} \mid \text{“tiene 128 MB de RAM”}) = (20/50) / (30/50)$$

en el que 30/50 resulta ser la probabilidad de tener 128 MB de RAM y, por tanto,

$$P(\text{“lleva Patinum1”} \mid \text{“tiene 128 MB de RAM”}) = \\ = (20 / 50) / P(\text{“tiene 128 MB de RAM”})$$

Ahora nos queda averiguar qué representa el numerador: puesto que hay veinte ordenadores que tienen Patinum1 y 128 MB de RAM, 20/50 es la probabilidad de que un ordenador tenga Patinum1 y también 128 MB de RAM, con lo que finalmente obtenemos:

$$P(\text{“lleva Patinum1”} \mid \text{“tiene 128 MB de RAM”}) = \\ = P(\text{“lleva Patinum1”} \cap \text{“tiene 128 MB de RAM”}) / P(\text{“tiene 128 MB de RAM”})$$

Observación

Fijaos en que si $P(B) = 0$, no tiene sentido definir la probabilidad $P(A \mid B)$.

Este resultado motiva la fórmula siguiente, válida para calcular probabilidades condicionadas. Sean A y B dos sucesos, de manera que $P(B) > 0$; entonces:

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

Cálculo de probabilidades condicionadas

En el ejemplo de la RAM y los Patinum tenemos que $P(\text{“tiene 128 MB”} \mid \text{“lleva Patinum22”}) = P(\text{“tiene 128 MB”} \cap \text{“lleva Patinum22”}) / P(\text{“lleva Patinum22”}) = (10/50) / (20/50) = 1/2$.

6.1. Relación entre probabilidad condicionada y probabilidad de la intersección

Recordad que no tenemos ninguna fórmula para determinar la probabilidad de la intersección de dos sucesos; en cambio, a partir de la fórmula de la probabilidad condicionada podemos obtener el resultado siguiente: dados dos sucesos cualesquiera A y B, con $P(A) > 0$ y $P(B) > 0$, entonces:

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

Esta propiedad se demuestra de forma fácil, simplemente aislando $P(A \cap B)$ en la definición de $P(A \mid B)$ y $P(B \cap A) = P(A \cap B)$ en la definición de $P(B \mid A)$.

7. Independencia de sucesos

Comenzaremos por una definición intuitiva del concepto de independencia para poder justificar más adelante la definición formal.

Dos sucesos A y B son **independientes** cuando el hecho de que ocurra uno no altera la probabilidad de que ocurra el otro.

La formalización de esta idea a partir de la probabilidad condicionada da lugar a la llamada **regla del producto de probabilidades**. Dos sucesos A y B son independientes si y sólo si:

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

Ahora podemos preguntarnos qué relación tiene esta definición de independencia con la probabilidad condicionada: pues bien, es fácil demostrar que para determinar si dos sucesos son independientes, tenemos tres vías diferentes.

Si $P(A) > 0$ y $P(B) > 0$, las afirmaciones siguientes son equivalentes:

1) Los sucesos A y B son independientes.
2) $P(A \cap B) = P(A) \cdot P(B)$
3) $P(A \mid B) = P(A)$
4) $P(B \mid A) = P(B)$

Demostración de la regla del producto de probabilidades

Las afirmaciones **1** y **2** son equivalentes por definición (es la regla del producto).

Ahora supongamos que se da la afirmación **2**, según la cual $P(A \cap B) = P(A) \cdot P(B)$; por tanto:

$$P(A \mid B) = P(A \cap B)/P(B) = P(A) \cdot P(B)/P(B) = P(A)$$

con lo que hemos demostrado la afirmación **3**.

Por otra parte, si $P(A \mid B) = P(A)$, tenemos que $P(A \cap B)/P(B) = P(A)$ y, por tanto, $P(A \cap B) = P(A) \cdot P(B)$, con lo que demostramos que la afirmación **3** implica la **2**.

Que la afirmación **2** es equivalente a la **4** se hace de la misma forma, observando que:

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

Ejemplo de sucesos independientes con un dado perfecto

Volvamos ahora al ejemplo del dado perfecto y consideremos los sucesos A = "sacar un número par", B = "sacar un número mayor que 3" y C = "sacar un número mayor que 4". Nos preguntamos si los sucesos A y B son independientes; para hacerlo, calculamos:

$P(A) = 3/6 = 1/2$ y
$P(A \mid B) = P(A \cap B)/P(B) = P(\text{"par mayor que 3"}) / P(\text{"mayor que 3"}) =$
$= P(\{4, 6\})/P(\{4, 5, 6\}) = (2/6) / (3/6) = 2/3$

Independencia de sucesos (I)

Para ver si dos sucesos son independientes calculamos $P(A)$, $P(B)$ y $P(A \cap B)$.

- Si $P(A \cap B) = P(A) \cdot P(B)$, entonces son independientes.
- Si $P(A \cap B) \neq P(A) \cdot P(B)$, entonces no son independientes.

Independencia de sucesos (II)

a) Calculamos $P(A)$ y $P(A \mid B)$; si $P(A) = P(A \mid B)$, entonces los sucesos son independientes; si $P(A) \neq P(A \mid B)$, entonces no lo son.

b) Calculamos $P(B)$ y $P(B \mid A)$; si $P(B) = P(B \mid A)$, entonces los sucesos son independientes; si $P(B) \neq P(B \mid A)$, entonces no lo son.

Ejemplo de sucesos independientes

Consideremos dos sucesos cualesquiera tales que $P(A) = 0{,}5$, $P(B) = 0{,}5$ y $P(A \cap B) = 0{,}25$; los sucesos A y B son independientes, ya que:

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

Interpretación de dos sucesos no independientes

Los sucesos A y B no son independientes, ya que dentro de B la proporción de pares es 2/3, mientras que del total de resultados la proporción de pares es 1/2; es decir, dentro del suceso B la probabilidad de A es mayor que con respecto al total de resultados; por eso A y B no son independientes.

Puesto que $P(A) = 1/2 \neq P(A/B) = 2/3$, los sucesos A y B no son independientes. En cambio:

$$P(A \mid C) = P(A \cap C) \;/\; P(C) = P(\text{“par mayor que 4”}) \;/\; P(\text{“mayor que 4”}) =$$
$$= P(\{6\}) \;/\; P(\{5, 6\}) = (1/6) \;/\; (2/6) = 1/2$$

Dado que $P(A) = 1 / 2 = P(A \mid C)$, los sucesos son independientes.

Ejemplo de procesos no independientes lanzando un dado dos veces

Continuamos con los dados, pero ahora lanzamos un dado dos veces. Consideremos los sucesos A = "el valor máximo de las dos tiradas es 1"; B = "la suma de los valores de las dos tiradas es un número par". Es fácil ver que $P(A) = 1 / 36$, ya que hay 36 resultados posibles y sólo en uno –el caso (1,1)– el máximo vale 1. Para calcular la $P(B)$, construimos el espacio muestral, en el que destacamos los resultados de B:

$$\begin{matrix} \underline{(1\ 1)} & (1\ 2) & \underline{(1\ 3)} & (1\ 4) & \underline{(1\ 5)} & (1\ 6) \\ (2\ 1) & \underline{(2\ 2)} & (2\ 3) & \underline{(2\ 4)} & (2\ 5) & \underline{(2\ 6)} \\ \underline{(3\ 1)} & (3\ 2) & \underline{(3\ 3)} & (3\ 4) & \underline{(3\ 5)} & (3\ 6) \\ (4\ 1) & \underline{(4\ 2)} & (4\ 3) & \underline{(4\ 4)} & (4\ 5) & \underline{(4\ 6)} \\ \underline{(5\ 1)} & (5\ 2) & \underline{(5\ 3)} & (5\ 4) & \underline{(5\ 5)} & (5\ 6) \\ (6\ 1) & \underline{(6\ 2)} & (6\ 3) & \underline{(6\ 4)} & (6\ 5) & \underline{(6\ 6)} \end{matrix}$$

Claramente, $P(B) = 18 / 36 = 1 / 2$. Ahora bien:

$$P(A \mid B) = P(A \cap B)/P(B) = (1/36) \;/\; (1/2) = 1/18$$

Había que esperar este valor, ya que dentro de B (que tiene 18 elementos) sólo hay un resultado que pertenezca a A. Dado que $P(A \mid B) \neq P(A)$, podemos concluir que A y B no son independientes.

Interpretación de dos sucesos independientes

En el caso del suceso C, la proporción de pares dentro de C es 1 / 2, igual que la proporción de pares con respecto al total; por tanto, el hecho de que ocurra C no afecta a la probabilidad de A, o lo que es lo mismo, A y C son independientes.

En resumen, y si nos fijamos en la relación entre los valores de $P(A)$ y $P(A \mid B)$, tenemos las posibilidades siguientes:

a) $P(A \mid B) > P(A)$; en este caso, podemos decir que el suceso B favorece que ocurra el suceso A, ya que la probabilidad de A sabiendo que hemos obtenido un resultado de B es mayor que la probabilidad de A solo.

b) $P(A \mid B) < P(A)$; en este caso podemos decir que el suceso B dificulta que ocurra el suceso A, ya que la probabilidad A disminuye si se da un resultado de B.

c) $P(A \mid B) = P(A)$; en este caso, el hecho de que ocurra un resultado de B no afecta a la probalilidad de A. Por tanto, ambos sucesos no se afectan mutuamente: son independientes.

8. Resumen

Esta sesión está dedicada íntegramente a definir el concepto de probabilidad y a deducir sus consecuencias más importantes. Se comienza trabajando sobre la noción de frecuencia relativa de sucesos para justificar la presentación de tres axiomas (las propiedades P1, P2 y P3), con las que quedan fijadas las características que debe tener una función de probabilidad. Una vez fijados los axiomas, se extraen las primeras consecuencias de la relación entre la probabilidad y las operaciones entre sucesos (unión, complementario...), considerando especialmente aquellas propiedades necesarias posteriormente en el desarrollo de la teoría.

Los axiomas de la probabilidad determinan cómo podemos asignar probabilidad a sucesos complejos a partir de sucesos más simples, pero también cabe preocuparse de la asignación de probabilidad a los sucesos más simples de todos: los resultados del experimento aleatorio; para hacerlo, hay que considerar dos casos, según si el espacio es uniforme o no:

1) Si el espacio es uniforme, es decir, si todos los resultados tienen la misma probabilidad, resulta que la probabilidad de un resultado es 1 dividido por el número de resultados posibles; como consecuencia, obtenemos la llamada regla de Laplace, según la cual la probabilidad de un suceso se obtiene dividiendo el número de resultados favorables al suceso por el número de resultados posibles.

2) En caso de que el espacio no sea uniforme, la probabilidad de los sucesos se aproxima por las llamadas probabilidades empíricas, que se corresponden al valor al que tienden las frecuencias relativas de los sucesos a medida que vamos repitiendo el experimento aleatorio.

En la segunda parte de la sesión se examina la noción de probabilidad condicionada, concepto que recoge y formaliza la posible influencia de un suceso en otro. Si suponemos que tenemos dos sucesos A y B, $P(A \mid B)$ denota la probabilidad de obtener un resultado de A sabiendo que se ha obtenido un resultado de B. En caso de que $P(A \mid B) = P(A)$, es decir, si el hecho de que se dé un resultado de B no afecta a la probabilidad del suceso A, diremos que A y B son independientes. Finalmente, se caracteriza la independencia de sucesos mediante la regla del producto de probabilidades, según la cual los sucesos A y B son independientes si $P(A \cap B) = P(A) \cdot P(B)$.

Ejercicios

1. Hemos fabricado dieciséis unidades de un determinado producto. Hemos obtenido diez buenos, cuatro con pequeños defectos y dos muy defectuosos. Los comerciantes sólo aceptan los artículos buenos, pero la compañía KaBaC sólo rechaza los muy defectuosos.

a) Si escogemos un artículo al azar, ¿cuál es la probabilidad de que sea aceptado por el comerciante?
b) Si escogemos dos, ¿cuál es la probabilidad de que los dos sean rechazados por la compañía KaBaC? ¿Y de que sea aceptado exactamente uno por el comerciante?
c) Si escogemos tres, ¿cuál es la probabilidad de que el comerciante no acepte ninguno?

2. Póquer
La baraja de póquer consta de cincuenta y dos cartas con cuatro palos (corazones, diamantes, tréboles y picas) y trece cartas de cada palo, numeradas así: as, 2, .., 10, *J*, *Q* y *K*. Cada vez se reparten cinco cartas, lo que se denomina *una mano de cinco cartas*.

a) ¿Cuál es la probabilidad de obtener una mano con un único as?
b) ¿Cuál es la probabilidad de obtener al menos un as?

3. Las alarmas
En una fábrica la probabilidad de que el sistema de alarma 1 falle es del 20%, la probabilidad de que falle el sistema de alarma 2 es del 10% y la probabilidad que fallen los dos al mismo tiempo, del 4%. ¿Cuál es la probabilidad de que:

a) al menos uno de los dos funcione?
b) funcionen los dos?

4. Defectos en cadena
El proceso de fabricación de un objeto pasa por dos cadenas independientes. La probabilidad de adquirir un defecto en la primera cadena es de 0,001 y en la segunda, del 0,0001. Calculad la probabilidad de que un objeto sea defectuoso.

5. Lanzamos una moneda perfecta tres veces:

a) ¿Cuál es la probabilidad de obtener un total de dos caras en las tres tiradas?
b) ¿Cuál es la probabilidad de obtener al menos una cara?
c) ¿Cuál es la probabilidad de obtener dos caras consecutivas, pero no tres caras?
d) Calculad la probabilidad de obtener dos caras consecutivas, pero no tres caras, sabiendo que, efectivamente, han salido dos caras.
e) ¿Los sucesos "obtener dos caras consecutivas pero no tres caras" y "obtener dos caras" son independientes?

f) Dad un ejemplo de dos sucesos relacionados con las tres monedas que sean independientes; justificad vuestra respuesta.

6. Lanzamos una moneda trucada tres veces. Repetid el problema anterior suponiendo que la moneda está trucada, de modo que la probabilidad de obtener cara en cada lanzamiento es 0,8.

7. Demostrad, utilizando las propiedades conocidas de la probabilidad, lo siguiente:

a) $P(A^C \mid B) = 1 - P(A \mid B)$
b) $P(B \mid A) = P(A \mid B) \cdot P(B)/P(A)$

8. Un estudio llevado a cabo entre los estudiantes de la universidad preguntaba, entre otras variables, el sexo y si alguna vez se había visitado el Valle de Nuria. Una vez procesadas las respuestas, se obtuvo lo siguiente:

- Respondieron 70 hombres y 30 mujeres.
- En total, 60 de los encuestados habían visitado el Valle de Nuria.
- 42 hombres habían visitado Nuria.

Se pide:

a) Calculad la probabilidad de que un encuestado seleccionado al azar sea hombre.
b) Calculad la probabilidad de que un encuestado seleccionado al azar sea hombre y haya visitado el Valle de Nuria.
c) Escogemos un cuestionario al azar y resulta ser el de un hombre. ¿Cuál es la probabilidad de que en este cuestionario se haya respondido "no" a la pregunta "¿Ha estado en el Valle de Nuria?".
d) Escogemos un cuestionario al azar y resulta ser el de una persona que afirma haber visitado el Valle de Nuria. ¿Cuál es la probabilidad de que este cuestionario corresponda a una mujer?
e) En vista de los resultados de la encuesta, ¿los sucesos "ser hombre" y "haber visitado el Valle de Nuria" son independientes? Interpretad el resultado.

Solucionario

1. Estamos en una situación en la que todos los posibles resultados tienen la misma probabilidad. Así, para calcular la probabilidad de un suceso, utilizaremos la fórmula:

a) El número total de resultados es 16 y el número a favor del suceso es 10:

$$\frac{\text{Número de resultados a favor del suceso}}{\text{Número total de resultados}}$$

Por tanto, P (“aceptado por el comerciante”) = 10/16.

b) En este caso el número total de resultados es el binomio:

$$\binom{16}{2} = 120$$

(es como escoger dos elementos de 16). Los resultados en los que ambos sean rechazados por la compañía KaBaC son únicamente uno (tenemos que coger los dos productos muy defectuosos) y, por tanto:

$$P\text{ (“los dos rechazados por KaBaC”)} = \frac{1}{\binom{16}{2}} = 0{,}00833$$

Los resultados a favor de que el comerciante acepte exactamente uno son 10 * 6 = 60 (tenemos que escoger uno entre los diez buenos y uno de entre los seis restantes). Por tanto:

$$P\text{ (“uno aceptado por el comerciante”)} = \frac{60}{\binom{16}{2}} = 0{,}5$$

c) Con un razonamiento análogo, obtenemos:

$$P\text{ (“comerciante que no acepte ninguno de los tres”)} = \frac{\binom{6}{3}}{\binom{16}{3}} = 0{,}0357$$

2. Póquer

a) Dado que todas las manos son equiprobables, tenemos que dividir el total de manos en las que hay un único as por el total de manos, con lo que obtenemos el resultado siguiente:

$$\frac{\binom{48}{4} \cdot 4}{\binom{52}{5}} = 0{,}29947$$

b) Será 1 menos la probabilidad de no tener ningún as; evidentemente:

$$P\text{ (“no obtener ningún as”)} = \frac{\binom{48}{5}}{\binom{52}{5}} = 0{,}6588$$

ya que para no tener ningún as, debemos escoger cinco cartas entre las cuarenta y ocho que no son as. Así pues:

$$P\text{ (“obtener al menos un as”)} = 1 - P\text{ (“no obtener ningún as”)} = \\ = 1 - 0{,}6588 = 0{,}3412$$

3. Las alarmas

Definimos F_1 = "la alarma 1 falla" y F_2 = "la alarma 2 falla". Sabemos que $P(F_1) = 0,2$; $P(F_2) = 0,1$ y $P(F_1 \cap F_2) = 0,04$. Entonces:

a)

$$P(\text{“que al menos uno de los dos funcione”}) =$$
$$= 1 - P(\text{“que ninguno funcione”}) = 1 - P(F_1 \cap F_2) = 1 - 0,04 = 0,96$$

b)

$$P(\text{“que funcionen los dos”}) = 1 - P(\text{“que alguno no funcione”}) =$$
$$= 1 - P(F_1 \cup F_2) = 1 - (P(F_1) + P(F_2) - P(F_1 \cap F_2)) =$$
$$= 1 - (0,2 + 0,1 - 0,04) = 0,74$$

4. Defectos en cadena

Definimos D_1 = "defecto en la cadena 1" y D_2 = "defecto en la cadena 2". Sabemos que:

$$P(D_1) = 0,001;\ P(D_2) = 0,0001 \text{ y } P(D_1 \cap D_2) = P(D_1) \cdot P(D_2) = 0,0000001$$

ya que, al ser las cadenas independientes, D_1 y D_2 son sucesos independientes. Entonces:

$$P(\text{“objeto defectuoso”}) =$$
$$= P(D_1 \cup D_2) = P(D_1) + P(D_2) - P(D_1 \cap D_2) = 0,001 + 0,0001 - 0,000001$$

5. Lanzamos una moneda perfecta tres veces.

En total tenemos ocho posibilidades: {ccc, cc+, c+c, +cc, c++, +c+, ++c, +++}; calcularemos las probabilidades y después comprobaremos el resultado buscando los sucesos favorables; observad que los ocho resultados tienen la misma probabilidad: 1/8.

a) $\dfrac{\binom{3}{2}}{8} = 0,37$; ya que de las tres tiradas dos tienen que ser caras. También se puede calcular haciendo:

$$\frac{\text{Card (“en total dos caras”)}}{8} = \frac{\text{Card}(\{\text{cc+, c+c,+cc}\})}{8} = \frac{3}{8}$$

b) La probabilidad de obtener al menos una cara es:

$$1 - \text{P(“no obtener ninguna cara”)} =$$
$$= 1 - \frac{\text{Card (“no obtener ninguna cara”)}}{8} = 1 - \frac{\text{Card}(\{+++\})}{8} = 1 - \frac{1}{8}$$

c) La probabilidad de obtener dos caras consecutivas, pero no tres caras, es:

$$\frac{\text{Card}(\{\text{cc+,+cc}\})}{8} = \frac{2}{8}$$

d) La probabilidad que nos piden en el enunciado es:

P ("dos caras consecutivas, pero no tres caras" | "han salido dos caras") =
= P ("dos caras consecutivas, pero no tres caras" $\cap$ "han salido dos caras) /
/ P ("han salido dos caras") = (2/8) / (3/8) = 2/3.

e) Hemos visto que:

P ("dos caras consecutivas, pero no tres caras"|"han salido dos caras") = 2/3

pero

P ("dos caras consecutivas, pero no tres caras") = 2/8

y, por tanto, los sucesos no son independientes.

f) Por ejemplo, consideremos los sucesos U = "la tercera vez sale cara" y F = "la primera vez sale cara"; tenemos que P(U) = 1/2 y P(F) = 1/2; también tenemos que P(U $\cap$ F) = = 2/8 = 1/4. Dado que P(U $\cap$ F) = 2/8 = 1/4 = P(U) · P(F), llegamos a la conclusión de que los sucesos son independientes.

6. Lanzamos una moneda trucada tres veces:
Calcularemos las probabilidades de cada uno de los posibles resultados:

$$P(\text{CCC}) = 0{,}8^3 = 0{,}512$$

ya que, puesto que las tiradas son independientes:

P(CCC) = P ("primera es cara" $\cap$ "segunda es cara" $\cap$ "tercera es cruz") =
= P (primera es cara) · P (segunda es cara) · P (tercera es cara) = $0{,}8^3$;

$$P(\text{CC+}) = 0{,}8^2 \cdot 0{,}2$$

Ya que las tiradas son independientes:

P(CC+) = P ("primera es cara" $\cap$ "segunda es cara" $\cap$ "tercera es cruz") =
= P ("primera es cara") · P ("segunda es cara") · P ("tercera es cruz") = $0{,}8^2 \cdot 0{,}2$.

Siguiendo este razonamiento, tenemos la tabla siguiente:

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Resultado	CCC	CC+	C+C	+CC	C++	+C+	++C	+++
Probabilidad	$0{,}8^3$	$0{,}8^2 \cdot 0{,}2$	$0{,}8^2 \cdot 0{,}2$	$0{,}8^2 \cdot 0{,}2$	$0{,}8 \cdot 0{,}2^2$	$0{,}8 \cdot 0{,}2^2$	$0{,}8 \cdot 0{,}2^2$	$0{,}2^3$

- La probabilidad de obtener dos caras será la suma de las probabilidades de los sucesos 2, 3 y 4; es, pues: $0{,}8^2 \cdot 0{,}2 + 0{,}8^2 \cdot 0{,}2 + 0{,}8^2 \cdot 0{,}2 = 0{,}384$.

- La probabilidad de obtener al menos una cara será la suma de las probabilidades de los resultados 1, 2, 3, 4, 5, 6, 7 o uno menos la probabilidad del resultado 8, es decir, será: $1 - 0{,}2^3 = 0{,}992$.

- Dos caras consecutivas, pero no tres caras, se corresponden con los sucesos 2 y 4; la suma de sus probabilidades es: $0{,}8^2 \cdot 0{,}2 + 0{,}8^2 \cdot 0{,}2 = 0{,}256$.

- Ahora debemos calcular:

P ("dos caras consecutivas, pero no tres caras"$\cap$"han salido dos caras") /
/ P ("han salido dos caras") = 0,256/0,384 = 0,67.

- Hemos visto que:

P ("dos caras consecutivas, pero no tres caras" | "han salido dos caras") = 0,67

pero P ("dos caras consecutivas, pero no tres caras") = 0,256 y, por tanto, los sucesos no son independientes.

- Consideremos los sucesos del ejercicio anterior:

U = "la tercera vez sale cara" y F = "la primera vez sale cara"

Tenemos que:

$$P(U) = 0{,}8^3 + 0{,}8^2\, 0{,}2 + 0{,}8^2\, 0{,}2 + 0{,}8\; 0{,}2^2 = 0{,}8$$

Es evidente que $P(F) = 0{,}8$; también tenemos que $P(U \cap F) = 0{,}8^3 + 0{,}8^2 \cdot 0{,}2 = 0{,}64$. Dado que $P(U \cap F) = 0{,}64 = P(U) \cdot P(F)$, llegamos a la conclusión de que los sucesos son independientes.

7.

a) Si aplicamos la definición, tenemos que: $P(A^C \mid B) = P(A^C \cap B) \;/\; P(B)$; por otro lado, sabemos que $P(B) = P(B \cap A^C) + P(A \cap B)$ y, así, $P(B \cap A^C) = P(B) - P(A \cap B)$. Por tanto:

$$P(A^C|B) = \frac{P(B) - P(A \cap B)}{P(B)} = 1 - \frac{P(A \cap B)}{P(B)} = 1 - P(A|B)$$

b) $P(A|B) \cdot \dfrac{P(B)}{P(A)} = \dfrac{P(A \cap B)}{P(B)} \cdot \dfrac{P(B)}{P(A)} = \dfrac{P(A \cap B)}{P(A)} = \dfrac{P(B \cap A)}{P(A)} = P(B|A)$

8.

a) En total respondieron cien personas, de las cuales setenta eran hombres. Por tanto:

$$P(\text{"hombre"}) = 70/100$$

b) Dado que hay cuarenta y dos hombres que han visitado el Valle de Nuria, P ("hombre" $\cap$ "ha visitado Nuria") = 42/100.

c) Nos pide que calculemos la probabilidad siguiente:

P ("no ha visitado el Valle de Nuria" | "hombre") =
= P ("no ha visitado el Valle de Nuria" $\cap$ "hombre") / P ("hombre") =
= (28/100) / (70/100) = 0,4

d) Nos piden:

P ("mujer" | "ha visitado el Valle de Nuria") =
P ("mujer" $\cap$ "ha visitado el Valle de Nuria") / P ("ha visitado el Valle de Nuria") = (18/100) / (60/100) = 0,3

e) Podemos calcular las probabilidades siguientes:

P ("ha visitado el Valle de Nuria") = 60/100 = 0,6 y
P ("ha visitado el Valle de Nuria" | "hombre") = (42/100) / (70/100) = 0,6

Puesto que estas probabilidades son iguales, los sucesos son independientes, y al ser sucesos independientes, podemos afirmar que el hecho de ser hombre ni favorece ni dificulta el que una persona haya visitado el Valle de Nuria.

El teorema de Bayes

En esta sesión presentaremos el teorema de Bayes, un resultado importantísimo para entender las relaciones causa-efecto entre diferentes sucesos. Comenzaremos trabajando con particiones del espacio muestral y relacionándolas con la probabilidad condicionada.

1. Particiones

Comenzaremos por definir una partición de un conjunto cualquiera.

> Dado un conjunto cualquiera E, una **partición** de E es una colección de subconjuntos de E disyuntos, tales que su unión es el conjunto E.

Utilidad de las particiones

Como tendremos oportunidad de ver más adelante, el hecho de trabajar con dos particiones permite estudiar situaciones en las que se estudian consecutivamente dos características y ver cómo la primera influye en la segunda (y también cómo la segunda influye en la primera).

Es decir, $\{A_1, A_2, ..., A_m\}$ es una partición de E si se verifican las condiciones siguientes:

1) $A_i \cap A_j = \emptyset$ para todo i, j, y $\neq j$
2) $A_1 \cup \ldots \cup A_m = E$

Sobre un mismo conjunto podemos tener diferentes particiones según diferentes conceptos.

Algunas posibles particiones

1) Si consideramos el conjunto O de los ordenadores de nuestra empresa, podemos obtener una partición según su procesador, de manera que, por ejemplo,

$$O = \{\text{“ordenadores con PIII”}\} \cup \{\text{“ordenadores con PIV”}\}$$

(suponiendo que cada ordenador tiene sólo un procesador y que sólo tenemos dos tipos de procesador: PIII y PIV).

También tenemos una partición según la RAM que tienen instalada, según la cual, por ejemplo, $O = \{\text{“64 MB”}\} \cup \{\text{“128 MB”}\} \cup \{\text{“264 MB”}\}$ (siempre suponiendo que sólo tenemos estas posibilidades).

2) Si consideramos el conjunto E de los estudiantes de informática de Cataluña, podemos obtener diferentes particiones:

a) Según el sexo: $E = \{\text{“hombres”}\} \cup \{\text{“mujeres”}\}$

b) Según el número de hijos:

$$E = \{\text{“no tienen ningún hijo”}\} \cup \{\text{“tienen 1 hijo”}\} \cup \{\text{“tienen 2 hijos”}\} \cup \\ \cup \{\text{“tienen 3 hijos”}\} \cup \{\text{“tienen 4 hijos”}\} \cup \{\text{“tienen más de 4 hijos”}\}$$

c) Según si les gustan los cebollinos:

$E = \{\text{“estudiantes a los que les gustan los cebollinos”}\} \cup \{\text{“estudiantes a los que no les gustan los cebollinos”}\}$

2. Teorema de las probabilidades totales

El **teorema de las probabilidades totales** afirma que si $\{A_1, A_2, ..., A_m\}$ es una partición del espacio muestral Ω y B es cualquier suceso, tenemos que:

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + P(B \cap A_3) + ... + P(B \cap A_m)$$

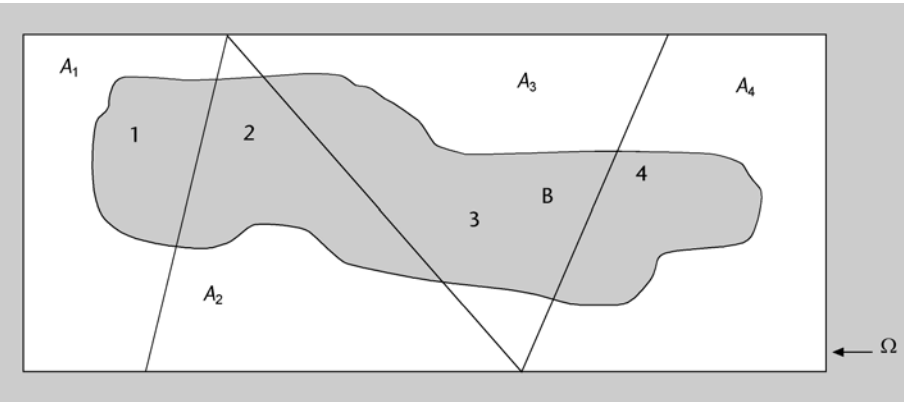
o lo que es lo mismo:

$$P(B) = P(B|A_1)\cdot P(A_1) + P(B|A_2)\cdot P(A_2) + P(B|A_3)\cdot P(A_3) + ... + P(B|A_m)\cdot P(A_m)$$

Demostración del teorema de las probabilidades totales

La demostración del resultado del teorema de las probabilidades es puramente conjuntista. Es fácil ver que dado que $\{A_1, A_2, ..., A_m\}$ es una partición del espacio muestral, tenemos que $B = (B \cap A_1) \cup (B \cap A_2) \cup ... \cup (B \cap A_m)$, y que dado que $B \cap A_i$ y $B \cap A_j$ son siempre disyuntos (ya que A_i y A_j lo son), tenemos, por la probabilidad de la unión de conjuntos disyuntos que $P(B) = P(B \cap A_1) + P(B \cap A_2) + ... + P(B \cap A_m)$. Puesto que $P(B \cap A_i) = P(B \mid A_i)\ P(A_i)$, tenemos la segunda fórmula del teorema.

Gráficamente el teorema de las probabilidades totales se entiende muy fácilmente (supongamos que la partición consta de cuatro conjuntos): primero dibujamos el espacio muestral Ω y los conjuntos de la partición, conjuntos que se reparten unos con otros todos los puntos del espacio muestral, sin que ningún punto esté simultáneamente en dos de los conjuntos; ahora representamos un suceso B arbitrario; este suceso se puede descomponer en la unión de los resultados de B que están en cada uno de los A_i.



Representación gráfica del teorema

Los elementos de B están repartidos entre los conjuntos $A_1, A_2, ..., A_m$: cada elemento de B está en uno y sólo en uno de los A_i. La región 1 es $A_1 \cap B$, la 2, $A_2 \cap B$,...

Ejemplos de aplicación del teorema de las probabilidades totales

a) Supongamos que dentro del colectivo de estudiantes de la UOC los chicos fumadores representan el 15% del total y que las chicas fumadoras representan el 12% del total. ¿Qué porcentaje de alumnos de la UOC fuma? En este ejemplo consideramos B = "fumar" y la partición $\{A_1, A_2\}$, donde A_1 = "hombre" y A_2 = "mujer". Por aplicación del teorema de las probabilidades totales y dado que {"hombre", "mujer"} es una partición del conjunto de alumnos, tenemos que:

$$P(\text{"fumar"}) = P(\text{"fumar"} \cap \text{"hombre"}) + P(\text{"fumar"} \cap \text{"mujer"}) = 0{,}15 + 0{,}12 = 0{,}27$$

b) En otra universidad el 54% de los estudiantes son hombres y el 46%, mujeres; por otra parte, el 30% de los hombres son fumadores, mientras que de las mujeres el 25% son fumadoras. ¿Qué porcentaje de alumnos de esta universidad fuma? Observad que en este caso disponemos de información sobre el porcentaje de fumadores dentro de cada sexo, es decir, disponemos de las probabilidades condicionadas de fumar por cada sexo, P (“fuma” | “hombre”) = 0,3, P (“fuma” | “mujer”) = 0,25. Por aplicación del teorema de las probabilidades totales y puesto que {“hombre”, “mujer”} es una partición del conjunto de alumnos, tenemos que:

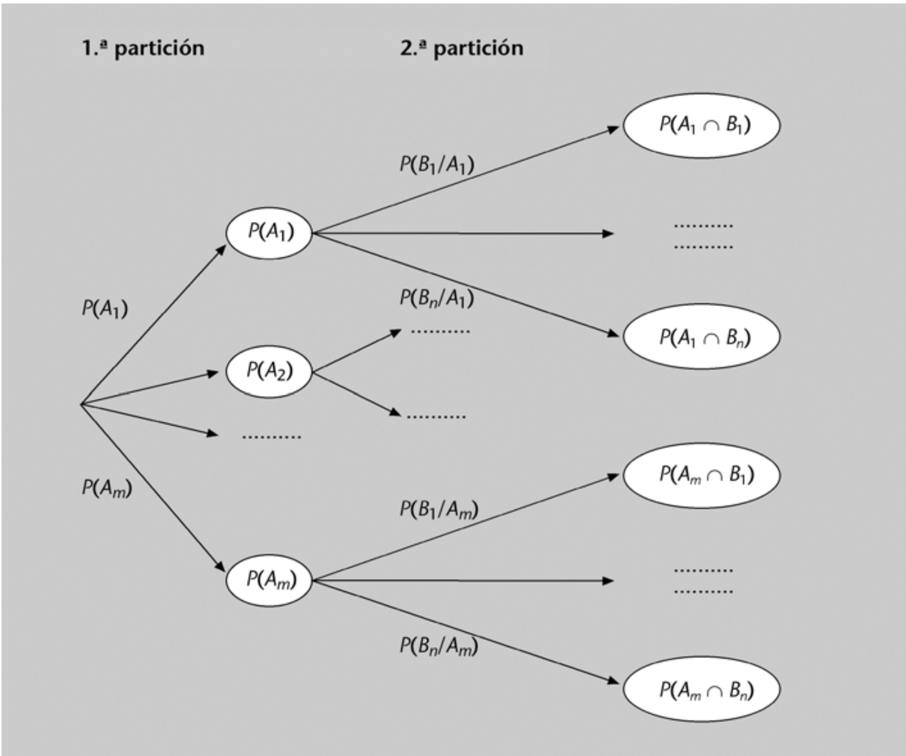
P (“fumar”) = P (“fumar” | “hombre”) · P (“hombre”) + P (“fuma” | “mujer”) · P (“mujer”) = 0,30 · 0,54 + 0,25 · 0,46 = 0,277

3. Árboles de probabilidad y probabilidad condicionada

Orden de las particiones

Si disponemos de dos particiones, consideraremos normalmente como primera partición aquella que contiene los sucesos que se dan antes en el tiempo, o bien aquellos que se pueden considerar causas de los sucesos de la segunda partición.

Supongamos que ahora disponemos de dos particiones $\{A_1, A_2, ..., A_m\}$ y $\{B_1, B_2, ..., B_n\}$ del espacio muestral; una manera muy útil de representar las particiones y las probabilidades de la forma $P(A_i \mid B_j)$ y $P(A_i \cap B_j)$ para todos los posibles pares (A_i, B_j) consiste en dibujar el árbol de probabilidad de estas dos particiones, árbol que presenta la forma siguiente:



en el que, como se puede observar:

a) En cada nodo tenemos la probabilidad de estar situados en éste.

b) En cada rama tenemos la probabilidad de pasar por la rama, suponiendo que estemos situados en el nodo en el que ésta comienza.

Además:

a) En cada nodo terminal tenemos la probabilidad de la intersección, que se puede calcular como el producto de las probabilidades de las ramas que hay que coger para llegar al nodo, ya que $P(A_i \cap B_j) = P(A_i) \cdot P(B_j \mid A_i)$.

b) La suma de las probabilidades de las ramas que salen de un mismo nodo es 1 (esto se debe a que tanto las A_i como las B_j son particiones del espacio muestral).

c) La suma de los valores de todos los nodos terminales ha de ser 1.

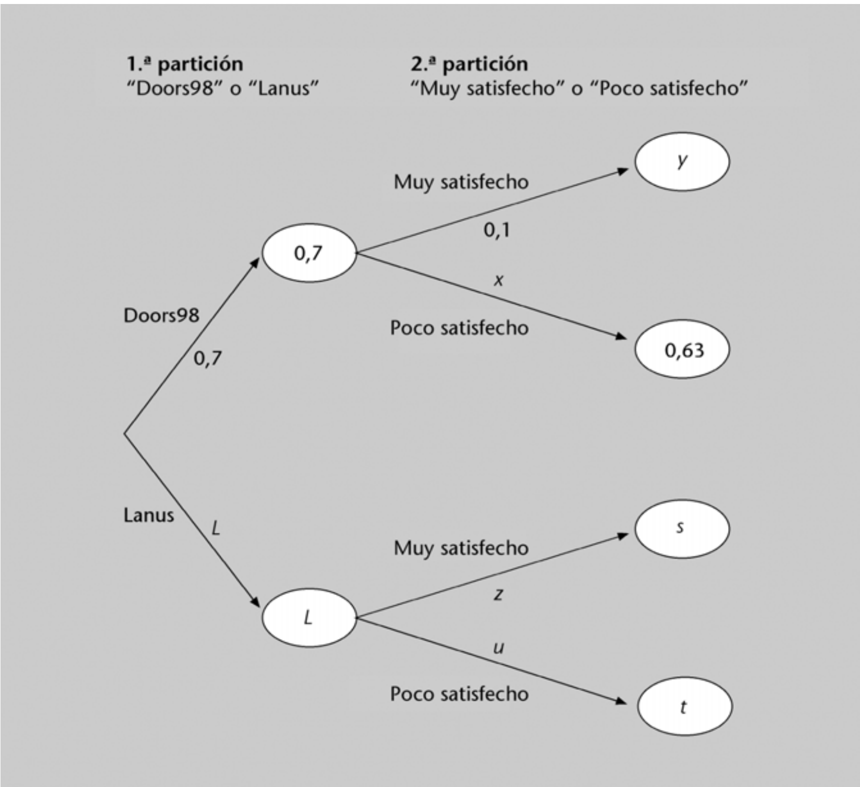
d) Por el teorema de las probabilidades totales tenemos que $P(B) = P(B \cap A_1) + \ldots + P(B \cap A_m)$. Por tanto, la probabilidad de B es la suma de las probabilidades de todos los nodos terminales a los que la rama que llega es B.

Ejemplo: Lanus frente a Doors98

Llevamos a cabo un estudio sobre el grado de satisfacción de los usuarios con el sistema operativo con el que trabajan. En nuestra empresa sólo tenemos ordenadores con Doors98 y Lanus y sólo se podía contestar "Muy satisfecho" y "Poco satisfecho". El espacio muestral son los usuarios que están distribuidos en dos particiones: {"Doors98", "Lanus"}, si consideramos el sistema operativo y {"Muy satisfecho", "Poco satisfecho"}, si consideramos su grado de satisfacción. Supongamos que de la encuesta se desprende que:

- P ("Doors98") = 0,7
- P ("Muy satisfecho" | "Doors98") = 0,1
- P ("Doors98" $\cap$ "Poco satisfecho") = 0,63
- P ("Muy satisfecho") = 0,34

En este caso concreto conviene considerar como primera partición el sistema operativo, ya que pensamos que puede ser la causa de la satisfacción o insatisfacción de los usuarios. Con lo que tenemos el árbol siguiente con algunas incógnitas que hay que encontrar:

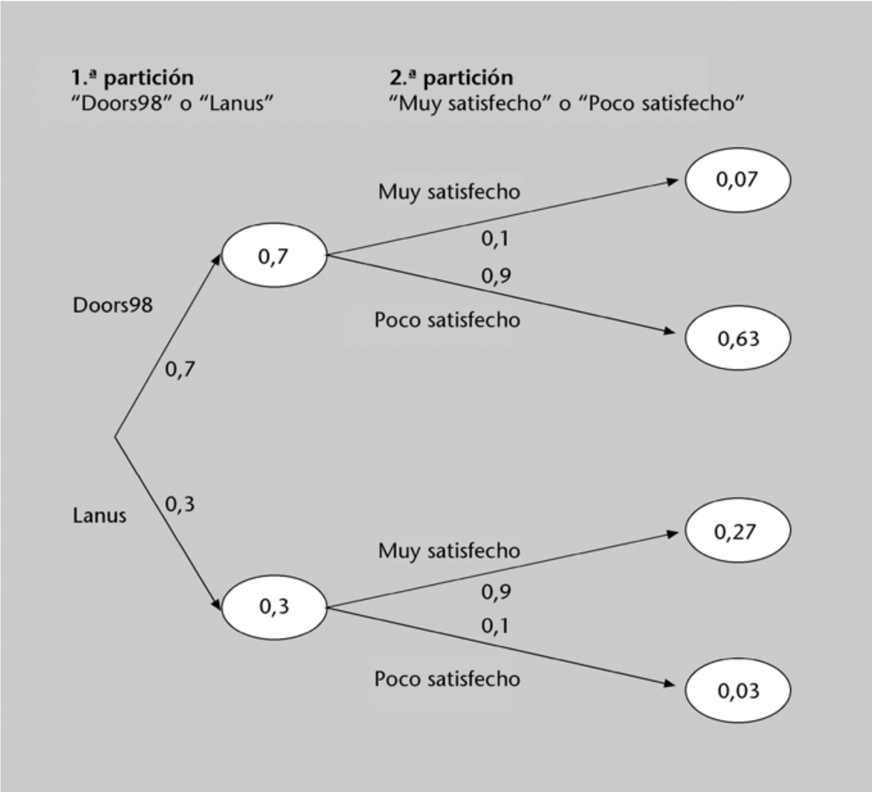


Ahora tenemos que:

- $L = 0,3$, ya que $L + 0,7 = 1$
- $y = P(\text{"Doors98"} \cap \text{"Muy satisfecho"}) = P(\text{"Muy satisfecho"} \mid \text{"Doors98"}) \cdot P(\text{"Doors 98"}) =$
$= 0,1 \cdot 0,7 = 0,07$

- Dado que s e y son los nodos terminales tales que la rama que llega a ellos es "Muy satisfecho", tenemos que $s + y = P(\text{"Muy satisfecho"}) = 0{,}34$ y, por tanto, $s = 0{,}34 - 0{,}07 = 0{,}27$.
- Puesto que $z \cdot L = s$, tenemos que $z = s/L = 0{,}27/0{,}3 = 0{,}9$.
- Ahora tenemos que $u = 0{,}1$ (ya que $u + z = 1$) y que $t = L \cdot u = 0{,}3 \cdot 0{,}1 = 0{,}03$.

Finalmente, el árbol queda de la manera siguiente:



Así, podemos ver que el uso de "Lanus" causa el efecto de mucha satisfacción en el 90% de los casos, mientras que, por ejemplo, el uso de "Doors98" sólo causa mucha satisfacción en un 10% de los casos.

Árboles como los presentados en esta sección permiten considerar las relaciones entre la primera partición y la segunda en términos de causas y efectos: la primera partición (usar "Doors98" o "Lanus") son las posibles causas de los efectos recogidos en la segunda partición ("estar muy satisfecho" o "estar poco satisfecho").

4. Tablas de contingencia

Un mecanismo muy útil para describir y encontrar relaciones de dependencia o independencia entre sucesos son las llamadas **tablas de contingencia.**

Formalmente tenemos, como en los árboles de probabilidad, dos particiones $\{A_1, A_2, ..., A_m\}$ y $\{B_1, B_2, ..., B_n\}$ del espacio muestral y construimos una tabla como ésta, tabla que contiene como filas los sucesos de una partición y como columnas, los sucesos de la otra partición; en la intersección de cada fila y columna aparece la probabilidad de la intersección de los correspondientes sucesos:

	A_1	A_2	...	A_m	Total
B_1	$P(B_1 \cap A_1)$	$P(B_1 \cap A_2)$	...	$P(B_1 \cap A_m)$	$P(B_1)$
B_2	$P(B_2 \cap A_1)$	$P(B_2 \cap A_2)$	...	$P(B_2 \cap A_m)$	$P(B_2)$
...					
B_n	$P(B_n \cap A_1)$	$P(B_n \cap A_2)$	...	$P(B_n \cap A_m)$	$P(B_n)$
Total	$P(A_1)$	$P(A_2)$	...	$P(A_m)$	1

En la última columna y en la última fila aparecen las llamadas *probabilidades marginales*, que son las probabilidades de cada suceso de las particiones involucradas; se pueden calcular sumando las probabilidades de la fila o de la columna correspondiente.

Independencia de dos sucesos

Para comprobar la independencia de dos sucesos, siempre se tiene que exhibir un argumento basado en el cálculo de las probabilidades de A, de B, de $A \cap B$, de $A \mid B$ o de $B \mid A$ y en la relación entre estas probabilidades.

Estas tablas permiten visualizar fácilmente la independencia de sucesos, ya que, por ejemplo, A_2 y B_2 son independientes si $P(A_2 \cap B_2) = P(A_2) \cdot P(B_2)$, es decir, si el producto de los marginales correspondientes es, efectivamente, la probabilidad de la intersección. Para calcular las probabilidades condicionadas, dividiremos la casilla correspondiente por el total de la fila (si condicionamos por el valor de la fila) o por el total de la columna (si condicionamos sobre el total de la columna).

Ejemplos de utilización de tablas de contingencia

a) Supongamos que a partir de los datos de los ordenadores de una empresa de la competencia obtenemos la tabla de contingencia siguiente, que relaciona el procesador (*A*1, *A*2) y la memoria RAM de las máquinas:

	64 MB	128 MB	Total
*A*1	10/50	20/50	30/50
*A*2	10/50	10/50	20/50
Total	20/50	30/50	1

A partir de la tabla podemos concluir lo siguiente:

- Los sucesos "*A*1" y "64 MB" de RAM no son independientes, ya que $(30/50) \cdot (20/50) \neq$ $\neq 10/50$
- La probabilidad de tener 128 MB de RAM sabiendo que tiene un "*A*2" es P("128 MB" $\cap$ $\cap$ "*A*2") / / P("*A*2") = (10/50) / (20/50)
- La probabilidad de tener 64 MB de RAM sabiendo que tiene un "*A*1" es P("64 MB" $\cap$ $\cap$ "*A*1") / / P("*A*1") = (10/50) / (30/50)

b) En el ejemplo Lanus frente a Doors98 podemos obtener la tabla de contingencia siguiente:

	Muy satisfecho	Poco satisfecho	Total
Doors98	0,07	0,63	0,7
Lanus	0,27	0,03	0,3
Total	0,34	0,66	1

En la que vemos claramente que ser usuario de Lanus y estar muy satisfecho no son independientes, ya que ($0{,}3 \cdot 0{,}34 \neq 0{,}27$). Cosa razonable, ya que, como hemos demostrado antes, usar Lanus aumenta el nivel de satisfacción del usuario.

5. El teorema de Bayes

A continuación consideraremos la misma situación de la que parte el teorema de las probabilidades totales, es decir, que disponemos de un suceso B y de una partición $\{A_1, A_2, ..., A_m\}$; de esta manera, podemos calcular la probabilidad de B como:

$$P(B) = P(B \mid A_1) \cdot P(A_1) + P(B \mid A_2) \cdot P(A_2) + ... + P(B \mid A_m) \cdot P(A_m).$$

Es decir, sabiendo las probabilidades de B condicionado a cada uno de los A_i y las probabilidades de los A_i, podemos obtener la probabilidad de B. Si queremos interpretar este resultado en términos de causa-efecto, tenemos que el suceso B se puede producir como efecto de m posibles causas ($A_1, ..., A_m$) y que para calcular la probabilidad de B, debemos hacer la suma de los productos de la probabilidad que se dé en cada una de las causas por la probabilidad que, dada la causa, ocurra efectivamente B.

Ahora “giraremos” la situación e intentaremos calcular la probabilidad de cada causa, sabiendo que se ha producido el efecto B. Esto nos puede parecer un tanto complicado, pero resulta que calcular las probabilidades de cada uno de los A_i (causas) condicionados a B (efecto) es bastante fácil, manipulando de forma adecuada las probabilidades.

El **teorema de Bayes** establece que si $\{A_1, A_2, ..., A_m\}$ es una partición de Ω y son conocidas las probabilidades de B condicionado a cada uno de los A_i y la probabilidad de cada uno de los A_i, podemos calcular la $P(A_i \mid B)$ de la manera siguiente:

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

$$= \frac{P(B|A_i)P(A_i)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3) + \ldots + P(B|A_m)P(A_m)}$$

Demostración del teorema de Bayes

Este teorema es muy fácil de demostrar dando los pasos siguientes:

1) Primero, la definición de probabilidad condicionada según la cual:

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

2) Después, sólo hay que aplicar el teorema de las probabilidades totales y sustituir $P(B)$ por $P(B \mid A_1) \cdot P(A_1) + P(B \mid A_2) \cdot P(A_2) + ... + P(B \mid A_m) \cdot P(A_m)$.

Aplicación del teorema de Bayes a un caso particular

En el caso de Lanus frente a Doors98, tenemos las probabilidades de cada nivel de satisfacción condicionadas por el sistema operativo que se utiliza. Es decir, conocemos las probabilidades de los supuestos efectos (“satisfecho” o “no satisfecho”) condicionados a las supuestas causas (usar un sistema u otro).

Ahora podemos preguntarnos las probabilidades de utilizar cierto sistema a partir del grado de satisfacción del usuario; es decir, nos preguntamos las probabilidades de cada causa (uso de un sistema u otro) condicionadas por los efectos sobre los usuarios (grado de satisfacción). En este caso tendremos que utilizar el teorema de Bayes. Supongamos que nos piden *P*("Doors98" | "está muy satisfecho"); para calcularla, tendremos que efectuar las operaciones siguientes (en las que abreviamos "Doors98" por D98, "Lanus" por *L* y "Muy satisfecho" por *S*:

$$P(D98|S) = \frac{P(S|D98) \cdot P(D98)}{P(S|D98) \cdot P(D98) + P(S|L) \cdot P(L)} = \frac{0,1 \cdot 0,7}{0,1 \cdot 0,7 + 0,9 \cdot 0,3} = 0,2058$$

Y, por tanto, dentro del grupo de los usuarios satisfechos, la probabilidad de usar Doors98 es de 0,2058.

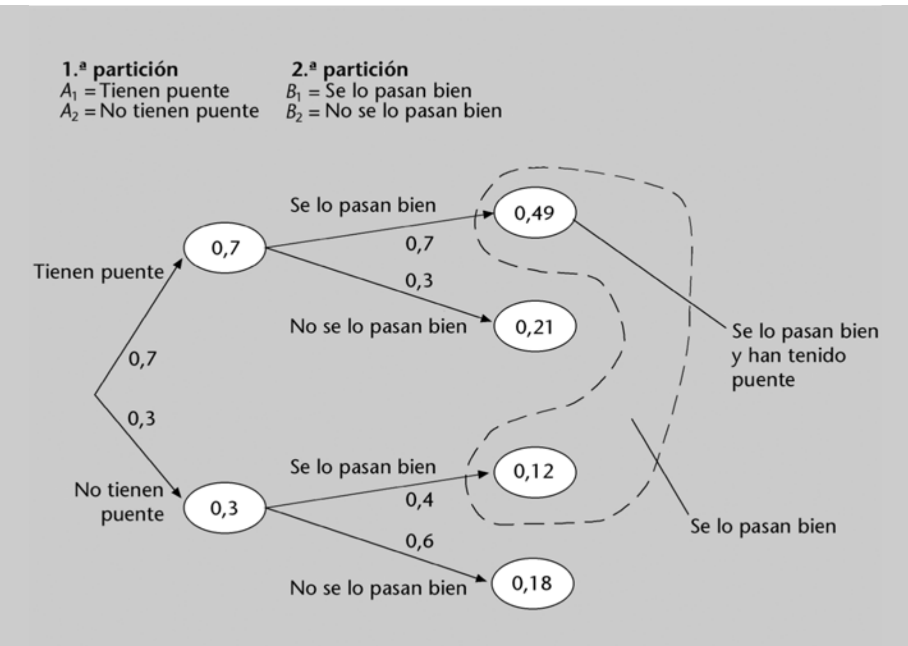
En este caso particular el teorema de Bayes permite invertir las relaciones causa-efecto, ya que es como si pudiésemos ir atrás en el tiempo y, partiendo del grado de satisfacción sobre el sistema informático, remontarnos a la probabilidad de la causa de esta satisfacción. En el ejemplo anterior, si escogemos a un usuario al azar y resulta estar satisfecho, tenemos una probabilidad de 0,2058 de que sea usuario de D98, mientras que tenemos una probabilidad de (1 – 0,2058) = 0,7942 de que sea usuario de "Lanus". Como podéis ver, ¡usar "Lanus" es una causa mucho más probable de satisfacción que usar D98!

6. El teorema de Bayes sobre un árbol de probabilidades

Si estamos trabajando con un árbol de probabilidades concreto, el teorema de Bayes tiene una formulación equivalente muy intuitiva; lo veremos con un ejemplo. Según la época del curso en el que estudiéis este material, es posible que estéis preparando el puente del 6 al 8 de diciembre. En este caso quizá os interesarían estos datos (ficticios) sobre cómo se lo pasa la gente en este puente:

a) El 70% de los catalanes tiene puente (o se lo coge).
b) De los catalanes que cogen puente, el 70% declara habérselo pasado bien.
c) De los catalanes que no cogen puente, el 40% declara habérselo pasado bien.
d) Sólo admitimos dos posibilidades: pasárselo bien o no.

El correspondiente árbol de probabilidades es el siguiente:



En este caso la primera partición es la que contiene los acontecimientos que ocurren temporalmente en primer lugar: primero se hace (o no se hace) puente y después se responde a la pregunta de si se lo han pasado bien. De alguna manera, ¡también estamos presuponiendo que la causa de pasárselo bien es haber hecho el puente!

Ahora nos preguntamos la probabilidad de que una persona haya tenido puente sabiendo que ha afirmado habérselo pasado bien ($P(A_1 \mid B_1)$). Si nos fijamos en el árbol, observamos que el 61% ($0{,}49 + 0{,}12 = 0{,}61$) dice que se lo ha pasado bien, pero sólo el 49% (0,49) ha tenido puente; por tanto, forzosamente tendremos que la probabilidad de haber disfrutado del puente sabiendo que se lo han pasado bien es de 80,32% ($0{,}49 / 0{,}61 = 0{,}8032$). Así pues, podemos pensar que en el 80,32% de los casos pasárselo bien es a causa del puente.

Es decir, la probabilidad P (“tiene puente | se lo ha pasado bien”) = P ($A_1 \mid B_1$) se calcula dividiendo la probabilidad del nodo “se lo pasan bien y han tenido puente” por la suma de las probabilidades correspondientes a “pasárselo bien”, es decir:

$$P(\text{tiene puente}|\text{se lo pasa bien}) =$$

$$\frac{\text{Probabilidad del nodo "tienen puente y se lo pasan bien"}}{\begin{array}{c}\text{Suma de las probabilidades de los nodos correspondientes}\\ \text{a pasárselo bien}\end{array}} =$$

$$= \frac{0{,}49}{0{,}49 + 0{,}12}$$

Y, en general, en un árbol de probabilidades en el que aparecen las probabilidades $P(B \mid A_i)$, la $P(A_i \mid B)$ se calcula como:

$$P(A_i|B) = \frac{\text{Probabilidad del nodo correspondiente a } A_i \cap B}{\text{Suma de las probabilidades de los nodos correspondientes a } B}$$

Cálculo de probabilidades por el teorema de Bayes

En el ejemplo del puente del 6 al 8 de diciembre, ¿cuáles son las probabilidades...

a) de tener puente sabiendo que no se lo han pasado bien?
b) de no tener puente sabiendo que no se lo han pasado bien?
c) de no tener puente sabiendo que se lo han pasado bien?
d) de que los sucesos “tener puente” y “pasárselo bien” sean independientes?

a)

$$P(\text{"tiene puente"}|\text{"no se lo pasa bien"}) =$$

$$= \frac{P(\text{"puente"} \cap \text{"no se lo pasa bien"})}{P(\text{"no se lo pasa bien"})} = \frac{0{,}21}{0{,}21 + 0{,}18} = 0{,}5384$$

Es decir, la probabilidad de que una persona haya tenido puente sabiendo que no se lo ha pasado bien es de 0,53; dicho de otro modo, el 53% de los que no se lo han pasado bien ha tenido puente (quizá estaban en una retención de tráfico).

Los puentes y las causas

En términos de causas y efectos, podemos entender tener puente o no tenerlo como causa de los efectos pasárselo bien o no. El teorema de Bayes permite encontrar las probabilidades de que los efectos provengan de alguna de las posibles causas. Por ejemplo, *P*(“tener puente” | “se lo ha pasado bien”) nos indica hasta qué punto tener puente se puede considerar como causa de pasárselo bien, ya que nos dice cuál es el porcentaje de quienes se lo han pasado bien que han tenido puente.

b)

$$P(\text{"no tiene puente"}|\text{"no se lo pasa bien"}) =$$

$$= \frac{P(\text{"no tiene puente"} \cap \text{"no se lo pasa bien"})}{P(\text{"no se lo pasa bien"})} = \frac{0{,}18}{0{,}21+0{,}18} = 0{,}4615$$

c) P(“no tiene puente” | “se lo ha pasado bien”) = 0,12 / (0,49 + 0,12) = 0,1967.

d) Los sucesos no son independientes, ya que:

P(“tiene puente” | “se lo ha pasado bien”) = 80,32%

que es mayor que la probabilidad de tener puente (70%); es decir, el hecho de pasárselo bien (efecto) hace que aumente la probabilidad de haber tenido puente (causa).

Acabaremos esta sesión con un fragmento de un libro de John Allen Paulos que esperamos que os ayude a entender mejor los cálculos y la interpretación que se pueden hacer del teorema de Bayes. Es recomendable que reescribáis el ejemplo en forma de árbol y que comprobéis los cálculos:

“Una elaboración interesante a partir del concepto de probabilidad condicional es el conocido teorema de Bayes, que fue demostrado por primera vez por Thomas Bayes en el siglo XVIII, y constituye la base del siguiente resultado, un tanto sorprendente, con importantes consecuencias para los análisis de SIDA o la detección del consumo de drogas.

Supongamos que haya un análisis para detectar el cáncer con una fiabilidad del 98%; es decir, si uno tiene cáncer el análisis dará positivo el 98% de las veces, *y* si no lo tiene, dará negativo el 98% de las veces. Supongamos además que el 0,5% de la población –una de cada doscientas personas– padece verdaderamente cáncer.

Imaginemos que uno se ha sometido al análisis y que su médico le informa en tono pesimista que ha dado positivo. ¿Hasta qué punto ha de deprimirse esa persona? Lo sorprendente del caso es que dicho paciente ha de mantenerse prudentemente optimista. El porqué de este optimismo lo encontramos al determinar la probabilidad condicional de que uno tenga cáncer sabiendo que el análisis ha dado positivo.

Supongamos que se hacen 10.000 pruebas de cáncer. ¿Cuántas de ellas darán positivo? En promedio 9,50 de estas 10.000 personas (el 0,5% de 10.000) tendrán cáncer, y como el 98% de ellos darán positivo, tendremos 49 análisis positivos. Por otra parte, el 2% de las 9.950 personas restantes, que no padecen cáncer, también darán positivo, con un total de 199 análisis positivos (0,02 · 9.950 = 199). Así, del total de 248 positivos (199 + 49 = = 248), la mayoría son falsos positivos, y la probabilidad condicional de padecer el cáncer sabiendo que se ha dado positivo es sólo 49/248, ¡aproximadamente un 20%! (Hay que comparar este porcentaje relativamente bajo con la probabilidad de dar positivo en el supuesto de que se tenga efectivamente el cáncer que, por hipótesis, es del 98%)”.

John Allen Paulos (1990). *El hombre anumérico*. Tusquets Editores (Metatemas, 20).

7. Resumen

En esta sesión se introducen dos resultados importantísimos para entender la probabilidad y cómo podemos aplicarla a multitud de situaciones: el teorema de las probabilidades totales y el teorema de Bayes. Para enunciar estos resultados, es necesario trabajar con particiones; para facilitar el seguimiento de la teoría y la visualización de los resultados, se introducen los árboles de probabilidad y las tablas de contingencia.

Los resultados de esta sesión son muy técnicos y resumir su contenido informalmente resulta muy difícil; de todos modos, podemos decir que el teorema de las probabilidades totales permite calcular la probabilidad de un suceso sabiendo la probabilidad del suceso en todas las posibles circunstancias en las que se puede dar. El teorema de Bayes permite calcular ciertas probabilidades condicionadas, probabilidades esenciales para entender muchos resultados; este teorema admite una presentación muy sencilla a partir de ciertos árboles de probabilidad.

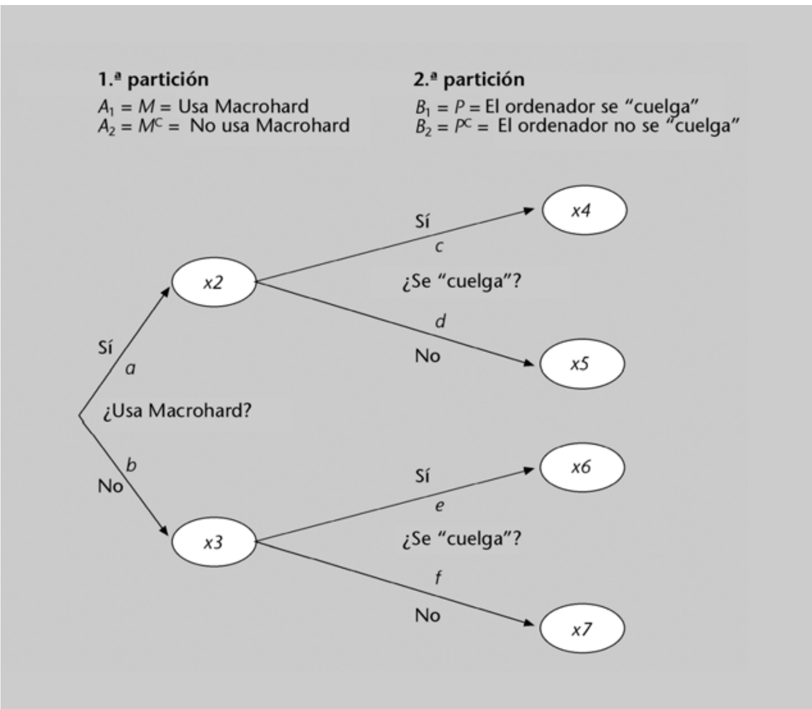
Ejercicios

1. Un amigo ha llevado a cabo un estudio sobre la satisfacción de los usuarios de diferentes entornos informáticos y ha resumido la información obtenida en forma de tabla de contingencia. La transmisión de la tabla por fax no ha sido demasiado buena y nos ha llegado con algunos "agujeros", que han sido sustituidos por letras mayúsculas:

Grado de satisfacción					
	Nada	**Poco**	**Bastante**	**Mucho**	**Total**
Linux	0,05	0,15	*A*	*B*	0,7
Windows	0,15	0,05	*C*	*D*	0,3
Total	0,2	0,2	0,3	0,3	1

a) ¿Cuál es la probabilidad de que un usuario de Linux esté poco satisfecho con el entorno?
b) ¿Cuál es la probabilidad de que una persona sea usuaria de Linux y responda que no está nada satisfecha?
c) Los sucesos "ser usuario de Windows" y "estar nada satisfecho", ¿son independientes?
d) Calculad el valor de *A*, si mi amigo me confirma telefónicamente que los sucesos "ser usuario de Linux" y "estar bastante satisfecho" son independientes.
e) Con el valor de *A* calculado antes, encontrad los valores de *B*, *C* y *D* y reconstruid la tabla de contingencia original.

2. Dado el árbol de probabilidades siguiente:



en el que si los sucesos son M = “Usa Macrohard Phrase” y P = “Se cuelga el ordenador”, sabemos lo siguiente:

- $P(M) = 0{,}8$
- $P(P^C \mid M) = 0{,}1$
- $P(P) = 0{,}74$

Se pide:

a) Acabad de rellenar el árbol de probabilidades (valores de $x2$, $x3$, $x4$, $x5$, $x6$, $x7$ y a, b, c, d, e, f).
b) ¿Los sucesos M y P son independientes?
c) ¿Cuánto vale $P(\mathrm{M}^C \cap P)$?

3. Se instala un programa antivirus en un ordenador. La probabilidad de que el ordenador tenga un virus detectable por el antivirus es de 0,2. Si el ordenador tiene el virus, la probabilidad de que el antivirus lo detecte es de 0,9. Si el ordenador no tiene el virus, la probabilidad de que el antivirus dé un mensaje de existencia de virus es de 0,02.

a) Demostrad que la probabilidad de que el antivirus detecte un virus (puede existir o no) es 0,196.
b) Calculad la probabilidad de que el ordenador no tenga virus y haya aparecido un mensaje de existencia de virus.
c) Calculad la probabilidad de que, si ha aparecido un mensaje de existencia de virus, el ordenador no lo tenga realmente.
d) Calculad la probabilidad de que el ordenador tenga el virus y el antivirus no lo detecte.
e) Calculad la probabilidad de que aunque no haya salido ningún mensaje de existencia de virus, el ordenador lo tenga.

Solucionario

1.

a) P(“poco satisfecho con el entorno” | “usuario de Linux”) =
= P(“poco satisfecho” ∩ “usuario de Linux”) / P(“usuario de Linux”) =
= 0,15/0,7 = 0,2143

b) P(“nada satisfecho” ∩ “usuario de Linux”) = 0,05

c) P(“nada satisfecho” ∩ “usuario de Windows”) = 0,15 ≠
≠ P(“nada satisfecho”) · P(“usuario de Windows”) = 0,2 · 0,3

Por tanto, no son independientes.

d) Si son independientes, se deberá cumplir:

A = P(“bastante satisfecho” ∩ “usuario de Linux”) =
= P(“bastante satisfecho”) · P(“usuario de Linux”) = 0,7 · 0,3 = 0,21

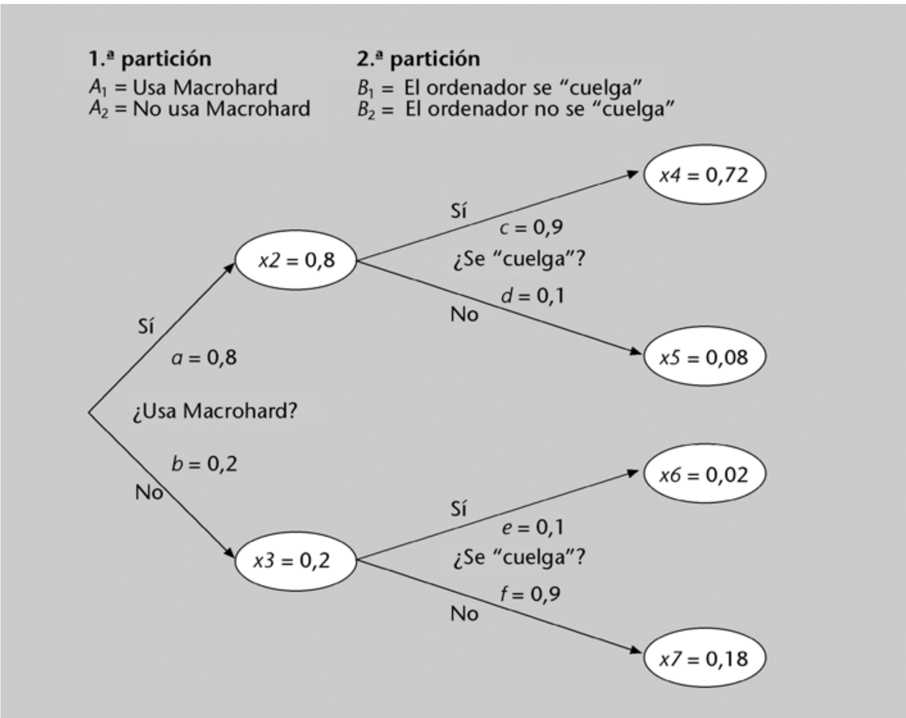
$$B = 0{,}7 - (0{,}05 + 0{,}15 + 0{,}21) = 0{,}29$$
$$C = 0{,}3 - 0{,}21 = 0{,}09;\ D = 0{,}3 - 0{,}29 = 0{,}01$$

2.

a) Observemos los cálculos siguientes:

- $a = P(M) = 0{,}8$
- $b = P(M^C) = 1 - P(M) = 1 - 0{,}8 = 0{,}2$
- $x2 = a = P(M) = 0{,}8$
- $x3 = b = P(M^C) = 1 - P(M) = 1 - 0{,}8 = 0{,}2$
- $c = P(P \mid M) = P(P \cap M) \ / \ P(M) = P(M \cap P) \ / \ P(M) = 0{,}72/0{,}8 = 0{,}9$
- $d = P(P^C \mid M) = 0{,}1$
- $x4 = P(M \cap P) = 0{,}72$
- $x5{=}P(M \cap P^C) = P(\mathrm{P}^C \mid M) \cdot P(\mathrm{M}) = 0{,}1 \cdot 0{,}8{=}\ 0{,}08;$
- $P(P) = 0{,}74 = P(P) = P(P \cap M) + P(P \cap M^C) = P(M \cap P) + P(M^C \cap P) =$ $= 0{,}72 + P(M^C \cap P)$ y, por tanto, $P(M^C \cap P) = 0{,}74 - 0{,}72 = 0{,}02$
  (o lo que es lo mismo, $x4 + x6 = 0{,}74$ y $x4 = 0{,}72$ implica $x6 = 0{,}02$)
- $x6 = P(M^C \cap P) = 0{,}02$
- $e = P(P \mid M^C) = P(M^C \cap P) \ / \ P(M^C) = 0{,}02/0{,}2 = 0{,}1$
- $f = 0{,}9$ (ja que $e + f = 1$)
- $x7 = P(M^C \cap P^C) = 0{,}18$ (también se puede hacer viendo que $x6 + x7 = 0{,}2$)

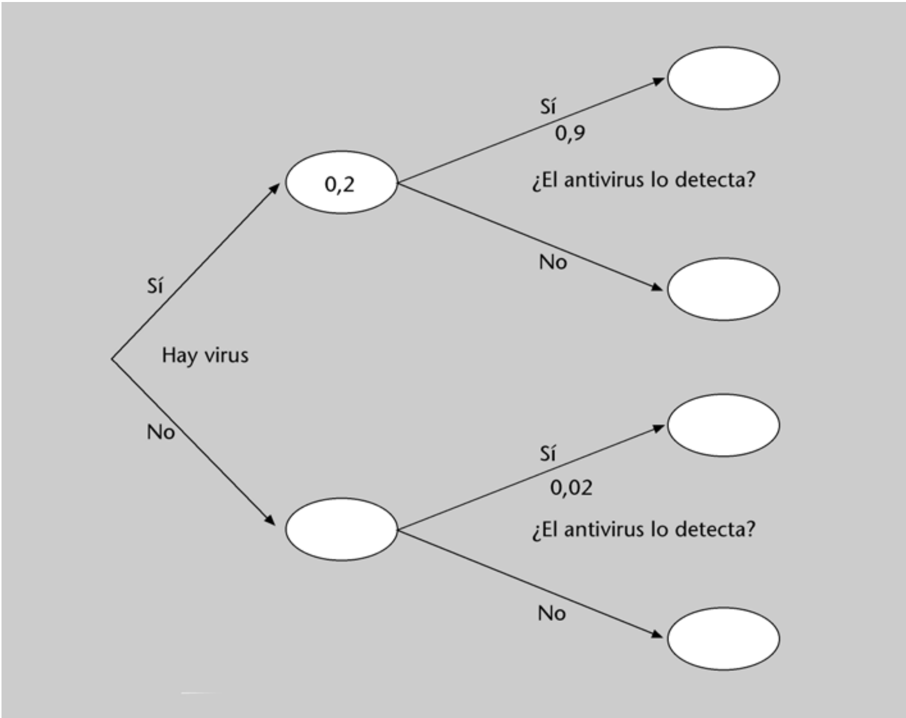
y el árbol completo queda de la manera siguiente:



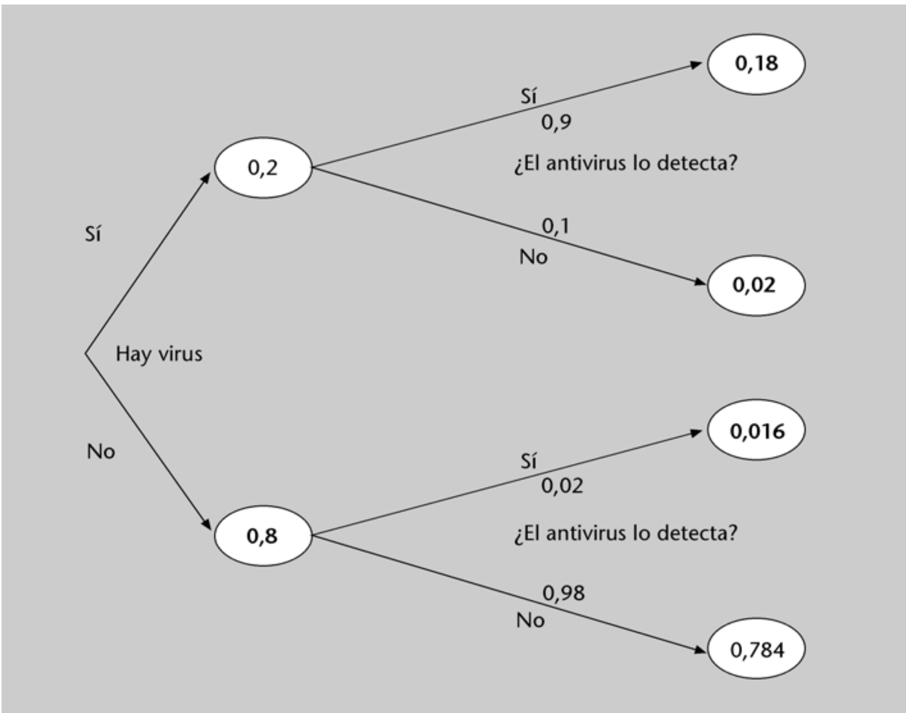
b) Dado que $P(P \mid M) = 0{,}9$ y $P(P) = 0{,}74$, está claro que $P(P \mid M) = 0{,}9 \neq P(P)$ y, por tanto, P y M no son independientes. También se puede hacer comprobando que $P(M \cap P) \neq \neq P(M) \cdot P(P)$.

c) $P(M^C \cap P)$ es la probabilidad de que "no use Macrohard Phrase" y "se "cuelgue" el ordenador". Esta probabilidad corresponde a $\mathrm{x}6 = P(M^C \cap P) = 0{,}02$.

3. Podemos considerar que este árbol en el que "hay virus" significa "hay virus detectable").



Decimos V = "el ordenador tiene virus detectable por el antivirus'' y D = "el antivirus detecta un virus''. Entonces, sabemos por el enunciado $P(V) = 0{,}2$; P("detecte el virus si hay virus") = $P(D \mid V) = 0{,}9$ y P("detecte el virus si no hay virus") = $P(D \mid V^C) = 0{,}02$. Inmediatamente, también sabemos que $P(D^C \mid V) = = 1 - 0{,}9 = 0{,}1$. Con esto podemos completar el árbol de la manera siguiente:



Ahora podemos contestar a las preguntas siguientes:

a) Por el teorema de las probabilidades totales:

$$P(\text{“el antivirus detecta un virus”}) =$$
$$= P(\text{“detecta y hay virus”} \cap \text{“detecta y no hay virus”}) =$$
$$= P(D) = P(D \cap V) + P(D \cap V^C) =$$
$$= P(D \mid V)\ P(V) + P(D \mid V^C)\ P(V^C) = (0{,}9 \cdot 0{,}2) + (0{,}02 \cdot 0{,}8) =$$
$$= 0{,}18 + 0{,}016 = 0{,}196$$

(es la suma de las probabilidades de los nodos a los que llega una rama del tipo **sí** lo detecta).

b)

$$P(\text{“el ordenador no tiene virus”} \cap \text{“el antivirus detecta un virus”}) =$$
$$= P(V^C \cap D) = P(D \mid V^C) \cdot P(V^C) = 0{,}02 \cdot 0{,}8 = 0{,}016$$

Es decir, con una probabilidad del 0,016 el ordenador no tiene virus, pero el antivirus da mensaje de existencia de virus.

c) Por el teorema de Bayes y el apartado **a**, tenemos que:

$$P(V^C \mid D) = P(V^C \cap D)\ /\ P(D) = 0{,}016/0{,}196 = 8{,}1633 \cdot 10^{-2} = 8{,}16\%$$

Es decir, el 8,16% de los casos en los que aparece un mensaje de existencia de virus, el antivirus se equivoca, ya que no hay virus.

d) $P(D^C \cap V) = P(D^C \mid V) \cdot P(V) = 0{,}1 \cdot 0{,}2 = 0{,}02 = 2\%$

Este caso sí es preocupante: ¡en el 2% de los casos el ordenador tiene virus y el antivirus no lo encuentra!

e) $P(V \mid D^C) = P(V \cap D^C)\ /\ P(D^C) = 0{,}02\ /\ (1 - 0{,}196) = 2{,}4876 \cdot 10^{-2} = 2{,}48\%$

Este caso también es preocupante: ¡el 2,48% de las veces que el antivirus no da ningún mnsaje de virus, efectivamente tenemos uno!