Curso a Distancia Thales-CICA-Web: ESTIA
Estadística Interactiva en el Aula.
Laboratorio Virtual de Estadística.

Contenidos Teóricos Tema 3.
Cálculo de Probabilidades.

A. Gámez, L.M. Marín, R. Rodríguez y S. Fandiño

Diciembre - 2005

Índice General

3 Cálculo de probabilidades **2**
- 3.1 Introducción a la teoría de la probabilidad 2
- 3.2 Definiciones de Probabilidad 3
  - 3.2.1 Fenómenos aleatorios 3
  - 3.2.2 Relaciones y Operaciones con sucesos 4
  - 3.2.3 Propiedades de las operaciones entre sucesos 5
  - 3.2.4 Definición frecuentista de probabilidad 5
  - 3.2.5 Definición clásica de probabilidad 6
  - 3.2.6 Propiedades de la probabilidad. 7
  - 3.2.7 Definición axiomática de probabilidad. 8
- 3.3 Recursos de interés para el cálculo de probabilidades de sucesos 10
  - 3.3.1 Regla de multiplicación 10
  - 3.3.2 Diagramas de árbol 11
  - 3.3.3 Combinatoria 11
  - 3.3.4 De lo particular a lo general 15
  - 3.3.5 La probabilidad geométrica 16
- 3.4 Probabilidad condicionada. 18
  - 3.4.1 Independencia de un par de sucesos 18
  - 3.4.2 Independencia de más de dos sucesos 20
- 3.5 Teorema de la Probabilidad Total. 20
- 3.6 Teorema de Bayes. 23

Capítulo 3

Cálculo de probabilidades

3.1 Introducción a la teoría de la probabilidad

Analizamos en primer lugar el uso del término “probable” en el lenguaje cotidiano. Supongamos que llaman a mi puerta y digo: “es muy probable que sea mi hermano”. ¿Cuál es el sentido de esta frase? Muestra una falta de información sobre la identidad de la persona que llama a la puerta. Por otra parte indica que, aunque no sé si es mi hermano o no el que llama a la puerta, mi opinión es bastante más favorable a la primera opción que a la segunda, manifestando por tanto un cierto grado de preferencia entre las opciones posibles. Esta preferencia esta fundamentada en algún tipo de información previa que puede ser, por ejemplo, que mi hermano viene a mi casa con mayor frecuencia que cualquier otra persona, o en otros datos que, aunque no sean definitivos para tomar una decisión segura, si son suficientes para decantarse a favor de una de las opciones posibles. Lo mismo ocurre si, al tirar un dado, comentamos que es menos probable sacar un cinco que no sacarlo. Mostramos un desconocimiento del resultado y una cuantificación de que nuestra opinión es más favorable a una de las opciones.

En esta situación estamos en la mayoría de los casos que ocurren a nuestro alrededor. No en vano se dice que *no hay nada seguro salvo la muerte*. Cuando un científico realiza un experimento o analiza un fenómeno natural, frecuentemente los resultados no pueden predecirse con certeza, aunque los experimentos u observaciones se hayan realizado en idénticas condiciones. Un ejemplo muy claro son los fenómenos meteorológicos. Es difícil hacer previsiones sobre la magnitud, intensidad y extensión de las lluvias, el porcentaje de humedad, la dirección y velocidad de los vientos, el valor de la temperatura máxima o mínima, etc. Lo mismo podemos decir sobre las consecuencias de estos fenómenos: volumen de agua en las presas, magnitud de los daños provocados por inundaciones, sequías, accidentes de tráfico, etc. A la imposibilidad de predecir con certeza los resultados de un fenómeno se le llama azar o aleatoriedad. Los fenómenos con esta característica se llaman fenómenos aleatorios (en principio esta palabra estaba relacionada exclusivamente con los juegos de azar. En latín alea significa suerte).

No obstante, el hecho de no tener certeza de los resultados que se obtendrán en cada prueba particular, no significa que no dispongamos de alguna información sobre estos. Así, aunque no podamos saber si una pareja determinada va a tener hijo o hija como primer retoño, si se sabe que el porcentaje de varones es aproximadamente del 51%. Una compañía de seguros no sabe a qué edad va a morir un cliente determinado, pero tiene información sobre la proporción de muertes por edades, sexo, etc. Esto le ayudará a decidir sobre el importe que debe exigir por las primas a los asegurados. El resultado de nuestro conocimiento sobre la mayor o menor frecuencia con que aparecen cada uno de los resultados posibles es lo que se pretende cuantificar con el concepto de Probabilidad.

3.2 Definiciones de Probabilidad

3.2.1 Fenómenos aleatorios

Diremos que un *experimento o prueba* es una acción que se realiza con el propósito de recoger algún tipo de observación sobre sus resultados. *Un experimento es aleatorio* cuando su resultado es impronosticable: Aunque el experimento se repita de la misma forma y bajo idénticas condiciones puede dar lugar a diferentes resultados. El concepto contrario de fenómeno aleatorio es el de *fenómeno determinista*, que se caracteriza por obtener los mismos resultados bajo idénticas condiciones. Un ejemplo típico de fenómeno determinista es la velocidad con que llega al suelo un móvil en caída libre, que según la Física Clásica, se obtiene con la fórmula $v = \sqrt{2gh}$. De esta forma el resultado sería único. Sin embargo no puede hablarse de fenómenos puramente deterministas. Supongamos que se desea medir la velocidad anterior experimentalmente. En ese caso, la medición viene influenciada por los instrumentos de medida, por la habilidad de los experimentadores, por la resistencia del aire, por las condiciones ambientales, por el lugar geográfico, etc. Todas estas condiciones influyen en los resultados de una forma más o menos desconocida, lo que contribuiría a obtener distintos resultados para el experimento consistente en medir la velocidad de caída de un móvil. El fenómeno sería, desde esta nueva óptica, un fenómeno aleatorio. Por contra, un ejemplo típico de experimento aleatorio es el lanzamiento de un dado. Como el movimiento de un dado se rige asimismo por las leyes de la Física, se comprende que si pudiéramos controlar perfectamente la posición inicial, así como las características de la fuerza de lanzamiento, deberíamos ser capaces de predecir el resultado, y por tanto se convertiría en un fenómeno determinista. Lo que varía en ambas ocasiones es el enfoque con que tratamos el problema. Lo que es cierto es que uno de estos enfoques, aleatorio o determinista, puede ser más adecuado para tratar el problema concreto que estemos intentando resolver.

Al conjunto de posibles resultados de un experimento aleatorio se le denomina *espacio muestral*, y lo denotaremos por Ω.

Llamaremos *suceso* a cualquier proposición formulada en relación con el experimento, y que en función del resultado del mismo pueda afirmarse categóricamente si ha ocurrido o no. Por ejemplo, si realizamos el experimento aleatorio “lanzar un dado”, el espacio muestral será $\Omega = \{1,2,3,4,5,6\}$. Ejemplos de sucesos asociados a este experimento pueden ser los siguientes: “sacar un 6”, “sacar un número par”, “sacar menos de tres”, ... Es directo comprobar que cada suceso puede identificarse con un subconjunto de Ω, el de los elementos que satisfagan la proposición. En el ejemplo anterior, los subconjuntos correspondientes a estos sucesos serían, respectivamente, $\{6\}$, $\{2,4,6\}$, $\{1,2\}$. *Los sucesos son subconjuntos del espacio muestral.* Los sucesos consistentes en sólo un elemento del espacio muestral suelen llamarse *sucesos elementales.* Se puede considerar el espacio muestral completo Ω como suceso, ya que es un subconjunto de sí mismo. Llamamos a este suceso *suceso seguro.*

3.2.2 Relaciones y Operaciones con sucesos

A veces conviene describir un suceso en relación con otros. Como los sucesos son subconjuntos pueden realizarse con ellos las operaciones y relaciones definidas entre conjuntos.

Se dice que el suceso A *implica el suceso* B $(A \subseteq B)$ si siempre que ocurre A en el resultado del experimento ocurre también B. Por ejemplo si A consiste en sacar un tres al tirar un dado, $A = \{3\}$, y B consiste en sacar una puntuación que sea múltiplo de 3, $B = \{3,\ 6\}$, siempre que se cumpla A, haya salido un tres al tirar el dado, también saldrá una puntuación múltiplo de 3, y por tanto se cumplirá también el suceso B.

Se dice que dos sucesos *A y B son iguales*, $A = B$, si el suceso A implica el suceso B y el suceso B implica el suceso A.

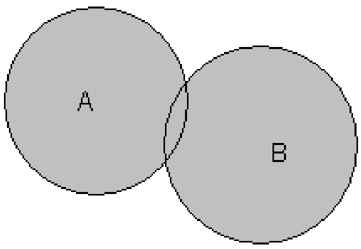
La unión de dos sucesos A y B es un suceso $C = A \cup B$ que se verifica si se verifica al menos uno de ellos (A o B o ambos)

La intersección de dos sucesos A y B es un suceso $C = A \cap B$ que se verifica si se verifican ambos sucesos (A y B). *Dos sucesos son incompatibles* si no pueden verificarse simultáneamente. En este caso su intersección es un suceso que no contiene ningún elemento del espacio muestral, ya que no se verifica nunca. Llamamos *suceso imposible* a este suceso y lo representamos con el signo $\emptyset$.

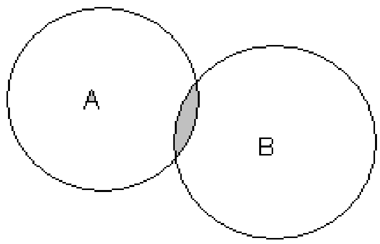
El suceso A' se llama *contrario*[1] *de* A si ocurre *siempre* que no ocurra A. En el caso de tirar un dado, el suceso contrario de $\{3,\ 6\}$ es $\{1,\ 2,\ 4\ ,5\}$. Se cumple que $A \cup A' = \Omega$ y que $A \cap A' = \emptyset$.

Se define *diferencia entre dos sucesos* $A - B = A \cap B'$, al suceso que se verifica si se cumple A pero no B.

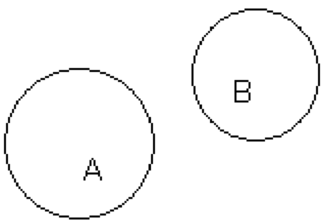
Es bastante útil representar los sucesos, al igual que los conjuntos por medio de diagramas de Venn. En las gráficas siguientes están representados los diagramas correspondientes a las operaciones y relaciones que acabamos de definir.



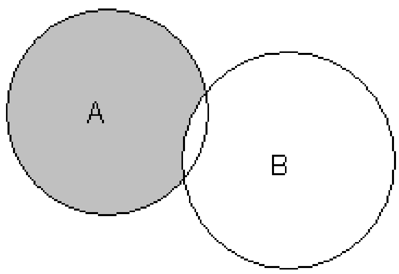
Unión de dos Sucesos



Intersección de dos Sucesos

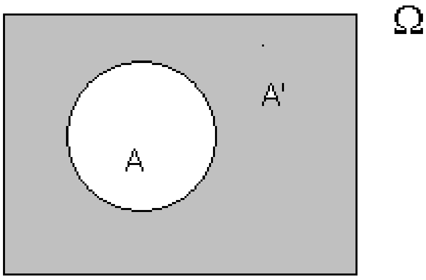


Sucesos incompatibles



Diferencia entre Sucesos

[1] También suele emplearse para el suceso contrario la notación A^c y $\overline{A}$



Sucesos Contrarios

3.2.3 Propiedades de las operaciones entre sucesos

Es inmediato comprobar las siguientes propiedades que son idénticas a las del Algebra de Boole de Conjuntos.

1. Conmutativa: $A \cup B = B \cup A$; $A \cap B = B \cap A$
2. Asociativa: $(A \cup B) \cup C = A \cup (B \cup C)$; $(A \cap B) \cap C = A \cap (B \cap C)$
3. Distributiva $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$; $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$
4. Doble complementación ó Involutiva: $(A')' = A$
5. Idempotente: $A \cup A = A$; $A \cap A = A$
6. Elemento neutro: $A \cup \emptyset = A$; $A \cap \Omega = A$
7. De absorción: $A \cup (A \cap B) = A$; $A \cap (A \cup B) = A$
8. Leyes de De Morgan: $(A \cup B)' = A' \cap B'$; $(A \cap B)' = A' \cup B'$

3.2.4 Definición frecuentista de probabilidad

Supongamos que repetimos n veces un experimento aleatorio, y que el suceso A ha ocurrido en n_A ocasiones. Entonces, la frecuencia relativa de ocurrencia del suceso A, tal como se ha definido en el tema anterior será:

$$fr(A) = \frac{n_A}{n}$$

Son evidentes las siguientes propiedades de la frecuencia:

1. La frecuencia relativa de un suceso está comprendida entre 0 y 1, puesto que se cumple que $0 \leq n_A \leq n$.
2. La frecuencia de la unión de dos sucesos incompatibles es la suma de las frecuencias de cada uno de estos sucesos.

   Si consideramos dos sucesos incompatibles A y B, es decir que $A \cap B = \emptyset$, se cumple que:

$$card(A \cup B) = car(A) + card(B) = n_A + n_B$$

Por tanto

$$fr(A \cup B) = \frac{n_A + n_B}{n} = \frac{n_A}{n} + \frac{n_B}{n} = fr(A) + fr(B)$$

La idea de probabilidad no es más que una generalización de dicha medida de frecuencia relativa, que podríamos considerar como el límite al que tiende $fr(A)$ conforme se vaya aumentando el valor de n. Es un hecho experimental que el valor de la frecuencia relativa se va estabilizando según aumentamos el número de pruebas. La existencia de dicho límite, ya era claramente intuida desde la antigüedad en los juegos de azar, sin embargo, este límite se basa en resultados experimentales o empíricos, y por tanto no da un procedimiento de cálculo para la probabilidad, sino sólo un valor aproximado de ésta. Por ejemplo, esta definición nos sirve para asignar un valor a la probabilidad de que una chincheta caiga boca arriba. No obstante, necesitamos una definición que posea un mayor rigor matemático.

3.2.5 Definición clásica de probabilidad

La teoría de la probabilidad, como hemos comentado en la sección anterior, tiene su origen en los juegos de azar y de ahí procede también la definición clásica de probabilidad, que está inspirada en la idea de frecuencia relativa de un suceso. Si consideramos las ocurrencias de cara o cruz en el lanzamiento de una moneda, podríamos hacer la experiencia de tirar la moneda una gran cantidad de veces y decidir que la frecuencia relativa obtenida es el valor aproximado de la probabilidad de cada cara, o también, podemos razonar que puesto que sólo hay dos resultados posibles y parece que la moneda es más o menos simétrica, la frecuencia de cara o cruz debe ser más o menos la misma y por tanto su valor debe ser aproximadamente $\frac{1}{2}$. Este razonamiento es el que se hace en la definición clásica de probabilidad. Que este razonamiento sea correcto sólo puede comprobarse con la experiencia, haciendo una gran cantidad de lanzamientos de una moneda bien construida. Aunque pueda parecer extraño, así se ha hecho históricamente.

Una probabilidad tiene que ser definida de forma que a cada suceso le corresponda un número, y que se cumplan las propiedades de la frecuencia relativa. Estos requisitos se cumplen en la siguiente definición que se conoce como Regla de Laplace.

Sea Ω un espacio muestral con un número finito de elementos. Si puede admitirse que todos los elementos del espacio muestral tienen la misma probabilidad entonces se define como probabilidad de un suceso S:

$$P(S) = \frac{\text{número de elementos del conjunto } S}{\text{número de elementos del espacio muestral } \Omega} = \frac{Card(S)}{Card(\Omega)}$$

Para aclarar el significado de esta fórmula la aplicamos en un par de ejemplos sencillos:

a) Calcular la probabilidad de sacar un número múltiplo de tres en el lanzamiento de un dado: El suceso cuya probabilidad queremos calcular es $S = \{3,\ 6\}$, y el espacio muestral está formado por todos los resultados posibles $\Omega = \{1, 2, 3, 4, 5, 6\}$. Si admitimos que todos estos sucesos tienen la misma probabilidad, la regla de Laplace nos da un valor para esta probabilidad:

$$P(S) = \frac{\text{número de elementos del conjunto } S}{\text{número de elementos del espacio muestral } \Omega} = \frac{2}{6} = \frac{1}{3}$$

b) Calcular la probabilidad de sacar un número menor que 4 en una baraja española (de 40 cartas): En este caso el espacio muestral está formado por todos los resultados posibles, las cuarenta cartas de la baraja. El suceso contiene 12 cartas, 3 por cada uno de los cuatro palos, luego $P(S) = \frac{12}{40} = \frac{3}{10}$

Es obvio que la probabilidad definida cumple las siguientes propiedades de la frecuencia relativa:

1. $0 \leq P(S)$
2. $P(\Omega) = 1$
3. Si A_1, A_2 son sucesos, cumpliendo que $A_1 \cap A_2 = \emptyset$, entonces se cumple que

$$P(A_1 \cup A_2) = P(A_1) + P(A_2)$$

3.2.6 Propiedades de la probabilidad.

De las anteriores propiedades pueden deducirse las que siguen

1. $P(\emptyset) = 0$.
2. Si A y B son dos sucesos cumpliendo $A \subset B$, entonces $P(B - A) = P(B) - P(A)$.
3. Si A y B son dos sucesos cumpliendo $A \subset B$, entonces $P(A) \leq P(B)$.
4. $P(A) \leq 1 \qquad \forall A$
5. $P(A') = 1 - P(A)$.
6. Si $A_1, A_2, \ldots, A_n$ son sucesos, cumpliendo que $A_i \cap A_j = \emptyset \qquad \forall i \neq j$, entonces se cumple que $P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$.
7. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
8. $P(\displaystyle\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i)$.

Demostraciones de las propiedades enunciadas:

1. $\emptyset = \emptyset \cup \emptyset \qquad$ aplicando el axioma $P(\emptyset) = P(\emptyset) + P(\emptyset) \Longrightarrow P(\emptyset) = 0$.
2. $B = (B - A) \cup A \Longrightarrow P(B) = P(B - A) + P(A)$.
3. Se deduce de la propiedad 2 y del axioma 1 ya que se cumple que $P(B - A) \geq 0$.
4. $A \subset \Omega$, por lo tanto, aplicando por la propiedad 3 se tiene que $P(A) \leq P(\Omega) = 1$.
5. Como $\Omega = A \cup A'$ y $A \cap A' = \emptyset$ puede aplicarse el axioma 3 que nos dará $P(A) + P(A') = P(\Omega) = 1$
6. Por inducción usando el axioma 3.
7. Se parte de la igualdad $A \cup B = A \cup (B - (A \cap B))$ y se aplica la propiedad 2.
8. Se demuestra por inducción sobre n partiendo de la propiedad 7.

La definición clásica de probabilidad tiene algunos problemas. En primer lugar sólo es válida para espacios muestrales finitos y además no queda claro cuando deberíamos entender que "los sucesos elementales son equiprobables", ya que lo que se está definiendo forma parte de la definición. Por ejemplo, en el caso del nacimiento de un niño, varón o mujer, aunque también hay únicamente dos resultados del experimento como en el caso de la moneda, no hay motivos evidentes para aplicar un principio de simetría, por lo que no podemos usar la regla de Laplace y sería más adecuado usar una definición frecuentista, usando el valor de la frecuencia relativa. En la actualidad se atribuye una probabilidad de 0.51 para el nacimiento de varón y de 0.49 para mujer.

Tampoco la fórmula de Laplace es apropiada para el siguiente caso: Supongamos que queremos calcular la probabilidad de sacar dos caras tirando dos monedas. Podíamos decir que el espacio muestral esta formado por tres elementos consistente en los tres resultados posibles: sacar dos caras: CC, sacar dos cruces: XX o sacar una cara y una cruz: CX. Como sólo uno de éstos, CC, forma parte del suceso cuya probabilidad queremos conocer, el número de elementos de S es 1. Si suponemos que los tres elementos del espacio muestral son igualmente probables la probabilidad del suceso CC sería $\frac{1}{3}$. Este resultado, al que se llega usando la fórmula de Laplace, aunque formalmente correcto, no está de acuerdo con la experiencia, que no da la misma frecuencia relativa a los tres sucesos. El motivo es que la composición CX, puede obtenerse de dos formas variando la moneda en la que aparece la cara y la cruz. La regla de Laplace nos daría un resultado de acuerdo con la experiencia si suponemos que el espacio muestral está formado por los resultados {CC, CX, XC, XX}, con lo que obtendríamos para el suceso sacar dos caras una probabilidad $\frac{1}{4}$, y que coincidiría aproximadamente con el valor para la frecuencia relativa del suceso, que obtendríamos experimentalmente realizando una gran cantidad de tiradas con las dos monedas.

A pesar de las limitaciones de la definición clásica, en la práctica se usa frecuentemente, teniendo cuidado de adaptar al caso real las consideraciones de equiprobabilidad que exige la formulación clásica. No obstante, muchas cuestiones y problemas no pueden adaptarse el modelo de definición de probabilidad de Laplace, ni al modelo de definición experimental como frecuencia relativa. Por ejemplo, ninguna de estas definiciones son muy adaptables para resolver preguntas como las siguiente: ¿cuál es la probabilidad de que estalle una tercera guerra mundial?, ¿cuál es la probabilidad de que un automóvil de un nuevo modelo dure más de 8 años?, ¿cuál es la probabilidad de que se encuentre una nueva galaxia en el presente siglo?....

3.2.7 Definición axiomática de probabilidad.

La definición axiomática de probabilidad, debida a Kolmogorov, es más general y se adapta a un conjunto más amplio de situaciones. Para poder dar dicha definición necesitamos disponer de un espacio muestral Ω sobre el que esté definida un σ-álgebra de sucesos. es decir imponemos una cierta estructura al conjunto de sucesos

Un *σ-álgebra de sucesos* es el conjunto de todos los sucesos de interés asociados a un experimento aleatorio. Denotamos por $\mathcal{A}$ al *σ-álgebra de sucesos*. Se verifica que

$$\mathcal{A} \subset \mathcal{P}(\Omega)$$

donde $\mathcal{P}(\Omega)$ representa el conjunto de las partes de Ω, esto es, el conjunto de todos los subconjuntos de Ω.

Al conjunto $\mathcal{A}$ le impondremos que cumpla las siguientes propiedades:

a) $\Omega \in \mathcal{A}$

b) las operaciones unión, intersección, y complementación de elementos de $\mathcal{A}$ realizadas un número finito (o infinito numerable) de veces, son internas en $\mathcal{A}$.

No siempre tiene que ocurrir que $\mathcal{A} = \mathcal{P}(\Omega)$, es decir que no es necesario que todos los subconjuntos de Ω sean sucesos. Por ejemplo, si lanzando un dado y estamos apostando a par o impar solamente, entonces

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

es el conjunto de interés. Puede comprobarse que este conjunto cumple las propiedades de un σ-álgebra de sucesos. Es conveniente observar que $\emptyset \in \mathcal{A}$, ya que es el complementario de Ω. Es decir, el conjunto vacío siempre será un suceso.

La definición axiomática de probabilidad, dada por Kolmogorov, sería la siguiente:

Sea Ω un espacio muestral y A un σ-álgebra de sucesos definida sobre Ω. Una probabilidad es una aplicación P cumpliendo:

1. $P : \mathcal{A} \longrightarrow [0,1]$
2. $P(\Omega) = 1$
3. Si $A_1, A_2, \ldots, A_n$ son sucesos, cumpliendo que $A_i \cap A_j = \emptyset \quad \forall i \neq j$, entonces se cumple que

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$$

Si la unión de sucesos es infinita numerable también se cumple:

$$P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i), \ i \in \aleph$$

Es de notar que esta definición contiene a la de Laplace como un caso particular: En efecto si Ω es un espacio muestral finito formado por n sucesos elementales $s_1, s_2, \ldots\, s_n$, que suponemos equiprobables, se tiene que la probabilidad de cualquier suceso sería $\frac{1}{n}$, ya que usando los axiomas 2 y 3 obtenemos

$$P(\Omega) = P\{s_1, s_2, ...s_n\} = P\{s_1\} + P\{s_2\} + ... + P\{s_n\} = 1$$

y como los sucesos tienen por hipótesis la misma probabilidad, se deduce que la probabilidad de cualquiera de estos sucesos elementales es $\frac{1}{n}$. La probabilidad de un suceso compuesto se obtendría sumando la probabilidad de cada uno de los sucesos elementales que lo formen. Por ejemplo:

$$P\{s_i, s_j, s_k\} = P\{s_i\} + P\{s_j\} + P\{s_k\} = \frac{1}{n} + \frac{1}{n} + \frac{1}{n} = \frac{3}{n} =$$
$$= \frac{\text{número de elementos del suceso}}{\text{número de elementos del espacio muestral}}$$

que sigue la regla de Laplace.

Es de notar que la definición axiomática no nos suministra un procedimiento de cálculo de la probabilidad de un experimento concreto, limitándose a indicar qué propiedades ha de cumplir una función P para que sea considerada una probabilidad en sentido matemático. También aquí se cumplen las propiedades enumeradas en la página 7, puesto que los axiomas son también las propiedades de la frecuencia relativa. Estas propiedades nos van a permitir hallar la probabilidad de cualquier suceso a partir de la probabilidad de unos cuantos sucesos. ¿Cómo definimos la probabilidad de estos cuantos sucesos? Con tal de que se cumplan los axiomas de la probabilidad, tenemos libertad en este sentido. La conveniencia de una forma u otra de dar valores concretos a la probabilidad de estos pocos sucesos viene determinado por el caso particular que estemos tratando y los valores numéricos de la probabilidad se deben asignar buscando un buen acuerdo entre el modelo teórico y el real o experimental.

Para asignar valores a las probabilidades de estos cuantos sucesos se emplean los siguientes enfoques: *frecuentista, clásico y subjetivo.*

El enfoque frecuentista esta basado en la interpretación de la probabilidad como límite de la frecuencia relativa de un suceso cuando el número de pruebas tiende a infinito. El enfoque clásico consiste en la aplicación de regla de Laplace, siempre que sea adaptable. El enfoque subjetivo asigna valores a la probabilidad de algunos sucesos en función de la opinión de determinadas personas, que se supone que son expertos en el tema. Esta última forma de asignar valores para la probabilidad se usa, por ejemplo, en las apuestas o en las quinielas.

Una vez asignados valores de probabilidad a algunos sucesos, a menudo los sucesos elementales, por medio de cualquiera de los enfoques indicados, debemos calcular las probabilidades de los otros sucesos de interés. Ahora podemos calcular indirectamente el valor para las probabilidades de sucesos más complejos, para los que quizá no es fácil asignar directamente valores de probabilidad. Se puede recurrir al uso de propiedades, como por ejemplo las indicadas en la página 7, que como hemos indicado también se cumplen si usamos la definición de Kolmogorov, u otras propiedades, como las que damos en el resto del capítulo.

3.3 Recursos de interés para el cálculo de probabilidades de sucesos

Si queremos calcular la probabilidad de un suceso usando la fórmula de Laplace:

$$\frac{\text{número de elementos del suceso}}{\text{número de elementos del espacio muestral}}$$

debemos poder contar el número de resultados posibles, *y equiprobables,* del experimento y también el número de elementos del suceso. Estos valores, a veces, no son fáciles de calcular si no se dispone de algunas reglas de conteo. A continuación damos algunas reglas que pueden ser útiles para este propósito.

3.3.1 Regla de multiplicación

Supongamos que el experimento aleatorio consiste en tirar una moneda y después un dado. La variable aleatoria que consideramos está formada por el resultado de la moneda, cara o cruz, y la puntuación del dado. Los posibles resultados del experimento podemos indicarla por medio de una tabla de doble entrada.

	1	2	3	4	5	6
C	C1	C2	C3	C4	C5	C6
X	X1	X2	X3	X4	X5	X6

Está claro que el número de elementos del espacio muestral se obtiene multiplicando el número de elementos obtenidos en la primera prueba (2 posiciones de la moneda) por el número de resultados obtenidos en la segunda prueba (6 posibles puntuaciones del dado). Así que el espacio muestral asociado a este experimento compuesto estaría formado por $2 \times 6 = 12$ elementos.

Si el experimento aleatorio consiste en tirar una moneda, tirar un dado y después sacar una carta de una baraja, el espacio muestral cuyos elementos constarían de tres resultados (por ejemplo *Cara, tres y sota de espadas*) estaría formado por $2 \times 6 \times 40 = 480$ elementos.

Ejemplo:

Calcular la probabilidad de sacar una puntuación par en el dado y una carta de oros en el experimento anterior.

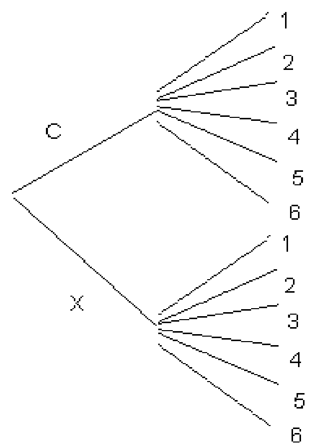
Ya hemos calculado el número de elementos del espacio muestral: 480 resultados posibles. Por la misma regla podemos calcular el número de elementos del suceso: *2 figuras de la moneda* $\times$ *3 puntuaciones pares para el dado* $\times$ *10 cartas de oros. Por lo tanto la probabilidad pedida sería:*

$$\text{Probabilidad} = \frac{\text{número de elementos del suceso}}{\text{número de elementos del espacio muestral}} = \frac{2 \times 3 \times 10}{480} = \frac{1}{8}.$$

3.3.2 Diagramas de árbol

Una representación en forma de diagrama de árbol es también útil para poner de manifiesto la regla de multiplicación:

Para el ejemplo de la moneda y el dado tal diagrama tomaría la forma de la figura



Recorriendo cada rama del árbol desde la raíz, obtenemos los distintos resultados del experimento.

3.3.3 Combinatoria

A continuación incluimos algunas nociones de combinatoria que pueden resultar útiles para acometer algunos problemas de probabilidad. La combinatoria estudia las posibles agrupaciones, cumpliendo ciertas condiciones, que pueden formarse con un número finito de elementos. A veces, nos sirve de ayuda para evaluar la probabilidad siguiendo la regla de Laplace ya que, en ocasiones, permite evaluar el número total de elementos equiprobables pertenecientes al espacio muestral y/o cuántos de éstos cumplen la propiedad característica del suceso cuya probabilidad queremos calcular.

Variaciones

Las variaciones de n elementos tomados de m en m son los distintos grupos con m elementos que se pueden formar eligiendo m elementos de los n totales.

a) Una variación es distinta de otra si se distingue en algún elemento, o si se distingue en el orden de los elementos entre sí.

b) Cada elemento sólo puede aparecer como máximo una vez dentro de cada variación.

Por lo tanto, las variaciones son todas las agrupaciones ordenadas de m objetos seleccionados entre los n elementos de un conjunto dado.

Por ejemplo, si se toma una baraja con cuarenta cartas, cada una de las distintas formas en que se pueden repartir 2 cartas, teniendo en cuenta cuál ha sido la primera y cuál la segunda, es una variación de las 40 cartas tomadas en grupos de dos. El número de pares de cartas, puede obtenerse por medio del siguiente razonamiento: La primera carta puede darse de 40 maneras diferentes, y la segunda de 39, ya que queda una carta menos. Por lo tanto el número de pares de cartas diferentes serán 40×39. En general para obtener el número de variaciones diferentes de n elementos tomados de m en m que se denota por Vn, m, se usa la fórmula siguiente:

$$V_{n,m} = n \times (n-1) \times (n-2) \times ... \times [n-(m-1)]$$

Por ejemplo: $V_{5,3} = 5 \times 4 \times 3 = 60$.

Permutaciones

Las permutaciones de n elementos son las variaciones de estos n elementos tomados de n en n. En este caso una permutación sólo puede distinguirse de otra en el orden en que están ordenados sus elementos, ya que todas ellas contendrán los n elementos del conjunto completo. Por tanto, las permutaciones son las distintas formas en que se pueden ordenar los n elementos de un conjunto. Supongamos ahora que tenemos cuatro cartas y que las señalamos con las letras a,b,c,d. ¿De cuántas formas pueden ordenarse? La primera carta puede colocarse de 4 formas, la segunda sólo de tres, porque una de ellas ya se ha tomado en la primera extracción, la tercera de dos formas. La última sólo puede ser la que nos quede. Por lo tanto hay $4 \times 3 \times 2 \times 1 = 24$ ordenaciones diferentes para estas cuatro cartas. A continuación detallamos cada una de estas 24 permutaciones. Comenzamos escribiendo las variaciones con un único elemento:

a, b, c, d

Ahora con dos elementos:

ab, ac, ad,
ba, bc, bd
ca, cb, cd
da, db, dc

Con tres elementos:

abc, abd	acb, acd	adb,a dc
bac, bad	bca, bcd	bda, bdc
cab, cad	cba, cbd	cda, cdb
dab, dac	dba, dbc	dca, dcb

Para formar las permutaciones añadimos ahora el elemento que falta en cada una de las anteriores

abcd, abdc	acbd, acdb	adbc, adcb
bacd, badc	bcad, bcda	bdac, bdca
cabd, cadb	cbad, cbda	cdab, cdba
dabc, dacb	dbac, dbca	dcab, dcba

El número de permutaciones de n elementos se denota por P_n y se calcula con la expresión

$$P_n = V_{n,n} = n\times(n-1)\times(n-2)\times...\times[n-(n-1)] = n\times(n-1)\times(n-2)\times...\times3\times2\times1 = n!$$

Esta última expresión, $n!$, se conoce con el nombre de *factorial de n,* y como se ve representa el producto de todos los enteros positivos de 1 a n. Por conveniencia de cálculo se define $0! = 1$.

Se puede comprobar que se cumple

$$V_{n,m} = n\times(n-1)\times n\,(-2)\times ...\times(n-(m-1)) = \frac{P_n}{P_{n-m}} = \frac{n!}{(n-m)!}$$

Combinaciones

Si consideramos como iguales las variaciones que tengan exactamente los mismos elementos y como distintas aquellas cuyos elementos estén colocados en cierto orden, los grupos distintos forman las combinaciones. En concreto:

Las combinaciones de n elementos tomados de m en m, $C_{n,m}$, son los distintos grupos con m elementos que se pueden formar eligiendo m elementos de los n totales. Una combinación es distinta de otra sí:

a) Se distingue en algún elemento.

b) Cada elemento sólo puede aparecer como máximo una vez dentro de cada combinación.

Siguiendo con el ejemplo del reparto de dos cartas de cuarenta. En el caso de las variaciones hemos tenido en cuenta el orden en que nos han dado las cartas, pero en algunos juegos no importa el orden en que se reciben éstas cartas. Si suponemos formadas las posibles variaciones de 40 cartas tomadas de dos en dos, observamos que la variación AB, y la variación BA forman ahora una única combinación, ya que no vamos a tener en cuenta el orden de reparto.

En el ejemplo detallado en el párrafo anterior, en el que se consideraban las variaciones de las 4 cartas tomados de 2 en 2, podemos observar que cada pareja de dos letras aparece repetida en dos variaciones. Por lo tanto

$C_{4,2} = \frac{V_{4,2}}{2} = \frac{V_{4,2}}{P_2} = \frac{4\times3}{2\times1} = 6$

Estas seis combinaciones posibles son:

ab, ac, ad,
bc, bd
cd

En general, el número de combinaciones de n elementos tomados de m en m se escribe $C_{n,m}$, y su valor está dado por la siguiente fórmula:

$$C_{n,m} = \frac{V_{n.m}}{P_m} = \frac{V_{n.m}P_{n-m}}{P_m P_{n-m}} = \frac{n!}{(n-m)!m!} = \binom{n}{m}$$

Los valores $C_{n,m} = \binom{n}{m}$ suelen conocerse con el nombre de *números combinatorios.*

Variaciones con repetición

Las variaciones de n elementos tomados de m en m son los distintos grupos con m elementos que se pueden formar eligiendo m elementos de los n totales. Se debe cumplir que:

a) Cada elemento puede repetirse hasta m veces dentro de cada variación con repetición

b) Una variación con repetición se distingue de otra por que contenga elementos distintos, por el orden en que están colocados estos elementos o por el número de veces que esté repetido cada elemento.

En este caso m puede ser mayor que n.

Por ejemplo, si consideramos las distintas variaciones con repetición que pueden formarse a partir de los diez dígitos, tomados en grupos de tres, obtenemos los números desde 000, hasta el 999, es decir $V'_{10,3}$ $=10^3$ $=1000$ variaciones con repetición.

En general representamos el número de variaciones con repetición como $V'_{n,m} = n^m$

Permutaciones con repetición

Si queremos calcular las distintas ordenaciones de un conjunto de n elementos, algunos de los cuales son indistinguibles entre sí como ocurre con las 8 letras de la palabra PAPANATA en la que aparece, 4 veces la letra A, dos veces la P, y una vez cada una de las letras N y T, aplicamos la siguiente expresión:

$$P'_{8;\,4,2,1,1} = \frac{8!}{4!2!1!1!}$$

Para escribir esta expresión se parte del número total de permutaciones que se obtendrían si, de forma artificial, suponemos distintos los elementos iguales de entre los n elementos de partida. En el caso del ejemplo, podemos distinguir las letras iguales con subíndices: $P_1A_1P_2A_2N_1A_3T_1A_4$. Se divide el número total de permutaciones por el número de veces que cada una de ellas aparecería repetida si igualáramos las letras, suprimiendo los subíndices.

En general, el número de variaciones con repetición de n elementos entre los que hay m grupos de elementos diferentes, cada uno de ellos con $n_1,\ n_2, ..., n_m$ elementos iguales siendo $n = n_1 + n_2 + ... + n_m$, se obtienen con la expresión:

$$P'_{n;\,n_1,n_2,..,n_m} = \frac{n!}{n_1!n_2!...n_m!}$$

Combinaciones con repetición

Si tenemos un conjunto con n elementos distintos y admitimos que cada combinación puede tener elementos repetidos, obtenemos las combinaciones con repetición. En concreto:

Las combinaciones con repetición de n elementos tomados de m en m, $C'_{n,m}$, son los distintos grupos con m elementos que se pueden formar eligiendo m elementos de los n totales.

a) Una combinación es distinta de otra si se distingue en algún elemento.

b) Cada elemento puede repetirse hasta un total de m veces dentro de cada combinación.

En este caso m puede ser mayor que n.

La expresión de cálculo del número de combinaciones con repetición es:

$$C'_{n,m} = C_{n+m-1,m} = \binom{n+m-1}{m} = \frac{(n+m-1)!}{m!(n-1)!}$$

Las combinaciones con repetición de las dos letras {a, b} en grupos con cuatro elementos son:

$$C'_{2,4} = \frac{(2+4-1)!}{4!(2-1)!} = 5$$

Escribimos estas 5 combinaciones con repetición:

a a a a   b b b b   a a a b   a a b b   a b b b

Como puede verse en el ejemplo no se tiene en cuenta los cambios de orden para formar diferentes combinaciones. Por ejemplo la combinación con repetición “aaab” es la misma que la combinación “aaba”.

3.3.4 De lo particular a lo general

Cuando tenemos un problema bastante complejo, un procedimiento bastante usado es comenzar resolviendo casos particulares de este problema, que debido a su simplicidad, sean más fáciles de resolver. Esperamos que su resolución nos arroje alguna luz que nos permita vislumbrar la solución del caso general. También este método puede ser útil para calcular probabilidades de sucesos. Lo ilustramos con el siguiente ejemplo.

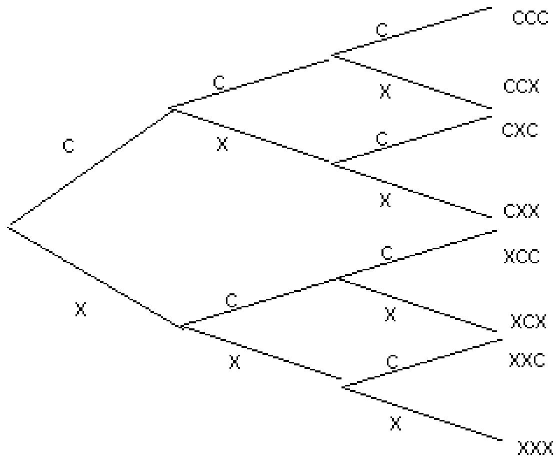
Ejemplo: Se realiza una tirada de n monedas ($n \geq 3$). Calcular la probabilidad de obtener exactamente tres caras.

Aquí se nos pide una expresión general en función de n.

Comenzamos por el caso particular más sencillo: $n = 3$. Usamos la fórmula de Laplace:

$$\text{Probabilidad} = \frac{\text{número de elementos del suceso}}{\text{número de elementos del espacio muestral}}$$

Calculamos el denominador. Para ello necesitamos contar todos los resultados posibles en las tres monedas. Usamos un diagrama de árbol



Usando la regla de multiplicación, el número de elementos del espacio muestral es $2 \times 2 \times 2 = 2^3 = 8$. Estos 8 resultados están especificados en la columna derecha del diagrama de árbol.

¿Cuántos sucesos de entre ellos cumplen que tienen tres caras? Solamente el resultado CCC. Por tanto la probabilidad de obtener exactamente tres caras tirando tres monedas es $\frac{1}{8}$.

Nos lo ponemos ahora un poco más difícil. Consideraremos que tiramos cuatro monedas.

El número de elementos del espacio muestral lo calculamos, análogamente al caso anterior. Éste número es ahora $2 \times 2 \times 2 \times 2 = 2^4 = 16$.

No es demasiado difícil describir los casos favorables al suceso: CCCX, CCXC, CXCC, XCCC. Observamos que son todos las permutaciones con repetición de 3 caras y una cruz. Usando la fórmula de las permutaciones con repetición de 4 elementos donde se repiten 3 de ellos comprobamos que efectivamente es así:

$$P'_{4;\,3,1} = \frac{4!}{3!1!} = 4$$

Por tanto la probabilidad de obtener tres caras en una tirada de cuatro monedas es $\frac{4}{16} = \frac{1}{4}$.

Resolvemos ahora el problema primitivo considerando n monedas. El número de elementos del espacio muestral sería 2^n, y el número de elementos con exactamente tres caras es $P'_{n;\,3,n-3}$

La probabilidad de obtener tres caras en una tirada de $n \geq 3$ monedas es:

$$\frac{P'_{n;\,3,n-3}}{2^n} = \frac{\frac{n!}{3!(n-3)!}}{2^n} = \frac{1}{6}n\,(n-1)\,(n-2)\,2^{-n}$$

3.3.5 La probabilidad geométrica

En algunas ocasiones la fórmula de Laplace no es adecuada. Consideremos el caso siguiente

Ejemplo: En un segmento de longitud 1 marcamos dos puntos interiores. Calcular la probabilidad de que los tres segmentos resultantes puedan formar un triángulo.

No parece que en este caso tenga sentido contar el número de elementos del espacio muestral, que estaría formado por todas las ternas de segmentos cuyas longitudes sumen uno. Evidentemente este conjunto es de cardinal infinito, así que la regla de Laplace no es aplicable en esta ocasión. Para calcular esta probabilidad se suele identificar el número de elementos del espacio muestral con el área del conjunto que lo represente. Es lo que haremos aquí para dar una solución a este problema:

Sean $x,\ y,\ z$ las longitudes de los tres segmentos. Como hemos dicho, el espacio muestral está formado por las ternas de longitudes que cumplen $x+y+z=1$. Las ternas que forman el suceso cuya probabilidad queremos calcular son los que, además, cumplan las tres condiciones siguientes:

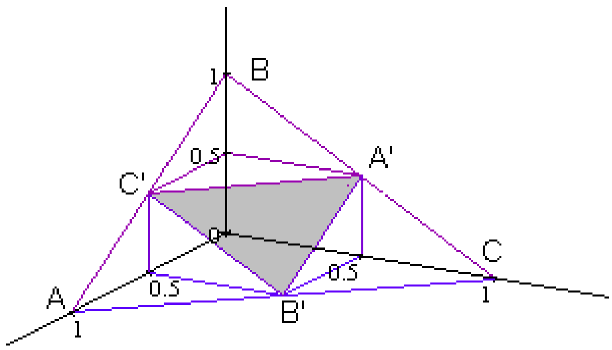
$$\begin{array}{l} x < y+z \\ y < x+z \\ z < x+y \end{array} \tag{3.1}$$

ya que cada lado de un triángulo ha de ser menor que la suma de los otros dos.

Teniendo en cuenta que $x+y+z=1$ y las expresiones 3.1, deducimos

$$\begin{array}{lcl} x < y+z = 1-x & \Longrightarrow & x < 0.5 \\ y < x+z = 1-y & \Longrightarrow & y < 0.5 \\ z < x+y = 1-z & \Longrightarrow & z < 0.5 \end{array}$$

En la siguiente figura se ha realizado una representación gráfica tridimensional de ambos conjuntos. Los puntos del triángulo ABC forman el espacio muestral (las longitudes de los segmentos son magnitudes positivas, por eso están representadas en el octante con las tres coordenadas positivas del plano $x+y+z=1$). Los puntos interiores del triángulo $A'B'C'$ son las ternas $x,\ y,\ z$ que cumplen el suceso. Verifican las condiciones necesarias: $x<0.5$, $y<0.5$, $z<0.5$, $x+y+z<1$.



Calculamos la probabilidad como cociente entre las dos áreas

$$\text{Probabilidad de que los segmentos de longitudes } x,\ y,\ z \text{ formen un triángulo} =$$
$$= \frac{\text{Área del triángulo } A'B'C'}{\text{Área del triángulo } ABC} = \frac{1}{4}$$

Esta enfoque de la probabilidad, basado en las medidas de la representación gráfica de los sucesos, se conoce con el nombre de *probabilidad geométrica*.

3.4 Probabilidad condicionada.

Con este concepto tratamos de registrar el cambio que experimenta la probabilidad de un suceso si aumenta la información de que disponemos sobre el resultado del experimento. Supongamos que tenemos el billete 347 de una lotería de 1000 números. Nuestra probabilidad de ganar el premio es, siguiendo la regla de Laplace $\frac{1}{1000}$. Pero si aparece un amigo nuestro y nos dice. “No me acuerdo en que número ha caído el premio, pero me acuerdo que acababa en 7”. ¿Cuál es ahora nuestra probabilidad de ganar el premio? Ya sólo hay 100 casos posibles, el de los números de tres cifras que terminen en 7, y nosotros tenemos uno de ellos, nuestra probabilidad es ahora $\frac{1}{100}$.

Supongamos que queremos calcular la probabilidad del suceso A sabiendo que se ha presentado el suceso B. Esto es lo que se denomina *probabilidad de A condicionada a B* y se denota como $P(A/B)$. La siguiente definición nos relaciona la probabilidad condicionada, también llamada “a posteriori” en función de las probabilidades “a priori”, definidas al principio cuando se carecía de información adicional.

Por definición $P(A/B) = \dfrac{P(A \cap B)}{P(B)}$

Así, en el ejemplo anterior, si A es el suceso que consiste en “que salga el número 347” y B el suceso que consiste en “que salga un número terminado en 7”

$$P(A/B) = P(\text{“que salga el número 347”/si “ha salido un número terminado en 7”}) =$$
$$= \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{1000}}{\frac{100}{1000}} = \frac{1}{100}$$

Del mismo modo, tendríamos que $P(B/A) = \dfrac{P(A \cap B)}{P(A)}$ de donde deducimos la fórmula para la probabilidad de la intersección de dos sucesos:

$$P(A \cap B) = P(A/B)P(B) = P(B/A)P(A) \tag{3.2}$$

Se cumple que $P(A'/B) = 1 - P(A/B)$.

3.4.1 Independencia de un par de sucesos

Dos sucesos se dicen independientes cuando $P(A/B) = P(A)$. Como $P(A \cap B) = P(A/B)P(B)$ si los sucesos A y B son independientes se cumple que $P(A \cap B) = P(A)P(B)$.

Recíprocamente, si $P(A \cap B) = P(A)P(B)$ despejando $P(A)$obtenemos:

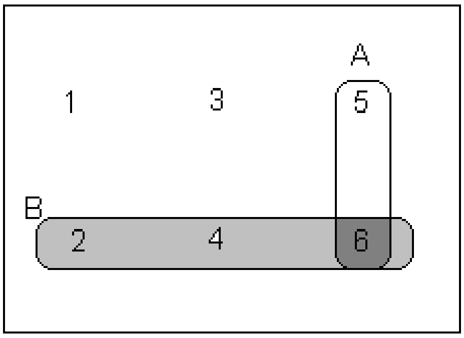
$$P(A) = \frac{P(A \cap B)}{P(B)} = P(A/B)$$

y por tanto A y B son independientes.

Ejemplos:

1. Lanzamos un dado, hallar la probabilidad de sacar un número mayor o igual que 5 sabiendo que se ha sacado un número par.

   El diagrama de Venn correspondiente al ejemplo es el de la siguiente figura



$A = \{$sacar un número mayor o igual que 5$\}$=$\{5{,}6\}$, $B = \{$sacar par$\}$=$\{2,\ 4,\ 6\}$, $A \cap B = \{6\}$

$P(A) = 2/6$; $P(B) = 3/6$; $P(A \cap B) = 1/6$.

Por tanto

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{6}}{\frac{3}{6}} = 1/3$$

Los sucesos A y B son independientes porque $P(A) = P(A/B) = 1/3$.

Otra forma de calcular $P(A/B)$ es usar la regla de Laplace considerando el suceso $B = \{2,\ 4,\ 6\}$ como universo o espacio muestral (en el diagrama de Venn se considera solamente lo que está dentro de B, marcado en gris, haciendo caso omiso el resto). El espacio muestral sólo tiene ahora 3 elementos y sólo uno de ellos, 6, verifica el suceso A.

$$P(A/B) = \frac{\text{número de veces que se cumple el suceso } A \text{ dentro de } B}{\text{número de elementos de } B} = \frac{1}{3}$$

No es raro que los alumnos confundan los sucesos independientes y los incompatibles. Aprovechamos este ejemplos para resaltar la diferencia entre estos conceptos. Como hemos dicho los sucesos A y B son independientes porque $P(A) = P(A/B) = \frac{1}{3}$. Sin embargo no son incompatibles. Dos sucesos son incompatibles cuando su intersección es el suceso imposible y por tanto no pueden verificarse simultáneamente. En este caso $A \cap B = \{6\} \neq \phi$. Cuando se obtenga un 6 al arrojar el dado se cumplirán simultáneamente los sucesos A y B.

2. Lanzamos un dado, hallar la probabilidad de sacar un número mayor o igual que 4 sabiendo que se ha sacado un número par. $A = \{$ sacar un cuatro, un cinco o un seis$\}$, B=$\{$ sacar par$\}$, $A \cap B = \{4, 6\}$

   $P(A) = 3/6$; $P(B) = 3/6$; $P(A \cap B) = 2/6$; $P(A/B) = 2/3$. Los sucesos A y B no son independientes.

3. En una empresa los trabajadores están distribuidos en dos plantas. En cada una de ellas los trabajadores entran en tres turnos, porque la maquinaria no puede quedar inactiva completamente. En la tabla siguiente aparece el número de trabajadores que tiene la empresa por planta y turno.

	Primer turno 0 a 8 horas	Segundo turno 8 a 16 horas	Tercer turno 16 a 24 horas	Total planta
Primera planta	250	150	200	600
Segunda planta	100	100	200	400
Total turno	350	250	400	Total = 1000

Si se selecciona un empleado al azar. Considerar los sucesos:

A, que se cumple si se elige un empleado de la primera planta.

B, que se cumple si se elige un trabajador del segundo turno (8 a 16).

C, que se cumple si se selecciona un trabajador del primer turno (de 0 a 8).

Una tabla de este tipo, en la que constan la frecuencia de las intersecciones de los sucesos considerados se conoce con el nombre de *tabla de contingencia* y puede ser de gran ayuda en la resolución de algunos problemas de probabilidad.

Estudiamos la independencia de los sucesos A y B:

$P(A) = \frac{600}{1000} = 0.6$; $P(A/B) = \frac{150}{250} = 0.6$. Como ambos valores son iguales, los sucesos A y B son independientes.

Estudiamos ahora la independencia de los sucesos A y C:

$P(A) = \frac{600}{1000} = 0.6$; $P(A/C) = \frac{250}{350}. = 0.71429$. Como estos valores no son iguales, los sucesos A y C no son independientes.

3.4.2 Independencia de más de dos sucesos

La probabilidad de la intersección de más de dos sucesos viene dada por la expresión

$P(A_1 \cap A_2 \cap ... \cap A_k) = P(A_1) \times P(A_2/A_1) \times P(A_3/A_1 \cap A_2) \times ... \times P(A_k/A_1 \cap A_2 \cap ... \cap A_{k-1})$

Se dice que los sucesos $A_1,\ A_2, ...,\ A_k$ son *independientes* si para todo subconjunto $\{A_{i_1},\ A_{i_2}, ...,\ A_{i_l}\}$ de $\{A_1,\ A_2, ...,\ A_k\}$, es decir que $\{A_{i_1},\ A_{i_2}, ...,\ A_{i_l}\} \subseteq \{A_1,\ A_2, ...,\ A_k\}$, se verifica:

$$P(A_{i_1} \cap A_{i_2} \cap ... \cap\ A_{i_l}) = P(A_{i_1}) \times P(A_{i_2}) \times ... \times P(A_{i_l})$$

3.5 Teorema de la Probabilidad Total.

Veremos en primer lugar un ejemplo que nos ayudara a entender el campo de aplicación del teorema de la probabilidad total:

Los trabajadores de una fábrica pueden ser directivos, obreros u oficinistas. Si se elige un trabajador al azar se sabe que:

1) La probabilidad de que sea un directivo $= P(A_1) = 0.10$.

2) La probabilidad de que sea un obrero $= P(A_2) = 0.75$.

3) La probabilidad de que sea oficinista $= P(A_3) = 0.15$.

4) La probabilidad de que sea mujer, si es un directivo $= P(B/A_1) = 0.20$.

5) La probabilidad de que sea mujer, si es un obrero $= P(B/A_2) = 0.45$.

6) La probabilidad de que sea mujer, si es un oficinista $= P(B/A_3) = 0.50$.

¿Cuál es la probabilidad de que un trabajador elegido al azar sea mujer $= P(B)$?

Esta última probabilidad es un ejemplo de la llamada probabilidad total. El espacio muestral es el conjunto total de trabajadores. Las probabilidades de ser mujer dentro de cada categoría de empleados de la fábrica son probabilidades condicionadas (la información suplementaria en cada caso es que se conoce la categoría profesional a la que pertenece el trabajador seleccionado).

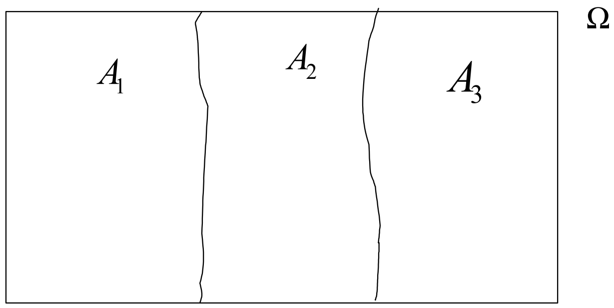
Un conjunto de sucesos A_1, A_2, ..., A_n se dicen que forman una *partición de* Ω ó también que forman un *sistema exhaustivo y excluyente*, si se cumple:

1. $\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup ... \cup A_n = \Omega$

2. $A_i \cap A_j = \emptyset \qquad \forall i \neq j$

   En el ejemplo estos sucesos podrían ser: A_1, que se cumple si la persona seleccionada pertenece a la categoría de directivos, A_2, que se cumple si la persona seleccionada pertenece a la de obreros, y A_3, que se cumple si la persona seleccionada pertenece a la categoría de oficinistas. Estos conjuntos no tienen elementos en común y su unión cubre el conjunto de todos los empleados de la fábrica, cumpliéndose por tanto las propiedades de una partición (ó de un sistema exhaustivo y excluyente).

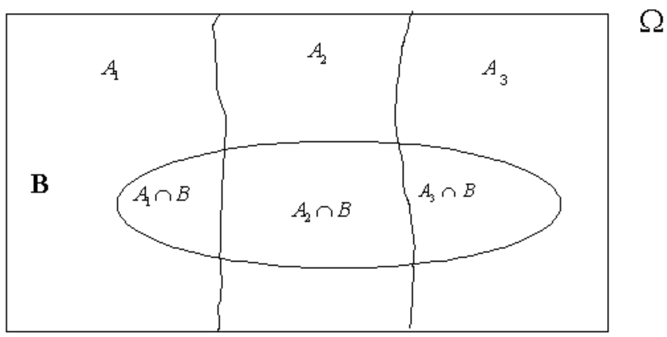
   El diagrama de Venn correspondiente a una partición del espacio muestral en tres sucesos puede presentar el siguiente aspecto.



El teorema de la probabilidad total tiene el siguiente enunciado:
Sea $A_1, A_2, \ldots, A_n$ una partición del espacio muestral. Sea B un suceso. Entonces se cumple:

$$P(B) = \sum_{i=1}^{n} P(A_i)P(B/A_i) = P(A_1)(B/A_1) + P(A_2)P(B/A_2) + ... + P(A_n)P(B/A_n).$$

La siguiente gráfica representa un diagrama de Venn apropiado para ilustrar el enunciado de este teorema:



El suceso B es el de *probabilidad total (sin añadir ninguna condición).* En el ejemplo, el suceso que se cumple si se selecciona una mujer de esta fábrica es de probabilidad total, en contraposición con $P(B/A_i)$ que serían probabilidades parciales (condicionadas por algún otro suceso). Por ejemplo $P(B/A_1) =$ 0.20 indica la probabilidad de que la persona seleccionada sea una mujer si sabemos que pertenece a la categoría de directivos.

Demostración del teorema de probabilidad total:

$$P(B) = P(\Omega \cap B) = P((A_1 \cup A_2 \cup ... \cup A_n) \cap B) =$$

aplicando la propiedad distributiva :

$$= P((A_1 \cap B) \cup (A_2 \cap B) \cup ... \cup (A_n \cap B)) =$$

De $A_i \cap A_j = \emptyset$ se deduce que $(A_i \cap B) \cap P(A_j \cap B) = \emptyset$. Por el segundo axioma de la probabilidad tenemos:

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

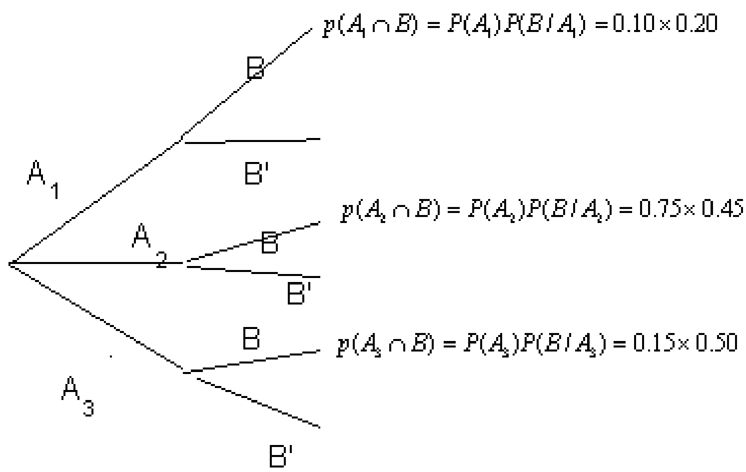
Usando ahora la fórmula 3.2 de la página 18 para la probabilidad de la intersección obtenemos:

$$= P(A_1)P(B/A_1) + P(A_2)P(B/A_2) + ... + P(A_n)P(B/A_n) = \sum_{i=1}^{n} P(A_i)P(B/A_i).$$

En el problema del ejemplo obtenemos que la probabilidad de seleccionar una mujer entre el total de los empleado de la fábrica es:

$$\begin{aligned} P(B) &= P(A_1)P(B/A_1) + P(A_2)P(B/A_2) + P(A_3)P(B/A_3) = \\ &= 0.10 \times 0.20 + 0.75 \times 0.45 + 0.15 \times 0.50 = 0.432\,5 \end{aligned}$$

En la siguiente figura está esquematizado el diagrama de árbol del problema junto con las probabilidades que le corresponden a las ramas en las que se cumple el suceso B.



3.6 Teorema de Bayes.

Este teorema se usa para calcular una probabilidad condicionada de una forma indirecta. Volviendo al ejemplo enunciado en el párrafo anterior, supongamos que una vez seleccionado el trabajador de la fábrica se nos informa que es una mujer. ¿Cuál es la probabilidad de que esta mujer sea directiva? La probabilidad de que la persona seleccionada fuera un directivo, sin ninguna información adicional es $P(A_1) = 0.10$, que llamamos probabilidad "a priori". Pero lo que queremos hallar es la probabilidad de que sea directivo sabiendo que se ha seleccionado una mujer. La probabilidad pedida es $P(A_1/B)$ que llamamos probabilidad "a posteriori", o sea que esta probabilidad viene influida por la información conocida sobre el resultado del experimento.

El teorema de Bayes se enuncia de la forma siguiente:

Sea $A_1, A_2, \ldots, A_n$ una partición de Ω. Sea B un suceso con $P(B) \neq 0$. Entonces se verifica:

$$P(A_k/B) = \frac{P(A_k \cap B)}{P(B)} = \frac{P(A_k)P(B/A_k)}{P(B)} = \frac{P(A_k)P(B/A_k)}{\sum_{i=1}^{n} P(A_i)P(B/A_i)}$$

Entonces, en el ejemplo anterior, la probabilidad de que un empleado sea directivo sabiendo que es mujer, $P(A_1/B)$, se calcularía aplicando el teorema de Bayes. En el denominador aparece la expresión de la probabilidad total, $P(B)$, que ya hemos calculado anteriormente. En el numerador aparecería el primer sumando de dicha probabilidad. Así tenemos que

$$\begin{aligned} P(A_1/B) &= \frac{P(A_1)P(B/A_1)}{P(A_1)P(B/A_1) + P(A_2)P(B/A_2) + P(A_3)P(B/A_3)} \\ P(A_1/B) &= \frac{0.10 \times 0.20}{0.10 \times 0.20 + 0.75 \times 0.45 + 0.15 \times 0.50} = \frac{0.0\,2}{0.\,432\,5} = 0.04624 \end{aligned}$$

es la probabilidad de seleccionar un directivo entre las mujeres trabajadoras de esta fábrica.

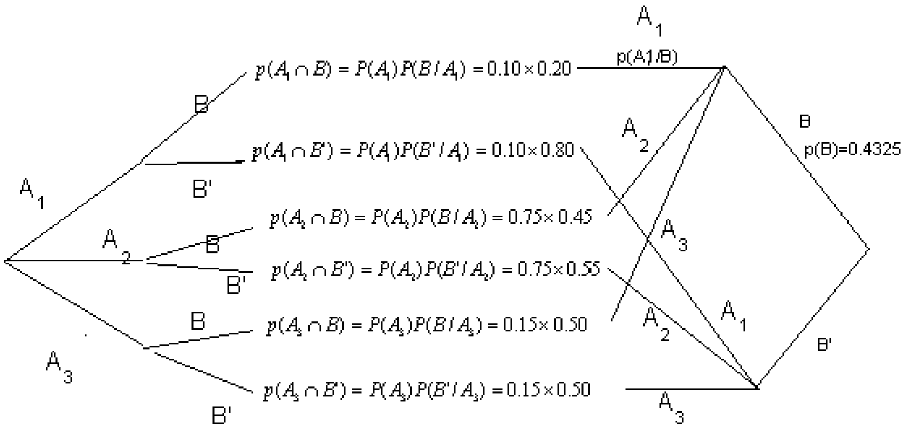
Se puede ilustrar el teorema de Bayes por medio de un diagrama de árbol. En la figura siguiente aparece representado el que corresponde al ejemplo. Las probabilidades de la intersección de los sucesos A_1 y B deben ser la misma si se obtienen usando las ramas del árbol de la derecha o de la izquierda. El caso del ejemplo está ilustrado en las ramas superiores de ambos árboles:

Rama de la izquierda: $p(A_1 \cap B) = P(A_1)P(B/A_1) = 0.10 \times 0.20$

Rama de la derecha: $p(A_1 \cap B) = P(B)P(A_1/B) = 0.4325 \times P(A_1/B)$

Igualando y despejando obtenemos

$$P(A_1/B) = \frac{0.10 \times 0.20}{0.4325} = 0.04624$$



Si quisiéramos hallar $P(A_3/B')$, probabilidad de que el empleado seleccionado sea un oficinista si sabemos que es un hombre, usaríamos las ramas inferiores de ambos árboles:

$$\text{Rama de la izquierda: } p(A_3 \cap B') = P(A_3)P(B'/A_3) = 0.15 \times 0.50$$

$$\text{Rama de la derecha: } p(A_3 \cap B') = P(B')P(A_3/B') = 0.5675 \times P(A_1/B)$$

Igualando y despejando obtenemos

$$P(A_3/B') = \frac{0.15 \times 0.50}{0.5675} = 0.1322$$