TEMA 3 ELEMENTOS DE PROBABILIDAD

Pablo Buenestado

Curso 2020-2021 Otoño

Departamento de Matemáticas (UPC)

Índice

Esquema

El desarrollo de la teoría de la probabilidad fue financiada por apostadores en el siglo XVII, quienes contrataron a algunos matemáticos famosos para que calculasen la probabilidad correcta de ciertos juegos de azar.

Con el tiempo, la gente se dio cuenta de que los procesos científicos también son azarosos y desde entonces se han empleado métodos de probabilidad para estudiar el entorno físico.

Actualmente, la probabilidad constituye una gran rama de las matemáticas. Existen muchos libros al respecto y numerosos investigadores han dedicado bastante de su tiempo con el propósito de ampliar su desarrollo.

En este tema se presenta una introducción de los conceptos de probabilidad más relevantes para el estudio de la estadística.

Esquema

1. PROBABILIDAD
   - Introducción
   - Ideas básicas
     - Combinación de eventos
     - Eventos mutuamente excluyentes
     - Probabilidades
     - Axiomas de la probabilidad
     - Espacios muestrales con resultados igualmente probables
     - Regla de la suma
   - Probabilidad condicional e independencia
     - Eventos independientes
     - La regla de la multiplicación
     - Ley de la probabilidad total
     - Regla de Bayes
   - Métodos de conteo
     - Permutaciones
     - Combinaciones

Para realizar un estudio sistemático de probabilidad, se necesita cierta terminología.

Un experimento constituye un proceso con un resultado que no se puede predecir certeramente con anterioridad.

El hecho de lanzar una moneda al aire, arrojar un dado, medir el diámetro de un perno, pesar los contenidos de una caja de cereales, o medir la resistencia de una cuerda de pescar, son ejemplos de experimentos.

Con la finalidad de analizar un experimento en términos probabilísticos, se debe especificar sus posibles resultados.

Definición

Al conjunto de todos los posibles resultados de un experimento se le llama espacio muestral.

Utilizamos la letra S para definir al espacio muestral.

El espacio muestral puede ser finito o infinito.

Ejemplos de espacios muestrales finitos:

Al lanzar al aire una moneda se puede utilizar el conjunto {"cara", "cruz"} como el espacio muestral.

Al lanzar un dado de seis caras se puede usar al conjunto $\{1, 2, 3, 4, 5, 6\}$.

Algunos experimentos tienen espacios muestrales con un número infinito de resultados:

Por ejemplo, imaginemos que un buril[1] con diámetro de 10 mm hace perforaciones en una lámina de metal. Debido a las variaciones en el ángulo de la perforación y a los pequeños movimientos en la lámina de metal, los diámetros de los agujeros varían entre 10 y 10.2 mm.

Por tanto, para el experimento de perforación sería razonable un espacio muestral que esté en el intervalo (10.0, 10.2), o en notación de conjuntos, $\{x|10.0 < x < 10.2\}$. Obviamente, este conjunto contiene un número infinito de resultados.

[1] Instrumento usado principalmente por los grabadores para grabar metales o piedra que consiste en una barra prismática fina y puntiaguda de acero.(Punzón)

En muchos experimentos se puede escoger entre diversos espacios muestrales.

Por ejemplo, supongamos un proceso que produce clavos de acero cuyas longitudes varían entre 5.20 y 5.25 cm. Una opción obvia para el espacio muestral de la longitud de un clavo sería el conjunto $\{x|5.20 < x < 5.25\}$.

Sin embargo, si el objetivo fuera simplemente determinar si el clavo es demasiado corto, demasiado largo o está dentro de ciertos límites específicos, una buena elección sería que el espacio muestral fuera {demasiado corto, demasiado largo, dentro de las especificaciones}.

Con frecuencia, al estudiar experimentos, se está interesado en un subconjunto particular de resultados.

Por ejemplo, se puede tener interés en la probabilidad de que un dado caiga en un número par. El espacio muestral para el experimento es {1, 2, 3, 4, 5, 6} y el correspondiente a que caiga en un número par es el subconjunto {2, 4, 6}.

En el ejemplo del buril usado para perforar, se puede tener interés en la probabilidad de que un hueco tenga un diámetro menor a 10.1 mm. Esto último corresponde al subconjunto $\{x|10.0 < x < 10.1\}$.

Existe un nombre especial para el subconjunto de un espacio muestral:

Definición

Un subconjunto de un espacio muestral se denomina evento o suceso.

Observemos que para cualquier espacio muestral, el conjunto vacío $\emptyset$ es un evento, como lo es todo el espacio muestral.

Se dice que un evento ocurrió si el resultado del experimento es alguno de los resultados en el evento.

Por ejemplo, si un dado cae en el número 2, habrán ocurrido los eventos $\{2, 4, 6\}$ y $\{1, 2, 3\}$, junto con cualquier otro evento que contenga el resultado "2".

Ejemplo

Un ingeniero eléctrico tiene en su mano dos cajas de resistores[2], cada una con cuatro de éstos. Los resistores de la primera caja están etiquetados con 10 Ω (ohms), pero, de hecho, sus resistencias son de 9, 10, 11 y 12 Ω. Los resistores de la segunda caja tienen la etiqueta de 20 Ω, pero sus resistencias son de 18, 19, 20 y 21 Ω. El ingeniero elige un resistor de cada caja y determina la resistencia de cada uno.

Sea A el evento para el cual el primer resistor tiene una resistencia mayor a 10, sea B el evento en el que el segundo resistor tiene una resistencia menor a 19 y sea C el evento en el cual la suma de las resistencias es igual a 28.

Se pide determinar un espacio muestral para este experimento y especificar los subconjuntos que corresponden a los eventos A, B y C.

[2] Componente electrónico diseñado para introducir una resistencia eléctrica determinada entre dos puntos de un circuito eléctrico. (Resistencia)

Solución

Un buen espacio muestral es el conjunto de pares ordenados en el que el primer elemento representa la resistencia del primer resistor (primera caja) y el segundo elemento constituye la del segundo resistor (segunda caja).

Así, el espacio muestral S es:

$$S = \{(9,18),(9,19),(9,20),(9,21),(10,18),(10,19),(10,20),(10,21),$$

$$(11,18),(11,19),(11,20),(11,21),(12,18),(12,19),(12,20),(12,21)\}$$

Los eventos A, B y C están dados por:

$$A = \{(11,18),(11,19),(11,20),(11,21),(12,18),(12,19),(12,20),(12,21)\}$$

$$B = \{(9,18),(10,18),(11,18),(12,18)\}$$

$$C = \{(9,19),(10,18)\}$$

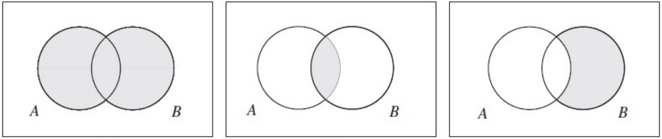
Combinación de eventos

Con frecuencia se construyen eventos mediante la combinación de eventos más sencillos. Debido a que aquéllos son subconjuntos de espacios muestrales, es usual emplear la notación de conjuntos para describir los eventos construidos de esta forma. A continuación se repasará la notación necesaria.

- La unión de dos eventos A y B, se denota por $A \cup B$, es el conjunto de resultados que pertenecen ya sea a A o B, o a ambos. Esto es, $A \cup B$ significa "A o B". Por tanto, el evento $A \cup B$ se presenta siempre que ocurre A o B (o ambos).
- La intersección de dos eventos A y B se denota como $A \cap B$; es decir, constituye el conjunto de resultados que pertenece tanto a A como a B. Por consecuencia, $A \cap B$ significa "A y B". Por consiguiente, el evento $A \cap B$ se presenta siempre que A y B ocurren.
- El complemento o complementario de un evento A se denota por A^c (o $\bar{A}$), es el conjunto de resultados que no pertenecen a A. Es decir, A^c significa "no A". Por consiguiente, el evento A^c se presenta siempre que no ocurra A.

Los eventos se pueden mostrar gráficamente con los diagramas de Venn. Dónde el espacio muestral es el recuadro, en el que se dibujan los sucesos A y B.

La figura muestra sombreados los eventos a) $A \cup B$, b) $A \cap B$ y c) $B \cap A^c$.



a) b) c)

Ejemplo (Continuación)

A partir de los eventos definidos en el ejemplo de los resistores, determinemos $B \cup C$

Solución

El evento $B \cup C$ contiene todos los resultados que pertenecen a B o a C, o a ambos. Por tanto,

$$B \cup C = \{(9, 18), (10, 18), (11, 18), (12, 18), (9, 19)\}$$

Observemos que el suceso $(10, 18)$ forma parte tanto de B como de C y en la unión $B \cup C$ lo especificamos una sola vez.

Ejemplo (Continuación)

A partir de los eventos definidos en el ejemplo de los resistores, determinemos $A \cap B^c$

Solución

El evento B^c contiene los resultados en el espacio muestral que no pertenecen a B.

$$B^c = \{(9,19),(9,20),(9,21),(10,19),(10,20),(10,21),$$

$$(11,19),(11,20),(11,21),(12,19),(12,20),(12,21)\}$$

De ahí, el evento $A \cap B^c$ contiene los resultados que pertenecen a A y no pertenecen a B. Por consiguiente,

$$A \cap B^c = \{(11,19),(11,20),(11,21),(12,19),(12,20),(12,21)\}$$

Eventos mutuamente excluyentes

Existen ciertos eventos que nunca se presentan simultáneamente.

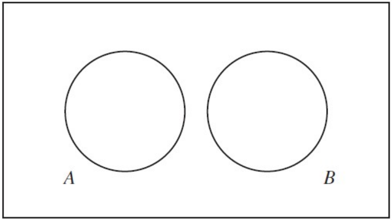
Por ejemplo, es imposible que una moneda que se lance al aire caiga a la vez en ”cruz” y ”cara”, al igual que un clavo de acero sea al mismo tiempo demasiado largo y corto.

A eventos de este tipo se les llama mutuamente excluyentes o disjuntos.

Definición

Se dice que los eventos A y B son mutuamente excluyentes o disjuntos si no tienen resultados en común.

El diagrama de Venn de la figura muestra los eventos A y B mutuamente excluyentes.



Ejemplo (Continuación)

A partir de los eventos definidos en el ejemplo de los resistores, si se realiza el experimento, ¿es posible que los eventos A y B ocurran al mismo tiempo?

¿Qué pasa con B y C? ¿A y C?

¿Qué par de eventos es mutuamente excluyente?

Solución

Si el resultado es (11, 18) o (12, 18), entonces tanto el evento A como el B ocurren.

Si el resultado es (10, 18), entonces ocurren los eventos B y C.

Es imposible que ocurran al mismo tiempo A y C, ya que estos eventos son mutuamente excluyentes al no tener ningún resultado en común.

Probabilidades

Todo evento en un espacio muestral tiene una probabilidad de ocurrir.

Intuitivamente, la probabilidad es una medida cuantitativa de qué tan probable es que ocurra un evento.

Formalmente hablando, hay varias interpretaciones de la probabilidad; la primera que se adoptará es que la probabilidad de un evento representa la proporción de veces que se presentaría el evento a largo plazo, si el experimento se repitiera una y otra vez.

Con frecuencia se usa la letra P para representar la probabilidad. Por tanto, cuando se lanza una moneda al aire la notación "$P(\text{"cara"}) = 1/2$" significa que la probabilidad de que la moneda caiga en "cara" es igual a $1/2$.

Resumen

Dado un experimento y cualquier evento A:

- La expresión $P(A)$ denota la probabilidad de que ocurra el evento A.
- $P(A)$ constituye la proporción de veces que se presenta el evento A en el tiempo, si es que el experimento se realizara una y otra vez.

En muchas situaciones, la única forma de calcular la probabilidad de un evento es repetir el experimento muchas veces y determinar la proporción de veces que ocurre.

Por ejemplo, si se deseara calcular la probabilidad de que un tablero de circuitos impresos fabricado por cierto proceso esté defectuoso, usualmente se necesitaría producir cierta cantidad de tableros y probarlos para determinar la proporción de los defectuosos.

En algunos casos, las probabilidades se pueden determinar si se conoce la naturaleza física del experimento.

Por ejemplo, si se sabe que la forma de un dado es casi igual a la de un cubo perfecto y que su masa está distribuida aproximadamente en forma homogénea, se puede suponer que cada una de sus seis caras tiene la misma probabilidad de salir cuando se lanza el dado.

Una vez que se han encontrado las probabilidades de ciertos eventos mediante el conocimiento científico o la experiencia, se puede calcular matemáticamente las probabilidades de otros eventos.

Por ejemplo, si se ha calculado a través de la experimentación que la probabilidad de que un tablero de circuitos impresos esté defectuoso es de 0.10, se puede calcular que la probabilidad de que un tablero no esté defectuoso es de 0.90.

Como otro ejemplo, supongamos que los clavos de acero producidos por determinado proceso no cumplen con la longitud especificada, ya sea porque son demasiado cortos o demasiado largos. Al medir gran cantidad de clavos, se calculó que la probabilidad de que uno de ellos sea demasiado corto es de 0.02 y que la probabilidad de que otro sea demasiado largo es de 0.03. Entonces puede calcularse que la probabilidad de que un clavo no cumpla con la especificación es de 0.05.

En la práctica, los científicos e ingenieros calculan las probabilidades de ciertos eventos con base en el conocimiento científico y la experiencia, y posteriormente utilizan reglas matemáticas para calcular las estimaciones de las probabilidades de otros eventos.

Más adelante, en el presente tema, se explican algunas de estas reglas y se muestra cómo utilizarlas.

Axiomas de la probabilidad

El tema de la probabilidad se basa en tres reglas de sentido común, conocidas como axiomas.

Axiomas de la probabilidad

1. Sea S un espacio muestral. Entonces
$$P(S) = 1$$
2. Para cualquier evento A,
$$0 \leq P(A) \leq 1$$
3. Si A y B son eventos mutuamente excluyentes, entonces
$$P(A \cup B) = P(A) + P(B)$$
De forma más general, si $A_1, A_2, \ldots$ son eventos mutuamente excluyentes, entonces $P(A_1 \cup A_2 \cup \ldots) = P(A_1) + P(A_2) + \ldots$

Veamos que los tres axiomas en realidad concuerdan con el sentido común.

El primero establece que el resultado de un experimento siempre está en el espacio muestral. Esto es obvio, puesto que, por definición, el espacio muestral contiene todos los resultados posibles del experimento.

El segundo dice que la frecuencia a largo plazo de cualquier evento siempre se encuentra entre 0 y 100%.

Un ejemplo que demuestra al tercer axioma, que ya se analizó anteriormente, es el del proceso que produce clavos de acero, en donde la probabilidad de que un clavo sea demasiado corto es de 0.02 y la de que un clavo es demasiado largo es de 0.03. El tercer axioma establece que la probabilidad de que el clavo sea demasiado corto o muy largo es $0.02 + 0.03 = 0.05$.

Presentamos dos reglas sencillas que son útiles para calcular probabilidades. Estas reglas son intuitivamente obvias y también pueden comprobarse a través de los axiomas.

Consecuencias de los Axiomas de Probabilidad

- Para cualquier evento A, se cumple que

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

- Si $\emptyset$ denota el espacio vacío, entonces

$$P(\emptyset) = 0$$

La ecuación $P(A^c) = 1 - P(A)$ establece que la probabilidad de que un evento no ocurra es igual a 1 menos la probabilidad de que ocurra. Por ejemplo, si existe una probabilidad de 40% de que llueva, hay una probabilidad de 60% de que no llueva.

La ecuación $P(\emptyset) = 0$ establece que es imposible que un experimento no tenga ningún resultado.

Ejemplo

El objetivo de una prueba de tiro consiste de un blanco rodeado por dos anillos concéntricos.

Se dispara un proyectil hacia el objetivo.

La probabilidad de que pegue en el blanco es de 0.10, la de que atine en el anillo interior es de 0.25 y la de que acierte en el anillo exterior es de 0.45.

¿Cuál es la probabilidad de que el proyectil acierte en el objetivo?

¿Cuál es la probabilidad de que no pegue en este último?

Solución

Pegar en el blanco, acertar en el anillo interior y atinar en el anillo exterior son eventos mutuamente excluyentes, ya que es imposible que más de uno de éstos ocurra a la vez.

Por tanto, utilizando el axioma 3,

$$P(\text{pega en el objetivo}) = P(\text{blanco}) + P(\text{anillo interior}) + P(\text{anillo exterior})$$

$$P(\text{pega en el objetivo}) = 0.1 + 0.25 + 0.45 = 0.80$$

Ahora se puede calcular la probabilidad de que el proyectil no pegue en el objetivo utilizando la ecuación relativa al complementario:

$$P(\text{no pega en el objetivo}) = 1 - P(\text{pega en el objetivo}) = 1 - 0.80 = 0.20$$

Ejemplo

La tabla presenta las probabilidades del número de veces en que el sistema de cierta computadora se ”caerá” en el transcurso de una semana.

Número de casos	Probabilidad
0	0.60
1	0.30
2	0.05
3	0.04
4	0.01

Sea A el evento de que haya más de dos ”caídas” durante la semana, y B el evento de que el sistema se ”caerá” por lo menos una vez.

Determinemos el espacio muestral.

Después precisemos los subconjuntos del espacio muestral que corresponden a los eventos A y B.

Posteriormente determinemos $P(A)$ y $P(B)$.

Solución

El espacio muestral del experimento es el conjunto $S = \{0, 1, 2, 3, 4\}$.

Los eventos son $A = \{3, 4\}$ y $B = \{1, 2, 3, 4\}$.

Para encontrar $P(A)$, advertimos que A constituye el evento en que se presenten tres "caídas" o que haya cuatro "caídas". Los eventos "que se presenten tres caídas" y "que ocurran cuatro caídas" son mutuamente excluyentes. Por tanto, mediante el axioma tres, se concluye que

$$P(A) = P(\text{ocurran tres "caídas" o se presenten cuatro "caídas"})$$

$$P(A) = P(\text{ocurran tres "caídas"}) + P(\text{sucedan cuatro "caídas"})$$

$$P(A) = 0.04 + 0.01 = 0.05$$

Solución (Continuación)

Calculamos $P(B)$ en dos formas.

Primero, observemos que B^c es el evento de que no haya ninguna caída. Por tanto, utilizando la ecuación relativa al complementario

$$P(B) = 1 - P(B^c) = 1 - P(\text{no ocurran ”caídas”}) = 1 - 0.60 = 0.40$$

En la segunda forma para calcular $P(B)$, observemos que B es el evento de que haya una ”caída” o de que sucedan dos o que ocurran tres ”caídas” o haya cuatro de éstas. Estos eventos son mutuamente excluyentes. Por consiguiente, utilizando el axioma tres, se concluye que

$$P(B) = P(\text{una ”caída”}) + P(\text{dos ”caídas”}) + P(\text{tres ”caídas”}) + P(\text{cuatro ”caídas”})$$

$$P(B) = 0.30 + 0.05 + 0.04 + 0.01 = 0.40$$

En el ejemplo de la computadora calculamos las probabilidades de los eventos $A = \{3, 4\}$ y $B = \{1, 2, 3, 4\}$ sumando las probabilidades de los resultados de cada uno de los eventos:

$$P(A) = P(3) + P(4)$$

$$P(B) = P(1) + P(2) + P(3) + P(4)$$

En general, este método funciona. Dado que cualesquiera dos resultados en un espacio muestral son mutuamente excluyentes, la probabilidad de que cualquier evento contenga un número finito de resultados se puede determinar mediante la suma de las probabilidades de los resultados que incluyen al evento.

Consecuencias de los Axiomas de Probabilidad

Si A es un evento y $A = \{E_1, E_2, \ldots, E_n\}$, entonces

$$P(A) = P(E_1) + P(E_2) + \cdots + P(E_n)$$

Espacios muestrales con resultados igualmente probables

En algunos experimentos se puede construir un espacio muestral en el cual todos los resultados sean igualmente probables.

Un ejemplo sencillo es el lanzamiento de un dado, en el cual el espacio muestral es $\{1, 2, 3, 4, 5, 6\}$ y cada uno de estos resultados tiene una probabilidad de $1/6$.

Otro tipo de experimento que tiene resultados igualmente probables es la selección aleatoria de un elemento a partir de una población de elementos. Se puede suponer que los elementos en la población son los resultados en un espacio muestral y que cada elemento tiene la misma probabilidad de ser seleccionado.

Espacios muestrales con resultados igualmente probables

Una población a partir de la cual se muestrea un elemento en forma aleatoria constituye un espacio muestral con resultados igualmente probables.

Si un espacio muestral contiene N resultados igualmente probables, la probabilidad de cada resultado es $1/N$.

Esto es así porque la probabilidad de todo el espacio muestral debe ser 1 y esta probabilidad se divide equitativamente entre los N resultados.

Si A representa un evento que contiene k resultados, entonces se puede determinar $P(A)$ mediante la suma de las probabilidades de los k resultados, de tal forma que $P(A) = k/N$.

Resumen

Si S es un espacio muestral que contiene N resultados igualmente probables y si A es un evento que contiene k resultados, entonces

$$P(A) = k/N$$

Ejemplo

Un troquel de extrusión[3] se utiliza para producir varillas de aluminio.

Existen ciertas especificaciones para la longitud y diámetro de las varillas.

Para cada una de éstas, la longitud puede ser demasiado corta, demasiado larga o estar bien y el diámetro se puede clasificar en muy delgado, muy grueso o estar bien.

En una población de mil varillas, el número de ellas en cada clase es:

	Diámetro		
Longitud	Muy delgado	Está bien	Muy grueso
Demasiado corta	10	3	5
Está bien	38	900	4
Demasiado larga	2	25	13

Se toma una varilla aleatoriamente a partir de esta población. ¿Cuál es la probabilidad de que sea demasiado corta?

[3] Proceso utilizado para crear objetos con sección transversal definida y fija.

Solución

Se considera que cada una de las mil varillas representa un resultado en un espacio muestral.

Cada uno de los mil resultados tiene la misma probabilidad.

Se resolverá el problema contando el número de resultados que corresponde al evento.

El número de varillas que son demasiado cortas es $10 + 3 + 5 = 18$.

Dado que el número total de varillas es mil,

$$P(\text{demasiado corta}) = \frac{18}{1000}$$

Regla de la suma

Si A y B son eventos mutuamente excluyentes, entonces

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

Esta regla se puede generalizar para abarcar el caso en el que A y B no sean mutuamente excluyentes.

En el ejemplo siguiente se muestra este razonamiento.

Ejemplo

Con referencia al ejemplo de las varillas, si se selecciona aleatoriamente una varilla, ¿cuál es la probabilidad de que sea demasiado corta o demasiado gruesa?

Solución

Primero, resolveremos este problema al contar el número de resultados que corresponde al evento. En la siguiente tabla se circuló la cantidad de varillas que son demasiado gruesas y el número de varillas que son muy cortas aparecen en recuadros. Observemos que hay cinco varillas que son muy cortas y demasiado gruesas.

	Diámetro		
Longitud	Muy delgado	Está bien	Muy grueso
Demasiado corta	$\boxed{10}$	$\boxed{3}$	$\boxed{5}$
Está bien	38	900	4
Demasiado larga	2	25	13

De los mil resultados, el número de varillas que son demasiado cortas o muy gruesas es $10 + 3 + 5 + 4 + 13 = 35$. Por consiguiente,

$$P(\text{demasiado corta o muy gruesa}) = \frac{35}{1000}$$

Ahora se resolverá el problema de tal forma que conduzca a un método más general.

En el espacio muestral hay $10 + 3 + 5 = 18$ varillas que son demasiado cortas y $5 + 4 + 13 = 22$ varillas muy gruesas.

Pero si se trata de encontrar el número de varillas que sean demasiado cortas o muy gruesas al sumar $18 + 22$, se obtiene un número muy grande (40 en vez de 35).

La razón es que hay cinco varillas que son demasiado cortas y muy gruesas y éstas se cuentan dos veces.

No obstante, se puede resolver el problema al sumar 18 y 22, pero entonces se le debe restar cinco para corregir el doble conteo.

Se replantea este razonamiento al utilizar probabilidades:

$$P(\text{demasiado corta}) = \frac{18}{1000}$$
$$P(\text{muy gruesa}) = \frac{35}{1000}$$
$$P(\text{demasiado corta y muy gruesa}) = \frac{5}{1000}$$

Construimos la probabilidad buscada de la siguiente manera:

$P(\text{demasiado corta o muy gruesa}) =$
$P(\text{demasiado corta}) + P(\text{muy gruesa}) - P(\text{demasiado corta y muy gruesa})$

En consecuencia,

$$P(\text{demasiado corta o muy gruesa}) = \frac{18}{1000} + \frac{22}{1000} - \frac{5}{1000} = \frac{35}{1000}$$

El método del ejemplo anterior es válido para cualesquiera dos eventos en un espacio muestral.

En general, para determinar la probabilidad de que ocurran cualesquiera de los dos eventos, se suman las probabilidades de los eventos y después se resta la probabilidad de que ambos ocurran.

Resumen

Sean A y B cualesquiera eventos, entonces

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

Observemos que si A y B son mutuamente excluyentes, entonces $P(A \cap B) = 0$, por lo que, en este caso la expresión anterior se reduce al axioma 3.

Esquema

Un espacio muestral contiene todos los resultados posibles de un experimento.

A veces se obtiene algo de información adicional acerca de un experimento que indica que los resultados provienen de cierta parte del espacio muestral.

En este caso, la probabilidad de un evento está basada en los resultados de esa parte del espacio muestral.

Una probabilidad que se basa en una parte de un espacio muestral se llama probabilidad condicional o probabilidad condicionada.

Analicemos esta idea a través de algunos ejemplos.

En el ejemplo en el que se analizó una muestra de mil varillas de aluminio, para cada varilla, la longitud se clasifica como demasiado corta, demasiado larga o está bien y el diámetro se clasifica como muy delgado, muy grueso o está bien.

Esas mil varillas constituyen un espacio muestral en el que cada varilla tiene la misma probabilidad de ser seleccionada.

El número de varillas en cada categoría se presenta en la tabla siguiente.

	Diámetro		
Longitud	Muy delgado	Está bien	Muy grueso
Demasiado corta	10	3	5
Está bien	38	900	4
Demasiado larga	2	25	13

De las mil varillas, 928(=3+900+25) satisfacen la especificación de diámetro.

Por tanto, si se selecciona una varilla,

$$P(\text{diámetro está bien}) = 928/1000 = 0.928$$

A esta circunstancia se le llama probabilidad incondicional, ya que se basa en todo el espacio muestral.

Ahora supongamos que se toma una varilla, se mide su longitud y se encuentra que satisface la especificación. ¿Cuál es la probabilidad de que el diámetro también satisfaga la especificación?

La clave para calcular esta probabilidad es darse cuenta de que el saber que la longitud satisface la especificación reduce el espacio muestral del que se seleccionó la varilla.

La tabla siguiente presenta esta idea.

	Diámetro		
Longitud	Muy delgado	Está bien	Muy grueso
Demasiado corta	-	-	-
Está bien	38	900	4
Demasiado larga	-	-	-

Una vez que se conoce que se satisface la especificación de la longitud, se sabe que esa varilla será una de las 942(=38+900+4) en el espacio muestral que se presenta en la tabla anterior.

De las 942 varillas de este espacio muestral, 900 satisfacen la especificación del diámetro.

Por tanto, si se sabe que la varilla satisface la especificación de longitud, la probabilidad de que la varilla satisfaga la especificación del diámetro es: 900/942.

Se dice que la probabilidad condicional de que la varilla satisfaga la especificación de un diámetro dado que satisface la especificación de longitud es igual a 900/942 y se escribe

$$P(\text{diámetro está bien} \mid \text{longitud está bien}) = 900/942 = 0.955$$

Observemos que la probabilidad condicional $P(\text{diámetro está bien} \mid \text{longitud está bien})$ es diferente de la probabilidad incondicional $P(\text{diámetro está bien})$, que se calculó para todo el espacio muestral de $\frac{3+900+25}{1000} = 0.928$,

$$P(\text{diámetro está bien}) = \frac{928}{1000} = 0.928$$

Ejemplo

Calcula la probabilidad condicional
$P(\text{diámetro está bien} \mid \text{longitud demasiada larga})$. ¿Ésta es la misma que la probabilidad incondicional $P(\text{diámetro está bien})$?

Solución

La probabilidad condicional
$P(\text{diámetro está bien} \mid \text{longitud demasiada larga})$ se calcula bajo la suposición de que la varilla es demasiado larga. Esto último reduce el espacio muestral a los 40 elementos que se muestran resaltados en la tabla siguiente.

	Diámetro		
Longitud	Muy delgado	Está bien	Muy grueso
Demasiado corta	10	3	5
Está bien	38	900	4
Demasiado larga	2	25	13

Solución (Continuación)

De los 40 resultados, 25 satisfacen la especificación de diámetro. Por tanto

$$P(\text{diámetro está bien} \mid \text{longitud demasiada larga}) = \frac{25}{40} = 0.625$$

La probabilidad incondicional $P(\text{diámetro está bien})$ se calcula con base en todos los mil resultados en el espacio muestral y es igual a $928/1000 = 0.928$.

$$P(\text{diámetro está bien}) = \frac{928}{1000} = 0.928$$

En este caso, la probabilidad condicional es diferente de la probabilidad incondicional.

Analicemos la solución del ejemplo anterior detenidamente. Encontramos que

$$P(\text{diámetro está bien} \mid \text{longitud demasiada larga}) = \frac{25}{40}$$

En la respuesta $25/40$, el denominador, 40, representa el número de resultados que satisfacen la condición de que la varilla es demasiado larga, mientras que el numerador, 25, representa el número de resultados que satisfacen ambas condiciones, que la varilla es demasiado larga y que su diámetro está bien.

Si dividimos tanto al numerador como al denominador de esta respuesta entre el número de resultados en todo el espacio muestral, que es de mil, obtenemos

$$P(\text{diámetro está bien} \mid \text{longitud demasiada larga}) = \frac{25/1000}{40/1000}$$

Ahora 40/1000 representa la probabilidad de que se satisface la condición de que la varilla es demasiado larga. Esto es,

$$P(\text{longitud demasiada larga}) = \frac{40}{1000}$$

La cantidad 25/1000 representa la probabilidad de que se satisfacen tanto la condición de que la varilla es demasiado larga y de que el diámetro está bien. Esto es,

$$P(\text{diámetro está bien y longitud demasiada larga}) = \frac{25}{1000}$$

Ahora se puede expresar la probabilidad condicional como

$$P(\text{diámetro está bien} \mid \text{longitud demasiada larga}) =$$

$$= \frac{P(\text{diámetro está bien y longitud demasiada larga})}{P(\text{longitud demasiada larga})}$$

Este razonamiento se puede ampliar para construir una definición de la probabilidad condicional que es válida para cualquier espacio muestral.

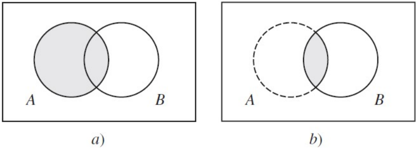
Definición

Sean A y B eventos con $P(B) \neq 0$. La probabilidad condicional de A dado B es

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

Prestemos atención a que las probabilidades del término de la derecha de la ecuación anterior hacen referencia a todo el espacio muestral, mientras que la probabilidad condicionada, término izquierdo de la expresión, tiene por espacio muestral solo el suceso B, o sea la condición.

La figura siguiente presenta diagramas de Venn para ilustrar la idea de la probabilidad condicional.



a) El diagrama representa la probabilidad incondicional $P(A)$. $P(A)$ se muestra al considerar el evento A en proporción con todo el espacio muestral, el cual se representa por el rectángulo.

b) El diagrama representa la probabilidad condicional $P(A|B)$. Puesto que se sabe que ocurre el evento B, ahora éste será el espacio muestral. Para que el evento A ocurra, el resultado debe estar en la intersección $A \cap B$. Por tanto, la probabilidad condicional $P(A|B)$ se muestra al considerar la intersección $A \cap B$ en proporción con todo el evento B.

Eventos independientes

Algunas veces el conocimiento de que un evento ha ocurrido no cambia la probabilidad de que ocurra otro.

En este caso las probabilidades condicional e incondicional son las mismas y se dice que los eventos son independientes.

A continuación se presenta un ejemplo.

Ejemplo

Si una varilla de aluminio se selecciona del espacio muestral que se presenta en la tabla

	Diámetro		
Longitud	Muy delgado	Está bien	Muy grueso
Demasiado corta	10	3	5
Está bien	38	900	4
Demasiado larga	2	25	13

Determinemos $P(\text{demasiado larga})$ y $P(\text{demasiado larga} \mid \text{muy delgada})$. ¿Estas probabilidades son diferentes?

Solución

$$P(\text{demasiado larga}) = \frac{40}{1000} = 0.040$$

$$P(\text{demasiado larga} \mid \text{muy delgada}) = \frac{P(\text{demasiado larga y muy delgada})}{P(\text{demasiado larga})}$$

$$P(\text{demasiado larga} \mid \text{muy delgada}) = \frac{2/1000}{50/1000} = \frac{2}{50} = 0.040$$

La probabilidad condicional y la probabilidad incondicional son las mismas.

La información de que la varilla es muy delgada no cambia la probabilidad de que la varilla es demasiado larga.

El ejemplo anterior muestra que el conocimiento de que un evento ocurre a veces no cambia la probabilidad de que ocurra otro evento.

En estos casos, se dice que los dos eventos son independientes.

El evento de que una varilla es demasiada larga y el evento de que una varilla es muy delgada son independientes.

A continuación se presenta una definición más precisa del término, tanto en palabras como en símbolos.

Definición

Dos eventos A y B son independientes si la probabilidad de cada uno es la misma ocurra o no el otro evento.

En símbolos:

si $P(A) \neq 0$ y $P(B) \neq 0$, entonces A y B son independientes si

$$P(B|A) = P(B) \text{ o, de manera equivalente, } P(A|B) = P(A)$$

Si A y B son independientes, entonces los siguientes pares de eventos son también independientes:

$$A \text{ y } B^c$$
$$A^c \text{ y } B$$
$$A^c \text{ y } B^c$$

La regla de la multiplicación

Algunas veces se conoce $P(A|B)$ y se desea encontrar $P(A \cap B)$.

Se puede obtener un resultado que sea útil para este propósito al multiplicar ambos lados de la ecuación $P(A|B) = \frac{P(A \cap B)}{P(B)}$ por $P(B)$.

Esto conduce a la regla de la multiplicación.

La regla de la multiplicación

Si A y B son dos eventos con $P(B) \neq 0$, entonces

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

Si A y B son dos eventos con $P(A) \neq 0$, entonces

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

Si $P(A) \neq 0$ y $P(B) \neq 0$ entonces las expresiones anteriores son válidas.

Cuando dos eventos son independientes, entonces

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

así la regla de la multiplicación se simplifica:

La regla de la multiplicación para eventos independientes

Si A y B son eventos independientes, entonces

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

Este resultado se puede ampliar para cualquier número de eventos. Si $A_1, A_2, \ldots, A_n$ son eventos independientes, entonces para cada colección $A_{j1}, A_{j2}, \ldots, A_{jm}$ de eventos

$$P(A_{j1} \cap A_{j2} \cap \cdots \cap A_{jm}) = P(A_{j1}) \cdot P(A_{j2}) \cdots P(A_{jm})$$

En particular:

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1) \cdot P(A_2) \cdots P(A_n)$$

Ejemplo

Un vehículo tiene dos motores: uno principal y otro auxiliar. El componente del motor falla sólo si fallan ambos motores. La probabilidad de que el motor principal falle es de 0.05 y la de que el motor auxiliar falle es de 0.10. Supongamos que los motores principal y auxiliar funcionan de manera independiente. ¿Cuál es la probabilidad de que el componente del motor falle?

Solución

La probabilidad de que el componente del motor falle es la probabilidad de que ambos motores fallen. Por tanto,

$$P(\text{componente del motor falla}) =$$

$$= P(\text{motor principal falla y motor auxiliar falla})$$

Solución (Continuación)

Puesto que los motores son independientes, se puede usar la ecuación $P(A \cap B) = P(A) \cdot P(B)$:

$$P(\text{motor principal falla y motor auxiliar falla}) =$$
$$= P(\text{motor principal falla}) \cdot P(\text{motor auxiliar falla})$$

En consecuencia:

$$P(\text{motor principal falla y motor auxiliar falla}) = 0.10 \cdot 0.05 = 0.005$$

$$P(\text{componente del motor falla}) = 0.005$$

Ejemplo

Un sistema contiene dos componentes, A y B. Ambos componentes deben funcionar para que el sistema trabaje. La probabilidad de que el componente A falle es de 0.08 y de que falle el componente B es de 0.05. Supongamos que los dos componentes funcionan de manera independiente, ¿cuál es la probabilidad de que el sistema funcione?

Solución

La probabilidad de que el sistema funcione es la probabilidad de que ambos componentes funcionen. Por tanto,

$$P(\text{funciona el sistema}) = P(\text{funciona } A \text{ y funciona } B)$$

Puesto que los componentes funcionan de manera independiente,

$$P(\text{funciona } A \text{ y funciona } B) = P(\text{funciona } A) \cdot P(\text{funciona } B)$$

Solución (Continuación)

Atendiendo a que

$$P(\text{funciona } A) = 1 - P(\text{falla } A)$$

$$P(\text{funciona } B) = 1 - P(\text{falla } B)$$

Entonces,

$$P(\text{funciona } A \text{ y funciona } B) = [1 - P(\text{falla } A)] \cdot [1 - P(\text{falla } B)]$$

$$P(\text{funciona } A \text{ y funciona } B) = [1 - 0.08] \cdot [1 - 0.05] = 0.874$$

Así,

$$P(\text{funciona el sistema}) = 0.874$$

La regla de la multiplicación para eventos independientes ($P(A \cap B) = P(A) \cdot P(B)$) indica cómo calcular las probabilidades cuando se sabe que los eventos son independientes, pero no son generalmente de mucha ayuda para determinar si dos eventos son realmente independientes.

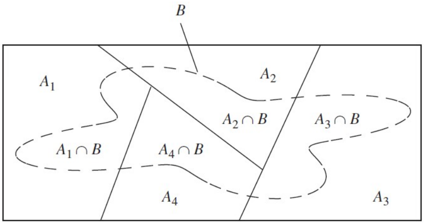
En la mayoría de los casos, la mejor manera de determinar si los eventos son independientes es comprendiendo el proceso que los produce. He aquí algunos ejemplos:

- Se tira dos veces un dado. Es razonable creer que el resultado de la segunda tirada no se vea afectado por el resultado de la primera. Por tanto, conocer el resultado de la primera tirada no ayuda a predecir el resultado de la segunda. Las dos tiradas son independientes.
- Cierta reacción química se realiza dos veces, utilizando equipos diferentes cada vez. Es razonable creer que el producto de una reacción no afectará la producción de la otra. En este caso los productos son independientes.

- Una reacción química se realiza dos veces sucesivamente, utilizando el mismo equipo. En este caso, no es sensato suponer que las producciones son independientes. Por ejemplo, una producción baja en la primera realización podría indicar que hay más residuos de lo normal. Lo anterior podría tender a hacer que la producción en la siguiente realización fuese más alta. Por tanto, conocer la primera producción puede ayudar a predecir la producción en la segunda realización. No podemos asegurar independencia.
- Los elementos en una muestra aleatoria simple se pueden tratar como independientes, a menos que la población sea finita y la muestra consista en más de un 5% de la población.

Ley de la probabilidad total

Exponemos la ley de la probabilidad total a partir del espacio muestral S que se presenta en la figura siguiente:



Este espacio muestral contiene los eventos A_1, A_2, A_3 y A_4. Éstos son mutuamente excluyentes, ya que ningún par de sucesos coinciden. Son también exhaustivos, ello significa que su unión abarca todo el espacio muestral.

$$S = A_1 \cup A_2 \cup A_3 \cup A_4$$

Dicho de otra forma, cada resultado en este espacio pertenece a uno y sólo uno de los eventos A_1, A_2, A_3, A_4.

El evento B es cualquier evento. En la figura anterior, cada uno de los eventos A_i que intersectan a B, forman los eventos $A_1 \cap B$, $A_2 \cap B$, $A_3 \cap B$ y $A_4 \cap B$.

Es obvio de la figura que los eventos $A_1 \cap B$, $A_2 \cap B$, $A_3 \cap B$ y $A_4 \cap B$ son mutuamente excluyentes y que abarcan a B.

Cada resultado en B pertenece a uno y sólo uno de los eventos $A_1 \cap B$, $A_2 \cap B$, $A_3 \cap B$ y $A_4 \cap B$. Por lo que

$$B = (A_1 \cap B) \cup (A_2 \cap B) \cup (A_3 \cap B) \cup (A_4 \cap B)$$

En términos de probabilidad obtenemos que

$$P(B) = P((A_1 \cap B) \cup (A_2 \cap B) \cup (A_3 \cap B) \cup (A_4 \cap B))$$

Como se trata de la unión de eventos mutuamente excluyentes llegamos a la expresión

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + P(A_3 \cap B) + P(A_4 \cap B)$$

Debido a que $P(A_i \cap B) = P(B|A_i) \cdot P(A_i)$ para cada A_i,

$$P(B) = P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2) + P(B|A_3) \cdot P(A_3) + P(B|A_4) \cdot P(A_4)$$

Esta expresión es un caso especial de la ley de la probabilidad total, restringida al caso donde hay cuatro eventos mutuamente excluyentes y exhaustivos.

La intuición detrás de la ley de la probabilidad total es muy simple:

- Los eventos A_1, A_2, A_3, A_4 parten al evento B en piezas.
- La probabilidad de B se encuentra sumando las probabilidades de todas las piezas.

Se podría dibujar nuevamente la figura previa para tener cualquier número de eventos A_i. Esto conduce al caso general de la ley de la probabilidad total.

Ley de la probabilidad total

Si $A_1, \ldots, A_n$ son eventos mutuamente excluyentes y exhaustivos y B es cualquier evento, entonces

$$P(B) = P(A_1 \cap B) + \cdots + P(A_n \cap B)$$

De manera equivalente, si $P(A_i) \neq 0$ para cada A_i,

$$P(B) = P(B|A_1) \cdot P(A_1) + \cdots + P(B|A_n) \cdot P(A_n)$$

Ejemplo

Clientes que compran cierta marca de automóvil pueden pedir un motor en cualquiera de sus tres tamaños. De todos los automóviles vendidos, el 45% tiene el motor más pequeño, el 35% tamaño mediano y el 20% más grande. Los automóviles en una prueba de emisiones en los primeros dos años de su compra fallan el 10% con el motor más pequeño, el 12% los de tamaño mediano y el 15% los de motor más grande.

¿Cuál es la probabilidad de que un automóvil elegido aleatoriamente pueda fallar en una prueba de emisiones dentro de los dos primeros años?

Solución

Sea B el evento de que un automóvil falle en una prueba de emisiones dentro de los dos primeros años. Sea A_1 el evento que es un automóvil con un motor pequeño, A_2 el evento que un automóvil tiene un motor mediano y A_3 el evento que un automóvil tiene un motor grande.

Solución (Continuación)

Entonces $P(A_1) = 0.45$, $P(A_2) = 0.35$ y $P(A_3) = 0.20$.

La probabilidad de que un automóvil falle una prueba, dado que tiene un motor pequeño, es de 0.10. Es decir,

$$P(B|A_1) = 0.10$$

De manera similar,

$$P(B|A_2) = 0.12$$

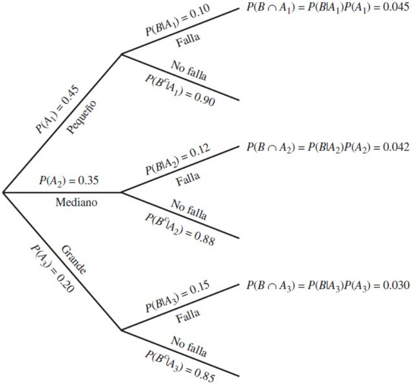
$$P(B|A_3) = 0.15$$

Por la ley de probabilidad total:

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

$$P(B) = 0.10 \cdot 0.45 + 0.12 \cdot 0.35 + 0.15 \cdot 0.20 = 0.117$$

A veces los problemas de probabilidad se resuelven con el uso de diagramas de árbol.



El diagrama de árbol anterior representa nuestro ejemplo.

Hay tres ramas principales en el árbol, que se corresponden con los tres tamaños de motor (A_1, A_2, A_3).

Las probabilidades de los tamaños de motor se colocan en sus ramas respectivas.

Al final de cada rama principal hay dos ramas secundarias, que representan los eventos de falla (B) y no falla(B^c).

Las probabilidades condicionales de falla y no falla, dado el tamaño del motor, se colocan en las ramas secundarias.

Al multiplicar a lo largo de cada una de las ramas que corresponden al evento $B =$ falla, se obtienen las probabilidades $P(B|A_i) \cdot P(A_i)$. Al sumar estas probabilidades se obtiene $P(B)$, como se quería.

Regla de Bayes

Si A y B son dos eventos, normalmente se cumple que $P(A|B) \neq P(B|A)$.

La regla de Bayes proporciona una fórmula que permite calcular una de las probabilidades condicionales si se conoce la otra.

Para ver cómo funciona, supongamos que se conoce $P(B|A)$ y que se desea calcular $P(A|B)$.

Se inicia con la definición de probabilidad condicional:

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

Ahora sustituimos la relación $P(A \cap B) = P(B|A) \cdot P(A)$ en nuestra expresión:

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

La ecuación anterior es esencialmente la regla de Bayes.

Cuando se escribe la regla de Bayes, la expresión $P(B)$ del denominador se reemplaza con una expresión más complicada obtenida por lo general de la ley de la probabilidad total.

De modo específico, puesto que los eventos A y A^c son mutuamente excluyentes y exhaustivos, la ley de la probabilidad total muestra que

$$P(B) = P(B|A) \cdot P(A) + P(B|A^c) \cdot P(A^c)$$

Al conjugar las 2 expresiones anteriores obtenemos la regla de Bayes.

La regla de Bayes

Sean A y B eventos con $P(A) \neq 0$, $P(A^c) \neq 0$ y $P(B) \neq 0$. Entonces

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

También se puede obtener una versión más general de la regla de Bayes al considerar una colección $A_1, \ldots, A_n$ de eventos mutuamente excluyentes y exhaustivos y utilizando la ley de la probabilidad total sustituyendo a $P(B)$.

La regla de Bayes

Sean $A_1, \ldots, A_n$ eventos mutuamente excluyentes y exhaustivos con $P(A_i) \neq 0$ para cada A_i. Sea B cualquier evento con $P(B) \neq 0$. Entonces

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

Ejemplo

Continuando con el ejemplo anterior, se elige aleatoriamente un registro de una prueba de emisiones con falla. ¿Cuál es la probabilidad de que éste sea un automóvil con un motor pequeño?

Solución

Sea B el evento que un automóvil falla en una prueba de emisiones.

Sea A_1 el evento que un automóvil tiene un motor pequeño, A_2 el evento de que lo tiene mediano y A_3 que su motor es grande. Eventos mutuamente excluyentes y exhaustivos.

Se desea encontrar $P(A_1|B)$.

Recordemos las probabilidades del ejemplo:

$$P(A_1) = 0.45,\ P(A_2) = 0.35,\ P(A_3) = 0.20$$

$$P(B|A_1) = 0.10,\ P(B|A_2) = 0.12,\ P(B|A_3) = 0.15$$

Solución (Continuación)

Por la regla de Bayes,

$$P(A_1|B) = \frac{P(B|A_1) \cdot P(A_1)}{P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2) + P(B|A_3) \cdot P(A_3)}$$

$$P(A_1|B) = \frac{0.10 \cdot 0.45}{0.10 \cdot 0.45 + 0.12 \cdot 0.35 + 0.15 \cdot 0.20} = 0.385$$

Esquema

Cuando se calculan probabilidades, algunas veces se necesita determinar el número de resultados en un espacio muestral.

En esta sección se describirán diversos métodos con ese propósito.

La regla básica, que se conoce como principio fundamental de conteo, se presenta por medio del siguiente ejemplo.

Ejemplo

Cierto tipo de automóvil se encuentra disponible en tres colores: rojo, azul o verde, y puede tener un motor grande o pequeño. ¿De cuántas maneras puede un comprador elegir un automóvil?

Solución

Hay tres opciones de color y dos opciones de motor. Una lista completa de las opciones se muestra en la siguiente tabla de 3 columnas y 2 filas, es decir una tabla de **3x2**. El número total de opciones es $3 \cdot 2 = 6$.

	Rojo	Azul	Verde
Grande	Rojo, Grande	Azul, Grande	Verde, Grande
Pequeño	Rojo, Pequeño	Azul, Pequeño	Verde, Pequeño

Al generalizar el ejemplo, si hay n_1 elecciones de color y n_2 elecciones de motor, una lista completa de elecciones se puede escribir como una tabla $n_1 \mathbf{x} n_2$, por lo que el número total de elecciones es $n_1 \cdot n_2$.

Resumen

Si una operación se puede realizar en n_1 maneras y si para cada una de esas maneras se puede realizar una segunda operación en n_2 maneras, entonces el número total de maneras en que se realizan las dos operaciones es $n_1 \cdot n_2$.

Este razonamiento del principio fundamental del conteo de estados se puede ampliar para cualquier número de operaciones.

El principio fundamental del conteo

Supongamos que se pueden realizar k operaciones. Si hay n_1 maneras de realizar la primera operación y si para cada una de esas maneras hay n_2 maneras de realizar la segunda operación y si para cada una de esas elecciones de esas maneras de realizar las dos primeras operaciones hay n_3 maneras de realizar la tercera operación y así sucesivamente, entonces el número total de maneras de realizar la secuencia de las k operaciones es $n_1 \cdot n_2 \cdots n_k$.

Ejemplo

Cuando se hace un pedido de cierto tipo de computadora, hay tres elecciones de disco duro, cuatro de la cantidad de memoria, dos de la tarjeta de video y tres de monitor. ¿En cuántas maneras se puede solicitar una computadora?

Solución

El número total es $3 \cdot 4 \cdot 2 \cdot 3 = 72$.

Permutaciones

Una permutación constituye un ordenamiento de un conjunto de elementos.

Por ejemplo, hay seis permutaciones de las letras A, B, C:

ABC
ACB
BAC
BCA
CAB
CBA

Con solamente tres elementos, es fácil determinar el número de permutaciones, sólo con hacer una lista de todas ellas.

Pero con un gran número de elementos esto último no sería factible.

El principio fundamental del conteo se puede usar para determinar el número de permutaciones de cualquier conjunto de elementos.

Por ejemplo, se puede determinar el número de permutaciones de un conjunto de tres elementos de la siguiente manera:

- Hay tres elecciones para colocar el primer elemento.
- Después de que se hace la elección, hay dos elecciones restantes para el elemento del segundo lugar.
- Entonces queda una elección para el elemento del último lugar.
- Por tanto, el número total de maneras de ordenar tres objetos es

$$3 \cdot 2 \cdot 1 = 6$$

Este razonamiento se puede generalizar.

El número de permutaciones de un conjunto de n elementos es

$$n \cdot (n-1) \cdot (n-2) \cdots 3 \cdot 2 \cdot 1$$

Éste es el producto de los enteros del 1 al n.

Este producto se puede escribir con el símbolo $n!$, que se lee "n factorial".

Definición

Para cualquier entero positivo n,

$$n! = n \cdot (n-1) \cdot (n-2) \cdots 3 \cdot 2 \cdot 1$$

También se define $0! \equiv 1$.

Permutaciones

El número de permutaciones de n objetos es $n!$

Ejemplo

Cinco personas están en la hilera de un cine. ¿En cuántas maneras diferentes se pueden ordenar?

Solución

El número de permutaciones de un conjunto de cinco personas es

$$5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$$

.

A veces se está interesado en contar el número de permutaciones de los subconjuntos de cierto tamaño elegidos de un conjunto más grande. A estas permutaciones alguna bibliografía las denomina variaciones. Lo vemos en el ejemplo siguiente.

Ejemplo

Cinco salvavidas están disponibles para la guardia de un sábado por la tarde. Hay tres estaciones salvavidas. ¿De cuántas maneras se pueden elegir y organizar los salvavidas entre las estaciones?

Solución

Se usa el principio fundamental del conteo.

Hay cinco maneras de elegir a un salvavidas para que ocupe la primera estación, luego cuatro de elegir a un salvavidas para que ocupe la segunda estación y por último tres para elegir un salvavidas que ocupe la tercera estación.

El número total de permutaciones de los tres salvavidas elegidos entre los cinco es

$$5 \cdot 4 \cdot 3 = 60$$

El razonamiento usado para resolver el ejemplo anterior se puede generalizar.

El número de permutaciones de k objetos elegidos de un grupo de n objetos es

$$n \cdot (n-1) \cdots (n-k+1)$$

Esta expresión se puede simplificar utilizando la notación factorial:

$$n \cdot (n-1) \cdots (n-k+1) = \frac{n \cdot (n-1) \cdots (n-k+1) \cdot (n-k) \cdot (n-k-1) \cdots 2 \cdot 1}{(n-k) \cdot (n-k-1) \cdots 2 \cdot 1}$$

$$n \cdot (n-1) \cdots (n-k+1) = \frac{n!}{(n-k)!}$$

Variaciones

El número de permutaciones de k objetos elegidos de un grupo de n elementos es

$$\frac{n!}{(n-k)!}$$

Combinaciones

En algunos casos, cuando se elige un conjunto de elementos de un conjunto más grande, no se tiene en cuenta el orden de los elementos elegidos; sólo se consideran los elementos que se eligen.

Por ejemplo, puede que no importe qué salvavidas ocupe cada estación; puede que sólo sea importante la elección de tres salvavidas.

A cada grupo distinto de elementos que se puede seleccionar, sin importar el orden, se le llama combinación.

A continuación se mostrará cómo determinar el número de combinaciones de k elementos elegidos de un conjunto de n objetos.

Se mostrará el razonamiento con el resultado del ejemplo anterior.

En ese ejemplo se mostró que hay 60 permutaciones de tres elementos elegidos entre cinco.

Al denotar a los elementos por A, B, C, D, E, a continuación se presenta una lista de las 60 permutaciones.

$$ABC, ABD, ABE, ACD, ACE, ADE, BCD, BCE, BDE, CDE$$

$$ACB, ADB, AEB, ADC, AEC, AED, BDC, BEC, BED, CED$$

$$BAC, BAD, BAE, CAD, CAE, DAE, CBD, CBE, DBE, DCE$$

$$BCA, BDA, BEA, CDA, CEA, DEA, CDB, CEB, DEB, DEC$$

$$CAB, DAB, EAB, DAC, EAC, EAD, DBC, EBC, EBD, ECD$$

$$CBA, DBA, EBA, DCA, ECA, EDA, DCB, ECB, EDB, EDC$$

Las 60 permutaciones anteriores están ordenadas en diez columnas de seis permutaciones cada una.

Dentro de cada columna, los tres elementos son los mismos y la columna contiene las seis permutaciones diferentes de esos tres elementos.

Por tanto, cada columna representa una combinación distinta de tres elementos elegidos entre cinco y hay diez combinaciones de ese tipo.

En consecuencia, la lista muestra que el número de combinaciones de tres elementos elegidos entre cinco se puede encontrar al dividir el número de permutaciones de los tres elementos elegidos, o sea $5!/(5-3)!$, por el número de permutaciones de los tres elementos, que es $3!$.

En resumen, el número de combinaciones de los tres elementos elegidos es $\frac{5!}{3!(5-3)!}$.

Con frecuencia el número de combinaciones de k elementos elegidos de n se denota por el símbolo $\binom{n}{k}$.

El razonamiento utilizado para deducir el número de combinaciones de los tres elementos elegidos se puede generalizar para deducir una expresión para $\binom{n}{k}$.

Combinaciones

El número de combinaciones de k elementos elegidos de un grupo de n elementos es

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Elegir una combinación de k elementos de un conjunto de n divide a los n elementos en dos subconjuntos: k que fueron elegidos y $n-k$ que no fueron elegidos.

Ejemplo

A cierto evento asisten 30 personas y se elegirá aleatoriamente a cinco para recibir premios. Estos últimos son iguales, así que el orden en que se elige a las personas no es importante. ¿Cuántos grupos diferentes de cinco personas se puede elegir?

Solución

En virtud de que el orden de las cinco personas elegidas no es importante, se tiene que calcular el número de combinaciones de cinco elegidas entre 30. Esto es

$$\binom{30}{5} = \frac{30!}{5!25!} = \frac{30 \cdot 29 \cdot 28 \cdot 27 \cdot 26}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 142506$$