TEMA 3 ELEMENTOS DE PROBABILIDAD

Pablo Buenestado

Curso 2020-2021 Otoño

Departamento de Matemáticas (UPC)

Índice

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(diámetro está bien | longitud demasiada larga). ¿Ésta es la misma que la probabilidad incondicional P(diámetro está bien)?

Solución

La probabilidad condicional
P(diámetro está bien | 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.

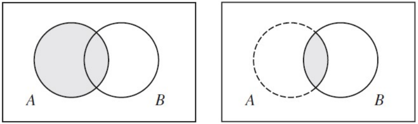
Definición

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) *b*)

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.

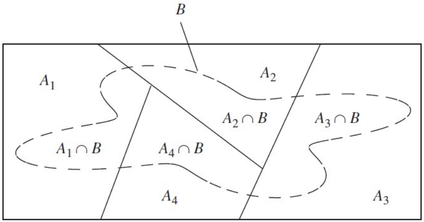
En la mayoría de los casos, la mejor manera de determinar si los eventos son independientes es comprendiendo el proceso que los produce. He aquí algunos ejemplos:

- 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,

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

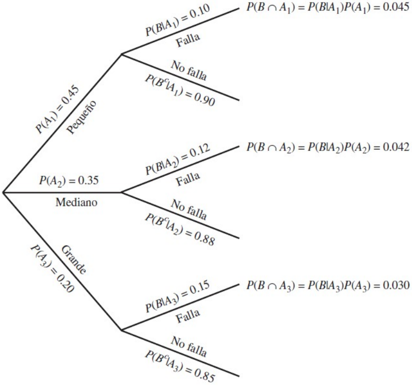
$$P(B|A_3) = 0.15$$

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,

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

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