

Clase 3: Probabilidad Condicional y Eventos Independientes.

Harvard Statistics 110: Probability

Resumen

En esta ocasión estudiaremos la probabilidad condicional o aquella que calculamos asumiendo que ya manejamos algo de información. También veremos su vínculo con el concepto de independencia de eventos y resolveremos algunos ejercicios.

1. Probabilidad Condicional.

Cuando evaluamos la posibilidad de que ocurra un hecho, suele ser más habitual hacerlo manejando algo de información. Esta puede ser desde una genuina creencia hasta evidencia rigurosa. Para incluir a cualquiera de ellas en nuestro cálculo matemático, usamos la fórmula de la llamada probabilidad condicional.

1.1. Definición de la probabilidad condicional.

Considere el siguiente caso.

Suponga que estamos jugando con un grupo de personas a lanzar dos dados al mismo tiempo mientras tenemos los ojos vendados. Si, aún estando ciegos, decimos correctamente los números obtenidos, ganamos.

En nuestro turno y luego de tirar los dados, apostamos mentalmente a que salieron los valores (6, 2). Ahora que tenemos un conocimiento básico de teoría de conjuntos, establecemos que el evento $A \subseteq S$ (S = espacio muestral) es:

$$A = \{\text{Obtener el par } (6, 2)\}$$

Para tener una idea de nuestras posibilidades de ganar, aprovechamos de calcular la proba-

bilidad de que ocurra A como:

$$P(A) = \frac{1}{36}$$

ya que S es finito y asumimos que los dados no están cargados, de manera que todos los resultados son equiprobables.

Como $P(A) \approx 0.03$, optamos por tomarnos un poco más de tiempo para pensar mejor la apuesta. Sin embargo, una de las personas lo nota y decide darnos una pista: “La suma de los dos números que obtuviste es par”. Puesto que esta afirmación agrupa a algunos resultados del experimento, lo expresamos como un evento $B \subseteq S$.

$$B = \{\text{La suma de los dos números obtenidos es par}\}$$

La pista nos ayuda a sacar dos conclusiones. La primera es que $(6, 2) \in B$. Además, dado a que A es un singleton conformado solo por ese elemento, entonces $A \subset B$. En segundo lugar, esta información permite reducir el espacio muestral a solo B (i.e, se descartan todos los resultados de B^c), donde $|B| = 18$.

Así, calculamos la probabilidad de A dado que tenemos conocimiento de B , $P(A | B)$, como:

$$P(A | B) = \frac{|A|}{|B|} = \frac{1}{18}$$

$P(A | B)$ recibe el nombre de **probabilidad condicional** y se lee como “la probabilidad de A dado B ”. Si bien la que calculamos nos da el resultado correcto, se restringe solo a ese caso. No nos sirve si $|A| > |B|$, ya que incluiríamos los resultados en A que no están en B . Una mejor opción es considerar, más bien, el conjunto $A \cap B$.

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

Si lo usamos en el ejemplo, el resultado sigue siendo el mismo ya que $A \cap B = \{(6, 2)\}$.

$$P(A | B) = \frac{|A \cap B|}{|B|} = \frac{1}{18}$$

Ahora, esta ecuación se restringe solo al contexto de la definición clásica de las probabilidades. No obstante, en 1933, Andrey Kolmogorov en *Foundations of the Theory of Probability* estableció la siguiente generalización de $P(A | B)$ que es usada hasta el día de hoy (1956: 6):

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \quad (\text{con } P(B) > 0)$$

Ejercicio 1. Antes de salir de su hogar nota que no encuentra sus llaves. Suponga que tiene un 80 % de certeza de que están en uno de los dos bolsillos de una chaqueta, con un 40 % de posibilidad de que esté en el izquierdo y un 40 % en el derecho. ¿Cuál es la probabilidad de que estén en el derecho si, al buscarlas, no las encuentra en el izquierdo?

Solución. Definamos los eventos:

$$I = \{\text{Las llaves están en el bolsillo izquierdo}\}$$

$$D = \{\text{Las llaves están en el bolsillo derecho}\}$$

Así, las probabilidades dadas en el ejercicio son:

$$P(I) = P(D) = 0.4 \quad \text{y} \quad P(I \cup D) = 0.8$$

y nos preguntan $P(D \mid I^c)$. Para calcular esta probabilidad condicional necesitamos conocer tanto a $P(I^c)$ como a $P(D \cap I^c)$. La primera corresponde a:

$$P(I^c) = 1 - P(I) = 0.6$$

En cuanto a $P(D \cap I^c)$, observemos que las llaves solo pueden estar en uno de los dos bolsillos de la chaqueta, lo que implica que I y D son mutuamente excluyentes¹. Por lo tanto,

$$P(D \cap I^c) = P(D) = 0.4$$

De este modo,

$$P(D \mid I^c) = \frac{P(D \cap I^c)}{P(I^c)} = \frac{2}{3} \approx 0.666$$

Ejercicio 2. En la televisión darán los seis números ganadores de una lotería y, para participar, usted compró un boleto con los valores 1, 14, 15, 20, 23 y 27. Estos serán elegidos al azar y sin reemplazo desde una tómbola con treinta bolas etiquetadas del 1 al 30. Para ganar, los suyos deben coincidir con los que se mencionen independiente del orden de aparición.

Suponga que, al empezar a escuchar los números, se apaga el televisor, pero después supo que uno de los mencionados fue el 15. Calcule la probabilidad de que su boleto sea el ganador.

¹Incluso lo podemos ver en $P(I \cup D) = P(I) + P(D) - P(I \cap D) = 0.4 + 0.4 - 0 = 0.8$

Solución. Comencemos estableciendo los eventos:

$A = \{\text{El boleto con los números 1, 14, 15, 20, 23 y 27 es el ganador}\}$

$B = \{\text{El boleto ganador tiene el número 15}\}$

En ese sentido, buscamos calcular $P(A \mid B)$ y, para ello, debemos encontrar los valores de $P(A \cap B)$ y $P(B)$. Dado que los resultados de la lotería son equiprobables y a que el espacio muestral es finito, calcularemos ambas probabilidades con la definición clásica.

A partir de lo señalado en el ejercicio, el espacio muestral S consiste de todas las combinaciones de seis de los treinta números elegidos al azar y sin repetición de la tómbola. Por lo tanto, su tamaño puede calcularse como:

$$|S| = \binom{30}{6} = \frac{30!}{6! \cdot 24!}$$

Luego, veamos que $A \cap B = A$ porque el número 15 está en nuestro boleto. Esto implica que $|A \cap B| = 1$ y, por tanto, que:

$$P(A \cap B) = \frac{|A \cap B|}{|S|} = \frac{1}{\binom{30}{6}} = \frac{6! \cdot 24!}{30!}$$

Por otra parte, el evento B consiste de todos los boletos donde ya salió el número 15, independiente de su posición. Como la selección de las bolas es sin reemplazo, entonces:

$$|B| = \binom{29}{5} = \frac{29!}{5! \cdot 24!}$$

De este modo,

$$P(B) = \frac{|B|}{|S|} = \frac{29!}{5! \cdot 24!} \cdot \frac{6! \cdot 24!}{30!} = \frac{1}{5}$$

Lo que conlleva a que:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{(6! \cdot 24!)/30!}{1/5} \approx 8.421 \cdot 10^{-6}$$

Es decir, hay alrededor de un 0.0008421 % de posibilidad de que nuestro boleto sea el ganador solo por saber que se mencionó el número 15.

Una ventaja de la definición de Kolmogorov de la probabilidad condicional, es que permite calcularla en relación al espacio muestral sin importar si lo restringimos. Esto se puede ver,

por ejemplo, al aplicar la definición clásica de las probabilidades en ella.

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

Otra ventaja de la definición de $P(A \mid B)$, es que de ella se puede desprender una fórmula para calcular $P(A \cap B)$ al multiplicarla por $P(B)$.

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

Del mismo modo, si $P(A) > 0$, $P(B \cap A)$ se puede obtener mediante $P(B \mid A) = \frac{P(B \cap A)}{P(A)}$.

$$P(B \cap A) = P(A) \cdot P(B \mid A) \iff P(A) > 0$$

Puesto que la intersección es una operación conmutativa, entonces $A \cap B = B \cap A$ y, por consiguiente, $P(A \cap B) = P(B \cap A)$. Esto implica que:

$$P(A \cap B) = P(A) \cdot P(B \mid A) \quad (\text{con } P(A) > 0)$$

Por lo tanto, si $P(A) > 0$ y $P(B) > 0$ con A y B siendo dos eventos, entonces:

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

Para n eventos A_i , con $i = 1, \dots, n$, la intersección de todos ellos nos lleva a la siguiente igualdad que se deriva de la de $P(A \cap B)$, conocida como el **teorema general del producto**²:

$$P\left(\bigcap_{i=1}^n A_i\right) = P(A_1) \cdot P(A_2 \mid A_1) \cdot P(A_3 \mid A_1 \cap A_2) \cdot \dots \cdot P(A_n \mid A_1 \cap A_2 \cap \dots \cap A_{n-1})$$

La definición de $P(A \mid B)$ es una función de probabilidad porque cumple con los axiomas de Kolmogorov. Para $A, B \subseteq S$ y $P(B) > 0$,

1. $P(A \mid B) \geq 0$
2. $P(S \mid B) = 1$
3. Para $\{A_k\}_{k=1}^\infty \subseteq S$,

$$P\left(\bigcup_{k=1}^\infty A_k \mid B\right) = \sum_{k=1}^\infty P(A_k \mid B) \iff A_i \cap A_j = \emptyset \text{ cuando } i \neq j$$

²También llamada “regla general del producto” o “regla de la cadena” de las probabilidades.

El primer teorema se puede demostrar intuitivamente. La definición de $P(A \mid B)$ se aplica solo si $P(B) > 0$. Por otra parte, $0 \leq P(A \cap B) \leq 1$. Debido a estas razones, $P(A \mid B)$ siempre será mayor o igual a cero.

En cuanto al segundo teorema, puede ser demostrado desde su probabilidad condicional.

$$P(S \mid B) = \frac{P(S \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$$

Para el tercer teorema, considere los eventos $\{A_k\}_{k=1}^{\infty} \subseteq S$ con $A_i \cap A_j = \emptyset$ para todo $i \neq j$ y que $P(B) > 0$ donde $B \subseteq S$. Entonces,

$$P\left(\bigcup_{k=1}^{\infty} A_k \mid B\right) = \frac{P([A_1 \cup A_2 \cup \dots] \cap B)}{P(B)}$$

En el numerador del lado derecho se puede distribuir la intersección sobre las uniones³.

$$P\left(\bigcup_{k=1}^{\infty} A_k \mid B\right) = \frac{P([A_1 \cap B] \cup [A_2 \cap B] \cup \dots)}{P(B)} = \frac{P(\bigcup_{k=1}^{\infty} [A_k \cap B])}{P(B)}$$

Como todos los $[A_k \cap B]$ son disjuntos entre sí, por el tercer axioma de Kolmogorov se obtiene:

$$P\left(\bigcup_{k=1}^{\infty} A_k \mid B\right) = \frac{\sum_{k=1}^{\infty} P(A_k \cap B)}{P(B)}$$

Así, puesto que $P(A_k \cap B) = P(A_k \mid B) \cdot P(B)$ para $k = 1, 2, \dots$, se concluye que:

$$P\left(\bigcup_{k=1}^{\infty} A_k \mid B\right) = \frac{\sum_{k=1}^{\infty} P(A_k \mid B) \cdot P(B)}{P(B)} = \frac{P(B) \cdot \sum_{k=1}^{\infty} P(A_k \mid B)}{P(B)} = \sum_{k=1}^{\infty} P(A_k \mid B)$$

Otras igualdades fácilmente demostrables de la definición de la probabilidad condicional son:

$$P(A \mid A) = 1 \quad \text{y} \quad P(\emptyset \mid B) = 0$$

³Es una de las propiedades distributivas vista en la Clase 1: Para tres conjuntos A, B y C ,

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

1.2. Interpretaciones.

La probabilidad condicional $P(A \mid B)$ puede ser entendida como una función que asigna una probabilidad a la ocurrencia de A luego de saber que sucedió B .

En el contexto donde el espacio muestral S es finito y sus elementos son equiprobables, $P(A \mid B)$ puede ser interpretado como el valor de la posibilidad de que una fracción de los resultados de B también sean parte de A . Al saber que B ocurrió, este evento pasa a tomar el papel de S y, como consecuencia de aquello, $P(A)$ se **actualiza** a $P(A \mid B)$.

Otra interpretación de $P(A \mid B)$ que es usada en estadística, es aquella que denota la probabilidad de que la hipótesis que se está investigando sea verdadera luego de haber observado una cantidad determinada de evidencia, donde:

- $A \implies$ La hipótesis.
- $B \implies$ La evidencia (por ejemplo, datos).

Es decir, en esta interpretación de $P(A \mid B)$ se asume que la evidencia es verdadera (porque se observó) y, con ella, se intenta ver qué tan posible es que la hipótesis que se está poniendo a prueba también lo sea. De ésta surgen dos conceptos fundamentales:

- **Probabilidad previa** (*prior probability*): La probabilidad de que la hipótesis sea verdadera antes de recoger evidencia. Es decir, $P(A)$.
- **Probabilidad posterior** (*posterior probability*): La probabilidad de que la hipótesis sea verdadera después de haber obtenido evidencia. Es decir, $P(A \mid B)$.

Por lo tanto, la probabilidad posterior es una **actualización** de la probabilidad previa luego de manejar evidencia relacionada a la hipótesis que se está investigando.⁴

Un **error lógico** aparentemente común que cometemos las personas es plantear a $P(A \mid B)$ como $P(B \mid A)$. Es decir, se suele caer en la confusión de que:

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

Sin embargo, sí es posible encontrar una relación entre ambas probabilidades condicionales mediante un importante teorema que estudiaremos en la siguiente sección.

⁴En ocasiones también se señala que representa una actualización de nuestras creencias previas.

1.3. Teorema de Bayes.

En esta sección se abordará una consecuencia de la definición de probabilidad condicional, conocida como el teorema de Bayes.

1.3.1. Derivando el teorema.

Sean A y B dos eventos. En la sección 1.1 vimos que si $P(A) > 0$ y $P(B) > 0$, entonces:

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

Al despejar a $P(A | B)$, se obtiene lo siguiente:

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

Esta forma de expresar a $P(A | B)$ se conoce comúnmente como el **teorema de Bayes**, en referencia al matemático y reverendo inglés Thomas Bayes⁵.

El teorema de Bayes permite establecer un vínculo entre $P(A | B)$ y $P(B | A)$ simplemente al despejar a una de las dos.

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

La relación que se ve arriba es una ventaja porque, en ocasiones, una de las probabilidades condicionales es más fácil de obtener que la otra. Por lo tanto, se puede optar por calcular a la primera para después resolver la de nuestro interés.

En estadística, con A siendo una muestra de datos de una población y B un parámetro de esta última, $P(B | A)$ recibe el nombre de **función de verosimilitud** (*likelihood function*). Sin embargo, se debe tener en cuenta que no es una función de probabilidad. El matemático Ronald Fisher (quien planteó este concepto) discutió aquello de manera detallada⁶.

La igualdad del teorema de Bayes también puede ser expresado como una cuota de probabilidades o mediante una partición del espacio muestral. Ambas son estudiadas a continuación.

⁵A dos años de su muerte en 1761, Richard Price (matemático y amigo suyo) lo ayudó a publicar un trabajo titulado *An Essay towards solving a Problem in the Doctrine of Chances* en la revista *Philosophical Transactions of the Royal Society of London* (1763). Si bien no era el tema principal de su ensayo, el reverendo planteó dentro de sus proposiciones lo que simbólicamente conocemos como la probabilidad condicional. Por lo tanto, se puede decir que el “teorema de Bayes” es, más bien, una derivación de la primera.

⁶Ver, por ejemplo, sus artículos *On the mathematical foundations of theoretical statistics* (1922) donde, además, define el método de estimación de máxima verosimilitud, y en *Inverse Probability* (1930).

1.3.2. Expresada como cuota.

Una cuota (*odds*) es la razón entre que ocurra un evento y que el mismo no suceda. Es decir,

$$q(A) = \frac{P(A)}{P(A^c)}$$

Desde su definición es claro que $q(A)$ **no mide la probabilidad de que ocurra A** , sino más bien compara a esta última con la de que falle en suceder. Por ejemplo, suponga que la probabilidad de que, en una carrera de autos, gane el único de color rojo (el evento A) es:

$$P(A) = \frac{2}{3}$$

En este caso, su cuota $q(A)$ será:

$$q(A) = \frac{\frac{2}{3}}{1 - \frac{2}{3}} = \frac{2}{1}$$

Por lo tanto, la cuota de que gane la carrera el auto rojo es de 2 a 1 en su favor.

Al despejar a $P(A)$ en $q(A) = P(A)/P(A^c)$, se obtiene que:

$$P(A) = q(A) \cdot P(A^c) = q(A) \cdot [1 - P(A)] = q(A) - q(A)P(A)$$

Veamos que $P(A) = q(A) - q(A)P(A)$ es lo mismo que:

$$q(A) = P(A) + q(A)P(A) = P(A) \cdot [1 + q(A)]$$

La igualdad de arriba nos lleva a una fórmula para convertir una cuota $q(A)$ a $P(A)$.

$$P(A) = \frac{q(A)}{1 + q(A)}$$

Ahora digamos que la probabilidad de que suceda un evento A está condicionado a que, anteriormente, ocurrió B . La cuota q , entonces, puede expresarse como:

$$q(A \mid B) = \frac{P(A \mid B)}{P(A^c \mid B)}$$

Centrémonos solo en el lado derecho de esta igualdad. Si asumimos que $P(B) > 0$, es posible aplicar el teorema de Bayes tanto en el numerador como en el denominador de esta fracción.

Con un poco de álgebra elemental se obtiene lo siguiente:

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

Esta igualdad también puede ser expresada como:

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

La igualdad de arriba es el teorema de Bayes expresado como una cuota de probabilidades. Su lado izquierdo recibe el nombre “cuota posterior”, mientras que $P(A)/P(A^c)$ es la “cuota previa” y el factor que multiplica a esta última, $P(B | A)/P(B | A^c)$, se conoce como la “cuota de verosimilitud”.

Veamos, también, que al aplicar la fórmula para convertir una cuota a la probabilidad del evento en el teorema de Bayes que acabamos de ver, se obtiene que:

$$P(A | B) = \frac{q(A | B)}{1 + q(A | B)} = \frac{P(A | B)}{P(A^c | B) + P(A | B)} = \frac{P(A) \cdot P(B | A)}{P(A^c) \cdot P(B | A^c) + P(A) \cdot P(B | A)}$$

Si volvemos a la fórmula inicial del teorema de Bayes, de la igualdad de arriba se puede desprender que $P(B) = P(A^c) \cdot P(B | A^c) + P(A) \cdot P(B | A)$. En la siguiente sección explotaremos este hecho mediante la ley de la probabilidad total, que nos lleva a la forma más usada de este teorema.

1.3.3. La ley de la probabilidad total y el teorema de Bayes.

La **ley de la probabilidad total** señala que si el espacio muestral S es una partición de los eventos $\{A_k\}_{k=1}^n$, con cada $P(A_k) > 0$, entonces la probabilidad de que suceda otro evento B cualquiera es:

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

Demostración. El evento B puede ser expresado como:

$$B = B \cap S = B \cap \left(\bigcup_{k=1}^n A_k \right)$$

Por la propiedad distributiva, el evento B también corresponde a:

$$B = \bigcup_{k=1}^n (B \cap A_k)$$

En ese sentido, la probabilidad de que ocurra B es:

$$P(B) = P\left(\bigcup_{k=1}^n (B \cap A_k)\right)$$

Para $i \neq j$, $(B \cap A_i) \cap (B \cap A_j) = \emptyset$ ya que los eventos A_k son mutuamente excluyentes. Así, por el axioma 3 de las probabilidades se tiene que:

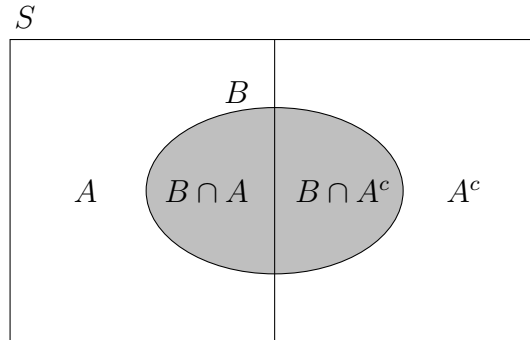
$$P(B) = \sum_{k=1}^n P(B \cap A_k)$$

Como las $P(A_k) > 0$, entonces $P(B \cap A_k) = P(A_k) \cdot P(B \mid A_k)$. Al reemplazar esta igualdad en el lado derecho de la de arriba, se concluye que:

$$P(B) = \sum_{k=1}^n P(A_k) \cdot P(B \mid A_k) \quad (\text{Q. E. D})$$

La intuición detrás de la ley de la probabilidad total es que la probabilidad de un evento puede ser calculada como **la suma de sus partes**, que corresponden a las intersecciones entre este conjunto y los que constituyen al espacio muestral como una partición.

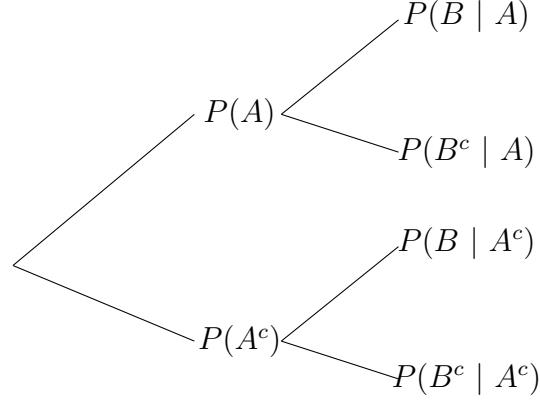
En el siguiente diagrama de Venn se observa la ley de probabilidad total a partir de los eventos A , A^c y B de un espacio muestral S formado como una partición entre A y A^c .



Del diagrama se deduce que $B = (B \cap A) \cup (B \cap A^c)$. Como $(B \cap A) \cap (B \cap A^c) = \emptyset$, entonces:

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

Un método visual para aplicar la ley de la probabilidad total es el diagrama de árbol. Como se observa a continuación, primero se dibujan las probabilidades de los eventos de la partición y luego las condicionales de B y B^c .



Para calcular los productos seguimos las ramas del diagrama de árbol. Por ejemplo, a partir de $P(A)$ podemos obtener $P(A) \cdot P(B | A)$ y $P(A) \cdot P(B^c | A)$, donde $P(B^c | A) = 1 - P(B | A)$.

Si estamos usando el diagrama de árbol y la ley de probabilidad total para calcular $P(B)$, entonces nos centramos en las probabilidades de los productos que contengan al evento B , que en este caso son $P(A) \cdot P(B | A)$ y $P(A^c) \cdot P(B | A^c)$. Finalmente, las sumamos.

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

Ahora considere un evento A_j de los $\{A_k\}_{k=1}^n$ que forman a S como una partición, con $0 \leq j \leq n$. La probabilidad $P(A_j | B)$ se puede calcular mediante el teorema de Bayes como:

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

Dada la característica de los A_k y de que las $P(A_k) > 0$, se puede aplicar la ley de la probabilidad total para establecer que $P(B) = \sum_{k=1}^n P(A_k) \cdot P(B | A_k)$. Al reemplazar esta suma en el denominador del lado derecho del teorema de Bayes, se obtiene que:

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

Esta forma del teorema de Bayes permite calcular a $P(A_j | B)$ mediante casos más sencillos. Sin embargo, dicha facilidad dependerá de cómo se decida “cortar” el espacio muestral. También es posible terminar con n probabilidades más complejas de resolver.

Ejercicio 3. Suponga que existe la posibilidad de hacerse gratuitamente un examen para

cierta enfermedad y usted decide realizárselo porque es un proceso rápido y no deja secuelas.

Al llegar al centro donde realizan el examen, le entregan un folleto en el que se informa su nivel de precisión: Si padece la enfermedad, hay un 90 % de probabilidad de que resulte positivo. Esta se reduce a un 10 % si no la ha contraído.

Los datos indican que solo hay una posibilidad de 10000 de que padezca la enfermedad. Calcule la probabilidad de que la haya adquirido si, días después de haberse realizado el examen, se entera que resultó positivo.

Solución. Puesto que el experimento de este ejercicio se centra en el examen médico, definamos a los eventos como:

$$E = \{\text{Tener la enfermedad}\}, \quad E^c = \{\text{No tener la enfermedad}\}, \quad T = \{\text{Examen positivo}\}$$

Ahora consideremos la información dada en el ejercicio.

$$P(T | E) = 0.9, \quad P(T | E^c) = 0.1, \quad P(E) = \frac{1}{10000} = 0.0001$$

Lo que se pide calcular es la probabilidad de tener la enfermedad luego de saber que el examen salió positivo. Es decir, $P(E | T)$, donde:

$$P(E | T) = \frac{P(E) \cdot P(T | E)}{P(T)}$$

El espacio muestral de este ejercicio podemos definirlo como una partición entre E (tener la enfermedad) y E^c (no tener la enfermedad). Como el evento T también es un subconjunto de este conjunto, se puede establecer que:

$$T = (T \cap E) \cup (T \cap E^c)$$

Así, por la ley de la probabilidad total, $P(T) = P(E) \cdot P(T | E) + P(E^c) \cdot P(T | E^c)$. Como $P(E^c) = 1 - P(E) = 0.9999$, entonces:

$$P(T) = (0.0001 \cdot 0.9) + (0.9999 \cdot 0.1) \approx 0.1001$$

Por lo tanto, la probabilidad de padecer la enfermedad dado que el examen salió positivo es:

$$P(E | T) = \frac{P(E) \cdot P(T | E)}{P(T)} = \frac{0.0001 \cdot 0.9}{0.1001} \approx 0.001$$

Es decir, la posibilidad de que tengamos la enfermedad aumentó de 1 de 10000 a casi 1 de 1000 (o a 10 de 10000) solo por saber que el examen resultó positivo.

2. Independencia.

A partir de la definición de las probabilidades condicionales se puede formalizar el concepto de independencia entre eventos. En esta sección profundizamos en ello.

2.1. Eventos independientes.

En la Clase 1, cuando se revisó el Problema del Cumpleaños, uno de los supuestos señalados fue que dichas fechas de las r personas eran independientes. Es decir, se asumió que conocer aquello no debe afectar la probabilidad que se buscaba calcular. Formalicemos esta idea.

Si dos eventos A y B son independientes, simbólicamente debe implicar que:

$$P(A | B) = P(A) \quad \text{y} \quad P(B | A) = P(B)$$

En otras palabras, $P(A)$ no se ve afectado por saber que B sucedió previamente cuando este y el evento A son independientes. Lo mismo con $P(B)$ dado que A ocurrió.

Apliquemos la definición de las probabilidades condicionales en los lados izquierdos de $P(A | B) = P(A)$ y $P(B | A) = P(B)$, con $P(B) > 0$ y $P(A) > 0$, respectivamente.

$$\frac{P(A \cap B)}{P(B)} = P(A) \quad \text{y} \quad \frac{P(B \cap A)}{P(A)} = P(B)$$

Luego, multipliquemos por $P(B)$ y $P(A)$ en ambas ecuaciones.

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

Dado a que $P(A \cap B) = P(B \cap A)$ y que $P(A) \cdot P(B) = P(B) \cdot P(A)$ ya que la multiplicación es conmutativa, se concluye que dos eventos A y B son **independientes** si:

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

Es decir, dos eventos A y B son independientes si la probabilidad de que ambos ocurran (no necesariamente al mismo tiempo) es igual al producto de sus probabilidades individuales.

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

$$(1) A^c \text{ y } B \quad (2) A \text{ y } B^c \quad (3) A^c \text{ y } B^c$$

Demostremos la afirmación (2). Si A y B^c son independientes, debe cumplirse entonces que $P(A \cap B^c) = P(A) \cdot P(B^c)$ o que $P(A | B^c) = P(A)$. Veamos esta última igualdad.

Asumiendo que $P(B^c) > 0$, podemos aplicar el teorema de Bayes en el lado izquierdo de $P(A | B^c) = P(A)$ para expresar esta ecuación como:

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

Puesto que $P(B^c | A) = 1 - P(B | A)$, la ecuación de arriba es igual a:

$$\frac{P(A) \cdot [1 - P(B | A)]}{P(B^c)} = P(A)$$

Además, $P(B | A) = P(B)$ ya que A y B son independientes.

$$\frac{P(A) \cdot [1 - P(B)]}{P(B^c)} = P(A)$$

Así, debido a que $P(B^c) = 1 - P(B)$, entonces:

$$\begin{aligned} \frac{P(A) \cdot P(B^c)}{P(B^c)} &= P(A) \\ P(A) &= P(A) \quad (\text{Q. E. D}) \end{aligned}$$

En cuanto a la demostración de $P(A \cap B^c) = P(A) \cdot P(B^c)$, se puede comenzar aplicando la propiedad 5 de las probabilidades⁷ para establecer que $P(A \cap B^c) = P(A) - P(B \cap A)$.

Ejercicio 4. Suponga que dos dados no cargados de seis lados son lanzados al mismo tiempo. Evalúe si los siguientes eventos son independientes:

$$A = \{\text{En el primer dado sale el número 4}\}$$

$$B = \{\text{La suma de los resultados de los dos dados es igual a seis}\}$$

⁷Estudiada en la Clase 2.

Solución. Dado la poca cantidad, consideremos los elementos de ambos eventos.

$$A = \{(4, i)\}, \text{ con } i = 1, 2, \dots, 6$$

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

Como $|A| = 6$, $|B| = 5$, $|S| = 36$ (S = el espacio muestral) y $A \cap B = \{(4, 2)\}$, entonces:

$$P(A) = \frac{6}{36} = \frac{1}{6}; \quad P(B) = \frac{5}{36}; \quad P(A \cap B) = \frac{1}{36}$$

Para ver si A y B son independientes, debemos evaluar si se cumple la siguiente igualdad:

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

$$\frac{1}{36} = \frac{1}{6} \cdot \frac{5}{36}$$

$$\frac{1}{36} \neq \frac{5}{216}$$

De este modo se concluye que A y B **no son eventos independientes**.

Ejercicio 5. Cien personas adultas son encuestadas en relación al nivel de deuda estudiantil que existe en la educación superior y si, actualmente, tienen algún hijo/a estudiando en alguna de esas instituciones. En la siguiente tabla se resumen sus respuestas:

Hijo/a estudiando	Nivel de deuda			Total
	Muy alta (A)	Más o menos adecuada (B)	Muy baja (C)	
Sí (D)	0.20	0.09	0.01	0.30
No (E)	0.41	0.21	0.08	0.70
Total	0.61	0.30	0.09	1

Evalúe cuál de los pares de eventos señalados a continuación son independientes:

$$(a) A \text{ y } D \quad (b) B \text{ y } D \quad (c) C \text{ y } D$$

Solución. Mediante la tabla del ejercicio se pueden derivar las probabilidades:

$$P(A) = 0.61 \quad P(B) = 0.30 \quad P(C) = 0.09 \quad P(D) = 0.30$$

$$P(A \cap D) = 0.20 \quad P(B \cap D) = 0.09 \quad P(C \cap D) = 0.01$$

A partir de estas probabilidades evaluemos las siguientes igualdades:

$$\begin{array}{lll}
 P(A \cap D) = P(A) \cdot P(D) & P(B \cap D) = P(B) \cdot P(D) & P(C \cap D) = P(C) \cdot P(D) \\
 0.20 = 0.61 \cdot 0.30 & 0.09 = 0.30 \cdot 0.30 & 0.01 = 0.09 \cdot 0.30 \\
 0.20 \neq 0.183 & 0.09 = 0.09 & 0.01 \neq 0.027
 \end{array}$$

Por lo tanto, solo los eventos B y D son independientes. Una interpretación de este resultado es que el hecho de tener un hijo/a en la educación superior no incidió (aparentemente) en quienes respondieron que el nivel de deuda estudiantil es relativamente el adecuado.

2.2. Eventos independientes y eventos mutuamente excluyentes.

Es un buen momento para dejar claro el siguiente punto:

Dos eventos independientes no son necesariamente mutuamente excluyentes.

La afirmación de arriba también se aplica en sentido contrario:

Dos eventos mutuamente excluyentes no son necesariamente independientes.

El concepto de independencia tiene que ver con que saber que un evento sucedió no cambia nuestra creencia de qué tan posible es que ocurra el de nuestro interés. En cambio, el de exclusividad mutua se centra en el hecho de que si uno ocurre, el otro no puede suceder.

Otra manera de verlo es a partir de las definiciones matemáticas de los conceptos de independencia y exclusividad mutua entre dos eventos A y B :

Es posible que $P(A \cap B) = 0$, pero al mismo tiempo que $P(A) \cdot P(B) \neq 0$

Por lo tanto, la moraleja es que debemos evitar igualar los conceptos de independencia y exclusividad mutua entre eventos, así como establecer alguna relación causal entre ellos.

2.3. Independencia de varios eventos.

Determinar que más de dos eventos son independientes no es un hecho que se deriva inmediatamente de la definición que vimos en la sección 2.1.

Tres eventos A , B y C son independientes si se cumplen las siguientes igualdades:

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

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

Por lo tanto, para determinar que A , B y C son independientes debemos comprobar que lo sean entre los pares que se pueden formar entre ellos y, luego, con los tres.

Para evaluar si cuatro eventos son independientes, se deben corroborar:

- Las $\binom{4}{2} = 6$ independencias pareadas entre ellos.
- Las $\binom{4}{3} = 4$ independencias de los grupos de tres eventos que se forman con los cuatro.
- $\binom{4}{4} = 1$ independencia entre los cuatro eventos en conjunto.

En total, deben cumplirse once independencias para concluir que cuatro eventos lo son.

Así, se generaliza que una colección de eventos $\{A_i\}_{i=1}^n$ son **mutuamente independientes** si, para cualquier subcolección $\{A_{i_j}\}_{j=1}^k$ de ellos, se cumple que:

$$P\left(\bigcap_{j=1}^k A_{i_j}\right) = \prod_{j=1}^k P(A_{i_j})$$

Es decir, los n eventos A_i son mutuamente independientes si son independientes para cualquier intersección que pueden formarse entre ellos.

2.4. Independencia condicional.

La independencia de eventos también puede estar condicionada a la ocurrencia de otro suceso.

Se dice que dos eventos A_1 y A_2 son **condicionalmente independientes** si, dado que un suceso B ocurrió, se cumple la igualdad:

$$P(A_1 \cap A_2 \mid B) = P(A_1 \mid B) \cdot P(A_2 \mid B)$$

Que A_1 y A_2 sean condicionalmente independientes dado B **no implica** que:

- También sean condicionalmente independientes dado B^c .
- A_1 y A_2 sean independientes.

Tampoco implica que si A_1 y A_2 son mutuamente independientes, entonces también son condicionalmente independientes a B .

En el siguiente ejercicio veremos el cuidado que debemos para asegurar la independencia condicional entre eventos.

Ejercicio 6. Suponga que una bebé llora solo si tiene hambre, si está aburrida o ambas. Considere los siguientes eventos junto con sus probabilidades:

Eventos		
Símbolo	Significado	Probabilidad
C	La bebé está llorando	$0 < c < 1$
H	La bebé tiene hambre	$0 < h < 1$
T	La bebé está aburrida	$0 < t < 1$

Asumiendo que los eventos H y T son independientes:

- (a) Calcule c en términos de h y t .
- (b) Obtenga las probabilidades $P(H \mid C)$, $P(T \mid C)$ y $P(H \cap T \mid C)$.
- (c) Evalúe si H y T son condicionalmente independientes dado C .

Solución (a). Dado que la bebé puede llorar porque tiene hambre, está aburrida o por ambas razones, podemos expresar al evento C como:

$$C = H \cup T$$

Así, se puede calcular la probabilidad de C como:

$$P(C) = P(H \cup T) = P(H) + P(T) - P(H \cap T)$$

Debido a que H y T son independientes, entonces $P(H \cap T) = P(H) \cdot P(T) = h \cdot t$. Además, puesto que $P(C) = c$, en consecuencia:

$$c = h + t - ht$$

Solución (b). Las tres probabilidades condicionales podemos expresarlas usando el teorema

de Bayes:

$$\begin{aligned}P(H \mid C) &= \frac{P(H) \cdot P(C \mid H)}{P(C)} = \frac{h \cdot P(C \mid H)}{c} \\P(T \mid C) &= \frac{P(T) \cdot P(C \mid T)}{P(C)} = \frac{t \cdot P(C \mid T)}{c} \\P(H \cap T \mid C) &= \frac{P(H \cap T) \cdot P(C \mid H \cap T)}{P(C)} = \frac{ht \cdot P(C \mid H \cap T)}{c}\end{aligned}$$

Observemos que $P(C \mid H) = P(C \mid T) = P(C \mid H \cap T) = 1$ porque, a partir de lo señalado en el ejercicio, si ocurre H , T o ambos eventos la bebé llorará. De este modo,

$$P(H \mid C) = \frac{h}{c} \quad P(T \mid C) = \frac{t}{c} \quad P(H \cap T \mid C) = \frac{ht}{c}$$

Solución (c). Utilicemos la definición de la independencia condicional para evaluar esta pregunta.

$$\begin{aligned}P(H \cap T \mid C) &= P(H \mid C) \cdot P(T \mid C) \\ \frac{ht}{c} &\neq \frac{h}{c} \cdot \frac{t}{c} = \frac{ht}{c^2}\end{aligned}$$

Esto nos muestra que H y T no son condicionalmente independientes dada la ocurrencia de C , a pesar que estos dos primeros eventos son independientes.

Una explicación del resultado del Ejercicio 6 (c) puede ser, por ejemplo, que si sabemos que la bebé lloró y no fue porque estaba aburrida, entonces podemos deducir que se debió a que tenía hambre. Es decir, dada su actitud, los eventos H y T son dependientes.