Probabilidad total y Teorema de Bayes

Hemos estudiado las probabilidades marginales en lecturas anteriores. Ahora estudiaremos otra forma de expresarlas, otro ángulo desde el cual podemos abordarlas. Tendrás otra herramienta para resolver problemas y podrás aplicarla a ejemplos concretos. En este punto nos detendremos a estudiar el teorema de la probabilidad total, que ya lo utilizamos intuitivamente en el diagrama de árbol.

También estudiaremos, especialmente, qué sucede en los árboles de decisión cuándo se conocen datos previos o probabilidades previas y se agrega información al problema. Esto hará que tengamos que revisar las probabilidades calculadas en un principio y podamos estimar otros resultados. Esto lo veremos a través del teorema de Bayes, el cual es de mucha importancia en la toma de decisiones.

¡Comencemos con la parte final de probabilidades!

Caso: el problema de la aerolínea Good Fly

El área de atención al cliente de la aerolínea Good Fly está haciendo una supervisión debido al retraso habitual que presentan algunos de sus vuelos nocturnos que salen del aeroparque de la Ciudad de Buenos Aires, su atención se centra en tres vuelos. Durante el último año, 7, 8 y 5 % de los pasajeros a Córdoba, Mendoza y Bariloche, respectivamente, han tenido que tomar otros vuelos.

Además el 55, 20 y 25 % de los pasajeros de los vuelos nocturnos de Good Fly toman, en el aeroparque de la Ciudad de Buenos Aires, vuelos a Córdoba, Mendoza y Bariloche, respectivamente. Cuál es la probabilidad de que un pasajero:

a) No haya podido tomar el vuelo original.

b) Que no pudo tomar el vuelo original y haya comprado un boleto para Córdoba.

c) Que no pudo tomar el vuelo original y haya comprado un boleto para Mendoza.

d) Que no pudo tomar el vuelo original y haya comprado un boleto para Bariloche.

Probabilidades marginales bajo dependencia estadística

Probabilidades marginales en una tabla y en un árbol

La probabilidad marginal de un evento es la probabilidad de que ese evento ocurra. Dijimos que es un evento simple, pero que, muchas veces, no es tan fácil de calcular cuando no es dato del problema. Además, el evento puede depender de otros o estar relacionado con otra serie de eventos.

Primero, veremos cómo se calcula en una tabla de contingencia y, luego, en un diagrama de árbol. Al finalizar con estos ejemplos, definiremos el **teorema de la probabilidad total** que incluye a las probabilidades marginales y conjuntas.

Comenzaremos por plantear el siguiente problema y resolverlo mediante una tabla de contingencias y un diagrama de árbol, como se muestra a continuación.

Ejemplo 1:

Un dermatólogo está analizando cierta enfermedad en la piel y la posible influencia de un factor de riesgo que es la cama solar. Para ello, se considera una muestra de 300 personas, de las cuales 100 están enfermas y 200, no. De las personas enfermas, 30 estuvieron expuestas a este factor de riesgo y, del total, son 100 las personas que estuvieron expuestas a este posible factor de riesgo. En base a estos datos responde:

a) ¿Cuál es la probabilidad de que una persona tenga la enfermedad?

b) Si se sabe que la persona se expuso a una cama solar, ¿cuál es la probabilidad de que tenga la enfermedad?

c) ¿Cuál es la probabilidad de que no tenga la enfermedad y se haya expuesto a este factor de riesgo?

Construyamos la tabla y el diagrama y respondamos las preguntas:

Expresión de una probabilidad marginal mediante una tabla de contingencia

Tabla 1: Tabla de contingencia para el ejemplo 1

exposición / estado	Si (S)	No (N)	Totales
Enfermo (E)	30	70	100
No enfermo (E´)	70	130	200
Totales	100	200	300

Fuente: elaboración propia

Descripción de la Tabla 1: en esta tabla se volcaron los datos del problema que presenta el ejemplo 1. En negrita están los datos, en color lo calculado de acuerdo con las reglas de construcción de la tabla. El nombre de cada evento está escrito en color, en los encabezados de las filas y columnas.

a) ¿Cuál es la probabilidad de que una persona tenga la enfermedad?

- Por tabla es la probabilidad marginal:

$$P(E) = \frac{100}{300} = \frac{1}{3}$$

- Por suma de intersecciones:

$$P(E) = \frac{30}{300} + \frac{70}{300} = \frac{100}{300} = \frac{1}{3}$$

Es decir, la probabilidad marginal puede expresarse como suma de las probabilidades conjuntas en las que figura el evento buscado:

$$P(E)=P(E\cap S)+P(E\cap N)$$

Podemos ver que el resultado es el mismo, solo queremos expresar la probabilidad marginal en función de los eventos que están relacionados con el evento E, porque es el que se tiene que tener en cuenta para responder la pregunta. Muchas veces, como veremos, necesitaremos esta fórmula especialmente en el teorema de Bayes. ¡¡¡Recuerda esta fórmula!!!

Las dos preguntas que faltan las resolveremos a modo de repaso, pero la importante para explicar lo que queríamos era la a)

b) Si se sabe que la persona se expuso a cama solar, ¿cuál es la probabilidad de que tenga la enfermedad?

$$P(E/S) = \frac{30}{100} = \frac{3}{10}$$

c) ¿Cuál es la probabilidad de que no tenga la enfermedad y se haya expuesto a este factor de riesgo?

$$P(E' \cap S) = \frac{70}{300} = \frac{7}{30}$$

Expresión de una probabilidad marginal mediante un diagrama de árbol

Recuerda que el diagrama de árbol es una alternativa a la tabla de contingencia y, según lo que se quiera calcular, es muy útil para visualizar las probabilidades condicionales y conjuntas.

- Primero analicemos si los eventos, estar enfermo o no, son dependientes con los eventos de exposición a un factor de riesgo o no. Pues, si son independientes, no se cumple la regla que vamos a enunciar.
  Tomemos los eventos E={la persona está enferma} y S={la persona se expuso a un factor de riesgo}
  Según el postulado de la independencia estadística debe darse que:

$$P(E/S) = P(E) \quad \text{o bien} \quad P(S/E) = P(S)$$

$$P(E/S) = \frac{30}{100} = \frac{3}{10} \text{ y } P(E) = \frac{100}{300} = \frac{1}{3}$$

$$P(E/S) \neq P(E)$$

<u>Por lo tanto, los eventos son dependientes.</u>

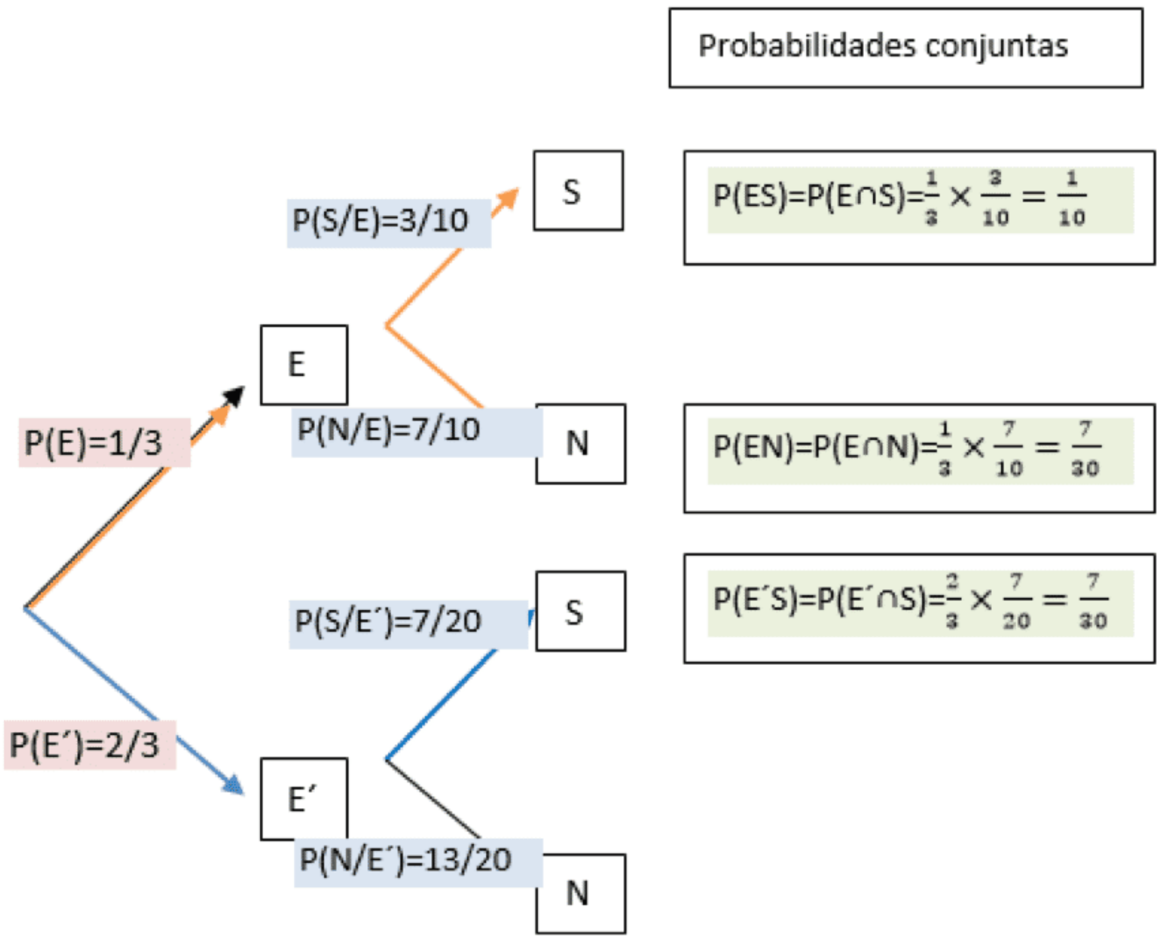
Ahora, construiremos el diagrama de árbol para el ejemplo 1.

Los cálculos para las probabilidades que van en las ramas pueden hacerse mirando la tabla 1, o bien por fórmulas.

Solo calcularemos las probabilidades que nos interesan.

Figura 1: Diagrama de árbol del ejemplo 1

Descripción de la figura 1: se muestra el diagrama de árbol del problema del ejemplo 1, en el que figuran probabilidades marginales, condicionales y conjuntas.



Fuente: elaboración propia

a) ¿Cuál es la probabilidad de que una persona tenga la enfermedad?

- Por árbol es la probabilidad marginal:

$$P(E) = \frac{1}{3}$$

- Por suma de intersecciones, solo sumaremos las ramas en las que intervenga el evento E. Calcularemos las probabilidades conjuntas y sumamos (ver árbol, las ramas color naranja)

$$P(E) = \frac{1}{10} + \frac{7}{30} = \frac{3+7}{30} = \frac{10}{30} = \frac{1}{3}$$

Es el mismo resultado obtenido anteriormente mediante la tabla. Observa que la probabilidad marginal del evento E también la podemos escribir:

$$P(E)=P(E\cap S)+P(E\cap N)$$

Ten en cuenta que esta fórmula, resulta igual a la deducida en la tabla de contingencia.

b) Si se sabe que la persona se expuso a una cama solar, ¿cuál es la probabilidad de que tenga la enfermedad?

$P(E \,/\, S)$, esta probabilidad condicional ya la calculamos mediante la tabla de contingencia y es igual a 0,3. En este caso, el árbol construido no nos da información acerca de esta probabilidad condicional, al haber comenzado por dos eventos, que llamamos previos.

Si te fijas en el diagrama de árbol la obtenida en las ramas del segundo nivel es la P(E ⁄ S), recíproca de la que se está pidiendo en este punto.

Dejaremos para más adelante este ítem y comprobaremos que hay otra forma de resolverlo sin tabla. Todo va a depender del enunciado del problema, este problema, en especial, te conviene resolverlo con tabla de contingencia pero cuando los datos del problema son solo probabilidades o porcentajes necesitarás ver otra técnica para poder resolverlo más fácilmente.

c) ¿Cuál es la probabilidad de que no tenga la enfermedad y se haya expuesto a este factor de riesgo? (multiplicamos las probabilidades que están sobre las ramas de color azul)

$$P(E' \cap S) = \frac{2}{3} \times \frac{7}{20} = \frac{7}{30}$$

Es el mismo resultado obtenido por tabla.

Ahora podremos enunciar el teorema de la probabilidad total, que daremos a continuación, centrados en las conclusiones de la probabilidad marginal que estudiamos en este ejemplo 1.

Teorema de la probabilidad total

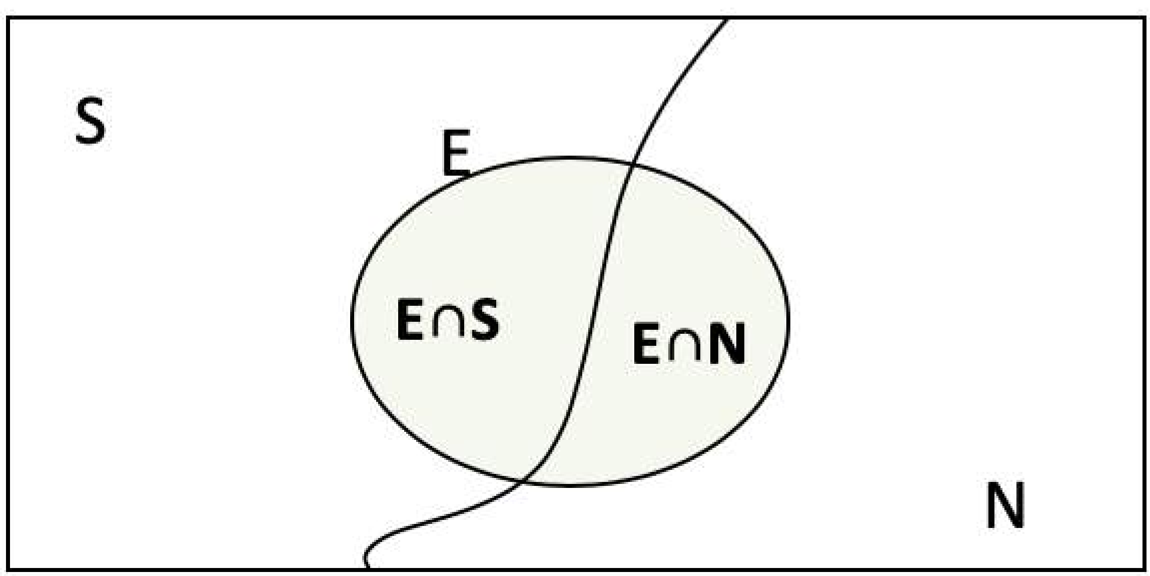
“Las probabilidades marginales en condiciones de dependencia estadística se calculan mediante la suma de las probabilidades de todos los eventos conjuntos en los que se presenta el evento sencillo” (Levin y Rubín, 2004, p. 155).

Para generalizar el ejemplo 1, observemos lo siguiente:

Si tomamos los eventos S={la persona se expuso a un factor de riesgo} y N{la persona no se expuso a un factor de riesgo}, vemos que ambos son mutuamente excluyentes y colectivamente exhaustivos, pero ambos están relacionados con otros eventos: E={la persona está enferma} y su complemento. Tomemos como ejemplo solo el evento E, dejemos el otro evento como el complemento de E. La situación se vería de la siguiente manera, en un diagrama de Venn, podemos particionar el espacio muestral de la siguiente manera:

Figura 2: representación con diagramas de Venn de los eventos que intervienen en el Teorema de la probabilidad total

Descripción de la figura 2: se presenta una situación entre cuatro conjuntos, S, N, E, E´, los dos primeros mutuamente excluyentes entre sí, pero ambos en intersección con los otros dos, que también son mutuamente excluyentes entre sí. La suma de las intersecciones son los eventos que nos interesan.



Fuente: elaboración propia

Observa que $(E \cap S) \cup (E \cap N) = E$

- El evento E, lo hemos tomado como información a priori.
- Los eventos S y N son la información nueva, la que se obtuvo después de investigar la influencia de los factores de riesgo.

Podemos interpretar que los eventos S y N son las causas (o las circunstancias) por las que puede ocurrir el evento E.

O dicho de otra forma, el estar o no enfermo depende de si la persona se expuso o no a los factores de riesgo.

Entonces, el teorema de la probabilidad total viene a decir que si el suceso **E** puede ocurrir por alguna de las causas: **S** o **N**, la probabilidad de que ocurra **E** es la suma de las probabilidades $P(E \cap S)$ y $P(E \cap N)$.

Pero, por definición de probabilidad conjunta y observando las ramas del árbol, las probabilidades conjuntas podemos escribirlas de las siguientes formas:

$P(E \cap S)=P(E).P(S/E)$ o también $P(S \cap E)=P(S).P(E/S)$

Y además:

$P(E \cap N)=P(E).P(N/E)$ o también $P(N \cap E)=P(N).P(E/N)$

Estas fórmulas nos dan más libertad de usarlas según la conveniencia, según el dato con el que contemos a priori, es decir, según la condición que tengamos.

åPor lo que podemos hacer es escribir otra forma el teorema de la probabilidad total:

$$P(E)=P(E \cap S)+P(E \cap N)$$

Sustituyendo:

$$\mathbf{P(E)=P(E).P(S/E)+P(E).P(N/E)}$$

O bien podríamos contar con la información sobre los eventos S y N por lo que las fórmulas quedan:

$$\mathbf{P(E)=P(S).P(E/S)+P(N).P(E/N)}$$

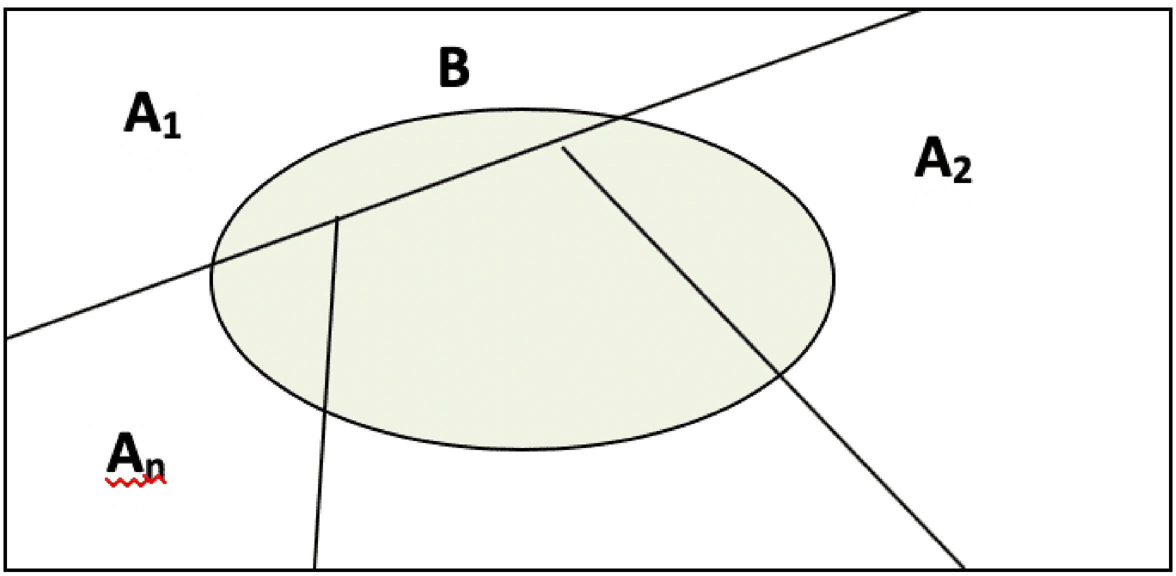
Recuerda que $P(E \cap S)=P(S \cap E)$

En general: Si tenemos eventos que son mutuamente excluyentes y colectivamente exhaustivos en el espacio muestral, por ejemplo: A1, A2, … An y, además, un nuevo evento B que tiene intersección con todos los eventos Ai del espacio muestral, podemos escribir:

$$P(\boldsymbol{B}) = P(A_1).P(B/A_1) + P(A_2).P(B/A_2) + \cdots P(A_n).P(B/A_n)$$

Figura 3: representación con diagramas de Venn de los eventos que intervienen en el teorema de la probabilidad total en forma general

Descripción de la figura 3: es una representación de lo que ocurre en el espacio muestral relacionada con eventos mutuamente excluyentes y conjuntos, para visualizar mejor el teorema de la probabilidad total. La suma de las intersecciones son los eventos que nos interesan.



Fuente: elaboración propia

¿Para qué sirve este teorema?

Es lo que te estarás preguntando en estos momentos…

Por eso te lo vamos a explicar para el ejemplo 1 y para el caso de la aerolínea Good Fly

EJEMPLO 1	CASO GOOD FLY

Supongamos que nos preguntan una probabilidad de otro evento marginal, por ejemplo:

¿Cuál es la probabilidad de que la persona se haya expuesto a un factor de riesgo?

- Por tabla es fácil, ya sabemos que es: P(S)=1/3 (es casualidad que coincida con la P(E))

- Pero en el diagrama de árbol no tenemos esa probabilidad, puesto que comenzamos a construirlo con las probabilidades marginales E y E´.

Pero, mediante el teorema de la probabilidad total, tenemos:

$P(S)=P(E\cap S) + P(E´\cap S) = P(E).P(S/E) + P(E´).P(S/E´)$,

todos estos son datos, entonces (ver imagen)
que es lo mismo que había dado por tabla.
Conclusiones:

- Calculamos la probabilidad marginal de dos formas distintas y dan lo mismo
- Como verás, es una forma alternativa de calcular las probabilidades marginales bajo dependencia estadística.
- Muchas veces se utiliza esta última fórmula, especialmente, cuando los datos son probabilidades o porcentajes. No siempre los problemas pueden adaptarse a una tabla de contingencia.
- Lo verás más claro en el caso de Good Fly.

$$P(S)=\frac{1}{3}\times\frac{3}{10}+\frac{2}{3}\times\frac{7}{20}=\frac{1}{10}+\frac{7}{30}=\frac{1}{3}$$

EJEMPLO 1	CASO GOOD FLY

Por ejemplo, este es un caso en el que conviene hacer un diagrama de árbol, porque tenemos un conocimiento previo de los pasajeros que toman esos vuelos y una información posterior, que es el porcentaje de personas que cambian el vuelo.
Además, si te fijas en los datos, son todos porcentajes, que tendremos que expresar en probabilidades y el árbol sale rápido y más fácil.
Respondamos solo la pregunta a) del caso:
a) ¿Cuál es la probabilidad de que una persona no haya podido tomar el vuelo original?
Primero definamos los eventos:

C= {pasajeros que viajan a Córdoba}.
M= {pasajeros que viajan a Mendoza}.
B= {pasajeros que viajan a Bariloche}.
O= {pasajeros que eligen otros vuelos}.
N= {pasajeros que viajaron en forma normal}.

Confeccionemos el árbol con los datos del problema.
Los datos en este problema son los siguientes:

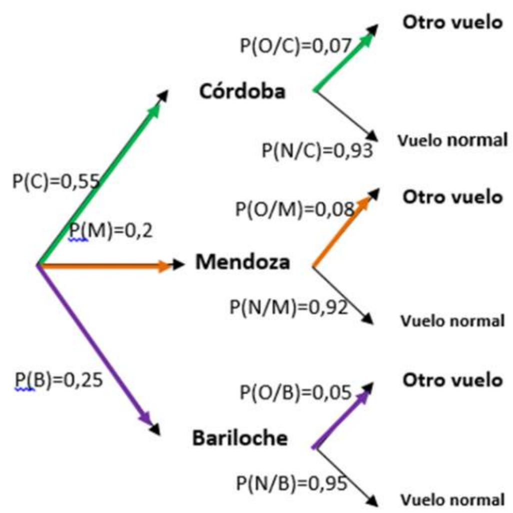
- Las probabilidades simples **P(C)**=0,55 **P(M)**=0,2 **P(B)**=0,25
- y las probabilidades condicionales: **P(O/C)**=0,07 **P(O/M)**=0,08 **P(O/B)**=0,05

Es más sencillo hacer el árbol dado que todos los datos se pueden volcar en él, sin hacer cálculos, en este caso

Los caminos que están marcados en colores son los que tenemos que multiplicar según indiquen las probabilidades de las ramas y luego sumarlos.

Figura 4: diagrama de árbol para el caso de la aerolínea Good Fly
Descripción de la Figura 4: se muestra el diagrama de árbol para el caso Good Fly. Observa que en las ramas están los datos del problema. Para calcular la probabilidad solicitada, solo hay que operar con las tres ramas que están involucradas con el evento O.
(ver imagen)

Es decir que los que han tenido que tomar otros vuelos representan el 6,7 % de los pasajeros.
Este razonamiento es más fácil porque, si seguimos las ramas del árbol, ni siquiera hace falta recordar la fórmula.



Fuente: elaboración propia

$$P(tomar\ otro\ vuelo) = P(O) = 0{,}55 \times 0{,}07 + 0{,}2 \times 0{,}08 + 0{,}25 \times 0{,}05 = 0{,}067$$

Un instituto de idiomas está capacitando a sus alumnos con dos métodos de enseñanza: A y B. El porcentaje de fracasos del método A, es 20 % y del B es 10 %. Sin embargo, como el método B es más caro, los alumnos inscriptos para el método B son el 30 %, y el resto se inscribió en el curso A. ¿Cuál es la probabilidad de que un alumno fracase en la adquisición de las habilidades necesarias?

- P(F)=P(A).P(F⁄(A)+P(B).P(F⁄B)= 0,17
- P(F)=P(F/A)+P(F⁄(B)=0,30

- [ ] $P(F)=P(F´/A)+P((F´)/(B)=0,70$

- [ ] $P(F´)=P(F´\cap A)+P(F´\cap B)=0,87$

SUBMIT

Teorema de Bayes

Importancia

Este teorema debe su importancia a que permite revisar y modificar las estimaciones sobre la probabilidad de un evento cuando se ha recabado información más actualizada sobre este.

Las probabilidades que se calculan luego de haber contado con la nueva información, reciben el nombre de **probabilidades posteriores o probabilidades revisadas.**

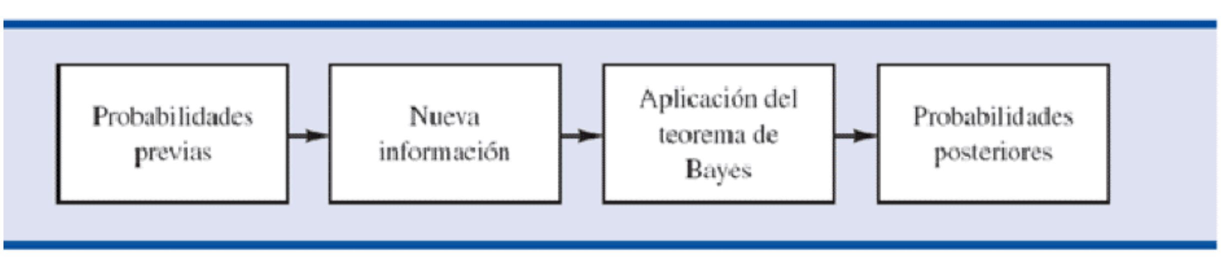
> El teorema de Bayes se usa mucho en la toma de decisiones. Las probabilidades previas suelen ser estimaciones subjetivas dadas por la persona que toma las decisiones. Se obtiene información muestral y se usan las probabilidades posteriores para emplearlas en la toma de decisiones. (Anderson, Sweeney y Williams, 2008, p. 176).

Es en este momento cuando la teoría de probabilidad cobra gran importancia, ya que tanto con el árbol de decisiones como con el teorema de Bayes, las inferencias van a ser más precisas.

La importancia del teorema de Bayes no reside únicamente en el cálculo sino en la capacidad de los operadores para predecir el futuro con la nueva información obtenida.

Figura 5: Revisión de la probabilidad utilizando el teorema de Bayes

Descripción de la figura 5: son los pasos que se visualizan en un problema en el que se aplica el teorema de Bayes.



Fuente: Anerson, Sweeney y Williams, 2008, p. 171.

Aplicación del teorema de Bayes

Antes de enunciar el teorema lo aplicaremos. En realidad, el ejemplo 1 no lo tendremos en cuenta en este caso, pues ya vimos que es más adaptable a presentarlo en una tabla de contingencia y calcular desde allí las probabilidades.

Recuerda que ahora hablamos de probabilidades a priori y probabilidades condicionales, y se le suma el condimento de poder tener información nueva, y poder así revisar las causas que originaron esas probabilidades.

La probabilidad condicional toma en cuenta información sobre la ocurrencia de un evento para predecir la probabilidad de otro.

Para este tipo de problemas, el diagrama de árbol es mucho más útil. La ventaja de tener probabilidades condicionales en el segundo nivel de ramas, nos permitirá ubicar fácilmente las probabilidades sin calcularlas cuando son datos o aplicar el teorema de Bayes para obtener la inversa.

Supongamos que contamos como dato con , pero, después de obtener cierta información o simplemente porque lo necesitamos, tengamos que calcular . Sabemos por experiencia que no son lo mismo. Entonces, se trata de encontrar una relación entre estas dos expresiones. De esto trata el teorema de Bayes.

Por lo tanto, hay varias razones para utilizar este teorema, entre las cuales están:

- el lenguaje con el que contamos a priori: generalmente, se trata de porcentajes o probabilidades de eventos simples: excluyentes y exhaustivos, fácilmente representables en un diagrama de árbol.
- También contamos con probabilidades condicionales como datos, si no es así, tendrás que calcularlas.
- También los datos pueden ser probabilidades conjuntas,
- En todos los casos que no tengas los datos, para volcarlos directamente en las ramas, tendrás que calcularlos mediante las reglas de la probabilidad estudiadas.

Comencemos resolviendo estos dos problemas, entre los que se encuentra el caso de la aerolínea Good Fly y, luego, generalizamos el teorema.

Ejemplo 2:

Una gran empresa de productos para el cuidado de la piel ha pautado con los canales de TV más importantes del país, tanto en aire como en cable, publicitar su producto.

De acuerdo con una encuesta realizada, asignaron probabilidades a los eventos siguientes:

- La probabilidad de que una persona compre el producto es de 0,20
- La probabilidad de que una persona recuerde haber visto la publicidad es de 0,40
- La probabilidad de que una persona compre el producto y recuerde haber visto la publicidad es de 0,12

Determina las siguientes probabilidades:

a) ¿Cuál es la probabilidad de que una persona compre el producto dado que recuerda haber visto la publicidad? ¿Ver la publicidad aumenta la probabilidad de que el individuo compre el producto? Si usted tuviera que tomar la decisión, ¿recomendaría que continuara la publicidad (suponiendo que los costos sean razonables)? Si una persona que no compra el producto de la empresa compra el de la competencia, ¿cuál sería su estimación de la participación de la empresa en el mercado? ¿Esperarías que continuando con la publicidad aumentara la participación de la empresa en el mercado? ¿Por qué si o por qué no?

b) La empresa probó, también, otra publicidad y los valores de probabilidad asignados fueron $P(C)=0{,}30$, $P(C \cap R)=0{,}10$. Determine $P(C/R)$ en el caso de esta otra publicidad. ¿Qué publicidad parece tener mejor efecto en la compra de los clientes? (Anerson, Sweeney y Williams, 2008, p. 182).

Primeramente, definamos los eventos que parecen estar involucrados en el problema:

C={la persona compre el producto}

C´={la persona no compre el producto}

R={la persona recuerda la publicidad}

R´={la persona no recuerda la publicidad}

Datos: $P(C)=0{,}20$ $P(R)=0{,}40$. $P(R \cap C)=0{,}12$

Veamos que se nos pide calcular:

a) $P(C / R)$ b) $P(C')$ c) $P(C / R)$ cambiando los datos

Como vemos, los datos son, en este caso, probabilidades simples y una probabilidad conjunta.

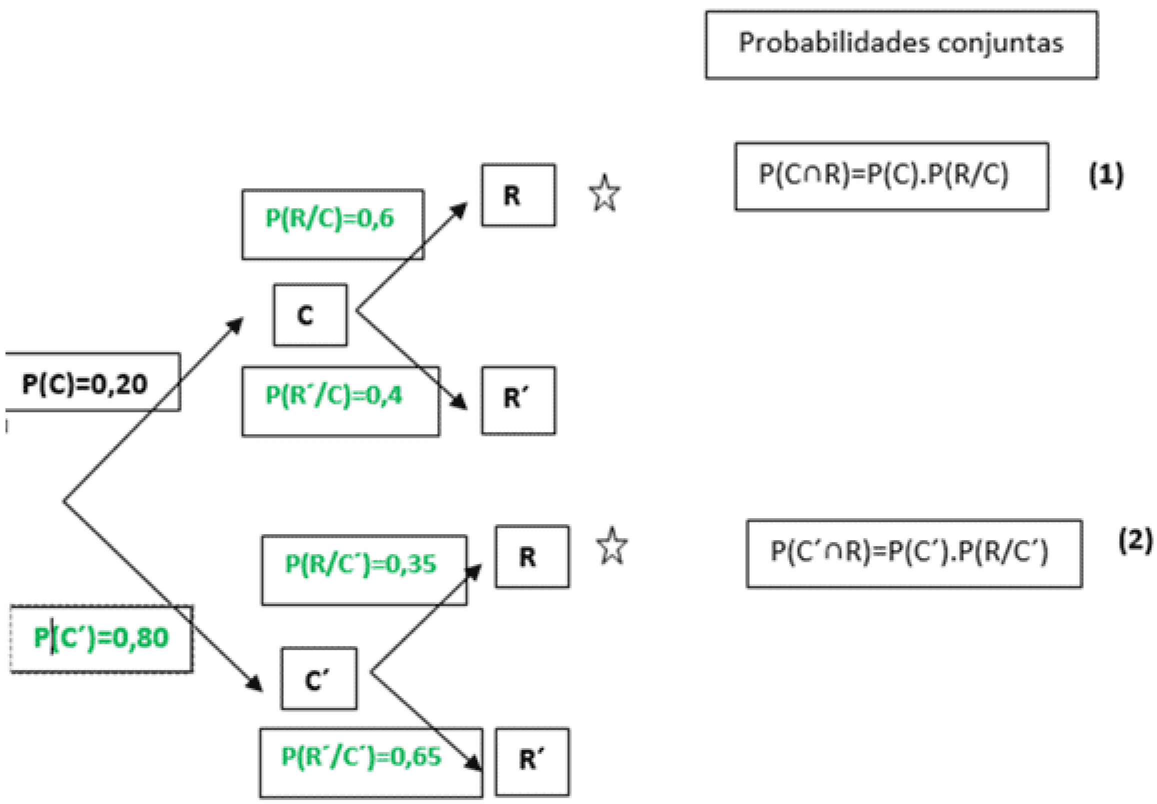
Comencemos a volcar los datos en un diagrama de árbol y a calcular lo que nos hace falta.

Como verás, acá no está todo tan servido, hay que hacer cálculos para poder completar lar ramas del árbol.

En el diagrama están con negrita las probabilidades que son datos y en color las probabilidades calculadas.

Figura 6: diagrama de árbol para el problema del ejemplo 2

Descripción de la Figura 6: en el diagrama se muestran sobre las ramas las probabilidades que son datos y las calculadas en base a ellos. La finalidad es responder las preguntas utilizando el teorema de Bayes. Así como obtener conclusiones debido a las revisiones realizadas.



Fuente: elaboración propia

De (1) obtenemos: $0{,}12=0{,}20 \times P(R/C)$. Entonces:

$$P(R/C)=\frac{0{,}12}{0{,}20}=0{,}6$$

De (2) obtenemos: $P(C' \cap R)=0{,}80 \,.\, P(R/C')$ el único dato que tenemos es $P(C')$

Pero $P(C' \cap R)$ podemos calcularla mediante el teorema de la probabilidad total, que son los caminos marcados por ☆ . En esos caminos se multiplican las ramas y luego se suman.

Entonces: $P(R)=P(C \cap R)+P(C' \cap R)$

Sustituimos en la ecuación (3)

$$0{,}40=0{,}12+P(C' \cap R)$$

Y obtenemos:

$$P(C' \cap R)=0{,}40-0{,}12=0{,}28$$

Ahora sí, sustituimos en (2): $0{,}28 = 0{,}8 \,.\, P(R/C')$. Entonces:

$$P(R/C')=\frac{0{,}28}{0{,}8}=0{,}35$$

Para asegurarnos de que hicimos bien los cálculos, tiene que verificarse el teorema de la probabilidad total para P(R)

$$P(R)=P(C).P(R/C)+P(C').P(R/(C'))$$

Sustituimos con las probabilidades obtenidas y escritas en el árbol:

$P(R)=0,2\times 0,6+0,8\times 0,35=0,4$

¡Que es lo que teníamos que obtener!

Hasta aquí, solo completamos el árbol, ahora respondamos las preguntas:

a) Calcular P(C/R). Pero tenemos la P(C/R) por lo tanto, tendremos que aplicar el teorema de Bayes con el beneficio de que la probabilidad marginal del denominador es dato.

$$P(C/R)=\frac{P(C).P(R/C)}{P(R)}=\frac{0,2\times 0,6}{0,4}=0,3$$

Fíjate que el producto del numerador es la multiplicación de las probabilidades que están en la rama superior del árbol.

Entonces la probabilidad que una persona, habiendo recordado la publicidad, haya comprado el producto es del 30 %

- Ver la publicidad aumenta las probabilidades de que la persona compre el producto.
- Yo aconsejaría que, desde este punto de vista, mantengan la publicidad, pues aumenta en un 10 % las compras.

b) La probabilidad de que una persona no compre el producto es 0,8. Si se supone que todas ellas compran en la competencia, la participación de la empresa en el mercado es baja.

La probabilidad de que una persona, habiendo recordado la publicidad, no haya comprado es:

$$P(C'/R) = \frac{P(C').P(R/C')}{P(R)} = \frac{0{,}8 \times 0{,}35}{0{,}4} = 0{,}7,$$

contra un 30 % que habiendo recordado la publicidad, compra el artículo.

La conclusión es que, no saque la publicidad, sino que la cambie por otra que haga que los clientes que la vieron compren más. Se trata de disminuir ese 70 % que ve la publicidad y no compra.

c) Te dejo este punto para que lo resuelvas solo, ya seguro que estás orientado. Cualquier duda consulte.

Caso Good Fly

Retomemos el caso de la aerolínea Good Fly. Ya hemos respondido a la pregunta a) mediante el teorema de la probabilidad total, pues se pedía un evento simple.

Quedan tres preguntas que podríamos describir de la siguiente manera:

a) P(C ⁄ O) Si sabemos que una persona ha tenido que tomar otro vuelo, ¿cuál es la probabilidad de que ese vuelo haya tenido como destino Córdoba?

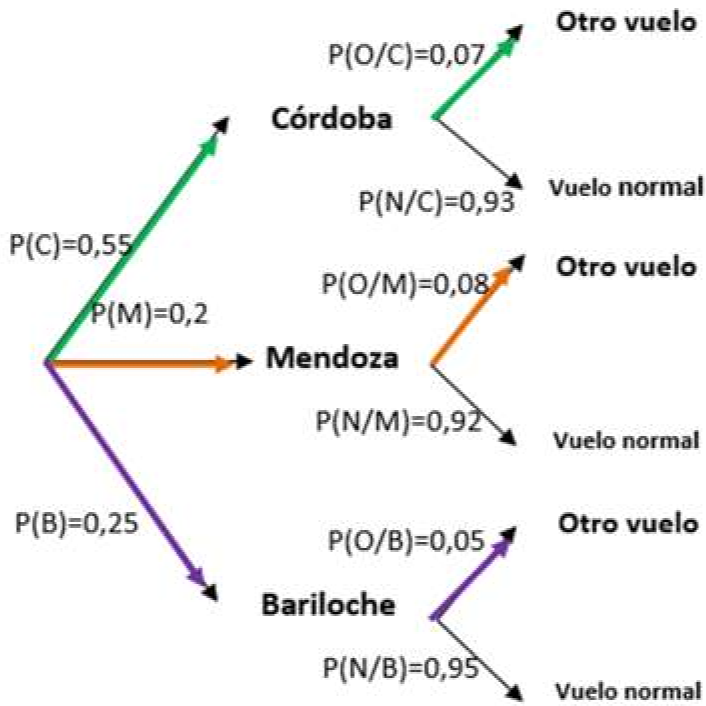
b) P(M ⁄ O) Si sabemos que una persona ha tenido que tomar otro vuelo, ¿cuál es la probabilidad de que ese vuelo haya tenido como destino Mendoza?

c) P(B ⁄ O) Si sabemos que una persona ha tenido que tomar otro vuelo, ¿cuál es la probabilidad de que ese vuelo haya tenido como destino Bariloche?

Con tres colores distintos se marcaron las ramas que nos interesan del árbol, de acuerdo a lo que se solicita en cada una de las preguntas.

Figura 7: diagrama de árbol para el problema del caso Good Fly

Descripción de la Figura 7: el árbol muestra los eventos a tener en cuenta para poder resolver el caso y en sus ramas las probabilidades condicionales, teniendo en cuenta los eventos previos. Las ramas que están en color son en las que debemos centrar nuestra atención.



Fuente: elaboración propia

Observa que lo que se está solicitando son las probabilidades recíprocas de los datos del problema que son las que tenemos en el segundo nivel del árbol.

Resolvamos las probabilidades por el teorema de Bayes, teniendo en cuenta el teorema de la probabilidad total para calcular la P(O). Solo tienes que observar las ramas del árbol.

b) $P(C/O) = \frac{P(C).P(O/C)}{P(O)} = \frac{0,55 \times 0,07}{0,55 \times 0,07 + 0,2 \times 0,08 + 0,25 \times 0,05} = \frac{0,0385}{0,067} = 0,5746$

c) $P(M/O) = \frac{P(M).P(O/M)}{P(O)} = \frac{0,2 \times 0,08}{0,55 \times 0,07 + 0,2 \times 0,08 + 0,25 \times 0,05} = \frac{0,016}{0,067} = 0,2388$

d) $P(B/O) = \frac{P(B).P(O/B)}{P(O)} = \frac{0,25 \times 0,05}{0,55 \times 0,07 + 0,2 \times 0,08 + 0,25 \times 0,05} = \frac{0,0125}{0,067} = 0,1866$

- Si sabemos que una persona ha tenido que tomar otro vuelo, la probabilidad de que ese vuelo haya tenido como destino Córdoba es 0,5746. El 57,46 % de las personas que tenían que viajar a Córdoba, tomaron otros vuelos.

- Si sabemos que una persona ha tenido que tomar otro vuelo, la probabilidad de que ese vuelo haya tenido como destino Mendoza es 0,2388. El 23,88 % de las personas que tenían que viajar a Mendoza, tomaron otros vuelos.

- Si sabemos que una persona ha tenido que tomar otro vuelo, la probabilidad de que ese vuelo haya tenido como destino Bariloche es 0,1866. El 18,66 % de las personas que tenían que viajar a Bariloche, tomaron otros vuelos.

Resolución del ejemplo 1 inciso b) pendiente

Revisa el ejemplo 1, recuerda que había quedado pendiente la resolución mediante el teorema de Bayes para la siguiente pregunta (ten presente el árbol del ejemplo 1):

b) Si se sabe que la persona se expuso a cama solar, ¿cuál es la probabilidad de que tenga la enfermedad?

- Por tabla de contingencia esta probabilidad es igual a 0,3.

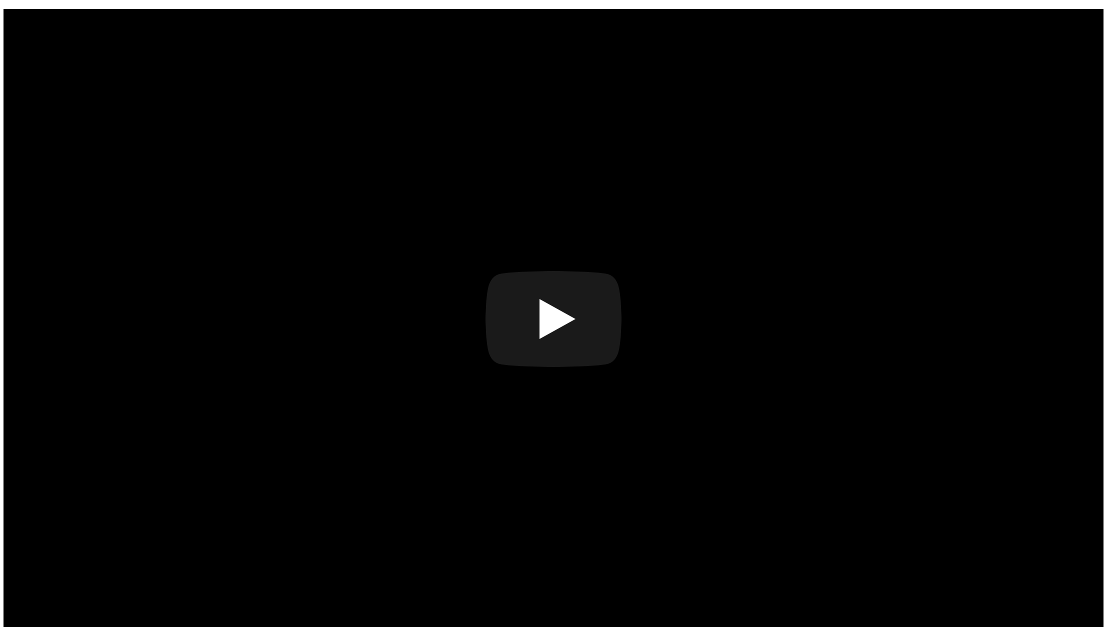
- Por teorema de Bayes:

$$P(E/S)=\frac{P(E).P(S/E)}{P(S)}=\frac{P(E).P(S/E)}{P(E).P(S/E)+P(E').P(S/E')}=\frac{\frac{1}{3}\times\frac{3}{10}}{\frac{1}{3}\times\frac{3}{10}+\frac{2}{3}\times\frac{7}{20}}=\frac{\frac{1}{10}}{\frac{1}{10}+\frac{7}{30}}=\frac{\frac{1}{10}}{\frac{10}{30}}=\frac{3}{10}$$

Por lo que la probabilidad de que esté enfermo si se expuso a un factor de riesgo es 0,3

Video 1: UDEM. Estadística para negocios. Teorema de Bayes

YOUTUBE



UDEM Estadística para negocios Teorema de Bayes(1)

En esta sección aprenderás a utilizar el teorema de Bayes.

VER EN YOUTUBE >

Fuente: Ingeniat [ingeniat]. (3 de marzo de 2011). UDEM Estadística para negocios Teorema de Bayes [YouTube]. Recuperado de https://www.youtube.com/watch?v=pokPfPLON5w

Descripción: en este video se muestra la utilidad del teorema de Bayes para las empresas. Parte de un ejemplo y lo resuelve. Muestra el diagrama de árbol, las probabilidades a priori y todos los elementos que requiere este teorema para poder ser aplicado. Este video es para que lo complementes con la lectura que acabamos de estudiar.

Fórmula general del teorema de Bayes

Partimos de la fórmula para la probabilidad condicional bajo dependencia estadística.

Recuerda las fórmulas de la probabilidad condicional:

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

Despejando $P(A\cap B)$›

$P(A\cap B)=P(B).P(A\diagup(B))$ **(1)** que es la probabilidad conjunta de los eventos A y B para eventos dependientes.

Pero también:

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

Despejando P(B∩A)

P(B∩A)=P(A).P(B⁄(A) **(2)** que también es la probabilidad conjunta de los eventos A y B para eventos dependientes.

Como los eventos conjuntos: A∩B y B∩A son iguales, pues se trata de que se den ambos eventos simultáneamente o sucesivamente, entonces:

P(A∩B)=P(B∩A

De las expresiones (1) y (2):

P(B).P(A⁄(B))=P(A).P(B⁄(A))

Y, finalmente, llegamos a la expresión de la probabilidad condicional que relaciona a una de ellas con su recíproca:

$$\boldsymbol{P(A/B)} = \frac{P(A).\,\boldsymbol{P(B/A)}}{P(B)}$$

Que es el teorema de Bayes para dos eventos.

El numerador no es más que la probabilidad conjunta entre A y B, cuando los eventos son dependientes.

Pero aún falta un paso imprescindible para poder expresar la fórmula definitiva del teorema de Bayes.

Hemos estudiado las probabilidades marginales bajo dependencia estadística y llegamos a la conclusión que una probabilidad marginal es la suma de probabilidades conjuntas. Estas probabilidades conjuntas están formadas por eventos mutuamente excluyentes y colectivamente exhaustivos. Entonces, en general, podemos escribir la fórmula de la probabilidad marginal de la siguiente manera:

$$P(\boldsymbol{B}) = P(A_1).\,P(B/A_1) + P(A_2).\,P(B/A_2) + \cdots P(A_n).\,P(B/A_n)$$

Siendo A_1, A_2, … A_n, mutuamente excluyentes y colectivamente exhaustivos en el espacio muestral, B un nuevo evento que tiene intersección con todos los eventos Ai del espacio muestral.

Si sustituimos la expresión de la probabilidad marginal de B, en la fórmula del teorema de Bayes:

$$P(A_i/B) = \frac{P(A_i).\boldsymbol{P(B/A_i)}}{P(A_1).\boldsymbol{P(B/A_1)} + P(A_2).\boldsymbol{P(B/A_2)} + \cdots P(A_n).\boldsymbol{P(B/A_n)}}$$

A_i es, en este caso, cualquiera de los eventos mutuamente excluyentes y colectivamente exhaustivos del espacio muestral, y B es un nuevo evento que tiene intersección con todos los A_i.

Esta es la fórmula general del teorema de Bayes que también es llamado: fórmula para la probabilidad de las causas.

En la actualidad, este concepto puede ampliarse –como dijimos– para revisar las probabilidades en base a nueva información y determinar la probabilidad de que un efecto en particular se deba a una causa específica. A veces, cuando los eventos son muchos se utiliza una tabla para organizar los datos, a éste método se lo llama **método tabular.**

Finalizamos con algunos comentarios de los autores del texto básico que nos parecieron muy oportunos y prácticos para este tema.

> El teorema de Bayes es un procedimiento formal que permite a los tomadores de decisiones combinar la teoría de probabilidad clásica con su mejor sentido intuitivo acerca de lo que es posible que ocurra. Advertencia: el valor real del teorema de Bayes no está en el álgebra sino en la habilidad de los administradores bien informados para hacer buenas predicciones del futuro. Sugerencia: en todas las situaciones en las que se use el teorema de Bayes, primero utilice todos los datos históricos disponibles y después (y solo entonces) agregue su propio juicio intuitivo al proceso. La intuición usada para hacer

predicciones acerca de cosas que ya están bien descritas, estadísticamente, está mal dirigida (Levin y Rubin, 2004, p. 162).

¿Cuáles son los elementos a tener en cuenta para saber si es conveniente utilizar el teorema de Bayes, siempre que sea posible?

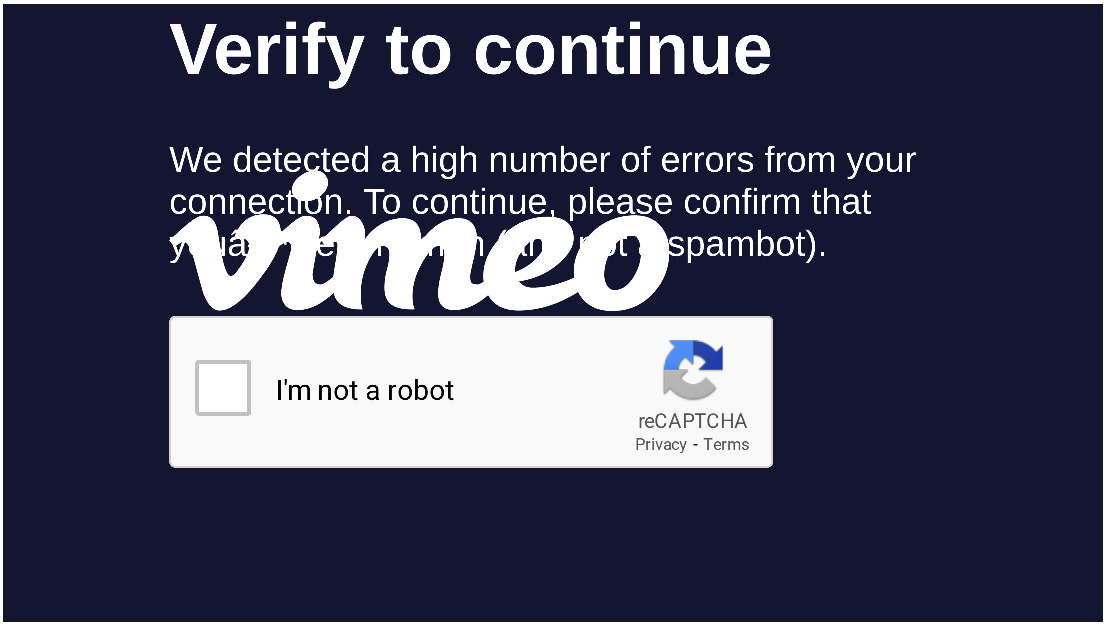
- [ ] contar con datos a priori, que son probabilidades simples: excluyentes y exhaustivas, entre otras.
- [ ] contar con datos que son probabilidades condicionales o conjuntas, entre otras.
- [ ] que se quiera calcular una probabilidad condicional que es inversa a la probabilidad condicional que se da como dato.
- [ ] que se desee saber la probabilidad de la causa si se tiene nueva información.
- [ ] que se quiera calcular una probabilidad marginal
- [ ] que se quiera calcular una probabilidad conjunta

SUBMIT

Video conceptual

Video 2: Teorema de BAYES

En este video se explica el teorema de Bayes, resolviendo un ejercicio con el método tabular, que te puede ser de mucha utilidad y practicidad.



Revisión del módulo

Hasta acá aprendimos

Probabilidad: concepto y terminología básica

En esta lectura has estudiado la definición de probabilidad, sus enfoques, los conceptos básicos de probabilidad que utilizarás en todo el módulo como: espacio muestral, experimento aleatorio, evento y tipos de eventos, según distintas clasificaciones. Has aprendido a graficar los eventos con diagramas de Venn y a definirlos. También estudiaste los axiomas de la probabilidad, tan importantes y aplicables en los distintos problemas.

Asignación de probabilidades

En esta lectura nos centramos en la asignación de probabilidades a los distintos tipos de eventos. Aprendiste la regla de la adición y su aplicación a eventos mutuamente excluyentes y no mutuamente excluyentes. Aprendiste a distinguir entre eventos dependientes e independientes y a aplicar la regla de la multiplicación. Diferenciaste entre los conceptos de probabilidad condicional, conjunta y marginal, y las fórmulas que las rigen.

Tablas de contingencia y árboles de decisión

Aquí aprendiste a graficar los eventos de dos maneras: en tablas de contingencia y árboles de decisión. Todas las situaciones estudiadas en la lectura 2 se pueden expresar mediante alguna de estas dos herramientas para visualizarlas mejor. Asignaste el valor de probabilidad a los eventos pero a través de herramientas gráficas.

Probabilidad Total y Teorema de Bayes

Sobre la base de lo estudiado en las lecturas anteriores, definimos el teorema de la probabilidad total, basado en las probabilidades marginales, y concluimos con el teorema de Bayes que utiliza probabilidades condicionales y marginales. Este teorema tiene una aplicación directa en la toma de decisiones.

Referencias

Anderson, D. R.; Sweeney, D. J. y Williams, T. A. (2008). *Estadística para Administración y Economía*. México: Ed. Cengage Learning Editores, S.A.

Ingeniat [ingeniat]. (3 de marzo de 2011). *UDEM Estadística para negocios Teorema de Bayes* [YouTube]. Recuperado de https://www.youtube.com/watch?v=pokPfPLON5w

Levin, R. y Rubin, D. (2012). *Estadística para Administración y Economía* (7.ma.ed.). México: Pearson.