Regla de Bayes

El Teorema o Regla de Bayes nos brinda un método para contestar algunas preguntas muy importantes. En su esencia, esta regla nos indica cuál información es necesaria tener y el método para invertir la condición cuando calculamos una probabilidad condicional: si A y B son eventos y conocemos $P(A \mid B)$, $P(B)$, $P(A \mid B^c)$, entonces podemos calcular $P(B \mid A)$. La necesidad de calcular este último valor a partir de la información disponible es imprescindible para entender las consecuencias de algunas de nuestras decisiones.

En esta sección estudiaremos varios ejemplos que luego generalizaremos para obtener formalmente la regla de Bayes. Esto ejemplos a su vez ilustrarán cuán importante y somunes son las situaciones donde es necesario usar esta regla. En muchas ocasiones nos hacemos preguntas sobre temas que no están, a primera vista muy relacionados entre sí. Sin embargo, luego de un análisis ponderado, comenzamos a notar que las respuestas o los métodos para obtenerlas guardan algunas relaciones entre sí.

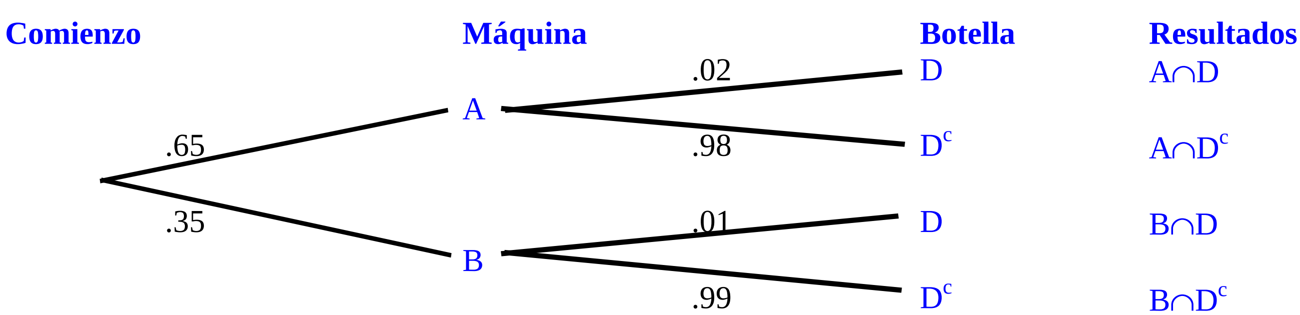
Ejemplo 1

Considera una fábrica de botellas que cuenta con dos máquinas para producir sus botellas. En esa fábrica se producen 10,000 botellas al día. La máquina A produce 6,500 botellas diarias de las cuales el 2% son defectuosas. La máquina B produce 3,500 botellas cada día de las cuales el 1% son defectuosas.

Pregunta

El inspector de calidad de la compañía selecciona una botella al azar y encuentra que está defectuosa. ¿Cuál es la probabilidad de que la botella haya sido producida por la máquina A?

Para visualizar mejor los datos, los organizamos en un diagrama de árbol. Denotamos por A el evento de que la botella seleccionada haya sido producida por la máquina A y por B el evento de que haya sido producida por la máquina B. El evento de que la botella seleccionada sea defectuosa se denota por D, su complemento D^c representa una botella que no es defecuosa.



La probabilidad de que una botella cualquiera haya sido producida por la máquina A es .65, pues de las 10,000 producidas, 6,500 son producidas por A. Nos interesa calcular $P(A \mid D)$, la cual no se puede obtener de forma directa de los datos o del árbol que los representa. Para esto recurrimos directamente a la definición de probabilidad condicional: $P(A \mid D) = P(A \cap D) / P(D)$.

Las cantidades $P(A \cap D)$ y $P(A)$ se pueden obtener del árbol. Para que una botella seleccionada al azar sea una defectuosa producida por la máquina A, debemos seleccionar primero la máquina A y de las botellas producidas allí seleccionar una defectuosa. Tenemos que $P(A \cap D) = P(A)\, P(D \mid A)$, lo que equivale a hacer la travesía en el árbol desde su raíz o comienzo hasta la hoja donde obtenemos el resultado $A \cap D$. Así $P(A \cap D) = .65 \times .02$.

Para encontrar P(D) debemos darnos cuenta que una botella defectuosa puede ser producida de la máquina A o de la B. Si examinamos las hojas del árbol, vemos que hay dos lugares donde obtenemos una botella defectuosa, $A \cap D$ o $B \cap D$. Esto equivale a hacer una travesía por uno de caminos en el árbol. Estos caminos son mutuamente excluyentes, pues si caminamos por uno no podemos estar caminando por el otro. Según se muestra en la figura de al lado, el evento $D = (A \cap D) \cup (B \cap D)$ y su probabilidad es entonces calculada $P(D) = P(A \cap D) + P(B \cap D)$.

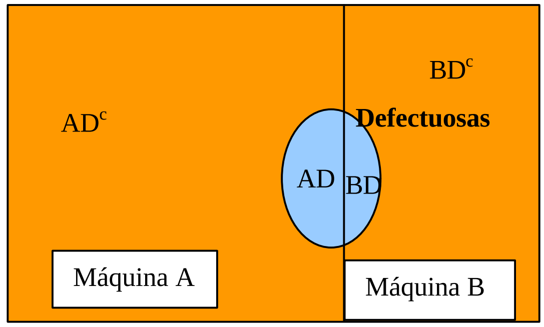


Figura 1 Partición de las botellas defectuosas de acuerdo a la máquina

El primero de estos términos $P(A \cap D)$ ya había sido calculado. El segundo se obtiene de forma similar. Obtenemos entonces que $P(B \cap D) = P(B)\, P(D \mid B)$. Uniendo estos resultados tenemos que $P(D) = P(A)\, P(D \mid A) + P(B)\, P(D \mid B)$. Finalmente podemos calcular la probabilidad deseada:

$$P(A \mid D) = \frac{\mathrm{P(A)\ P(\ D \mid A)}}{\mathrm{P(\ D \mid A)P(A) + P(D \mid B)P(B)}} = \frac{\frac{6{,}500}{10{,}000} \times .02}{\left(\frac{6{,}500}{10{,}000} \times .02\right) + \left(\frac{3{,}500}{10{,}000} \times .01\right)} = \frac{.013}{.013 + .0035} = .788$$

Esto quiere decir que una vez sabemos que una botella seleccionada al azar está defectuosa, la probabilidad de que haya sido producida por la máquina A es .788. Dicho de otra manera, de todas las botellas defectuosas produdidas, aproximadamente el 79% son producidas por la máquina A.

Pregunta

¿Cómo se puede explicar que la máquina A produzca el 79% de las botellas defectuosas?

Este hecho se debe a dos factores. El primero es que la máquina A produce casi el doble de botellas que la máquina B. Aún si la tasa de botellas defectuosas fuera la misma para ambas máquinas, por el mero hecho de producir un mayor número de botellas, la máquina A produciría casi el doble de defectuosas de la máquina B. El segundo factor es que la tasa de producción de defectuosas de la máquina A es el doble de la correspondiente de la máquina B. En este caso, aún si ambas máquinas produjeran la misma cantidad de botelllas, las producidas por la máquina A contendrían el doble de botellas defectuosas que las que vienen de la máquina B.

Ejemplo 2

El gobierno de Puerto Rico aprobó una ley para hacer obligatorio que los cerca de 200,000 empleados públicos se sometan a una prueba para detectar si son usuarios de drogas. Se estima que el 1% de los empleados públicos del país son usuarios de drogas. La prueba que se ofrece muestra un resultado positivo en el 98% de los casos en que se le administra a una persona que usa drogas, es decir, detecta el 98% de los usuarios de drogas. De manera similar, si la persona no usa droga alguna, la prueba arroja un resultado negativo en el 99% de los casos.

Pregunta

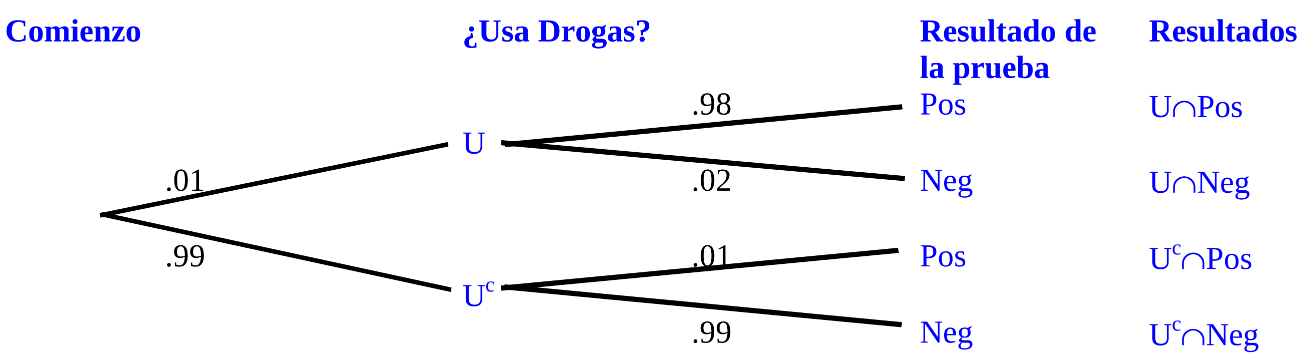
Se selecciona un empleado al azar, se le administra la prueba y se obtiene un resultado positivo. ¿Cuál es la probabilidad de que la persona sea un usuario de drogas?

En este ejemplo también podemos representar los datos mediante un diagrama de árbol. Para simplificar el trabajo, denotamos por U el evento de que la persona sea un usuario de drogas, por Pos el evento de que la prueba resulte positiva y por Neg el evento de que la prueba resulte negativa. En el lenguaje de la epdemiología, al por ciento de personas de la población de interés que poseen la característica deseada, en este caso usuarios de drogas, se le llama la **prevalencia**. En este ejemplo la prevalencia del uso de drogas es del 1%.

A la capacidad de una prueba para detectar aquellas personas que poseen la característica de interés se le llama la **sensitividad** de la prueba. La sensitividad describe el por ciento de personas cuyo resultado de la prueba sería positivo de entre aquellas que poseen la característica deseada, en este caso, usan drogas. La sensitividad de esta prueba para la detección de usuarios de drogas es 98%.

Otra medida es la **especificidad**. Esta indica el por ciento de personas cuyo resultado de la prueba sería negativo de entre aquellas que no poseen la característica deseada, en este caso, usan drogas. La especificidad de esta prueba para la detección de usuarios de drogas es 99%.

Usamos la notación que establecimos de estos eventos para escribir las probabilidades asociadas a ellos. Así, la prevalencia nos indica que si seleccionamos una persona al azar de entre los 200,000 empleados públicos, tenemos $P(U) = .01$. La sensitividad de la prueba nos dice que $P(Pos \mid U) = .98$ y de la misma manera la especificidad se traduce a $P(Neg \mid U^c) = .99$.[1]



La pregunta que nos interesa contestar es: ¿cuánto es $P(U \mid Pos)$? Al igual que antes la contestación a esta pregunta no se puede obtener de forma directa de los datos o del árbol que los representa. Recurrimos nuevamente a la definición de probabilidad condicional: $P(U \mid Pos) = P(U \cap Pos) / P(Pos)$.

[1] Nota que estas dos últimas probabilidades no se pueden sumar, ya que son valores obtenidos bajo condiciones distintas, la base de comparación no es la misma. En el primer caso comparamos la cantidad de resultados positivos del total de usuarios de drogas mientras que en el segundo comparamos la cantidad de resultados negativos del total de personas que no usan drogas.

Podemos usar el árbol para obtener las cantidades $P(U \cap Pos)$ y $P(Pos)$. Siguiendo el mismo proceso de antes, tenemos que $P(U \cap Pos) = P(U)\,P(Pos \mid U)$, lo que equivale a caminar por el árbol desde la raíz hasta la hoja donde obtenemos el resultado $U \cap Pos$. Así $P(U \cap Pos) = .01 \times .98$.

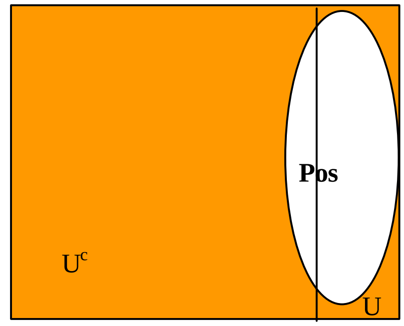


Figura 2 Partición de las personas de acuerdo al resultado de la prueba

Para encontrar $P(Pos)$ vemos que la prueba puede arrojar un resultado positivo cuando la persona es un usuario de drogas o en el caso en que no lo sea. Por esta razón hay dos caminos en el árbol donde obtenemos un resultado positivo de la prueba. Al igual que antes, estos caminos son mutuamente excluyentes. Según se muestra también en el diagrama de al lado, el evento $Pos = (U \cap Pos) \cup (U^c \cap Pos)$ y su probabilidad es $P(Pos) = P(U \cap Pos) + P(U^c \cap Pos)$.

El término $P(U \cap Pos)$ ya había sido calculado, el segundo se obtiene de forma similar. Obtenemos entonces que $P(U^c \cap Pos) = P(U^c)\,P(Pos \mid U^c)$. Uniendo estos resultados tenemos que $P(Pos) = P(U)\,P(Pos \mid U) + P(U^c)\,P(Pos \mid U^c)$. Finalmente podemos calcular la probabilidad deseada:

$$P(U \mid Pos) = \frac{\mathrm{P(U)\,P(\,Pos \mid U)}}{\mathrm{P(U)\,P(\,Pos \mid U) + P(U^c)\,P(Pos \mid U^c)}} = \frac{.01 \times .98}{(.01 \times .98) + (.99 \times .01)} = \frac{.0098}{.0098 + .0099} = .497$$

La contestación a la pregunta que hicimos es .497, es decir, la probabilidad de que una persona seleccionada al azar entre los 200,000 empleados sea un usuario de drogas si la prueba da positivo, es .497.

Pregunta

De la población a la que se administra la prueba, ¿cuántos resultados positivos esperarías observar?¿cuántos falsos positivos habría? ¿Cómo es posible que con una prueba que tiene una sensitividad y una especificidad tan altas, más de la mitad de los resultados positivos corresponden a personas que no son usuarios?

Esta última pregunta se puede contestar examinando cuidadosamente el numerador y el denominador de $P(U \mid Pos)$. Si no contamos con una prueba de mejor sensitividad y especificidad que ésta, ¿qué podemos hacer? Esto quiere decir que no podemos cambiar $P(Pos \mid U)$ ni $P(Neg \mid U^c)$. Sólo podemos trabajar con $P(U)$. Este valor sólo puede cambiar si cambiamos la población de la cual seleccionamos las personas a quienes se administrará la prueba.

Pregunta

Discute los costos asociados a ofrecer pruebas de drogas a 200,000 personas para detectar a 2,000 usuarios. ¿Hay costos no económicos? ¿Existe un plan para ayudar a las personas que son usuarios de drogas y a los que reciben un falso positivo? ¿Crees estas pruebas ofrecen una solución al problema social del uso de drogas? Supongamos que P(U) = .25, encuentra P(U | Pos). ¿Cómo lograrías este aumento en P(U) en la realidad?

Ejemplo 3

Tenemos una caja con 5 canicas, dos de ellas son rojas y las otras tres son azules. Se selecciona una canica al azar, sin mirarla la guardamos en el bolsillo. Luego seleccionamos otra canica al azar. Esta segunda canica era de color rojo. ¿Cuál es la probabilidad de que la primera canica haya sido también roja?

En la sección anterior resolvimos una situación similar imaginando que antes de seleccionar la primera canica, hemos mirado dentro de la caja y removido la canica que observaremos en nuestra segunda tentativa. En efecto, hemos reservado la segunda canica. La primera canica sólo puede ser seleccionada de entre las restantes 4 canicas, de las cuales 1 de ellas es roja. Por esta razón la probabilidad deseada es 1/4.

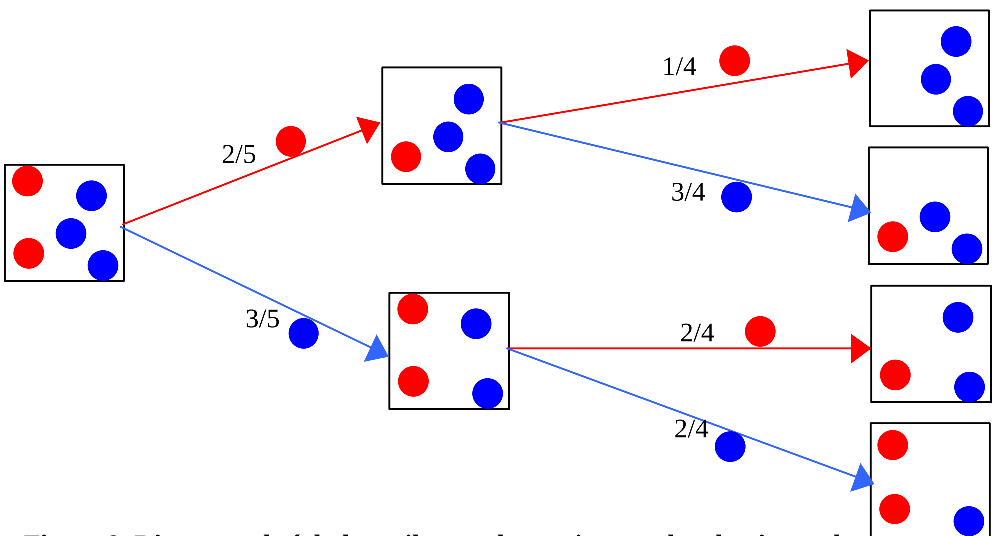


Figura 3. Diagrama de árbol que ilustra el experimento de seleccionar dos canicas de una caja

Como nuestro interés es encontrar un método formal para contestar estas preguntas, procedemos al análisis formal de la situación. Para facilitar el trabajo denotamos el evento de que la primera canica seleccionada es roja por R_1 y el evento de que la segunda sea roja por R_2. Hacemos lo propio para las canicas azules. Entonces la probabilidad que buscamos es $P(R_1 \mid R_2)$.

La definición de probabilidad condicional nos permite escribir $P(R_1 \mid R_2) = P(R_1 \cap R_2) / P(R_2)$. Para encontrar el numerador usamos nuevamente la definición de probabilidad condicional y escribimos $P(R_1 \cap R_2) = P(R_2 \mid R_1)P(R_1)$. Aquí hemos invertido la condición. Debemos condicionar en el evento que ocurre primero, pues es la forma natural de realizar el experimento.

Ahora calculamos el denominador. Para encontrar $P(R_2)$ descomponemos el evento R_2 en dos eventos disyuntos, tal como en la figura de al lado: $R_2 = (B_1 \cap R_2) \cup (R_1 \cap R_2)$. De esta manera obtenemos la probabilidad

$P(R_2) = P(B_1 \cap R_2) + P(R_1 \cap R_2)$.

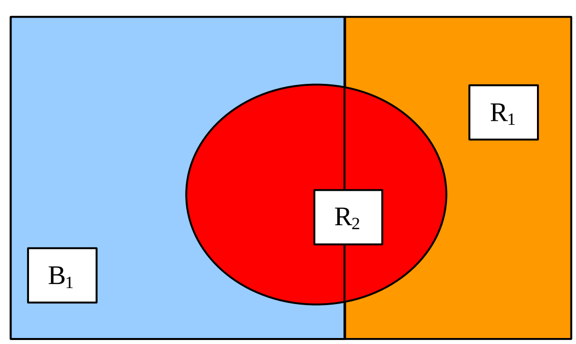


Figura 4 Partición del color de la segunda canica de acuerdo al color de la primera

Ya habíamos encontrado el segundo término de arriba, nos falta encontrar el primero. Para esto usamos nuevamente la definición de probabilidad condicional y obtenemos $P(B_1 \cap R_2) = P(R_2 \mid B_1)P(B_1)$. Ahora podemos escribir el denominador $P(R_2) = P(R_2 \mid R_1)P(R_1) + P(R_2 \mid B_1)P(B_1)$. Finalmente podemos escribir el resultado deseado: $P(R_1 \mid R_2) = \dfrac{P(R_2 \mid R_1)P(R_1)}{P(R_2 \mid R_1)P(R_1) + P(R_2 \mid B_1)P(B_1)} = \dfrac{\frac{1}{4}\times\frac{2}{5}}{\left(\frac{1}{4}\times\frac{2}{5}\right)+\left(\frac{2}{4}\times\frac{3}{5}\right)} = \dfrac{1}{4}.$

No debe sorprendernos que observamos el mismo resultado que antes. Esto sirve para verificar y justificar el razonamiento que hicimos para obtener la contestación.

Ejemplo 4

En el 1991 los contribuyentes de Puerto Rico sometieron un total de 1,320,600 planillas de contribución sobre ingresos al Departamento de Hacienda. Los datos se desglosan en la siguiente tabla por nivel de ingreso y si la planilla se somete conjunta o por separado.

Nivel de ingreso	*Planillas conjuntas rendidas (miles)*	*Planillas separadas rendidas (miles)*	*Total*
menos de 20,000	457.5	565.7	**1,023.2**
20,000 a 30,000	88.7	84.9	**173.6**
30,000 a 50,000	55.3	33.2	**88.5**
50,000 o más	20.1	15.2	**35.3**
Total	**621.6**	**699.0**	**1,320.6**

Fuente: Reforma Contributiva en Puerto Rico 1994. Estudio Técnico. Editorial UPR.

Tabla 1. Desglose de contribuyentes en Puerto Rico en el año 1991

Pregunta

El Secretario de Hacienda selecciona una planilla al azar.¿Cuál es la probabilidad de que la planilla haya sido sometida en forma conjunta si el nivel de ingreso en ella era menor de $20,000? ¿Cuál es la probabilidad de que el nivel de ingreso en ella era menor de $20,000 si la la planilla fue sometida en forma conjunta?

Para la primera pregunta, nos interesa conocer la probabilidad de que la planilla haya sido sometida en forma conjunta si el nivel de ingreso que refleja es menor de $20,000. Examinamos la primera fila de la Tabla 1. Vemos que de todas las planillas, 1,023,000 reflejan un ingreso menor de $20,000. De esas 457,500 fueron sometidas en forma conjunta, así la probabilidad deseada es: 475.5/1,023.0 = .46.

Para contestar la segunda pregunta es necesario comenzar examinando la columna correspondiente a las planillas que se sometieron en forma conjunta. El total que se indica al final de esa columna será nuestra base de comparación. Se sometieron 621,600 planillas conjuntas de las cuales 457,500 corresponden a planillas que además indicaron un ingreso menor de $20,000. Por lo tanto la probabilidad buscada es 475.5/621.6 = .76.

Si usamos los resultados obtenidos para describir la población de planillas recibidas, vemos que de las planillas que reflejaron un ingreso menor de $20,000, el 46% correspondian a planillas sometidas en forma conjunta. En el otro caso, de todas las planillas sometidas en forma conjunta, el 76% corresponde a planillas que reflejan un ingreso menor de $20,000. Estos dos porcentajes no son lo mismo ni significan lo mismo, reflejan bases de comparación distintas, lo que exige seamos muy cuidadosos en nuestro análisis.

Cuando tenemos una tabla con datos es muy fácil calcular estas probabilidades condicionales. Aunque en la práctica no usariamos el método de análisis que nos ofrece la regla de Bayes para analizar estas situaciones, procedemos a hacerlo para ilustrar su desarrollo. El fin que perseguimos es de presentar el análisis en una forma análoga a los problemas anteriores para de ahí obtener la regla deseada.

Denotemos por C el evento de que la planilla se somete en forma conjunta, por S el evento de que la planilla se somete por separado y por I el ingreso reflejado en la planilla.Usando probabilidad condicional podemos escribir así:

$$\mathrm{P}(\mathrm{C} \mid \mathrm{I} < \$20{,}000) = \frac{\mathrm{P}(\mathrm{C} \cap \{\mathrm{I} < \$20{,}000\})}{\mathrm{P}(\{\mathrm{I} < \$20{,}000\})} = \frac{\mathrm{P}(\mathrm{C} \cap \{\mathrm{I} < \$20{,}000\})}{\mathrm{P}\left(\{\mathrm{C} \cap \{\mathrm{I} < \$20{,}000\}\} \cup \{\mathrm{S} \cap \{\mathrm{I} < \$20{,}000\}\}\right)}.$$

$$= \frac{P\left(\mathrm{C} \cap \{\mathrm{I} < \$20{,}000\}\right)}{\mathrm{P}\left(\mathrm{C} \cap \{\mathrm{I} < \$20{,}000\}) + \mathrm{P}(\mathrm{S} \cap \{\mathrm{I} < \$20{,}000\}\right)}.$$

Podemos usar nuevamente la definición de probabilidad condicional para reescribir el denominador: $P(C \cap \{I < \$20{,}000\}) = P(I < \$20{,}000 \mid C)\, P(C)$ y $P(S \cap \{I < \$20{,}000\}) = P(I < \$20{,}000 \mid S)\, P(S)$. Así escribimos

$$\mathrm{P}(\mathrm{C} \mid \mathrm{I} < \$20{,}000) = \frac{P(I < \$20{,}000 \mid C)P(C)}{\mathrm{P}(\mathrm{I} < \$20{,}000 \mid \mathrm{C})\mathrm{P}(\mathrm{C}) + \mathrm{P}(\mathrm{I} < \$20{,}000 \mid \mathrm{S})P(S)}.$$

Pregunta

Representa estos datos usando un diagrama de Venn.

Obviamente si lo único que nos interesa es contestar las preguntas que nos hicimos en cada uno de los ejemplos, no hubiéramos procedido a desarrollar estas ideas. De hecho, en el ejemplo de las planillas de contribución y en el de las canicas era mucho mas fácil y directo resolverlos sin hacer alusión a esta metodología mas formal. En situaciones prácticas no lo haríamos. Sin embargo, para el descubrimiento y desarrollo de resultados generales en las matemáticas es necesario pasar por este proceso que nos ayuda a identificar los elementos importantes del problema.

Debemos calcular una probabilidad condicional cuyo valor no se puede obtener directamente de los datos. A través de la definición, expresamos esa probabilidad condicional en términos de otros eventos cuyas probabilidades sí conocemos o podemos calcular. Finalmente, descomponemos el evento deseado, sin condición alguna, en la unión de eventos disyuntos cuyas probabilidades conocemos y calculamos la probabilidad buscada. Antes de continuar debemos formalizar el resultado que nos facilita el cálculo de la probabilidad del evento descompuesto en partes disyuntas.

Teorema 1 (Fórmula de la probabilidad total)

Sea S un espacio muestral, P una medida de probabilidad en S y B un evento en S. Sea $A_1, A_2, \ldots, A_n$ una partición de S, es decir, eventos disyuntos tal que $S = \bigcup_{i=1}^{n} A_i$, entonces

$$P(B) = \sum_{i=1}^{n} P(A_i)P(B \mid A_i).$$

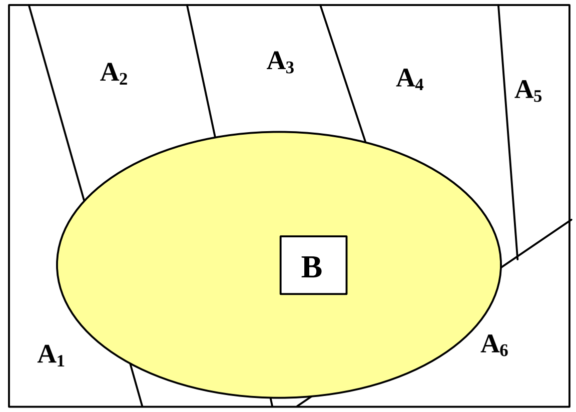


Figura 5 Partición del evento B

Prueba.

Podemos escribir, usando el hecho de que $A_1, A_2, \ldots, A_n$ es una partición de S, tenemos que

$B = B \cap S = B \cap (\bigcup_{i=1}^{n} A_i) = \bigcup_{i=1}^{n} (B \cap A_i)$. Usamos el hecho de que cada uno de los eventos $B \cap A_i$, i = 1, …, n son disyuntos y la definición de probabilidad condicional para calcular la probabilidad de B:

$$P(B) = P(\bigcup_{i=1}^{n} (B \cap A_i)) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(A_i)P(B \mid A_i).$$

En la figura de arriba vemos la partición de B. En este ejemplo, cada término $B \cap A_i$ i = 1, 2, 3, 4, 5, 6 corresponde a cada uno de los "pedacitos" en que dividimos a B. vemos en la figura que la intersección de A_6 con el evento B es vacía, por lo cual $P(B \cap A_6) = 0$. Para calcular la probabilidad de cada pedacito, usamos la definición de probabilidad condicional, así $P(B \cap A_i)=P(B \mid A_i)P(A_i)$, i = 1, 2, 3, 4, 5, 6.

Imaginemos que la figura 5 representa un tablero de dardos y que la probabilidad de caer en determiando región del tablero es igual a su área. Tiramos el dardo y sabemos que cayó en la región marcada por B, entonces, ¿cuál es la probabilidad de que haya caido en A_2? Para contestar esta pregunta necesitamos la regla de Bayes.

Teorema 2 (Regla de Bayes)

Sea S un espacio muestral, P una medida de probabilidad en S y B un evento en S. Sea $A_1, A_2, \ldots, A_n$ una partición de S, entonces para cada i = 1, 2, 3, ..., n tenemos

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{\sum_{j=1}^{n} P(A_j)P(B \mid A_j)}.$$

Prueba.

Dela definición de probabilidad condicional tenemos $P(A_i \mid B) = \frac{P(B \cap A_i)}{P(B)}$. Para calcular el numerador usamos la defínción de nuevo y obtenemos $P(B \cap A_i) = P(B \mid A_i)P(A_i)$. El denominador se obtiene aplicando la Fórmula de probabilidad total, así obtenemos el resultado deseado.

$$P(A_i \mid B) = \frac{P(B \cap A_i)}{P(B)} = \frac{P(A_i)P(B \mid A_i)}{P(B)} = \frac{P(A_i)P(B \mid A_i)}{\sum_{j=1}^{n} P(A_j)P(B \mid A_j)}.$$

Pregunta

Expresa los problemas presentados en los ejemplos 1,2, 3 y 4 en términos de la regla de Bayes. Indica a qué corresponden la partición y el evento B.

Problemas y ejercicios

1. Una fábrica tiene tres máquinas para producir bombillas. La máquina A produce el 35% del total de bombillas, la máquina B produce el 50% y la máquina C produce el 15% de las bombillas. Sin embargo, las máquinas no son perfectas, la máquina A daña el 10% de las bombillas que produce. La máquina B daña el 5% y la máquina C daña el 20%.
    a. Representa estos datos en un diagrama de árbol.
    b. La fábrica produce 10,000 bombillas sin defectos en un día. ¿Cuántas de éstas corresponden a la máquina A? ¿Cuántas daña en un día?
    c. Si seleccionamos una bombilla de la máquina C, ¿cuál es la probabilidad de que esté defectuosa?
    d. Luego de fabricadas, pero antes de probarlas, las bombillas se colocan juntas en un salón. Si se selecciona una bombilla al azar, ¿cuál es la probabilidad de que esté defectuosa?
    e. Si se comprueba que una bombilla está defectuosa, ¿cuál es la probabilidad de que provenga de la máquina B?

2. Una muestra de 400 adultos varones con angina de pecho son clasificados por peso y estatura como sigue:

	Peso (libras)			
Edad (años)	**130-149**	**150-169**	**170-189**	**190 o más**
30-39	10	20	20	40
40-49	10	15	50	70
50-59	5	15	50	40
60-69	5	10	15	25

Un individuo se selecciona al azar de entre los 400 participantes. Encuentra la probabilidad de que:
    a. tiene entre 40-49 años de edad
    b. está en el intervalo de 40-49 años y pesa 170-189 libras
    c. está en el intervalo 40-49 años ó entre 60-69 años
    d. está en el intervalo 30-39 o 50-59 años y pesa 150-169 libras
    e. pesa menos de 170 libras
    f. pesa menos de 190 libras y es mayor de 49 años
    g. pesa menos de 170 libras dado que es menor de 50 años
    h. ¿Son los eventos {tiene 60-69 años} , {pesa 130-149 libras} independientes? Explica.
    i. ¿Son los eventos {tiene 50-59 años} , {pesa 130-149 libras} mutuamente excluyentes? Explica.

3. Un estudio neurológico sobre la relación entre la presión sanguínea alta y la incidencia de derrame cerebral encontró que:
    a. para personas mayores de 70 años, el 10% tendrá un derrame dentro de los próximos cinco años
    b. de todos los pacientes de 70 años o más que han tenido un derrame, el 40% tenía presión alta
    c. para personas de 70 años o mas que no ha sufrido de derrame, el 20% tiene presión alta.

Un paciente de 74 años visita a su médico y éste le encuentra con la presión alta. ¿Cuál es la probabilidad de que sufra un derrame cerebral en los próximos cinco años?