El homenaje de hoy:

- Thomas Bayes 1702-1761, matemático y presbítero británico que descubrió la una fórmula para la probabilidad inversa.
- Este descubrimiento, que fundamenta una amplia rama de la estadística, fue publicado luego de su muerte por Richard Price

Bioestadística (Clase 4):

Fórmula de probabilidad total, fórmula de Bayes, Independencia

Ernesto Mordecki

CMAT, Facultad de Ciencias, Universidad de la República.

Uruguay

Contenidos de la clase

Probabilidad condicional

Fórmula de la probabilidad total

Fórmula de la probabilidad total

Fórmula de Bayes

Independencia de sucesos

Probabilidad condicional

Definición. Dados dos sucesos A y B, con $P(A) > 0$ definimos la probabilidad condicional de B dado A, mediante

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

Nota: Precisamos que $P(A) > 0$ porque aparece en el denominador

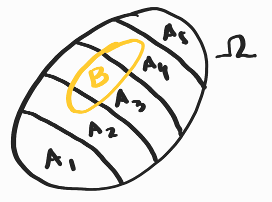
Es útil escribir la fórmula anterior como:

$$\boxed{P(B \cap A) = P(B \mid A) \times P(A).}$$

que llamamos la fórmula dos.

Fórmula de la probabilidad total

- Supongamos que el espacio Ω se particiona en los conjuntos $A_1, \ldots, A_n$ y tenemos un suceso de interés B:



- La fórmula de la probabilidad total nos permite calcular B mediante

$$\boxed{P(B) = P(B \mid A_1)P(A_1) + \cdots + P(B \mid A_n)P(A_n)}$$

que se llama fórmula de la probabilidad total.

Demostración

- Como los $A_1, \ldots, A_n$ parten el espacio, también parten el conjunto B, es decir

$$B = (B \cap A_1) \cup \cdots \cup (B \cap A_n)$$

- Como estos conjuntos son disjuntos (por serlo los A_i) aplicamos la suma de probabilidades:

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

- Finalmente a cada sumando le aplicamos la regla dos[1]

$$P(B) = P(B \mid A_1)P(A_1) + \cdots + P(B \mid A_n)P(A_n).$$

que es la fórmula buscada. $\square$

[1] $P(B \cap A) = P(B \mid A)P(A)$

Revisitamos el problema del penal[2].

- Penal para Uruguay. Tiran Suárez, Cavani, o un Tercer jugador. Sea

$$G = \{\text{se convierte el gol}\}$$

- Vamos a `https://www.transfermarkt.es/` (2008-2021):
    - Suárez convirtió 48 penales y erró 9 de los 57 pateados
    - Cavani convirtió 57 y erró 13 de los 70 pateados en el mismo período.
    - Otros jugadores convirtieron 20 y erraron 12 de los 32 pateados.

[2] Recalculando

- Estimamos probabilidades a partir de estas frecuencias:
- Para Suárez:
$$P(G \mid S) = \frac{48}{57} = 0{,}84,$$
- Para Cavani:
$$P(G \mid C) = \frac{57}{70} = 0{,}81,$$
- Para T (un tercer jugador)
$$P(G \mid T) = \frac{20}{32} = 0{,}63.$$

- Patearon penales por Uruguay entre 2009 y 2021: **S**uárez 18 veces, **C**avani 10 veces, un **T**ercer jugador 32 veces (total 60):
$$\Omega = S \cup C \cup T.$$

- Estimamos probabilidades para el tirador:
$$P(S) = \frac{18}{60}, \quad P(C) = \frac{10}{60}, \quad P(T) = \frac{32}{60}.$$

- Tenemos todos los ingredientes:
$$\begin{aligned} P(G) &= P(G \mid S)P(S) + P(G \mid C)P(C) + P(G \mid T)P(T) \\ &= \frac{48}{57} \times \frac{18}{60} + \frac{57}{70} \times \frac{10}{60} + \frac{20}{32} \times \frac{32}{60} = \mathbf{0.72} \end{aligned}$$

Fórmula de Bayes

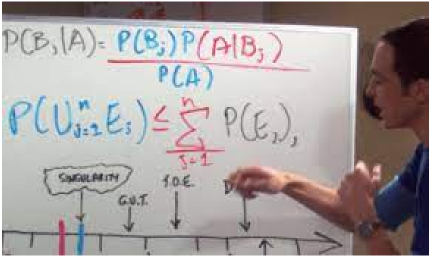
- Suponemos conocer $P(B \mid A)$ y queremos calcular $P(A \mid B)$.
- En el ejemplo anterior, sería: quien pateó el penal, dado que fue gol.
- Bayes hizo esta cuenta:

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

- Obtuvo su famosa fórmula

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

¿Lo conocen?



El físico Sheldon Cooper aplica la fórmula de Bayes, en el California Institute of Technology (Caltech)

- Está fórmula da origen a la estadística bayesiana
- $P(A)$ se llama probabilidad a priori
- $P(B \mid A)$ es la verosimilitud
- $P(A \mid B)$ se llama probabilidad a posteriori
- El cociente $\frac{P(B|A)}{P(B)}$ es el impacto de B en la probabilidad de A:

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

Ejemplo

- Supongamos que se tiró un penal y fue gol. ¿Cual es la probabilidad de que haya pateado Suárez?
- Es decir, queremos calcular $P(S \mid G)$.
- Aplicando la fórmula de Bayes:

$$P(S \mid G) = \frac{P(G \mid S)P(S)}{P(G)} = \frac{0{,}84}{0{,}72} \times 0{,}30 = 0{,}35$$

Fórmula general de Bayes

- La fórmula obtenida es

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

- Cuando no conocemos $P(B)$, si tenemos una partición de Ω con $A = A_1, \ldots A_n$, utilizamos en el denominador la fórmula de la probabilidad total. Obtenemos

$$\boxed{P(A_1 \mid B) = \frac{P(B \mid A_1)P(A_1)}{P(B \mid A_1)P(A_1) + \cdots + P(B \mid A_n)P(A_n)}}$$

Ejemplo[3]

Tenemos una fábrica que tiene una alarma que suena en caso de accidente ... Pero:

- La probabilidad de que haya un accidente es $0,1$.
- La probabilidad de que suene la alarma si hay un accidente es $0,97$
- La probabilidad de que suene si no hay accidente es $0,02$ (falsa alarma)
- Escuchamos sonar la alarma
- ¿cuál es la probabilidad de que no haya habido un accidente?

[3] Tomado de las notas del Curso de A. Cholaquidis

Identifiquemos los sucesos:

- I: hubo un acciente
- I^c: no hubo accidente
- A: sonó la alarma

Sabemos

- $P(I) = 0,1$ (hubo accidente)
- $P(A \mid I) = 0{,}97$ (sonó la alarma dado que hubo accidente)
- $P(A \mid I^c) = 0{,}02$ (sonó la alarma dado que no hubo accidente)
- Queremos calcular
$$P(I^c \mid A) = \frac{P(A \mid I^c)}{P(A)} P(I^c).$$

- Primero calculamos el denominador $P(A)$ por probabilidad total:

$$\begin{aligned} P(A) &= P(A \mid I)P(I) + P(A \mid I^c)P(I^c) \\ &= 0{,}97 \times 0{,}1 + 0{,}02 \times 0{,}9 = 0{,}115. \end{aligned}$$

- Ahora

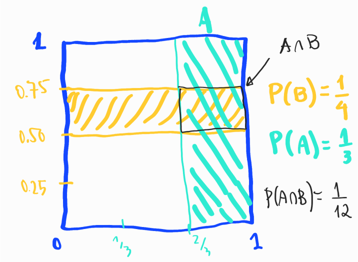
$$P(I^c \mid A) = \frac{P(A \mid I^c)}{P(A)} P(I^c) = \frac{0{,}02}{0{,}115} \times 0{,}9 = 0{,}157.$$

Sucesos independientes

- Decimos que dos sucesos A y B son independientes cuando se verifica

$$\boxed{P(A \cap B) = P(A) \times P(B).}$$

- La independencia es un concepto matemático.
- Si las probabilidades fuesen áreas, estos sucesos serían independientes:



Ejemplo

Supongamos que tiramos dos dados (el negro y el rojo)

- Veamos que el resultado del dado negro es independiente del dado rojo:
- Sea A el dado negro es par
- Sea B el dado rojo es menor que 3.

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	**(2,2)**	**(2,3)**	**(2,4)**	**(2,5)**	**(2,6)**
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	**(4,2)**	**(4,3)**	**(4,4)**	**(4,5)**	**(4,6)**
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	**(6,2)**	**(6,3)**	**(6,4)**	**(6,5)**	**(6,6)**

Cuadro: En negrita si el dado negro (el primero) es par. A la izquierda si el segundo dado es 1 o 2

- Tenemos $P(A) = \frac{18}{36} = \frac{1}{2}$.
- Además $P(B) = \frac{12}{36} = \frac{1}{3}$.
- Por último $P(A \cap B) = \frac{6}{36} = \frac{1}{6}$
- Como
$$P(A \cap B) = \frac{1}{6} = \frac{1}{2} \times \frac{1}{3} = P(A)P(B)$$
tenemos independencia de los sucesos A y B.

- Más en general, si N, R son los resultados respectivos de los dados del primer dado,
- El primer suceso es $N = k$ (el dado negro toma el valor k)
- El segundo suceso es $R = j$ (el dado rojo toma el valor j)
- Tenemos

$$P(N = k, R = j) = \frac{1}{36} = \frac{1}{6} \times \frac{1}{6} = P(N = k)P(R = j)$$

- Decimos que los resultados N y R son independientes.

Independencia y probabilidad condicional

- Supongamos que A y B son independientes, y que $P(A) > 0$. En ese caso tenemos

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

- La ocurrencia de A no modifica la probabilidad de B cuando son independientes.

Observaciones

- El suceso Ω es independiente de cualquier suceso A:

$$P(A \cap \Omega) = P(A) = P(A) \times 1 = P(A) \times P(\Omega)$$

- El suceso $\emptyset$ también es independiente de cualquier suceso A:

$$P(A \cap \emptyset) = P(\emptyset) = 0 = 0 \times P(A) = P(\emptyset) \times P(A)$$

- Si dos sucesos son disjuntos, y tienen probabilidad positiva, no pueden ser independientes:

$$0 = P(A \cap B) \neq P(A)P(B) > 0.$$

Mas observaciones

Si A, B son independientes, también lo son A, B^c, A^c, B y A^c, B^c. Por ejemplo

$$A = (A \cap B) \cup (A \cap B^c)$$

es una unión disjunta, y resulta

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

Despejando

$$P(A \cap B^c) = P(A) - P(A)P(B) = P(A)[1 - P(B)] = P(A)P(B^c)$$

nos da la independencia. Los otros resultados son análogos.