Probabilidad y Análisis de Datos

Daniel Fraiman

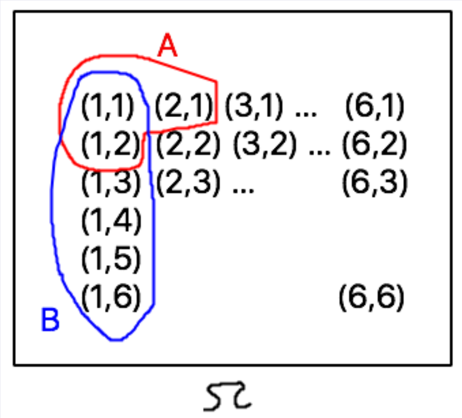
Maestría en Ciencia de Datos, Universidad de San Andrés

Probabilidad Condicional

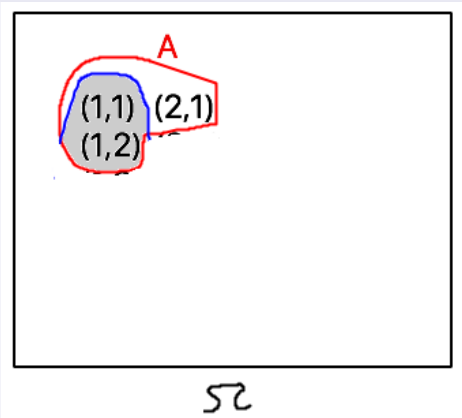
Probabilidad Condicional

Ejemplo

Tiro dos dados.
$A = \{$La suma es $<4\}$,
$B = \{$El primer dado es un $1\}$.
¿Cuál es la probabilidad de que el primer dado sea un 1 (B) sabiendo que la suma dio <4 (A)? $\equiv$ ¿$\mathbb{P}(B|A)$?



Probabilidad Condicional



¿Para el cálculo $\mathbb{P}(B|A)$ qué es lo que importa?

Probabilidad Condicional

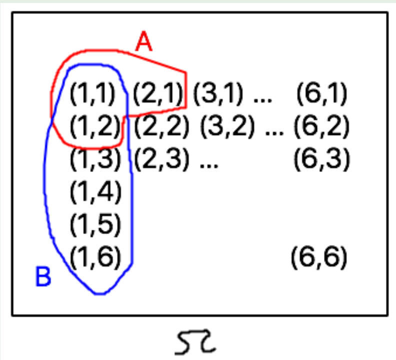
Probabilidad Condicional

Dado un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, $A \in \mathcal{F}$, con $\mathbb{P}(A) > 0$, definimos para todo $B \in \mathcal{F}$

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

Probabilidad Condicional: $\mathbb{P}(B|A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}$

Ejemplo: Tiro dos dados.



$A = \{\text{La suma es } {<}4\}$, $B = \{\text{El primer dado es un } 1\}$, ¿$\mathbb{P}(B|A)$?

- $\mathbb{P}(A) \stackrel{equip}{=} \frac{\#A}{\#\Omega} = \frac{3}{36}$
- $\mathbb{P}(B \cap A) \stackrel{equip}{=} \frac{\#A \cap B}{\#\Omega} = \frac{2}{36}$
- $\mathbb{P}(B|A) \stackrel{def}{=} \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} = \frac{2/36}{3/36} = 2/3$

Probabilidad Condicional

Probabilidad Condicional

La probabilidad condicional efectivamente es una probabilidad:

1. $\mathbb{P}(B|A) \geq 0$ para todo $B \in \mathcal{F}$.
2. $\mathbb{P}(\Omega|A) = 1$.
3. $\mathbb{P}\left(\bigcup_{i=1}^{\infty} B_i|A\right) = \sum_{i=1}^{\infty} \mathbb{P}(B_i/A)$ cuando los $B_i \cap B_j = \emptyset\ \forall i \neq j$.

Y por lo tanto también valen:

(a) $\mathbb{P}(\emptyset|A) = 0$

(b) $0 \leq \mathbb{P}(B|A) \leq 1$.

(c) Si $B_1 \subseteq B_2 \rightarrow \mathbb{P}(B_1|A) \leq \mathbb{P}(B_2|A)$.

(d) $\mathbb{P}\left(B^C|A\right) = 1 - \mathbb{P}(B|A)$.

(e) $\mathbb{P}(B_1 \cup B_2|A) = \mathbb{P}(B_1|A) + \mathbb{P}(B_2|A) - \mathbb{P}(B_1 \cap B_2|A)$.

Probabilidad Condicional

Ejemplo

La probabilidad de que mañana llueve y llegue tarde es 1/100. El servicio meteorológico anuncia que la probabilidad de que mañana llueva es 1/50.
¿Cuál es la probabilidad de que llegue mañana tarde si llueve?

$A = \{\text{mañana llueve}\}$, $B = \{\text{mañana llego tarde}\}$, $\mathbb{P}(A \cap B) = 1/100$ y $\mathbb{P}(A) = 1/50$.   ¿$\mathbb{P}(B|A)$?

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} = \frac{1/100}{1/50} = 1/2.$$

Probabilidad Condicional

Probabilidad Condicional

A partir de la definición de probabilidad condicional salen tres reglas:

1. Regla de multiplicación (útil para $\mathbb{P}(A \cap B)$).
2. Regla de probabilidad total (útil para $\mathbb{P}(B)$).
3. Regla de Bayes (útil para $\mathbb{P}(A|B)$ inversa).

Probabilidad Condicional

Regla de multiplicación

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

Ejemplo

El 60 % de los alumnos termina a tiempo la guía 1. El 80 % de los alumnos que no se atrasan con la primera guía realizan la guía 2 a tiempo. Mientras que el 70 % de los que no hicieron la guía 1 a tiempo, llegan a hacer la guía 2 a tiempo.

- ¿Qué porcentaje de los alumnos hace las guías 1 y 2 a tiempo?
- ¿Cuál es la probabilidad de que al elegir un alumno al azar este haya realizado las guías 1 y 2 a tiempo?

Regla de multiplicación: $\mathbb{P}(A \cap B) = \mathbb{P}(B|A)\,\mathbb{P}(A)$

Ejemplo

El 60 % de los alumnos termina a tiempo la guía 1. El 80 % de los alumnos que no se atrasan con la primera guía realizan la guía 2 a tiempo. Mientras que el 70 % de los que no hicieron la guía 1 a tiempo, llegan a hacer la guía 2 a tiempo.

¿Cuál es la probabilidad de que al elegir un alumno al azar este haya realizado las guías 1 y 2 a tiempo?

$A =$"Realiza la guía 1 a tiempo", $B =$"Realiza la guía 2 a tiempo"
Pregunta: ¿$\mathbb{P}(A \cap B)$?

Datos:

- $\mathbb{P}(A) = 0{,}6 \longleftrightarrow \mathbb{P}(A^C) = 0{,}4$
- $\mathbb{P}(B|A) = 0{,}8 \longleftrightarrow \mathbb{P}(B^C|A) = 0{,}2$
- $\mathbb{P}(B|A^C) = 0{,}7 \longleftrightarrow \mathbb{P}(B^C|A^C) = 0{,}3$

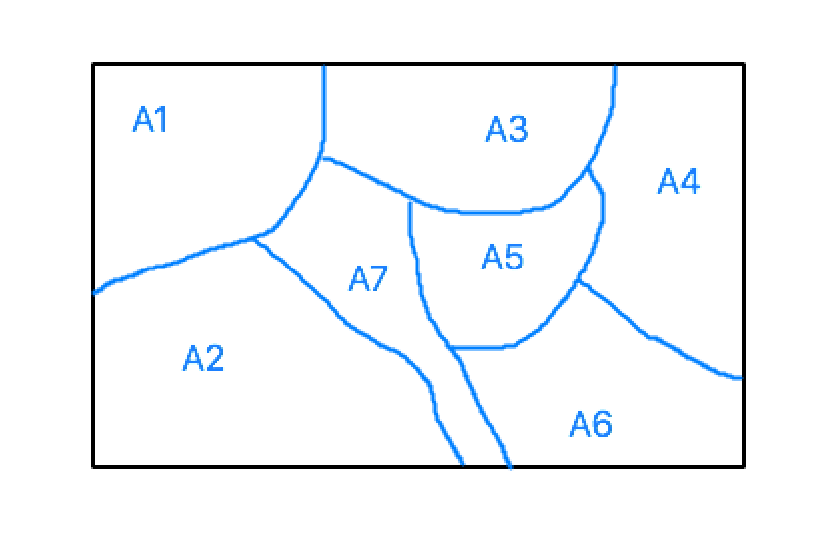
Regla de multiplicación: $\mathbb{P}(A \cap B) = \mathbb{P}(B|A)\,\mathbb{P}(A)$

$A =$"Realiza la guía 1 a tiempo", $B =$"Realiza la guía 2 a tiempo"
Pregunta: ¿$\mathbb{P}(A \cap B)$?

- Datos:
  - $\mathbb{P}(A) = 0{,}6 \longleftrightarrow \mathbb{P}(A^C) = 0{,}4$
  - $\mathbb{P}(B|A) = 0{,}8 \longleftrightarrow \mathbb{P}(B^C|A) = 0{,}2$
  - $\mathbb{P}(B|A^C) = 0{,}7 \longleftrightarrow \mathbb{P}(B^C|A^C) = 0{,}3$
- Respuesta: $\mathbb{P}(A \cap B) = \mathbb{P}(B|A)\,\mathbb{P}(A) = 0{,}8 \times 0{,}6 = 0{,}48$

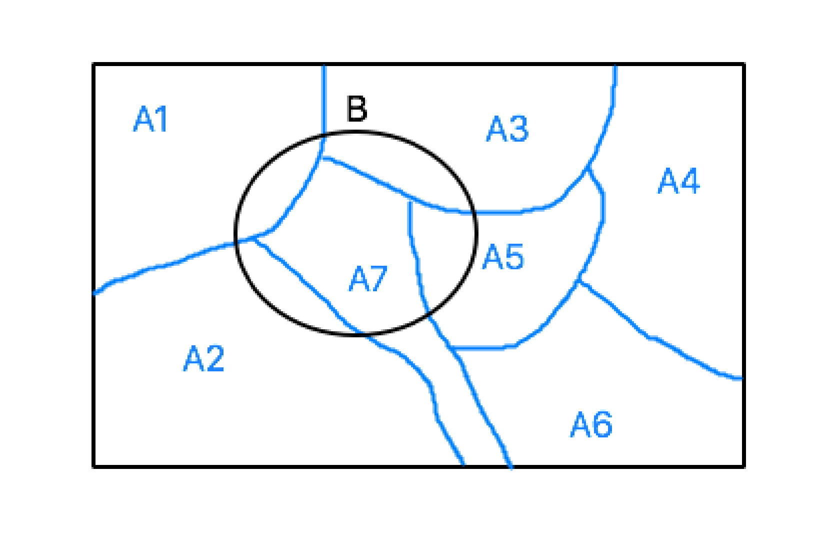
Probabilidad Condicional

Regla de probabilidad total



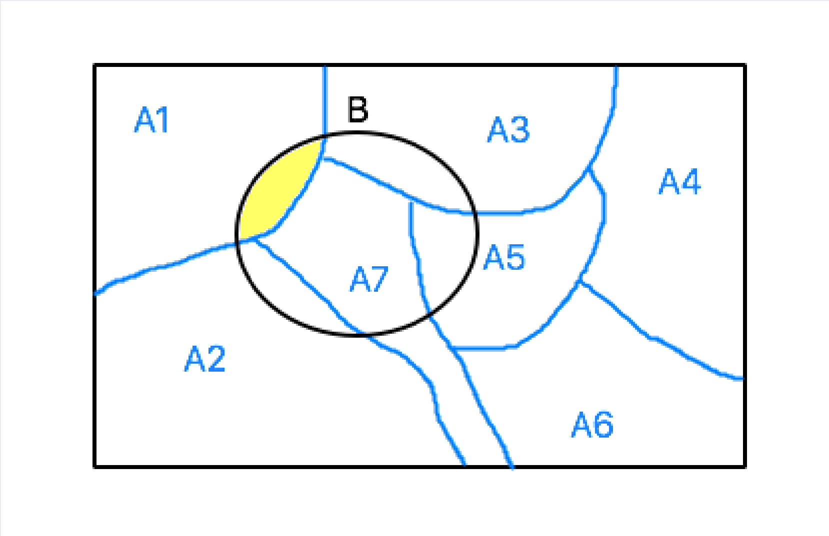
Probabilidad Condicional

Regla de probabilidad total



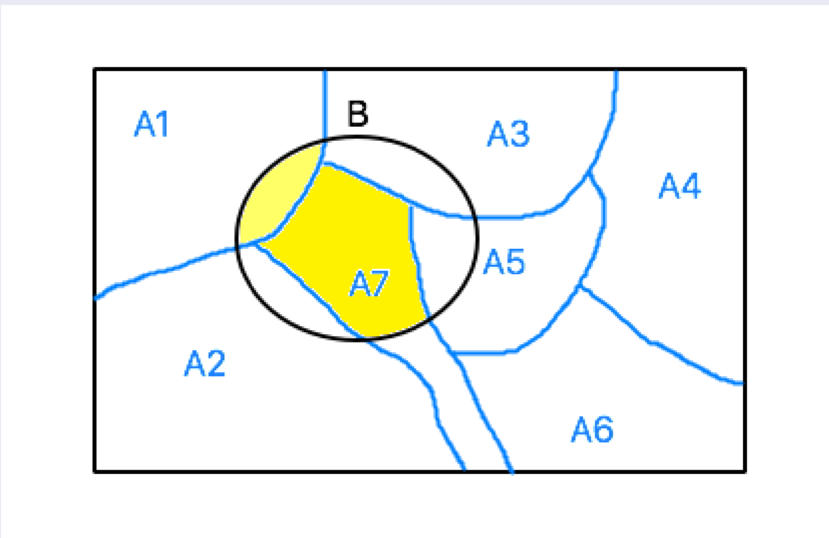
Probabilidad Condicional

Regla de probabilidad total



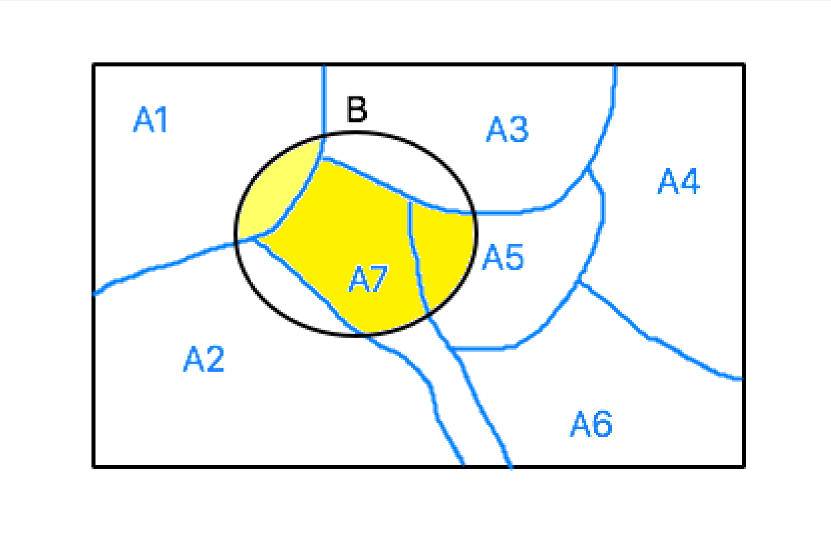
Probabilidad Condicional

Regla de probabilidad total



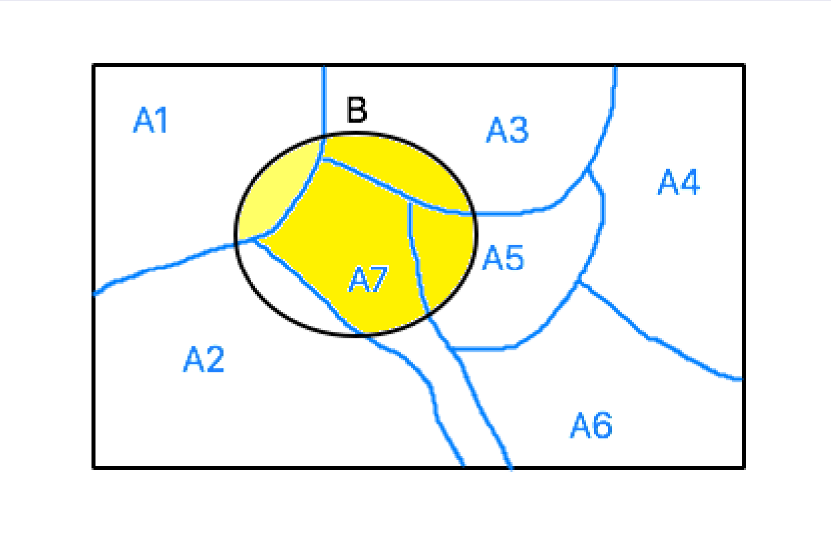
Probabilidad Condicional

Regla de probabilidad total



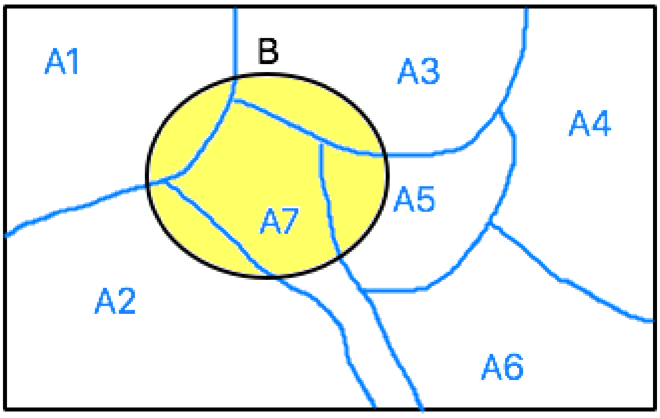
Probabilidad Condicional

Regla de probabilidad total



Probabilidad Condicional

Regla de probabilidad total

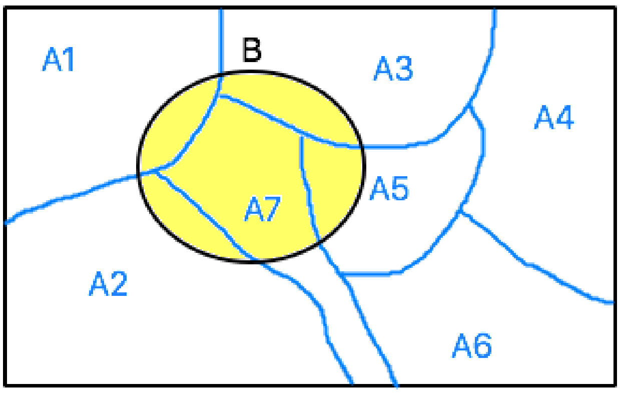


Probabilidad Condicional

Regla de probabilidad total

$P(B)=P(B \cap A1)+P(B \cap A2)+P(B \cap A3)+...+P(B \cap A7)$

$\boxed{P(B)=P(B|A1)P(A1)+P(B|A2)P(A2)+...+P(B|A7)P(A7)}$



Probabilidad Condicional

Regla de probabilidad total

Dado un evento B y A_1,A_2, …, A_n una partición disjunta de Ω ($\bigcup_{i=1}^{n} A_i = \Omega$ con $A_i \cap A_j = \emptyset$ para $i \neq j$) con $\mathbb{P}(A_i) > 0$ para $i = \{1, 2, \ldots, n\}$

$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(B|A_i)\,\mathbb{P}(A_i)$$

Regla de probabilidad total

Ejemplo

A =“Realiza la guía 1 a tiempo”, B =“Realiza la guía 2 a tiempo”
Pregunta: ¿$\mathbb{P}(B)$?

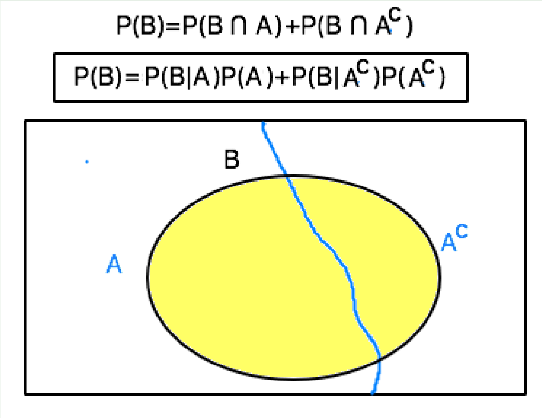
- Datos:
  - $\mathbb{P}(A) = 0{,}6 \longleftrightarrow \mathbb{P}(A^C) = 0{,}4$
  - $\mathbb{P}(B|A) = 0{,}8 \longleftrightarrow \mathbb{P}(B^C|A) = 0{,}2$
  - $\mathbb{P}(B|A^C) = 0{,}7 \longleftrightarrow \mathbb{P}(B^C|A^C) = 0{,}3$
- Respuesta:
  $\mathbb{P}(B) = \mathbb{P}(B|A)\,\mathbb{P}(A) + \mathbb{P}(B|A^C)\,\mathbb{P}(A^C) =$
  $0{,}8 \times 0{,}6 + 0{,}7 \times 0{,}4 = 0{,}76$

Regla de probabilidad total

Ejemplo

A =“Realiza la guía 1 a tiempo”, B =“Realiza la guía 2 a tiempo”
¿$\mathbb{P}(B)$?

$$\mathbb{P}(B) = \mathbb{P}(B|A)\,\mathbb{P}(A) + \mathbb{P}(B|A^C)\,\mathbb{P}(A^C) = 0{,}8\times0{,}6+0{,}7\times0{,}4 = 0{,}76$$



Probabilidad Condicional

Regla de Bayes (para problemas inversos)

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\,\mathbb{P}(A)}{\mathbb{P}(B)} \quad (\text{ojo! } \mathbb{P}(A) > 0 \text{ y } \mathbb{P}(B) > 0)$$

- No se puede condicionar a un evento que no puede ocurrir ($\mathbb{P}(B) = 0$)

Demostración.

$$\mathbb{P}(A|B) \overset{def.}{=} \frac{\mathbb{P}(B\cap A)}{\mathbb{P}(B)} \overset{R.multipl.}{=} \frac{\mathbb{P}(B|A)\,\mathbb{P}(A)}{\mathbb{P}(B)}$$

□

Regla de Bayes

Regla de Bayes

Dado un evento B con $\mathbb{P}(B) > 0$, y $A_1, A_2, \ldots, A_n$ una partición disjunta de Ω ($\bigcup_{i=1}^{n} A_i = \Omega$ con $A_i \cap A_j = \emptyset$ para $i \neq j$) con $\mathbb{P}(A_i) > 0$ para $i = \{1, 2, \ldots, n\}$

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(B|A_i)\,\mathbb{P}(A_i)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A_i)\,\mathbb{P}(A_i)}{\sum_{k=1}^{n} \mathbb{P}(B|A_k)\,\mathbb{P}(A_k)}$$

Regla de Bayes: $\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}$

$A =$"Realiza la guía 1 a tiempo", $B =$"Realiza la guía 2 a tiempo"
Pregunta: ¿$\mathbb{P}(A|B)$?

- Datos:
  - $\mathbb{P}(A) = 0{,}6 \longleftrightarrow \mathbb{P}(A^C) = 0{,}4$
  - $\mathbb{P}(B|A) = 0{,}8 \longleftrightarrow \mathbb{P}(B^C|A) = 0{,}2$
  - $\mathbb{P}(B|A^C) = 0{,}7 \longleftrightarrow \mathbb{P}(B^C|A^C) = 0{,}3$
- Respuesta: $\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)} = \frac{0{,}8 \times 0{,}6}{0{,}8 \times 0{,}6 + 0{,}7 \times 0{,}4} = \frac{0{,}48}{0{,}76} \approx 0{,}63$

Vision & Bayes

Independencia entre eventos

Independencia entre eventos

La independencia o no de los eventos (y las variables) es un tema bien relevante. Nos ayuda a clasificar un email como SPAM, a clasificar en forma automática a partir de una foto objetos, generar modelos para explicar ciertas variables,etc.

Independencia

Dos eventos A y B son independientes cuando uno no aporta información sobre el otro.

Independencia

Si $\quad \mathbb{P}(B|A) = \mathbb{P}(B) \quad$ entonces saber que ocurrió A no aporta información sobre las chances de que ocurra B.

Independencia entre eventos

Definición independencia

- Dos eventos A y B que pueden ocurrir son independientes si $\mathbb{P}(B|A) = \mathbb{P}(B)$.
- Dos eventos A y B que pueden ocurrir son dependientes si $\mathbb{P}(B|A) \neq \mathbb{P}(B)$.

Definición independencia equivalente

- Dos eventos A y B son independientes si $\mathbb{P}(B \cap A) = \mathbb{P}(B)\,\mathbb{P}(A)$.
- Dos eventos A y B son dependientes si $\mathbb{P}(B \cap A) \neq \mathbb{P}(B)\,\mathbb{P}(A)$.
- ¿Por qué esta definición es equivalente a la anterior?

Independencia entre eventos

Preguntas:

- Si A y B son disjuntos con $\mathbb{P}(A) > 0$ y $\mathbb{P}(B) > 0$, ¿son independientes o dependientes?
- Si $A \subseteq B$ y $B \neq \Omega$, ¿son independientes o dependientes?
- Si A y B son independientes, ¿A y B^C son independientes?
- Si A y B son independientes, ¿A^C y B^C son independientes?

Independencia entre eventos

Propiedades (las anteriores)

- Si A y B son disjuntos con $\mathbb{P}(A) > 0$ y $\mathbb{P}(B) > 0 \rightarrow A$ y B son dependientes.
- Si $A \subseteq B$ y $B \neq \Omega \rightarrow A$ y B son dependientes.
- Si A y B son independientes $\leftrightarrow A$ y B^C son independientes.
- Si A y B son independientes $\leftrightarrow A^C$ y B^C son independientes.

Independencia entre eventos

Ejemplo

Exp: Tiro 2 dados secuencialmente.
$A = \{$ El primer dado es un $4\}$, $\quad B = \{$ La suma es $6\}$.

¿A y B son independientes?

Ejemplo

Exp: Tiro 2 dados secuencialmente.
$A = \{$ El primer dado es un $4\}$, $\quad C = \{$ La suma es $7\}$.

¿A y C son independientes?

- Explique por qué B brinda información sobre A y C no la brinda.