UFM UNIVERSIDAD FRANCISCO MARROQUÍN
VERITAS • LIBERTAS • JVSTITIA

FACULTAD de CIENCIAS ECONÓMICAS

Repaso de probabilidad y estadística

Alan Reyes-Figueroa

Elements of Machine Learning (Aula 02) 13.enero.2025

Probabilidades

Construcción. Punto de partida: un experimento

- Resultado del experimento es $\omega \in \Omega \rightsquigarrow$ *espacio muestral.*
- Interés en ciertos eventos o subconjuntos de A. Llamamos $\mathcal{F}$ a la familia de estos eventos.
- Una probabilidad $\mathbb{P}$ es una función sobre ciertos eventos $\mathbb{P} : \mathcal{F} \mapsto \mathbb{R}$.

Ejemplo 1

Experimento: lanzar un dado.

$\Omega = \{1, 2, 3, 4, 5, 6\} = [1..6]$

Algunos eventos

Representación	Evento
$A_1 = \{2, 4, 6\}$	obtener un número par
$A_2 = \{3\}$	obtener 3
$A_3 = \{1, 2, 4, 5\}$	obtener un número no múltiplo de 3

Ejemplo 2
Experimento: lanzar dos dados.
$\Omega = \{(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),(2,1),\ldots,(5,6),(6,6)\}$
Probablemente aquí sea más simple representarlo como
$\Omega = \{(a,b) : a,b \in [1..6]\} = [1..6] \times [1..6]$

Algunos eventos

Representación	Evento
$A_1 = \{(1,6),(2,5),(3,4),\ldots,(6,1)\}$	que los dados sumen 7
$A_2 = \{(1,3),(3,1),\ldots,(6,3),(3,6)\}$	que aparezca al menos un 3

Otros espacios asociados: $\Omega_1 = [1..6]$, ¿Cuál es el mínimo de los dos dados?

Otros ejemplos (para pensar)

Especificar un espacio muestral para los siguientes experimentos:

a) Lanzar una moneda.
b) Lanzar una moneda hasta que aparezca “cruz”.
c) Distancia recorrida por un automóvil con un litro de gasolina.
d) Señal de radio que se recibe durante dos segundos.
e) Número de clientes acumulados en una caja de cobro si los clientes llegan a una tasa λ y la cada atiende a una tasa μ.
f) Juego entre tres jugadores: P, Q y R. El juego consiste en jugar partidas por parejas, comenzando P contra Q. Quien gane un partida juega con el otro jugador, hasta que uno de los jugadores gane dos partidas consecutivas, ganando entonces el juego.

Pregunta: ¿Cómo definir $\mathbb{P}$? ¿Cómo interpretarla?

Definición (Espacio de probabilidad)

*Un **espacio de probabilidad** es una estructura $(\Omega, \mathcal{F}, \mathbb{P})$, donde*

- *Ω es un conjunto (no vacío). Los elementos $\omega \in \Omega$ se llaman eventos.*
- *$\mathcal{F} \subseteq \Omega$ es una $\sigma-$álgebra.*
- *$\mathbb{P} : \mathcal{F} \to [0, 1]$ es una medida de probabilidad.*

Definición

*Una $\sigma-$**álgebra** $\mathcal{F}$ sobre un conjunto Ω es una colección de subconjuntos de Ω que satisface:*

- $\Omega \in \mathcal{F}$;
- $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$ *(es cerrada bajo complementos);*
- $A_i \in \mathcal{F}$, *para* $i = 1, 2, \ldots \Rightarrow \bigcup_i A_i \in \mathcal{F}$ *(es cerrada bajo uniones enum).*

Definición

*Una función $\mathbb{P} : \mathcal{F} \to [0, 1]$ es una **medida de probabilidad** si*

- $\mathbb{P}(\emptyset) = 0, \; \mathbb{P}(\Omega) = 1$;
- *para cualquier colección enumerable de eventos exclusivos $E_i \in \mathcal{F}$, vale*

$$\mathbb{P}\Big(\bigcup E_i\Big) = \sum \mathbb{P}(E_i) \;\; \textit{(enumerablemente aditiva)}.$$

Axiomas de la probabilidad, introducidos por Kolmogorov en 1933.

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de medida con $\mathbb{P}(E)$ la probabilidad de un evento $E \in \mathcal{F}$. Asumimos los siguientes supuestos para $\mathbb{P}$:

Axiomas

1. $\mathbb{P}(E) \geq 0, \forall E \in \mathcal{F}$ *(no-negativa).*

2. $\mathbb{P}(E)$ *es siempre finita, y* $\mathbb{P}(\Omega) = 1$ *(unitariedad).*

3. *Cualquier colección enumerable y mutuamente excluyente de eventos* $E_i \in \mathcal{F}$, *satisface*
$$\mathbb{P}\Big(\bigcup_{i=1}^{\infty} E_i\Big) = \sum_{i=1}^{\infty} \mathbb{P}(E_i), \qquad (\sigma\text{-aditiva}).$$

Propiedades

Si $\mathbb{P}$ es una medida de probabilidad sobre Ω, entonces

1. *(Monotonicidad) Si $A \subseteq B$ son eventos, entonces $\mathbb{P}(A) \leq \mathbb{P}(B)$.*
2. *(Conjunto vacío) $\mathbb{P}(\emptyset) = 0$.*
3. *(Complemento) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$, para todo evento $A \in \mathcal{F}$.*
4. *(Cotas para $\mathbb{P}$) Para todo evento $E \in \mathcal{F}$, $0 \leq \mathbb{P}(E) \leq 1$.*
5. *$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.*

Consecuencias

1. (Monotonicidad) Si $A \subseteq B$ son eventos en $\mathcal{F}$, entonces $\mathbb{P}(A) \leq \mathbb{P}(B)$.

   Prueba:
   Definamos $E_1 = A$, $E_2 = B - A$, y $E_i =$ para $i = 3, 4, \ldots$. Entonces, por σ-aditividad (axioma 3),
   $$\mathbb{P}(A) + \mathbb{P}(B - A) + \sum_{i \geq 3} \mathbb{P}(E_i) = \mathbb{P}(B).$$

   Como el lado izquierdo anterior es una suma de términos no-negativos (axioma 1), entonces
   $$\mathbb{P}(A) \leq \mathbb{P}(A) + \mathbb{P}(B - A) + \sum_{i \geq 3} \mathbb{P}(E_i) = \mathbb{P}(B).$$

Consecuencias

2. (Complemento) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$, para todo evento $A \in \mathcal{F}$.

   Prueba:
   A y $A^c = \Omega - A$ forman una partición de Ω. Por σ-aditividad (axioma 3) y el axioma 2

   $$\mathbb{P}(A) + \mathbb{P}(A^c) = \mathbb{P}(A \cup A^c) = \mathbb{P}(\Omega) = 1,$$

   luego $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

3. $\mathbb{P}(\emptyset) = 0$.

   Prueba:
   $\mathbb{P}(\emptyset) = \mathbb{P}(\Omega^c) = 1 - \mathbb{P}(\Omega) = 1 - 1 = 0$.

4. $0 \leq \mathbb{P}(E) \leq 1$, para todo evento E.

   Prueba:

   $\mathbb{P}(E) \geq 0$ por el axioma 1. Además, $E \subseteq \Omega$, y la monotonicidad de $\mathbb{P}$ implican $\mathbb{P}(E) \leq \mathbb{P}(\Omega) = 1$.

5. (Principio de Inclusión-Exclusión) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

   Prueba:
   Observe que $\mathbb{P}(A - B) + \mathbb{P}(A \cap B) = \mathbb{P}(A)$ (por aditividad). Luego $\mathbb{P}(A - B) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$. Similarmente, $\mathbb{P}(B - A) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$. Ahora, $A \cup B$ es la unión disjunta de $A - B$, $B - A$ y $A \cap B$. Por aditividad,

$$\mathbb{P}(A \cup B) = \mathbb{P}(A - B) + \mathbb{P}(B - A) + \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

Caso finito

Sea $\Omega = \{\omega_1, \omega_2, \ldots, \omega_k\}$.
Distribución de conteo o distribución uniforme: Corresponde a elegir un elemento al azar.
Para cada $A \subseteq \Omega$, se tiene

$$\mathbb{P}(A) = |A|/|\Omega| = |A|/k.$$

En particular, sin $A_i = \{\omega_i\}$, entonces

$$\mathbb{P}(\omega_i) = \mathbb{P}(\{\omega_i\}) = 1/k.$$

Caso general: Suponga que $\mathbb{P}(\omega_i) = \mathbb{P}(\{\omega_i\}) = p_i$, para $i = 1, 2, \ldots, k$. Entonces

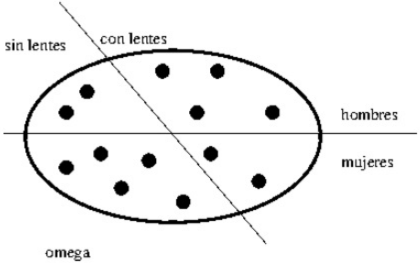
$$\mathbb{P}(A) = \sum_{\omega_i \in A} p_i$$

Interpretación de $\mathbb{P}$

Existen varias formas de interpretar probabilidades. Las 3 más comunes son:

- Probabilidades como límite de frecuencias relativas de ocurrencia de eventos(**enfoque frequentista**)
- Por medio de apuestas: probabilidades como creencias que pueden cambiar según se revela información (**enfoque bayesiano**)
- Sistema axiomático (Kolmogorov, 1933).

En áreas como computación e inteligencia artificial se han elaborado otros sistemas axiomáticos (fuzzy sets, Dempster-Shaffer, . . .) para modelas probabilidades.



Se elige una persona al azar.
¿Cuál es la probabilidad que sea una persona con lentes? $\frac{6}{13}$.

Alguien dice que es un hombre: ¿cuál es ahora la probabilidad que sea una persona con lentes? $\frac{2}{3}$.

Definición

Si $\mathbb{P}(B) > 0$, entonces la probabilidad condicional de A dado B se define como

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

Probabilidad condicional

Observaciones:

- $\mathbb{P}(\cdot|B)$ define una nueva función de probabilidad sobre el espacio $\Omega' = B$.
- En consecuencia, $\mathbb{P}(A^c|B) = 1 - \mathbb{P}(A|B)$.
- Observar que no hay ninguna relación directa entre $\mathbb{P}(A|B)$ y $\mathbb{P}(A|B^c)$.
- Siempre podemos escribir $\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\,\mathbb{P}(B)$.
  (Esto no requiere el supuesto que $\mathbb{P}(B) > 0$) ¿Por qué?

Ejemplo

Example No.	Color	Type	Origin	Stolen?
1	Red	Sports	Domestic	Yes
2	Red	Sports	Domestic	No
3	Red	Sports	Domestic	Yes
4	Yellow	Sports	Domestic	No
5	Yellow	Sports	Imported	Yes
6	Yellow	SUV	Imported	No
7	Yellow	SUV	Imported	Yes
8	Yellow	SUV	Domestic	No
9	Red	SUV	Imported	No
10	Red	Sports	Imported	Yes

Ejemplo

- ¿Cuál es la probabilidad de robo de un auto amarillo?
- ¿Cuál es la probabilidad de robo de un auto amarillo, dado que se sabe que es importado?
- ¿Cuál es la probabilidad de robo de un auto amarillo, dado que se sabe que es deportivo?
- ¿Cuál es la probabilidad de robo de un auto deportivo rojo?
- ¿Cuál es la probabilidad de robo de un auto deportivo rojo, dado que es importado?

Teorema (Ley de la probabilidad total, caso finito)

Dada una partición $\{B_i\}_{i=1}^n$ de Ω, tal que $\mathbb{P}(B_i) > 0$, $\forall i$, entonces

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A|B_i)\, \mathbb{P}(B_i).$$

<u>Prueba</u>: $\Omega = \bigcup B_i$, ya que es una partición. Luego

$$A = A \cap \Omega = A \cap \bigcup B_i = \bigcup (A \cap B_i),$$

y los $A \cap B_i$ forman una partición de A. Por el axioma de aditividad, y la definición de probabilidad condicional

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A \cap B_i) = \sum_{i=1}^{n} \mathbb{P}(A|B_i)\, \mathbb{P}(B_i).$$

Ejemplo

Se tienen tres cajas, cada una conteniendo 100 cartas:
La caja 1 contiene 75 cartas rojas, y 25 cartas azules,
la caja 2 contiene 60 cartas rojas, y 40 cartas azules,
la caja 3 contiene 55 cartas rojas, y 45 cartas azules.

Se elige una de las cajas al azar, y luego se elige una carta dentro de la caja seleccionada.
¿Cuál es la probabilidad de elegir una carta roja?

Ejemplo

<u>Solución</u>: Considere los eventos $A =$ elegir carta roja, y

$$E_1 = \text{elegir caja } 1, \; E_2 = \text{elegir caja } 2, \; E_3 = \text{elegir caja } 3.$$

Sabemos que $\mathbb{P}(A|E_1) = \dfrac{75}{100}$, $\mathbb{P}(A|E_2) = \dfrac{60}{100}$ y $\mathbb{P}(A|E_3) = \dfrac{55}{100}$.

Entonces, por la ley de probabilidad total

$$\begin{aligned}
\mathbb{P}(A) &= \mathbb{P}(A|E_1)\,\mathbb{P}(E_1) + \mathbb{P}(A|E_2)\,\mathbb{P}(E_2) + \mathbb{P}(A|E_3)\,\mathbb{P}(E_3) \\
&= \frac{75}{100} \cdot \frac{1}{3} + \frac{60}{100} \cdot \frac{1}{3} + \frac{55}{100} \cdot \frac{1}{3} \\
&= \frac{190}{300} = 0.6333
\end{aligned}$$

Teorema (Regla de Bayes)

Si $\mathbb{P}(A), \mathbb{P}(B) > 0$, *entonces*

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

<u>Prueba</u>: Por hipótesis, $\mathbb{P}(A), \mathbb{P}(B) > 0$, entonces

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

Despejando $\mathbb{P}(A \cap B)$ de la segunda ecuación, $\mathbb{P}(A \cap B) = \mathbb{P}(B|A)\,\mathbb{P}(A)$, y sustituyendo en la primera

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

Regla de Bayes

Example No.	Color	Type	Origin	Stolen?
1	Red	Sports	Domestic	Yes
2	Red	Sports	Domestic	No
3	Red	Sports	Domestic	Yes
4	Yellow	Sports	Domestic	No
5	Yellow	Sports	Imported	Yes
6	Yellow	SUV	Imported	No
7	Yellow	SUV	Imported	Yes
8	Yellow	SUV	Domestic	No
9	Red	SUV	Imported	No
10	Red	Sports	Imported	Yes

Regla de Bayes

Calcular la probabilidad $\mathbb{P}(\textbf{robado} = 1 \mid \textbf{rojo})$, mediante los siguientes métodos:

1. cálculo directo
2. usando la Regla de Bayes

UFM UNIVERSIDAD FRANCISCO MARROQUÍN

Conceptos derivados: Independencia

La idea de **independencia** es determinar si hay o no relación entre dos eventos A y B. En otras palabra, si al conocer A, cambia nuestro conocimiento sobre B (o al conocer B cambia nuestro conocimiento sobre A).

¿Cómo determinar esta relación? Comparar $\mathbb{P}(A|B)$ con $\mathbb{P}(A)$.

Definición

Si $\mathbb{P}(B) > 0$, decimos que A y B son ***independientes*** *si $\mathbb{P}(A|B) = \mathbb{P}(A)$.*

Definición

Dos eventos A y B son ***independientes*** *si, y sólo si,*

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

Ejemplo

Lanzamiento de dos dados D_1 y D_2. Consideremos los eventos $A = \{D_1 + D_2 \text{ es par}\}$, $B = \{D_1 < 5\}$, $C = \{D_1 \leq 3, D_2 \leq 3\}$.
Sabemos que $\mathbb{P}(A) = \frac{1}{2}$, $\mathbb{P}(A \cap B) = \frac{1}{2}$, $\mathbb{P}(A \cap C) = \frac{5}{9}$.

D_1 D_2	1	2	3	4	5	6
1	X		X		X	
2		X		X		X
3	X		X		X	
4		X		X		X
5						
6						

D_1 D_2	1	2	3	4	5	6
1	X		X			
2		X				
3	X		X			
4						
5						
6						

Luego, A y B son independientes; mientras que A y C no lo son.

UFM UNIVERSIDAD FRANCISCO MARROQUÍN

Referencias

- Kai-Lai Chung. *A Course in Probability Theory*.
- Lefebvre. *Basic Probability Theory with Applications*. Springer.