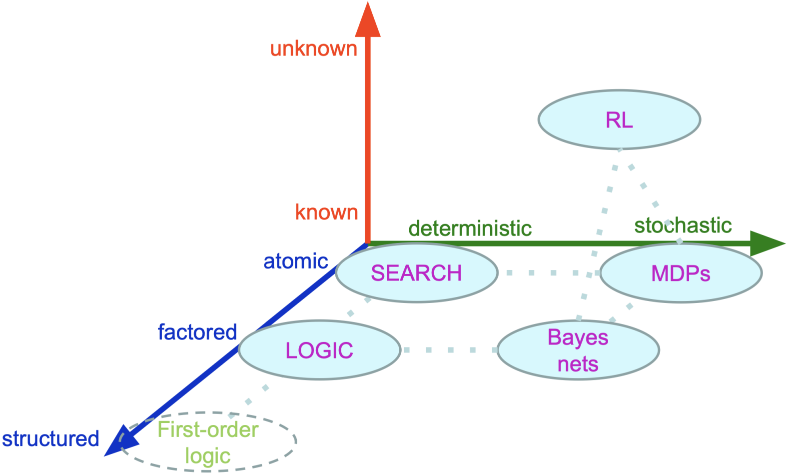
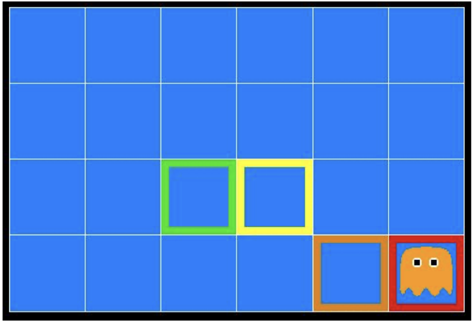
Repaso de Probabilidad

Alan Reyes-Figueroa

Inteligencia Artificial (Aula 27) 30.abril.2025

Probabilidades e IA





- Sensors are noisy, but we know P(Color(x,y) | DistanceFromGhost(x,y))

P(red | 3)	P(orange | 3)	P(yellow | 3)	P(green | 3)
0.05	0.15	0.5	0.3

Probabilidades

Construcción. Punto de partida: un experimento

- Resultado del experimento es $\omega \in \Omega$ $\rightsquigarrow$ *espacio muestral*.
- Interés en ciertos eventos A $\rightsquigarrow$ $\sigma-$*álgebra*
- Una probabilidad $\mathbb{P}$ es una función sobre ciertos eventos $\mathbb{P} : A \mapsto \mathbb{R}$.

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

Algunos eventos

Representación	Evento
$A_1 = \{(1,6),(2,5),(3,4),\dots,(6,1)\}$	que los dados sumen 7
$A_2 = \{(1,3),(3,1),\dots,(6,3),(3,6)\}$	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) 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.

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.*

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

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

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

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

Consecuencias

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

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

Ejemplo

1. Calcular la probabilidad que en un grupo de n personas hay al menos una que cumple años el 05 de abril.

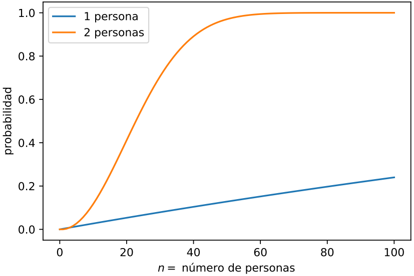
2. Calcular la probabilidad que en un grupo de n personas hay al menos dos personas que cumplen en el mismo día.

Ejemplo

n	probabilidad 1 persona 05/abril	probabilidad 2 personas mismo día
0	0	0
1	0.002739	0
5	0.013623	0.027135
10	0.027061	0.116948
20	0.053391	0.411438
30	0.079008	0.706316
40	0.103932	0.891231
50	0.128181	0.970373
60	0.151774	0.994122
70	0.174729	0.999159

Ejemplo

<u>Solución</u>:



$$\mathbb{P}(\text{alguien cumple años 17.enero}) = 1 - \left(\frac{364}{365}\right)^{n-1}, \quad n \geq 1.$$

$$\mathbb{P}(\text{dos personas cumplen años mismo día}) = 1 - \frac{365}{365} \cdot \frac{364}{365} \cdot \ldots \cdot \frac{365-(n-1)}{365}, \quad n \geq 2.$$

Caso $\Omega \subseteq \mathbb{R}^d$

Distribución uniforme:
Experimento: Elegir un número al azar de $[0, 2]$.
Tenemos $\Omega = [0, 2]$.
$A = [0, 1] \qquad \mathbb{P}(A) = 1/2.$
$B = [0.4, 1] \qquad \mathbb{P}(B) = 0.6/2 = 0.3.$

En general, para $A \subseteq \Omega$

$$\mathbb{P}(A) = \frac{\int_A dx}{\int_\Omega dx}.$$

¿Se puede calcular $\mathbb{P}$ siempre? No.

- Se requiere que $\int_\Omega dx < \infty$.
- Tenemos que limitarnos a conjuntos donde $\int_A dx$ existe.

Ejemplos

1. Se elige al azar un punto en un cuadrado con lado 4 cm. Calcula la probabilidad de que esté a una distancia menor a 1 cm. de alguna de las esquinas.

2. Dos estudiantes quieren ir a comer juntos. Se citan entre las 7 y las 8 de la noche y están dispuestos a esperar a lo más 10 minutos. ¿Cuál es la probabilidad de que puedan ir a comer si sus horas de llegada son uniformes entre las 7 y las 8?

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**)
  Este es el tipo de enfoque de probabilidad que más vamos a usar en IA.
- Sistema axiomático (Kolmogorov, 1933).

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

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.

Referencias

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