

REPASO DE PROBABILIDAD Y ESTADÍSTICA I

ALAN REYES-FIGUEROA

INTRODUCCIÓN A LA CIENCIA DE DATOS

(AULA 02) 14.ENERO.2021

Probabilidades

Construcción. Punto de partida: un experimento

- Resultado del experimento es $\omega \in \Omega \rightsquigarrow$ *espacio muestral*.
- Interés en ciertos eventos $A \rightsquigarrow$ σ -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 una 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 σ -álgebra.
- $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ es una medida de probabilidad.

Definición

Una σ -**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 $i = 1, 2, \dots \Rightarrow \bigcup_i A_i \in \mathcal{F}$ (es cerrada bajo uniones enum).

Definición

Una función $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ 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

$$\mathbb{P}\left(\bigcup E_i\right) = \sum \mathbb{P}(E_i) \text{ (enumerablemente aditiva).}$$

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*

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(E_i), \quad (\sigma\text{-aditiva}).$$

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 $E_i = \emptyset$ para $i = 3, 4, \dots$. 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 $E_i = \emptyset$ para $i = 3, 4, \dots$. 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 monotonidad 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,

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

Caso finito

Sea $\Omega = \{\omega_1, \omega_2, \dots, \omega_k\}$.

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 $i = 1, 2, \dots, k$.

Entonces

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

Ejemplo

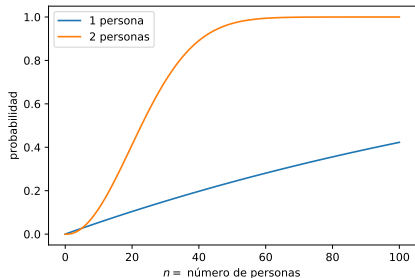
1. Calcula la probabilidad que en un grupo de n personas hay al menos una que cumple años el 14 de enero.
2. Calcula la probabilidad que en un grupo de n personas hay al menos dos personas que cumplen en el mismo día.

Ejemplo

n	probabilidad 1 persona 14/enero	probabilidad 2 personas mismo día
0	0	0
1	0.005479	0
5	0.027098	0.027135
10	0.053462	0.116948
20	0.104067	0.411438
30	0.151966	0.706316
40	0.197305	0.891231
50	0.240219	0.970373
60	0.280839	0.994122
70	0.319288	0.999159

Ejemplo

Solución:



$$\mathbb{P}(\text{alguien cumple años 14.enero}) = 1 - \left(\frac{1}{365}\right)^{n-1}, \quad n \geq 1.$$

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

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

Distribución uniforme:

Experimento: Elegir un número al azar de $[0, 2]$.

Tenemos $\Omega = [0, 2]$.

$$A = [0, 1] \quad \mathbb{P}(A) = 1/2.$$

$$B = [0.4, 1] \quad \mathbb{P}(B) = 0.6/2 = 0.3.$$

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 de uno 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?

- A partir del experimento elegir algo al azar.
- Probabilidades como límite de frecuencias relativas de ocurrencia (**enfoque frequentista**)
- Por medio de apuestas: probabilidades como creencias (base del **enfoque bayesiano**)
- Sistema axiomático (Kolmogorov, 1933).

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

- Kai-Lai Chung. *A Course in Probability Theory*.