Teoría de Probabilidad

Cristian Guarnizo-Lemus
cristianguarnizo@itm.edu.co

Maestria en Automatización y Control Industrial

Experimento aleatorio

- Se define un experimento aleatorio como aquel experimento cuyo resultado no se conoce con anterioridad.
- El enfoque matemático que se emplea para estudiar los resultados de un experimento aleatorio o de un fenónemo aleatorio se conoce como teoría de la probabilidad.

Contenido

Conjuntos

- Un conjunto se define como una colección de elementos.
- **Notación.** Letras mayúsculas denotan conjuntos $A, \ B, \ldots$. Letras minúsculas denotan miembros o elementos de un conjunto $a, \ b, \ldots$.
- El símbolo $\in$ denota pertenencia de un elemento a un conjunto. Por ejemplo, $a \in A$, se lee "el elemento a pertenece al conjunto A".
- El simbolo $\notin$ se lee como "no pertenece a".

Más conjuntos

- Conjunto vacío o nulo. Es un conjunto que no tiene elementos. Se denota como $\emptyset$.
- Conjunto universal. Es un conjunto que contiene a todos los conjuntos bajo consideración.
- Conjunto contable. Si los elementos del conjunto pueden ponerse en correspondencia uno a uno con los enteros.
- Conjunto finito. Conjunto contable con un número finito de elementos.

Relaciones entre conjuntos

- **Subconjunto.** $A \subset B$
  Cada elemento de A también es un elemento de B.

- **Igualdad de conjuntos.** Dos conjuntos son iguales si contienen exactamente los mismos elementos.

Relaciones entre conjuntos

- **Unión.** $A \cup B$.
  Conjunto de todos los elementos que pertenecen a A o pertenecen a B o a ambos. Si se tienen N conjuntos

$$A_1 \cup A_2 \cup \ldots A_N = \bigcup_{i=1}^{N} A_i$$

- **Intersección.** $A \cap B$
  Conjunto de todos los elementos que pertenecen a A y a B. Si se tienen N conjuntos

$$A_1 \cap A_2 \cap \ldots A_N = \bigcap_{i=1}^{N} A_i$$

Más relaciones entre conjuntos

- **Mutuamente excluyentes.** Se dice que dos conjuntos son mutuamente excluyentes si no tienen elementos comunes. Los conjuntos A y B son mutuamente excluyentes si

  $$A \cap B = \emptyset$$

  Los conjuntos $A_1,\ A_2, \ldots, A_N$ son mutuamente excluyentes si

  $$A_i \cap A_j = \emptyset,\ \forall i, j \text{ con } i \neq j$$

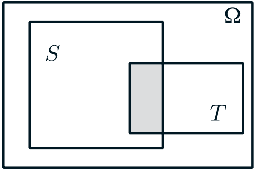
- **Complemento.** El complemento, $\overline{A}$, de un conjunto A relativo a S se define como el conjunto de todos los elementos de S que no están en A.

Álgebra de Conjuntos

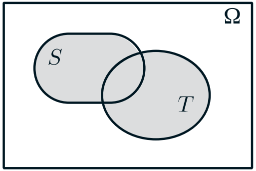
Las operaciones con conjuntos tienes varias propiedades, que son consecuencias elementales de las definiciones. algunos ejemplos son:

$$
\begin{aligned}
A \cup B &= B \cup A, & A \cup (B \cup C) &= (A \cup B) \cup C \\
A \cap (B \cup C) &= (A \cap B) \cup (A \cap C), & A \cup (B \cap C) &= (A \cup B) \cap (A \cup C) \\
\overline{(\overline{A})} &= A, & A \cap \overline{A} &= \emptyset \\
A \cup S &= S, & A \cap S &= A
\end{aligned}
$$

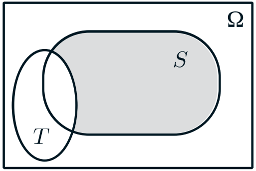
Ejemplos conjuntos



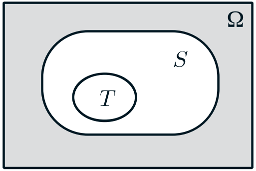
(a)



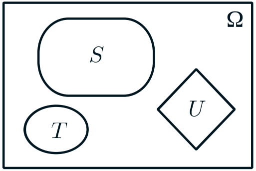
(b)



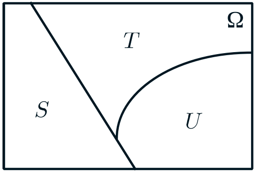
(c)



(d)



(e)



(f)

Contenido

Espacio muestral

- Desde el punto de vista de la teoría de probabilidad, los elementos del **conjunto universal** S son los resultados de un experimento.
- Un **experimento** es una secuencia de acciones que produce unos resultados.
- La totalidad de todos los posibles resultados que produce un experimento se conoce como el **espacio muestral**.
- El espacio muestral es entonces el conjunto universal S.
- Un **evento** se define como una colección de resultados de un experimento.
- Un evento es entonces un subconjunto (por ejemplo, A) del espacio muestral S.

Contenido

Preliminares

- La probabilidad de un evento A ($P(A)$) es un número que se asigna al evento A.
- Existen diferentes formas en las que probabilidades pueden asignarse a los subconjuntos de un espacio muestral.
- La forma coherente de hacerlo se realiza a través de la definición de una *medida de probabilidad*.

Medida de Probabilidad

- **Medida de probabilidad.** Consiste en asignar a los posibles eventos medibles del espacio muestral S un valor en el intervalo $[0, 1]$ que satisface
  1. (Normalización) $P(S) = 1$.
  2. (No-negativo) $P(A) \geq 0$ para todo $A \in S$.
  3. (Adición) Si $A_1, A_2, \ldots$ es una colección de eventos mutuamente excluyentes (ó disyuntos) en S, luego

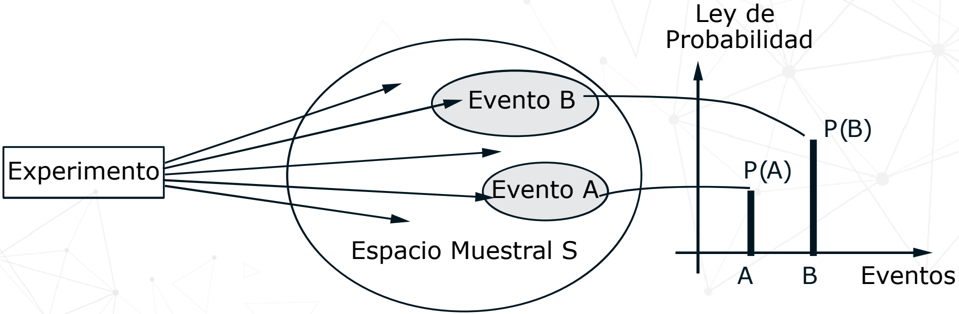
$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Espacio de Probabilidad

- Es posible asignarle a los resultados de un experimento aleatorio un espacio de probabilidad (S, P).
- **Ejemplo.** Una moneda, posiblemente parcializada, se tira al aire una vez. Se puede asumir que $S = \{H, T\}$ y que $\mathcal{F} = \{\emptyset,\ H,\ T,\ S\}$. Adicionalmente, una posible medida de probabilidad $P : \mathcal{F} \leftarrow [0, 1]$ está dada como

$$P(\emptyset) = 0, \quad P(H) = \rho, \quad P(T) = 1 - \rho, \quad P(S) = 1.$$

Espacio muestral



Ley de la probabilidad discreta

- Si el espacio muestral consiste en un número finito de posibles resultados, entonces la ley de la probabilidad esta especificada por las probabilidades de los eventos que consisten en un solo elemento. En particular, la probabilidad de cualquier evento $\{s_1, s_2, \ldots, s_n\}$ es la suma de las probabilidades de sus elementos:

$$P(\{s_1, s_2, \ldots, s_n\}) = P(s_1) + P(s_2) + \cdots + P(s_n).$$

Definición de probabilidad

- **Frecuencia relativa.** Un experimento aleatorio se repite n veces. Si el evento A ocurre n_A veces, luego su probabilidad $P(A)$ se define como el límite de la frecuencia relativa nA/n de la ocurrencia de A

$$P(A) \triangleq \lim_{n\to\infty} \frac{n_A}{n}$$

- **Definición clásica.** En esta definición, la probabilidad $P(A)$ de un evento se calcula como la relación entre el número de casos favorables, N_A, sobre el número de posibilidades, N,

$$P(A) \triangleq \frac{N_A}{N}.$$

Asume que todos los eventos son igualmente probables.

Leyes útiles de probabilidad

Dada una definición de probabilidad que satisfaga los tres axiomas de una medida de probabilidad, se pueden establecer las siguientes relaciones

1. $P(\emptyset) = 0$.
2. Para cualquier evento arbitrario A, $P(A) \leq 1$.
3. Si $A \cup \overline{A} = S$ y $A \cap \overline{A} = \emptyset$, $P(\overline{A}) = 1 - P(A)$.
4. Si $A \subset B$, luego $P(A) \leq P(B)$.
5. Si $A \cup B = P(A) + P(B) - P(A \cap B)$.
6. $P(A \cup B) \leq P(A) + P(B)$.
7. $P(A \cup B \cup C) = P(A) + P(\overline{A} \cap B) + P(\overline{A} \cap \overline{B} \cap C)$.

Leyes útiles de probabilidad

Si $A_1,\ A_2, \ldots,\ A_n$ son eventos aleatorios tales que

$$A_i \cap A_j = \emptyset,\ \forall i \neq j$$
$$A_1 \cup A_2 \cup \ldots A_n = S,$$

luego $P(A) = P(A \cap A_1) + P(A \cap A_2) + \ldots + P(A \cap A_n)$. Se dice que los conjuntos son mutuamente excluyentes y exhaustivos, si se cumplen las dos condiciones anteriores.

Contenido

1 Probabilidad

- Definiciones de conjuntos
- Espacio muestral
- Probabilidad de un evento aleatorio
- Probabilidad conjunta, marginal y condicional

Producto cartesiano

1. Sea E un experimento aleatorio que se compone de dos subexperimentos E_1 y E_2.
2. El espacio muestral S_1 asociado a E_1 se compone de las salidas $a_1,\ a_2,\ \ldots,\ a_{n_1}$.
3. El espacio muestral S_2 asociado a E_2 se compone de las salidas $b_1,\ b_2, \ldots,\ b_{n_2}$.
4. El espacio muestral S del experimento combinado es el producto Cartesiano de S_1 y S_2. Esto es

$$\begin{aligned} S &= S_1 \times S_2 \\ &= \{(a_i, b_j) : i = 1, \ldots, n_1; j = 1, \ldots, n_2\}. \end{aligned}$$

5. Se pueden definir medidas de probabilidad sobre S_1, S_2 y $S = S_1 \times S_2$. Si los eventos $A_1, A_2, \ldots, A_n$ están definidos sobre el subexperimento E_1, y los eventos $B_1, B_2, \ldots, B_m$ están definidos sobre el subexperimento E_2, el evento A_iB_j es un evento sobre el experimento total.

Probabilidad conjunta o marginal

1. **Probabilidad conjunta.** La probabilidad del evento $A_i \cap B_j$ se conoce como *probabilidad conjunta* y se denota $P(A_i \cap B_j)$ ó $P(A_iB_j)$.
2. **Probabilidad marginal.** Si los eventos $A_1, \ A_2, \ldots, \ A_n$ asociados con el subexperimento E_1 son mutuamente excluyentes y exhaustivos, luego

$$\begin{aligned} P(B_j) &= P(B_j \cap S) = P\left[B_j \cap (A_1 \cup A_2 \cup \ldots \cup A_n)\right] \\ &= \sum_{i=1}^{n} P(A_iB_j). \end{aligned}$$

La probabilidad $P(B_j)$ se conoce como la probabilidad marginal.

Probabilidad condicional

- Con frecuencia, la probabilidad de ocurrencia de un evento B podría depender de la ocurrencia de un evento relacionado A.
- La probabilidad condicional $P(B|A)$ se define como

$$p(B|A) \triangleq \frac{P(AB)}{P(A)}.$$

- Empleando la definición clásica de probabilidad,

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{N_{AB}}{N_A},$$

donde N_{AB} es el numero de casos favorables del evento $A \cap B$, y N_A el numero de casos favorables del evento A.

Probabilidad condicional

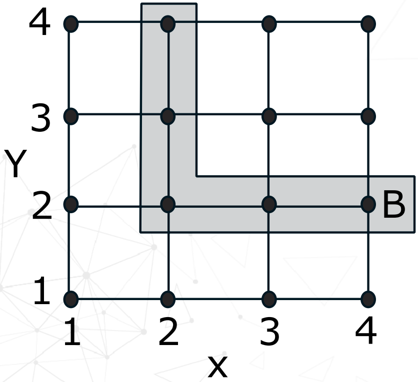
Ejemplo: Un dado justo de 4 lados es lanzado 2 veces y asumimos que todos los posibles resultados son igualmente probables. Sea X y Y los resultados del primer y segundo lanzamiento, respectivamente. Queremos determinar la probabilidad condicional $P(A|B)$ donde

$$A = \{\max(X, Y) = m\}, \quad B = \{\min(X, Y) = 2\}$$

y m toma cada uno de los valores 1, 2, 3, 4.

Probabilidad condicional

Todos los resultados igualmente probables



$$P(\{\max(X,Y)=m\}|B)=\begin{cases} 2/5 & \text{Si } m=3 \text{ o } m=4 \\ 1/5 & \text{Si } m=2 \\ 0 & \text{Si } m=1 \end{cases}$$

Relaciones entre prob. conj. marg. y cond.

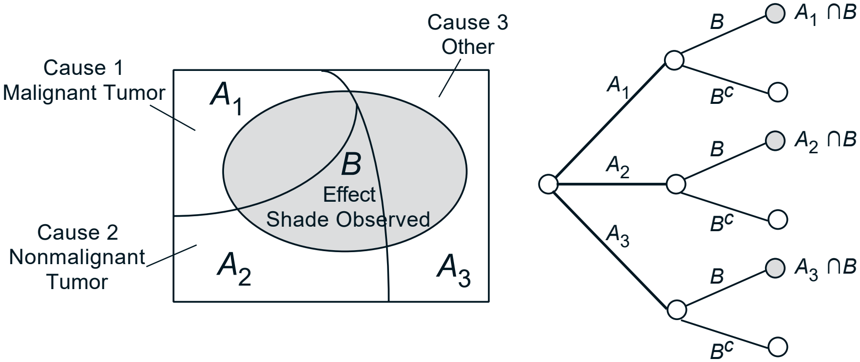
- $P(AB) = P(A|B)P(B) = P(B|A)P(A)$
- Si $AB = \emptyset$, luego $P(A \cup B|C) = P(A|C) + P(B|C)$.
- $P(ABC) = P(A)P(B|A)P(C|AB)$.
- Si $B_1, B_2, \ldots, B_m$ forman un conjunto de eventos mutuamente excluyentes y exhaustivos, luego

$$P(A) = \sum_{j=1}^{m} P(A|B_j)P(B_j).$$

Teorema de Bayes

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{j=1}^{m} P(A|B_j)P(B_j)}.$$

Teorema de Bayes - Ejemplo



$$P(A_i|B) = \frac{P(B|A_i)\,P(A_i)}{P(B|A_1)\,P(A_1) + P(B|A_2)\,P(A_2) + P(B|A_3)\,P(A_3)}, \quad i = 1, 2, 3$$

Teorema de Bayes - Aprendizaje Automático

Si tenemos un conjunto de parámetros $\boldsymbol{\theta} = \{\theta_1, \ldots, \theta_P\}$ que desciben los datos $\mathbf{y}$, podemos calcular la probabilidad de los parámetros dados los datos así:

$$P(\boldsymbol{\theta}|\mathbf{y}) = \frac{P(\mathbf{y}|\boldsymbol{\theta})P(\boldsymbol{\theta})}{\sum P(\mathbf{y}|\boldsymbol{\theta})P(\boldsymbol{\theta})},$$

ó, si los parámetros tienen una distribución continua

$$P(\boldsymbol{\theta}|\mathbf{y}) = \frac{P(\mathbf{y}|\boldsymbol{\theta})P(\boldsymbol{\theta})}{\int P(\mathbf{y}|\boldsymbol{\theta})P(\boldsymbol{\theta})}.$$

Independencia

Dos eventos A_i y B_j son independientes si

$$P(A_iB_j) = P(A_i)P(B_j), \text{ ó } P(A_i|B_j) = P(A_i).$$

Ejercicios

Demostrar que

1. Si $A_1 \subset A_2$ entonces $P(A_1) \leq P(A_2)$ y $P(A_2 - A_1) = P(A_2) - P(A_1)$.
2. $P(\overline{A}) = 1 - P(A)$.
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Referencias

- Basado en las presentaciones del Prof. Mauricio A. Álvarez, del curso “Procesos Estocásticos”.
- D.P. Bertsekas et al. “Introduction to Probability”, 2002.