Inteligencia Artificial 222
Introducción a la probabilidad
E. Barcenas

Definición 1 (Experimentos aleatorios)**.** Un experimento con diferentes resultados, incluso si es repetido en la misma manera, se llama un experimento aleatorio.

Definición 2 (Espacio muestra)**.** El conjunto de todos los posibles resultados de un experimento aleatorio, se llama el espacio muestra del experimento.

Ejemplo:

$$\begin{aligned}
S =& \mathbb{R}^+ = \{x \in \mathbb{R} \mid x > 0\} \\
S =& \{x \in \mathbb{N} \mid 10 < x < 20\} \\
S =& \{low, medium, light\} \\
S =& \{yes, not\}
\end{aligned}$$

Definición 3 (Espacio muestra discreto)**.** Un espacio muestra es discreto si y sólo si, es contable.

Ejemplos.

$$\begin{aligned}
S =& \{x \in \mathbb{N} \mid 10 < x < 20\} \\
S =& \{low, medium, light\} \\
S =& \{yes, not\} \\
S =& \{x \in \mathbb{N} \mid \exists y \in \mathbb{N} : 2y + 1 = x\} \\
S =& \{0, 1\}^\star
\end{aligned}$$

Definición 4 (Espacio muestra continuo)**.** Un espacio muestra S es continuo si y sólo si, existe una biyección $f : S \mapsto \mathbb{R}$.

Ejemplos.

$$\begin{aligned}
S =& \{x \in \mathbb{R} \mid 0 < x < 2\} \\
S =& \mathbb{R}^+ \times \mathbb{R}^+ \\
S =& \{x \mid x = |S'|, S' \subseteq \mathbb{N}\}
\end{aligned}$$

Definición 5 (Evento)**.** Un evento E es un subconjunto del espacio muestra S, es decir, $E \subseteq S$.

- La unión de dos eventos E_1, E_2, se define $E_1 \cup E_2 = \{x \mid x \in E_1 \text{ or } x \in E_2\}$.
- La intersección de dos eventos E_1, E_2, se define $E_1 \cap E_2 = \{x \mid x \in E_1 \text{ and } x \in E_2\}$.
- El complemento de un evento E en el espacio muestra S se define $E^c = S \setminus E = \{x \in S \mid x \notin E\}$.

Ejemplos. Considere el espacio muestra $S = \{yy, yn, ny, nn\}$. Los siguientes son eventos de S.

$E_1 = \{yy, yn, ny\}$, $E_2 = \{nn\}$, $E_3 = \emptyset$, $E_4 = S$, $E_5 = \{yn, ny, nn\}$,

$$E_1 \cup E_2 = S, \quad E_1 \cap E_5 = \{yn, ny\}, \quad E_1^c = \{nn\}.$$

Mas ejemplos. Considere $S = \mathbb{R}^+$, $E_1 = \{x \mid 1 \le x < 10\}$ y $E_2 = \{x \mid 3 < x < 118\}$, entonces

$$\begin{aligned} E_1 \cup E_2 &= \{x \mid 1 \le x < 118\} \\ E_1 \cap E_2 &= \{x \mid 3 < x < 10\} \\ E_1^c &= \{x \mid x \ge 10\} \\ E_1^c \cap E_2 &= \{x \mid 10 \le x < 118\} \end{aligned}$$

Algunas propiedades de los eventos.

Dos eventos A y B se dicen mutuamente exclusivos si y sólo si, su intersección es vacía, es decir, $A \cap B = \emptyset$.

$$\begin{aligned} (E^c)^c &= E \\ (A \cup B) \cap C &= (A \cap C) \cup (B \cap C) \\ (A \cap B) \cup C &= (A \cup C) \cap (B \cup C) \\ (A \cup B)^c &= A^c \cap B^c \\ (A \cap B)^c &= A^c \cup B^c \\ A \cup B &= B \cup A \\ A \cap B &= B \cap A \end{aligned}$$

Definición 6 (Probabilidad)**.** En un espacio muestra discreto, la probabilidad de un evento E, escrito $P(E)$, es igual a la suma de las probabilidades de sus resultados en E.

Ejemplo.

Un experimento aleatorio puede resultar en $\{a, b, c, d\}$, con probabilidades 0.1, 0.3, 0.5 y 0.1, respectivamente. Considere el evento A como $\{a, b\}$, $B = \{b, c, d\}$, y $C = \{d\}$. Entonces,

$$\begin{aligned} P(A) &= 0.1 + 0.3 = 0.4 \\ P(B) &= 0.3 + 0.5 + 0.1 = 0.9 \\ P(C) &= 0.1 \end{aligned}$$

También, $P(A^c) = 0.6$, $P(B^c) = 0.1$, $P(C^c) = 0.9$. Más aún, debido a $A \cap B = \{b\}$, entonces $P(A \cap B) = 0.3$. Debido a $A \cup B = \{a, b, c, d\}$, $P(A \cup B) = 0.1 + 0.3 + 0.5 + 0.1 = 1$. Y debido a $A \cap C = \emptyset$, entonces $P(A \cap C) = 0$.

Definición 7 (Axiomas de la probabilidad)**.** Considere los eventos E, E_1 y E del espacio muestra S de un experimento aleatorio.

- $P(S) = 1$
- $0 \leq P(E) \leq 1$
- Si $E_1 \cap E_2 = \emptyset$, entonces $P(E_1 \cup E_2) = P(E_1) + P(E_2)$.

Algunas propiedades.

$$P(\emptyset) = 0 \quad P(E^c) = 1 - P(E) \qquad \text{Si } E_1 \subseteq E_2, \text{ entonces } P(E_1) \leq P(E_2)$$

Ejercicios.

- Suponga que las placas de los vehículos están compuestas inicialmente por tres digitos $(0-9)$, seguidas de tres letras $(A-Z)$. Calcule la probabilidad de una determinada placa.
- Un mensaje puede seguir diferentes rutas a través de una red de servidores. En el primer paso, el mensaje puede llegar a cinco servidores, a partir de cada uno de estos servidores, el mensaje puede llegar a cinco servidores más, desde los cuales puede acceder a otros cuatro servidores.
  - Calcule la cantidad de rutas.
  - Si todas las rutas son igualmente probables, calcules la probabilidad de que el mensaje llegue a alguno de los cuatro servidores del tercer bloque.
- Demuestre lo siguiente.

  $$P(\emptyset) = 0 \quad P(E^c) = 1 - P(E) \qquad \text{Si } E_1 \subseteq E_2, \text{ entonces } P(E_1) \leq P(E_2)$$

- Una mezcla química es preparada correctamente por el 25% de los técnicos de un laboratorio, 70% de los técnicos la preparan con un error mínimo, y 5% con un error mayor.
  - Si un técnico es elegido aleatoriamente, ¿Cuál es la probabilidad de que prepare la mezcla sin error alguno?
  - Calcule la probabilidad de que el técnico la prepare con cualquier tipo de error.
- Considere las emisiones de tres fabricas clasificadas por su calidad. De la primera fábrica 22 muestras de emisiones cumplen con el mínimo, y 8 no lo hacen; 25 cumplen con el mínimo y 5 no, en el caso de la segunda fábrica; en cuanto a la tercera, 30 cumplen y 10 no. Considere A denota el evento de las muestras de emisiones de la primera fábrica, y B como el evento de una muestra cumple con el mínimo. Calcule las siguientes probabilidades.

  $$P(A) \qquad P(B) \qquad P(A^c) \qquad P(A \cap B) \qquad P(A \cup B) \qquad P(A^c \cup B)$$

Definición 8 (Regla de la adición)**.**

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Ejemplo. Considere la tabla abajo con el historial de producción de 949 semiconductores. Suponga que un semiconductor es elegido aleatoriamente. Considere H denota el evento de que el semiconductor contiene niveles altos de contaminación. C es el evento cuando el semiconductor se encuentra en el centro de una herramienta de pulverización.

$$P(H \cup C) = P(H) + P(C) - P(H \cap C) = \frac{1}{940}(358 + 626 - 112) = \frac{872}{940}$$

	Location in Sputtering Tool		
Contamination	Center	Edge	Total
Low	514	68	582
High	112	246	358
Total	626	314	

Ejemplo. En el mismo contexto del ejemplo anterior, considere E_1 el evento de que un semiconductor contiene 4 o más partículas contaminantes, E_2 es el evento de que un semiconductor se encuentra en la orilla de la herramienta.

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2) = 0.15 + 0.28 - 0.04 = 0.39$$

Number of Contamination Particles	Center	Edge	Totals
0	0.30	0.10	0.40
1	0.15	0.05	0.20
2	0.10	0.05	0.15
3	0.06	0.04	0.10
4	0.04	0.01	0.05
5 or more	0.07	0.03	0.10
Totals	0.72	0.28	1.00

Ejercicios.

- Si $P(A) = 0.3$, $P(B) = 0.2$, y $P(A \cap B) = 0.1$, determine las siguientes probabilidades:

$$P(A^c) \qquad P(A \cup B) \qquad P(A^c \cap B)$$
$$P(A \cap B^c) \qquad P((A \cup B)^c) \qquad P(A^c \cup B)$$

- Si A, B, y C son eventos mutuamente exclusivos, tal que $P(A) = 0.2$, $P(B) = 0.3$, and $P(C) = 0.4$, determine las siguientes probabilidades:

$$P(A \cup B \cup C) \qquad P(A \cap B \cap C) \qquad P(A \cap B)$$
$$P((A \cup B) \cap C) \qquad P(A^c \cap B^c \cap C^c)$$

Definición 9 (Probabilidad condicional)**.** La probabilidad condicional de un evento B dado otro evento A, escrita $P(B|A)$, se define

$$P(B|A) = \frac{P\left(A \cap B\right)}{P(A)}$$

para $P(A) > 0$.

Ejemplo. Considere la siguiente tabla.

		Surface Flaws		
		Yes (event F)	No	Total
Defective	Yes (event D)	10	18	38
	No	30	342	362
	Total	40	360	400

Entonces

$$P(D|F) = \frac{P(D \cap F)}{P(F)} = \frac{10/400}{40/400} = 10/40 = 0.25$$

Considere F' de partes libres de defectos. Entonces

$$P\left(D|F'\right) = \frac{P(D \cap F')}{P(F')} = \frac{18/400}{360/400} = 18/360 = 0.05$$

Ejemplo. Considere un lote con los objetos $\{a, b, c\}$. Si dos objetos son seleccionados aleatoriamente, entonces el espacio muestra es el siguiente:

$$\{ab, ac, ba, bc, ca, cb\}$$

Calcule la probabilidad de que b después de a.

Sea E_1 el evento de que a es seleccionado primero, entonces

$$E_1 = \{ab, ac\}$$

Sea E_2 el evento de que b es seleccionado segundo, entonces

$$E_2 = \{ab, cb\}$$

Entonces:

$$P(E_2|E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)} = \frac{1/6}{2/6} = 1/2$$

Ejercicios.

- Considere la siguiente tabla.

		length	
		excellent	good
surface	excellent	80	2
finish	good	10	8

Sea A el evento de que una muestra tenga excelente acabado, B el evento de que la muestra tenga excelente longitud.

- Determine: $P(A)$, $P(B)$, $P(A|B)$, y $P(B|A)$
- Determine la probabilidad de que si la parte seleccionada tiene excelente acabado, también tenga excelente longitud.
- Si la parte seleccionada tiene buena longitud, calcule la probabilidad de que el acabado sea excelente.

• Considere el experimento que consiste en el lanzamiento de una moneda en tres ocasiones sucesivas. Calcule la propabilidad de $P(A \mid B)$ donde

$$\begin{aligned} A =& \{\text{el resultado contiene más soles que águiles}\} \\ B =& \{\text{el primer lanzamiento resulta en sol}\} \end{aligned}$$

• Un dado de 4 lados es lanzado dos veces, asumimos que los 16 posibles resultados son igualmente posibles. Sean x y y el resultado del primer y segundo lanzamiento, resp. Calcule $P(A \mid B)$, donde

$$\begin{aligned} A =& \{(x, y) \mid \max(x, y) = m\} \\ B =& \{(x, y) \mid \min(x, y) = 2\} \end{aligned}$$

para todo m.

Definición 10 (Regla de la multiplicación)**.**

$$P(A \cap B) = P(B \mid A)P(A) = P(A \mid B)P(B)$$

Ejemplo. La probabilidad de que la bateria de un auto, supediatada a tener alta temperatura, tenga baja carga es de 0.7. La probabilidad de que la bateria tenga alta temperatura es 0.05 (evento T). La probabilidad de que la bateria tenga baja carga (evento C) y alta tempratura es:

$$P(C \cap T) = P(C \mid T)P(T) = 0.7 \cdot 0.05 = 0.035$$

Definición 11 (Regla de la multiplicación generalizada)**.**

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1) \prod_{i=1}^{n} P\left(A_{n+1-i} \,\middle|\, \bigcap_{i=1}^{n-1} A_i\right)$$

for $n > 1$.

Ejemplo. En una producción de 850 objetos, 50 son defectuosos. Si tres objetos son seleccionados de manera aleatoria, calcule la probabilidad de que el los primeros dos sean defectuosos y el tercero no lo sea (evento E). Eventos E_1 y E_2 denotan cuando el primer y segundo objeto son defectuosos, resp. E_3 denota cuando el tercer objeto no es defectuoso. Entonces:

$$P(E) = P(E_1)\, P(E_2|E_1)\, P(E_3|\,(E_2 \cap E_1)) = \frac{50}{850} \cdot \frac{49}{849} \cdot \frac{800}{848}$$

Definición 12 (Regla de la probabilidad total)**.** Sean E_i $(i = 1, \ldots, k)$ eventos mutuamente exclusivos y exhaustivos, entonces

$$P(B) = \sum_{i=1}^{k} P(B \cap E_i) = \sum_{i=1}^{k} P(B|E_i)P(E_i)$$

Ejemplo. Considere que las probabilidades de que un producto falle son 0.1, 0.01 y 0.001, asumiendo que el producto tiene altos, medios y bajos niveles de contaminación, respectivamente. 20% de los productos tienen niveles altos de contaminación, 30% niveles medios y 50% niveles bajos. Sea F el evento de que un producto falle, H, M y L los eventos concernientes a productos con niveles altos, medios y bajos de contaminación. Calcule la probabilidad de F.

$$P(F) = P(F|H)P(H) + P(F|M)P(M) + P(F|L)P(L) = 0.0235.$$

Ejercicios.

- Suponga que $P(A|B) = 0.4$ y $P(B) = 0.5$. Determine $P(A \cap B)$ y $P(B \cap A^c)$.
- En un torneo de ajedrez, la probabilidad de que ganar contra la mitad de los participantes en el torne es 0.3, 0.4 contra un cuarto de los participantes y 0.5 contra el otro cuarto. Calcule la probabilidad de ganar.

Teorema 1 (Bayes)**.**

$$P(A \mid B) = \frac{P(B|A)P(A)}{P(B)}$$

Ejemplo. Considere que las probabilidades de que un producto falle son 0.1, 0.01 y 0.001, asumiendo que el producto tiene altos, medios y bajos niveles de contaminación, respectivamente. 20% de los productos tienen niveles altos de contaminación, 30% niveles medios y 50% niveles bajos. Sea F el evento de que un producto falle, H, M y L los eventos concernientes a productos con niveles altos, medios y bajos de contaminación. Calcule la probabilidad de que un producto tenga altos niveles de contación, asumiendo que el producto falla.

$$P(H \mid F) = \frac{P(F|H)P(H)}{P(F)} = \frac{0.1(0.2)}{0.0235}$$

Ejercicios.

- Calcule $P(B|A)$, dado que $P(A|B) = 0.7$, $P(A) = 0.5$ y $P(B) = 0.2$.
- Un inspector de calidad detecta correctamente objetos defectuosos con 99% de precisión, y 0.5% de las veces clasifica incorrectamente objetos defectuosos. El %0.9 de la producción de la fábrica es defectuoso.

- Calcule la probabilidad de que el inspector detecte un objeto defectuoso.
- Si un objeto es clasificado libre de defectos, determine la probabilidad de que efectivamente lo esté.

Definición 13 (Clasificación Bayesiana)**.** Considere el espacio muestra compuesto por los siguientes vectores:

x_1	x_2	$\dots$	x_n	y
$a_{1,1}$	$a_{1,2}$	$\dots$	$a_{1,n}$	b_1
$a_{2,1}$	$a_{2,2}$	$\dots$	$a_{2,n}$	b_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$a_{m,1}$	$a_{m,2}$	$\dots$	$a_{m,n}$	b_m

El modelo de clasificación Bayesiana se define como sigue:

$$\hat{y} = \max\{P(y_i)P(x_1|y_i)P(x_2|y_i)\dots P(x_n|y_i) \mid i = 1, \dots, m\}$$

Ejemplo. Considere la siguiente base de datos.

Genero	Calificación
Terror	1
Acción	3
Drama	2
Drama	2
Acción	2
Terror	3
Terror	3
Drama	1
Acción	2

La calificación de una película de acción se puede predecir de la siguiente forma:

$$P(1)P(\text{Acción}|1) = 2/9 * 0 = 0$$
$$P(2)P(\text{Acción}|2) = 4/9 * 1/2 = 2/9$$
$$P(3)P(\text{Acción}|3) = 1/3 * 1/3 = 1/9$$

por lo tanto, la calificación de una película de acción será 2.

Ejemplo. Considere la siguiente base de datos.

Usuario	Genero	Calificación
F	Terror	1
M	Acción	3
F	Drama	2
M	Drama	2
F	Acción	2
M	Terror	3
F	Terror	3
M	Drama	1
F	Acción	2

Calcule la calificación que le pondría un usuario M a una película de Drama.

$$P(1)P(\mathrm{M}|1)P(\mathrm{Drama}|1) = 2/9 * 1/2 * 1/3 = 1/27$$
$$P(2)P(\mathrm{M}|2)P(\mathrm{Drama}|2) = 4/9 * 1/3 * 2/3 = 8/81$$
$$P(3)P(\mathrm{M}|3)P(\mathrm{Drama}|3) = 1/3 * 2/3 * 0 = 0$$

Entonces el usuario califica con 2 a la película de Drama.

Ejercicio. Implemente un algoritmo para el modelo de clasificación Bayesiana.