Bloque 2
Aprendizaje Automático

Tema 1:
Razonamiento Probabilístico

Tema 1- Razonamiento Probabilístico

Bibliografía

- S. Russell, P. Norvig. ***Artificial Intelligence. A modern approach.*** Prentice Hall, 4th edición, 2022 (Capítulo 3) http://aima.cs.berkeley.edu/
- S. Russell, P. Norvig. ***Artificial Intelligence. A modern approach.*** Prentice Hall, 3rd edición, 2010 (Capítulo 3)

1. Introducción

Considérese la acción y pregunta siguientes:

$$A_t = \text{SALIR AL AEROPUERTO } t \text{ MINUTOS ANTES DEL VUELO}$$

$$\text{¿ME PERMITE } A_{90} \text{ LLEGAR A TIEMPO?}$$

Es difícil decidir a partir de una respuesta del tipo:

A_{90} ME PERMITE LLEGAR A TIEMPO
SI NO HAY ATASCOS Y NO HAY PINCHAZOS Y . . .
MUCHAS OTRAS CONDICIONES DIFÍCILES DE GARANTIZAR

Es más fácil decidir a partir de una respuesta probabilística:

$$P(A_{90} \text{ ME PERMITE LLEGAR A TIEMPO}) = 0.9999$$

2. Representación probabilística

Supongamos una colección de datos sobre **desplazamientos o viajes en carretera** durante un período de tiempo en un área específica (dataset)

⇩

Ω → espacio muestral (muestras)

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	ω_{10}
Ω	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10

Espacio o modelo de probabilidad es un espacio muestral junto con una función $P : \Omega \rightarrow \mathbb{R}$ que asigna a cada $\omega \in \Omega$ un número real tal que:

$$0 \leq P(\omega) \leq 1; \qquad \sum_{\omega} P(\omega) = 1$$

P(V1)= 1/10 P(V2)= 1/10 P(V3)= 1/10 . . .

2. Representación probabilística

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	ω_{10}
Ω	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
	des	nub	nub	llu	nub	des	llu	des	nub	des
	dia	dia	noc	noc	dia	noc	dia	dia	noc	dia
	seg	seg	acc	acc	seg	seg	seg	seg	seg	acc

Características o elementos a considerar de cada muestra (viaje)

Espacio muestral: Desplazamientos por carretera (Ω). Factores a considerar:

- *Climatología* (C): *despejado* (DES), *nublado* (NUB), *lluvia* (LLU)
- *Luminosidad* (L): *dia* (DIA), *noche* (NOC)
- *Seguridad* (S): desplazamiento *seguro* (SEG), o con *accidente* (ACC)

2. Representación probabilística

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	ω_{10}
Ω	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
	des	nub	nub	llu	nub	des	llu	des	nub	des
	dia	dia	noc	noc	dia	noc	dia	dia	noc	dia
	seg	seg	acc	acc	seg	seg	seg	seg	seg	acc

Una **variable aleatoria** es una función del espacio muestral en algún rango (booleanos, rango de valores predeterminado, etc.)

Variables aleatorias: $C : \Omega \rightarrow \{\text{DES,NUB,LLU}\}$, $L : \Omega \rightarrow \{\text{DIA,NOC}\}$, $S : \Omega \rightarrow \{\text{SEG,ACC}\}$

C(V1)= des
C(V2)= nub

L(V1)= dia
S(V3)=acc

2. Representación probabilística

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	ω_{10}
Ω	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$
	des	nub	nub	llu	nub	des	llu	des	nub	des
	dia	dia	noc	noc	dia	noc	dia	dia	noc	dia
	seg	seg	acc	acc	seg	seg	seg	seg	seg	acc

Dada una variable aleatoria X, P induce una *distribución de probabilidad*: $P(X=x) \stackrel{\text{def}}{=} \sum_{\omega \in (X=x)} P(\omega)$

Probabilidad a priori o incondicional

P(C=des) = 1/10 + 1/10+ 1/10 + 1/10 = 4/10 $\longrightarrow$ P(des)

P(L=noc)= 1/10 + 1/10+ 1/10 + 1/10 = 4/10 $\longrightarrow$ P(noc)

P(S=acc)= 1/10 + 1/10+ 1/10 = 3/10 $\longrightarrow$ P(acc)

2. Representación probabilística

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	ω_{10}
Ω	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
	des	nub	nub	llu	nub	des	llu	des	nub	des
	dia	dia	noc	noc	dia	noc	dia	dia	noc	dia
	seg	seg	acc	acc	seg	seg	seg	seg	seg	acc

Probabilidad (incondicional o "a priori") de una variable aleatoria X:

$$P(X = x) \equiv P(x): \quad \sum_x P(x) = 1$$

P(des) + P(nub) + P(llu) = 1

P(dia) + P(noc) = 1

P(seg) + P(acc) = 1

2. Representación probabilística

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	ω_{10}
Ω	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
	des	nub	nub	llu	nub	des	llu	des	nub	des
	dia	dia	noc	noc	dia	noc	dia	dia	noc	dia
	seg	seg	acc	acc	seg	seg	seg	seg	seg	acc

Probabilidad conjunta de dos variables aleatorias X, Y:

$$P(X = x; Y = y) \equiv P(x, y) : \quad \sum_x \sum_y P(x, y) = 1$$

P(des,dia) = 3/10 P(des,noc) = 1/10

P(nub,dia) = 2/10 P(nub,noc) = 2/10

P(llu,dia) = 1/10 P(llu,noc) = 1/10

P(dia,seg) = 5/10

2. Representación probabilística

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	ω_{10}
Ω	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$
	des	nub	nub	llu	nub	des	llu	des	nub	des
	dia	dia	noc	noc	dia	noc	dia	dia	noc	dia
	seg	seg	acc	acc	seg	seg	seg	seg	seg	acc

Probabilidad conjunta de dos variables aleatorias X, Y:

$$P(X = x; Y = y) \equiv P(x, y) : \quad \sum_x \sum_y P(x, y) = 1$$

P(c,s)	des	nub	llu
seg	3/10	3/10	1/10
acc	1/10	1/10	1/10

Σ=1

P(d,s)	dia	noc
seg	5/10	2/10
acc	1/10	2/10

Σ=1

2. Representación probabilística

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	ω_{10}
Ω	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
	des	nub	nub	llu	nub	des	llu	des	nub	des
	dia	dia	noc	noc	dia	noc	dia	dia	noc	dia
	seg	seg	acc	acc	seg	seg	seg	seg	seg	acc

Probabilidad conjunta de más de dos variables: $P(x,y,z)$: $\sum_x \sum_y \sum_z P(x,y,z) = 1$

	dia			noc		
P(s,c,l)	des	nub	llu	des	nub	llu
seg	2/10	2/10	1/10	1/10	1/10	0
acc	1/10	0	0	0	1/10	1/10

Σ=1

2. Representación probabilística

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	ω_{10}
Ω	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
	des	nub	nub	llu	nub	des	llu	des	nub	des
	dia	dia	noc	noc	dia	noc	dia	dia	noc	dia
	seg	seg	acc	acc	seg	seg	seg	seg	seg	acc

Probabilidad condicional:

$$P(X = x \mid Y = y) \equiv P(x \mid y) : \quad \sum_x P(x \mid y) = 1 \quad \forall y$$

P(des |dia) = 3/6 = 1/2   P(nub |dia) = 2/6 = 1/3   P(llu |dia) = 1/6

P(dia | seg) = 5/7

3. Inferencia probabilística

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	ω_{10}
Ω	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
	des	nub	nub	llu	nub	des	llu	des	nub	des
	dia	dia	noc	noc	dia	noc	dia	dia	noc	dia
	seg	seg	acc	acc	seg	seg	seg	seg	seg	acc

La regla suma: $P(x) = \sum_y P(x,y)$

(también llamada *probabilidad marginal* o *marginalización* de la probabilidad conjunta)

P(des) = P(des,L) = P(des,dia) + P(des,noc) = 3/10 + 1/10 = 4/10

P(des) = P(des,S) = P(des,seg) + P(des,acc) = 3/10 + 1/10 = 4/10

P(acc) = P(acc,C)= P(acc,des)+P(acc,nub)+P(acc,llu)= 1/10 + 1/10 + 1/10 = 3/10

3. Inferencia probabilística

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	ω_{10}
Ω	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
	des	nub	nub	llu	nub	des	llu	des	nub	des
	dia	dia	noc	noc	dia	noc	dia	dia	noc	dia
	seg	seg	acc	acc	seg	seg	seg	seg	seg	acc

La regla producto: $P(x,y) = P(x)\,P(y \mid x) = P(y)\,P(x \mid y)$

P(acc,des) = P(des|acc) P(acc) = 1/3 · 3/10 = 1/10

P(acc,des) = P(acc|des) P(des) = 1/4 · 4/10 = 1/10

P(dia,seg) = P(dia|seg) P(seg) = 5/7 · 7/10 = 5/10

4. Independencia probabilística

La regla producto: $P(x,y) = P(x)\,P(y \mid x) = P(y)\,P(x \mid y)$

Dos variables **x** e **y** son independientes si:

$$P(x,y) = P(x)\,P(y) \quad \text{ó} \quad P(x \mid y) = P(x) \quad \text{ó} \quad P(y \mid x) = P(y)$$

Supongamos que tenemos otra variable aleatoria: JugarVideojuegos(J): $J \in \{0,1\}$

$$P(j,s,c,l) = P(j)\,P(s,c,l)$$

31 ☐ Sean C, L, S variables aleatorias que toman valores en {DES,NUB,LLU}, {DIA, NOC}, y {SEG,ACC}, respectivamente. Su probabilidad conjunta viene dada en la siguiente tabla:

s l c	SEG DIA DES	SEG DIA NUB	SEG DIA LLU	SEG NOC DES	SEG NOC NUB	SEG NOC LLU	ACC DIA DES	ACC DIA NUB	ACC DIA LLU	ACC NOC DES	ACC NOC NUB	ACC NOC LLU
$P(s, l, c)$	0.30	0.20	0.07	0.13	0.10	0.06	0.01	0.01	0.03	0.02	0.02	0.05

La probabilidad condicional $P(C = \text{LLU} | S = \text{ACC}, L = \text{DIA})$ es:

A) 0.60.
B) 0.03.
C) 0.05.
D) 0.02.

5. Teorema de Bayes

Regla de Bayes:

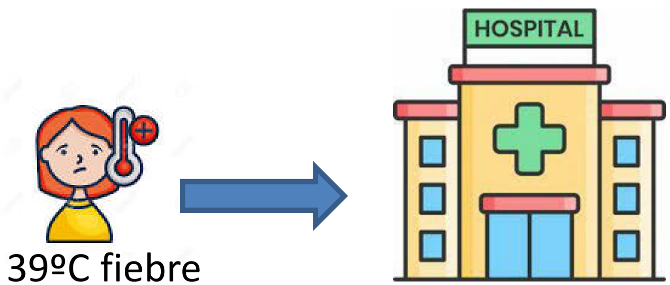
$$P(y \mid x) = \frac{P(x,y)}{P(x)} = \frac{P(y)\,P(x \mid y)}{P(x)} = \frac{P(y)\,P(x \mid y)}{\sum_{y'} P(y')\,P(x \mid y')}$$

(probabilidad de un evento o hipótesis **y**
tras la observación de la evidencia o muestra **x**)

probabilidad a posteriori (clasificación)

verosimilitud (de la muestra **x**)

(probabilidad de que ocurra o se dé la muestra **x** asumiendo que es cierta la estimación **y**)



39ºC fiebre

P(gripe|39) ??

$$P(\text{gripe}|39) = \frac{P(39|\text{gripe})\,P(\text{gripe})}{P(39)}$$

40 $\square$ Supóngase que tenemos dos cajas con 40 naranjas en la primera y 80 naranjas en la segunda. La primera caja contiene 9 naranjas Navelina y 31 Caracara. La segunda caja contiene tres veces más naranjas Navelina que Caracara. Ahora supóngase que se escoge una caja al azar, y luego una naranja al azar de la caja escogida. Si la naranja escogida es Navelina, la probabilidad P de que proceda de la primera caja es:

A) $0/4 \leq P < 1/4$.

B) $1/4 \leq P < 2/4$.

C) $2/4 \leq P < 3/4$.

D) $3/4 \leq P \leq 4/4$.

5. Teoría de la decisión

Minimizar el riesgo de error en la toma de decisiones

Sea $x \in \mathcal{X}$ un *hecho* o *dato* y sea $d \in \mathcal{D}$ una *decisión* que se toma para x.

1) Probabilidad de una decisión: P(d|x)

$$P(gripe|39) = \frac{P(39|gripe)\ P(gripe)}{P(39)} = 0.32$$

2) Probabilidad de error si tomamos la decisión d: $P_d(\text{error} \mid x) = 1 - P(d \mid x)$

$$P_{gripe}(error|39) = 1 - P(gripe|39) = 1 - 0.32 = 0.68$$

5. Teoría de la decisión

3) Decisión óptima (mejor decisión) que minimiza el riesgo de error

$$\forall x \in \mathcal{X} : \quad d^*(x) = \operatorname*{argmax}_{d \in \mathcal{D}} P(d \mid x)$$

D={resfriado, gripe, covid}

Regla de decisión de mínimo riesgo de error o de Bayes:

4) Probabilidad que maximiza la decisión óptima (máxima probabilidad asociada a una decisión)

$$\max_{d \in \mathcal{D}} P(d \mid x)$$

P(gripe|39) = 0.32
P(resfriado|39) = 0.47 → d=resfriado
P(covid|39) = 0.21

5. Teoría de la decisión

5) Mínima probabilidad de error para un dato **x**:

$$\forall x \in \mathcal{X}: \quad P_\star(\text{error} \mid x) = \min_{d \in \mathcal{D}} P_d(\text{error} \mid x) \; = \; 1 - \max_{d \in \mathcal{D}} P(d \mid x)$$

P(gripe|39) = 0.32
P(resfriado|39) = 0.47
P(covid|39) = 0.21

P(error|39) = 1 – 0.47 = 0.53

6) Mínimo riesgo global:

$$P_\star(\text{error}) = \sum_{x \in \mathcal{X}} P_\star(\text{error} \mid x)\, P(x)$$

P(error|39) P(39) + P(error|38) P(38) + P(error|37) P(37) + …

6 ☐ La valoración comercial de las 300 películas proyectadas en un cine durante el pasado año fue de *éxito* para 120 de ellas, y de *fracaso* para el resto. Se conoce las siguientes distribuciones de géneros de películas dada su valoración comercial:

g	Romance	Comedia	Intriga
$P(G = g \mid V = \text{Éxito})$	0.30	0.35	0.35
$P(G = g \mid V = \text{Fracaso})$	0.20	0.50	0.30

¿Cuál es la valoración comercial más probable para una película de intriga?

A) *Éxito*

B) *Fracaso*

C) Ambas valoraciones comerciales son equiprobables

D) No se puede determinar la valoración comercial con los datos disponibles

x → intriga
D= {éxito, fracaso}

P(éxito)=120/300= 0.4
P(fracaso)=180/300= 0.6

Valoración más probable = decisión óptima que minimiza el riesgo ⇒ regla Bayes

$$P(\text{éxito}|\text{intriga}) = \frac{P(\text{intriga}|\text{éxito})\ P(\text{éxito})}{P(\text{intriga})} = \frac{0.35 \cdot 0.4}{0.32} = 0.4375$$

P(intriga|éxito) P(éxito)+P(intriga|fracaso) P(fracaso)=0.35 · 0.4 + 0.30 · 0.6= 0.32

P(fracaso|intriga)=
1 – 0.4375 = 0.5625

Ejercicio *Flores iris* (solución en el boletín)

Un problema clásico de decisión consiste en clasificar flores de la familia *Iris* en tres clases: *setosa, versicolor* y *virgínica*, en base a los tamaños de sus pétalos y sépalos (y).

Para ello se han calculado sendos histogramas de las superficies de los pétalos de una muestra de 50 flores de cada clase. Normalizando estos histogramas, se ha estimado la siguiente distribución de tamaños de pétalos para cada clase (c):

	tamaño de los pétalos en cm^2											
$P(y \mid c)$	<1	1	2	3	4	5	6	7	8	9	10	>10
Setosa	0.90	0.10	0	0	0	0	0	0	0	0	0	0
Versicolor	0	0	0	0.20	0.30	0.32	0.12	0.06	0	0	0	0
Virgínica	0	0	0	0	0	0	0.08	0.12	0.24	0.14	0.20	0.22

Asumiendo que las clases son equiprobables, calcular:

a) Las probabilidades a posteriory $P(c \mid y)$, $c \in \{\text{SETO}, \text{VERS}, \text{VIRG}\}$, para una flor cuyo tamaño de pétalos es $y = 7$ cm^2

b) La decisión óptima de clasificación de esta flor y la probabilidad de que dicha decisión sea errónea

c) La mejor decisión y la correspondiente probab. de error para tamaños de pétalos $1, 2, \ldots, 10$ cm^2

d) La mínima probabilidad de error de decisión esperada para cualquier flor Iris; es decir, $P(\text{error})$

e) Repetir los calculos anteriores, asumiendo que las probabilidades a priori son:
$P(\text{SETO}) = 0.3,\ P(\text{VERS}) = 0.5,\ P(\text{VIRG}) = 0.2$

33 ☐ Sea un problema de clasificación en tres clases para objetos del tipo $\mathbf{x} = (x_1, x_2)^t \in \{0,1\}^2$, con las distribuciones de probabilidad de la derecha. ¿Cuál es el error de Bayes, ε^*, en este problema?

$\mathbf{x}$		$P(c \mid \mathbf{x})$			$P(\mathbf{x})$
x_1	x_2	$c=1$	$c=2$	$c=3$	
0	0	0.6	0.2	0.2	0.2
0	1	0.1	0.1	0.8	0.3
1	0	0.3	0.5	0.2	0.2
1	1	1/3	1/3	1/3	0.3

A) $\varepsilon^* < 0.2$.

B) $0.2 \leq \varepsilon^* < 0.4$.

C) $0.4 \leq \varepsilon^* < 0.7$.

D) $0.7 \leq \varepsilon^*$.

36 ☐ Sea un problema de clasificación en tres clases para datos del tipo $\mathbf{x}=(x_1, x_2)^t \in \{0,1\}^2$, con las distribuciones de probabilidad de la tabla. Indica en qué intervalo se halla el error del clasificador $c(\mathbf{x})$ dado en la tabla, ε:

$\mathbf{x}$		$P(c \mid \mathbf{x})$			$P(\mathbf{x})$	$c(\mathbf{x})$
x_1	x_2	$c=1$	$c=2$	$c=3$		
0	0	0.2	0.1	0.7	0.2	2
0	1	0.4	0.3	0.3	0	1
1	0	0.3	0.4	0.3	0.4	3
1	1	0.4	0.4	0.2	0.4	1

A) $\varepsilon < 0.25$.

B) $0.25 \leq \varepsilon < 0.50$.

C) $0.50 \leq \varepsilon < 0.75$.

D) $0.75 \leq \varepsilon$.

37 ☐ Dada la siguiente tabla de frecuencias conjuntas de las 3 variables de interés:

A	0	0	0	0	1	1	1	1
B	0	0	1	1	0	0	1	1
C	0	1	0	1	0	1	0	1
N(A,B,C)	124	28	227	175	126	222	23	75

¿Cuál es el valor de $P(A = 1 \mid B = 1, C = 0)$?

A) 0.023

B) 0.250

C) 0.092

D) 0.446

17 $\square$

Se tienen dos bolsas. La primera contiene 3 manzanas de color rojo y 5 de color verde; la segunda, 2 rojas, 2 verdes y 1 amarilla. Se escoge una bolsa al azar y, seguidamente, una manzana al azar de la misma. Supóngase que las bolsas tienen la misma probabilidad de ser escogidas y que, dada una bolsa cualquiera, sus manzanas también tienen idéntica probabilidad de ser escogidas. Si la manzana escogida es roja, ¿cuál es la probabilidad P de que sea de la primera bolsa?

A) $0.00 \leq P < 0.25$

B) $0.25 \leq P < 0.50$

C) $0.50 \leq P < 0.75$

D) $0.75 \leq P$

28 $\square$

Sea un problema de clasificación en dos clases, $c = 1, 2$, para objetos en un espacio de representación de 4 elementos, $E = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}$. La tabla de la derecha recoge las (verdaderas) probabilidades a posteriori $P(c \mid \mathbf{x})$, para todo c y $\mathbf{x}$; así como la (verdadera) probabilidad incondicional, $P(\mathbf{x})$, para todo $\mathbf{x}$. Asimismo, dicha tabla incluye la clase asignadada a cada $\mathbf{x} \in E$ por un cierto clasificador $c(\mathbf{x})$. Con base en el conocimiento probabilístico dado, la probabilidad de error de $c(\mathbf{x})$, ε, es:

A) $0/4 \leq \varepsilon < 1/4$.
B) $1/4 \leq \varepsilon < 2/4$.
C) $2/4 \leq \varepsilon < 3/4$.
D) $3/4 \leq \varepsilon \leq 4/4$.

	$P(c \mid \mathbf{x})$			
$\mathbf{x}$	$c = 1$	$c = 2$	$P(\mathbf{x})$	$c(\mathbf{x})$
$\mathbf{x}_1$	1	0	1/3	1
$\mathbf{x}_2$	3/4	1/4	1/4	1
$\mathbf{x}_3$	1/4	3/4	1/4	1
$\mathbf{x}_4$	1/2	1/2	1/6	2

41 $\square$ Sea un problema de clasificación en cuatro clases para datos del tipo $\mathbf{x} = (x_1, x_2)^t \in \{0, 1\}^2$, con las distribuciones de probabilidad de la tabla. Indica en qué intervalo se halla el error de Bayes, ε^*:

A) $\varepsilon^* < 0.40$.
B) $0.40 \leq \varepsilon^* < 0.45$.
C) $0.45 \leq \varepsilon^* < 0.50$.
D) $0.50 \leq \varepsilon^*$.

$\mathbf{x}$		$P(c \mid \mathbf{x})$				
x_1	x_2	$c=1$	$c=2$	$c=3$	$c=4$	$P(\mathbf{x})$
0	0	0.1	0.3	0.1	0.5	0
0	1	0.2	0.5	0.3	0	0.1
1	0	0.2	0.4	0.1	0.3	0.3
1	1	0.1	0.3	0.3	0.3	0.6

8 $\square$ Sean X, Y y Z tres variables aleatorias. Se dice que X e Y son *condicionalmente independientes* dada Z si y solo si

$$P(X = x, Y = y \mid Z = z) \;=\; P(X = x \mid Z = z)\, P(Y = y \mid Z = z) \qquad \text{para todo } x,\ y \text{ y } z.$$

Si se cumple esta igualdad, podemos calcular $P(Z = z \mid X = x, Y = y)$ como sigue:

A) $P(Z = z \mid X = x, Y = y) \;=\; \dfrac{P(X = x, Y = y, Z = z)}{P(X = x, Y = y)}$

B) $P(Z = z \mid X = x, Y = y) \;=\; \dfrac{P(Z = z)\, P(X = x, Y = y \mid Z = z)}{P(X = x, Y = y)}$

C) $P(Z = z \mid X = x, Y = y) \;=\; \dfrac{P(Z = z)\, P(X = x \mid Z = z)\, P(Y = y \mid Z = z)}{P(X = x, Y = y)}$

D) De las tres maneras anteriores.

18 $\square$ Sea x un objeto (vector de características o cadena de símbolos) a clasificar en una clase de C posibles. Indica cuál de los siguientes clasificadores *no* es de error mínimo:

A) $c(x) = \underset{c=1,\ldots,C}{\arg\max}\ P(x \mid c)$

B) $c(x) = \underset{c=1,\ldots,C}{\arg\max}\ P(x, c)$

C) $c(x) = \underset{c=1,\ldots,C}{\arg\max}\ \log P(x, c)$

D) $c(x) = \underset{c=1,\ldots,C}{\arg\max}\ P(c \mid x)$