Sistemas Inteligentes
Cuestiones y ejercicios del bloque 2, tema 1
Razonamiento probabilístico

Escola Tècnica Superior d'Informàtica
Dep. de Sistemes Informàtics i Computació
Universitat Politècnica de València

17 de noviembre de 2024

1. Cuestiones

1 $\boxed{B}$ (Examen de SIN del 15 de enero de 2014; examen del bloque 2; cuestión 1)
Dada la probabilidad conjunta de dos variables aleatorias X y Y, la probabilidad condicional $P(Y = y \mid X = x)$ se puede obtener mediante:

A) $P(y \mid x) \;=\; 1 \;/\; P(x, y)$
B) $P(y \mid x) \;=\; P(x, y) \;/\; \sum_{y'} P(x, y')$
C) $P(y \mid x) \;=\; \sum_{x'} P(x', y) \;/\; \sum_{y'} P(x, y')$
D) $P(y \mid x) \;=\; \sum_{x'} P(x', y) \cdot \sum_{y'} P(x, y')$

2 $\boxed{A}$ (Examen de SIN del 15 de enero de 2014; examen del bloque 2; cuestión 2)
En un problema de decisión binario ($D = \{0, 1\}$), sea y un hecho o dato y $d^\star(y) = 0$ la decisión de mínimo error para ese y. Identifica cuál de las siguientes expresiones determina *incorrectamente* la mínima probabilidad de error para dicho y:

A) $P_\star(\text{error} \mid Y = y) \;=\; 1 - P(D = 1 \mid Y = y)$
B) $P_\star(\text{error} \mid Y = y) \;=\; 1 - P(D = 0 \mid Y = y)$
C) $P_\star(\text{error} \mid Y = y) \;=\; P(D = 1 \mid Y = y)$
D) $P_\star(\text{error} \mid Y = y) \;=\; 1 - \text{máx}_d\, P(D = d \mid Y = y)$

3 $\boxed{D}$ (Examen de SIN del 15 de enero de 2014; examen del bloque 2; cuestión 3)
En un problema de diagnóstico diferencial entre *Gripe* y *Resfriado*, se sabe que la incidencia relativa de la *Gripe* con respecto al *Resfriado* es del 30 % y se conocen las siguientes distribuciones de temperaturas corporales:

$t(^oC)$	36	37	38	39	40
$P(T = t \mid D = \text{GRIPE})$	0.05	0.10	0.20	0.30	0.35
$P(T = t \mid D = \text{RESFR})$	0.10	0.30	0.40	0.15	0.05

La probabilidad a posteriori de que un paciente con 38º de fiebre tenga *Gripe* es:

A) mayor que 0.8
B) menor que 0.1
C) entre 0.3 y 0.6
D) menor que la probabilidad de que con esa temperatura tenga *Resfriado*

4 $\boxed{B}$ (Examen de SIN del 28 de enero de 2014; examen final; cuestión 2)
En un problema de diagnóstico diferencial entre *Gripe* y *Resfriado*, se sabe que la incidencia relativa de la *Gripe* con respecto al *Resfriado* es del 30 % y se conocen las siguientes distribuciones de temperaturas corporales:

$t(^oC)$	36	37	38	39	40
$P(T = t \mid D = \text{GRIPE})$	0.05	0.10	0.20	0.30	0.35
$P(T = t \mid D = \text{RESFR})$	0.10	0.30	0.40	0.15	0.05

$$P(\text{GRIPE} \mid 37) = \frac{\frac{30}{130}\,0.10}{\frac{30}{130}\,0.10 + \frac{100}{130}\,0.30} = \frac{1}{11}$$

El diagnóstico de minimo riesgo de error para un paciente con 37º de fiebre es:

A) *Gripe*
B) *Resfriado*
C) Hay un empate entre ambos diagnósticos
D) Las probabilidades dadas son incorrectas ya que no suman 1; por tanto no es posible hacer un diagnóstico.

5 $\boxed{\text{C}}$ Respecto a la regla de Bayes, ¿cuál de las siguientes expresiones no es correcta?

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

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

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

D) $P(x \mid y) \;=\; \dfrac{P(y \mid x)\; P(x)}{P(y)}$

6 $\boxed{\text{B}}$ 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* $P(V = \text{Fracaso} \mid \text{G=Intriga}) = 0.5625$
C) Ambas valoraciones comerciales son equiprobables
D) No se puede determinar la valoración comercial con los datos disponibles

7 $\boxed{\text{D}}$ En un problema de clasificación en tres clases ($C = \{a, b, c\}$), sea y un hecho o dato. La decisión óptima de clasificación para y es la clase A con una probabilidad a posteriori de 0.40. ¿Cuál de las siguientes afirmaciones es incorrecta?

A) $P(C = a \mid Y = y) \;\leq\; P(C = b \mid Y = y) + P(C = c \mid Y = y)$
B) $P_\star(\text{error} \mid Y = y) \;=\; P(C = b \mid Y = y) + P(C = c \mid Y = y)$
C) $P_\star(\text{error} \mid Y = y) \;=\; 1 - P(C = a \mid Y = y)$
D) $P_\star(\text{error} \mid Y = y) \;=\; 1 - \text{máx}_{d \in \{b,c\}}\, P(C = d \mid Y = y)$

8 $\boxed{\text{D}}$ 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.

9 $\boxed{\text{C}}$ En un problema de clasificación en tres clases ($C = \{a, b, c\}$), en el que se dispone de 100 muestras de la clase a, 100 muestras de la clase b y 100 muestras de la clase c, sea y un hecho o dato. La decisión óptima de clasificación para y es la clase a con una probabilidad a posteriori de 0.50. ¿Cuál de las siguientes afirmaciones es correcta?

A) $P(C = a \mid Y = y) \;>\; P(C = b \mid Y = y) + P(C = c \mid Y = y)$

B) $P(Y = y \mid C = a) \;=\; \dfrac{0.5\; P(C = a)}{P(Y = y)}$

C) $P(Y = y \mid C = a) \;=\; P(Y = y \mid C = b) + P(Y = y \mid C = c)$

D) Ninguna de las anteriores.

10 $\boxed{\text{A}}$ Un médico sabe que:

- La enfermedad de la meningitis causa rigidez de nuca en un 70 % de los casos.
- La probabilidad a priori de que un paciente tenga meningitis es de 1 / 100 000.
- La probabilidad a priori de que un paciente tenga rigidez de nuca es del 1 %.

Con base en el conocimiento anterior, la probabilidad P de que un paciente con rigidez de nuca tenga meningitis es:

A) $0.000 \leq P < 0.001$ $\quad P = P(m \mid r) = \frac{P(m)\,P(r|m)}{P(r)} = \frac{1/100\,000 \cdot 70/100}{1/100} = 0.0007$

B) $0.001 \leq P < 0.002$

C) $0.002 \leq P < 0.003$

D) $0.003 \leq P$

11 $\boxed{\text{D}}$ Considérese un problema de clasificación convencional, esto es, de C clases y objetos representados mediante vectores D-dimensionales de características reales. En términos generales, podemos decir que el problema será más difícil...

A) cuanto menor sean C y D.

B) cuanto menor sea C y mayor sea D.

C) cuanto mayor sea C y menor sea D.

D) cuanto mayor sean C y D.

12 $\boxed{\text{B}}$ Se tiene un problema de clasificación para el cual se han aprendido dos clasificadores diferentes, c_A y c_B. La probabilidad de error de c_A se ha estimado empíricamente, a partir de un cierto conjunto de 100 muestras de test, obteniéndose un valor de $\hat{p}_A = 0.10$ (10 %). La probabilidad de error de c_B se ha estimado análogamente, si bien en este caso se ha empleado un conjunto de test diferente, compuesto por 200 muestras, obteniéndose también un 10 % de error ($\hat{p}_B = 0.10$). Con base en estas estimaciones, podemos afirmar que, para un nivel de confianza del 95 %:

A) Los intervalos de confianza de $\hat{p}_A$ y $\hat{p}_B$ serán idénticos.

B) El intervalo de confianza de $\hat{p}_A$ será mayor que el de $\hat{p}_B$. $\quad I_A = \hat{p}_A \pm 1.96\sqrt{\frac{\hat{p}_A\,(1-\hat{p}_A)}{100}} = 0.10 \pm 0.06$

C) El intervalo de confianza de $\hat{p}_B$ será mayor que el de $\hat{p}_A$. $\quad I_B = \hat{p}_B \pm 1.96\sqrt{\frac{\hat{p}_B\,(1-\hat{p}_B)}{200}} = 0.10 \pm 0.04$

D) Los intervalos de confianza de $\hat{p}_A$ y $\hat{p}_B$ son en este caso irrelevantes ya que las tasas de error estimadas coinciden.

13 $\boxed{\text{D}}$ ¿Cuál de las siguientes expresiones es correcta?

A) $P(x,y) = \sum_z P(x)\,P(y)\,P(z)$.

B) $P(x,y) = \sum_z P(x)\,P(y \mid z)$.

C) $P(x,y) = \sum_z P(x \mid z)\,P(y \mid z)\,P(z)$.

D) $P(x,y) = \sum_z P(x,y \mid z)\,P(z)$. $\quad P(x,y) = \sum_z P(x,y,z) = \sum_z P(x,y \mid z)\,P(z)$

14 $\boxed{\text{A}}$ Un entomólogo descubre lo que podria ser una subespecie rara de escarabajo, debido al patrón de su espalda. En la subespecie rara, el 98 % de los ejemplares tiene dicho patrón. En la subespecie común, el 5 % lo tiene. La subespecie rara representa el 0.1 % de la población. La probabilidad P de que un escarabajo con el patrón sea de la subespecie rara es:

A) $0.00 \leq P < 0.05$. $\quad P = P(r \mid p) = \frac{P(r)\,P(p|r)}{P(p)} = \frac{P(r)\,P(p|r)}{P(r)\,P(p|r)+P(c)\,P(p|c)} = \frac{1/1000\cdot 98/100}{1/1000\cdot 98/100+999/1000\cdot 5/100} = 0.0192$

B) $0.05 \leq P < 0.10$.

C) $0.10 \leq P < 0.20$.

D) $0.20 \leq P$.

15 $\boxed{\text{C}}$ 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) $c(x) \;=\; \underset{c=1,\ldots,C}{\arg\max}\; \log_2\, P(c \mid x)$
B) $c(x) \;=\; \underset{c=1,\ldots,C}{\arg\max}\; \log_{10}\, P(c \mid x)$
C) $c(x) \;=\; \underset{c=1,\ldots,C}{\arg\max}\; a\, P(c \mid x) + b$ siendo a y b dos constantes reales cualesquiera
D) $c(x) \;=\; \underset{c=1,\ldots,C}{\arg\max}\; P(c \mid x)^3$

16 $\boxed{\text{C}}$ ¿Cuál de las siguientes expresiones es *incorrecta*?

A) $P(x \mid y) \;=\; \dfrac{P(x,y)}{\sum_z P(y \mid z)\; P(z)}$
B) $P(x \mid y) \;=\; \dfrac{P(x,y)}{\sum_z P(y,z)}$
C) $P(x \mid y) \;=\; \dfrac{\sum_z P(x,z)}{P(y)}$
D) $P(x \mid y) \;=\; \dfrac{P(y \mid x)\; P(x)}{P(y)}$

17 $\boxed{\text{B}}$ 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) $0.75 \leq P$

$$P = P(B=1 \mid C=r) = \frac{P(B=1)P(C=r|B=1)}{P(C=r)}$$
$$= \frac{P(B=1)P(C=r|B=1)}{P(B=1)P(C=r|B=1)+P(B=2)P(C=r|B=2)}$$
$$= \frac{\frac{1}{2}\cdot\frac{3}{8}}{\frac{1}{2}\cdot\frac{3}{8}+\frac{1}{2}\cdot\frac{2}{5}} = \frac{15}{31} = 0.4839$$

18 $\boxed{\text{A}}$ 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) $c(x) = \underset{c=1,\ldots,C}{\arg\max}\; P(c \mid x)$

19 $\boxed{\text{C}}$ Sean X y Y dos variables aleatorias, y sean $P(X,Y)$, $P(X \mid Y)$, $P(Y \mid X)$, $P(X)$ y $P(Y)$ las probabilidades conjunta, condicionales e incondicionales de esas variables. Indica cuál de las siguientes afirmaciones es *incorrecta*.

A) Tanto $P(X)$ como $P(Y)$ se pueden derivar a partir de $P(X,Y)$.
B) Tanto $P(X \mid Y)$ como $P(Y \mid X)$ se pueden derivar a partir de $P(X,Y)$.
C) Se puede obtener $P(Y \mid X)$ a partir de $P(X \mid Y)$ y $P(X)$, sin necesidad de conocer previamente $P(Y)$.
D) Se puede obtener $P(Y \mid X)$ a partir de $P(X \mid Y)$ y $P(Y)$, sin necesidad de conocer previamente $P(X)$.

20 $\boxed{\text{A}}$ ¿Cuál de las siguientes expresiones es *incorrecta*?
A) $\sum_y P(x \mid y) \;=\; 1,\ \forall x$
B) $\sum_x P(x \mid y) \;=\; 1,\ \forall y$
C) $\sum_x \sum_y P(x,y) \;=\; 1$
D) $\sum_x P(x \mid u) \;=\; \sum_y P(y \mid w),\ \forall u, w$

21 $\boxed{B}$ Se tienen dos almacenes de naranjas: 1 y 2. El 65 % de las naranjas se hallan en el almacén 1 y el resto en el 2. Se sabe que en el almacén 1 hay un 1 % de naranjas no aptas para el consumo; y un 3 % en el 2. Supóngase que se distribuye una naranja no apta para el consumo. ¿Cuál es la probabilidad P de que provenga del almacén 1?

A) $0.00 \leq P < 0.25$

B) $0.25 \leq P < 0.50$ $\quad P = P(A=1|C=0) = \frac{P(A=1)P(C=0|A=1)}{P(C=0)} = \frac{P(A=1)P(C=0|A=1)}{P(A=1)P(C=0|A=1)+P(A=2)P(C=0|A=2)} = 0.38$

C) $0.50 \leq P < 0.75$

D) $0.75 \leq P$

22 $\boxed{D}$ Sea $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$ un objeto dado mediante una secuencia de N vectores de características, el cual se quiere clasificar en una de C clases. Indica cuál de los siguientes clasificadores *sí* es de error mínimo ($\mathbf{x}_2^N$ denota $\mathbf{x}_2, \ldots, \mathbf{x}_N$):

A) $c(\mathbf{x}) = \underset{c=1,\ldots,C}{\arg\max}\ p(\mathbf{x}_1 \mid c)\, p(\mathbf{x}_2^N \mid \mathbf{x}_1)$

B) $c(\mathbf{x}) = \underset{c=1,\ldots,C}{\arg\max}\ p(\mathbf{x}_1, c)\, p(\mathbf{x}_2^N \mid \mathbf{x}_1)$

C) $c(\mathbf{x}) = \underset{c=1,\ldots,C}{\arg\max}\ p(\mathbf{x}_1 \mid c)\, p(\mathbf{x}_2^N \mid \mathbf{x}_1, c)$

D) $c(\mathbf{x}) = \underset{c=1,\ldots,C}{\arg\max}\ p(\mathbf{x}_1, c)\, p(\mathbf{x}_2^N \mid \mathbf{x}_1, c)$

23 $\boxed{D}$ Sea un problema de clasificación en cuatro clases equiprobables, $c = 1, 2, 3, 4$. Dado un objeto x, se sabe que el clasificador de Bayes lo asigna a la clase 1 y que su probabilidad a posteriori de pertenencia a dicha clase, $p(c = 1 \mid x)$, es igual a $1/3$. Con base en el conocimiento dado, indica cuál de las siguientes afirmaciones es correcta:

A) El objeto x puede clasificarse con una probabilidad de error menor que $1/3$.

B) $p(c = 1 \mid x) \ > \ p(c = 2 \mid x) + p(c = 3 \mid x) + p(c = 4 \mid x)$.

C) $p(x) > p(x \mid c = 1)$.

D) Ninguna de las anteriores.

24 $\boxed{D}$ ¿Cuál de las siguientes distribuciones de probabilidad *no puede* deducirse a partir de la prob. conjunta $P(x, y, z)$?:

A) $P(x \mid y)$.

B) $P(z \mid x, y)$.

C) $P(z)$.

D) Toda distribución en la que intervenga cualquier combinación de estas variables puede deducirse de $P(x, y, z)$.

25 $\boxed{C}$ Sea un problema de clasificación en cuatro clases, $C = \{a, b, c, d\}$, donde las cuatro clases son equiprobables, y sea y un hecho o dato. La decisión óptima de clasificación para y es la clase a con una probabilidad a posteriori de 0.30. ¿Cuál de las siguientes afirmaciones es correcta?

A) La probabilidad de error es menor que 0.50.

B) $P(C = a \mid Y = y) \ > \ P(C = b \mid Y = y) + P(C = c \mid Y = y) + P(C = d \mid Y = y)$.

C) $P(Y = y \mid C = a) \ = \ 0.3 \cdot P(Y = y) \ / \ 0.25$.

D) Ninguna de las anteriores.

26 $\boxed{C}$ Supóngase que tenemos dos cajas con 40 galletas cada una. La primera caja contiene 10 galletas de chocolate y 30 sin chocolate. La segunda caja contiene 20 galletas de cada tipo. Ahora supóngase que se escoge una caja al azar, y luego una galleta al azar de la caja escogida. Si la galleta escogida no es de chocolate, 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) $2/4 \leq P < 3/4$. $\quad P(C=1 \mid G=c) = \frac{P(C=1)\, P(G=c \mid C=1)}{P(C=1)\, P(G=c \mid C=1) + P(C=2)\, P(G=c \mid C=2)} = 0.6$

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

27 $\boxed{A}$ Sea $\mathbf{x} = (x_1, \ldots, x_D)^t$, $D > 1$, un objeto representado mediante un vector de características D-dimensional a clasificar en una de C clases. Indica cuál de los siguientes clasificadores *no* es de error mínimo:

A) $c(\mathbf{x}) = \arg\max_{c=1,\ldots,C}\ P(x_1 \mid c)\, P(x_2, \ldots, x_D \mid x_1, c)$

B) $c(\mathbf{x}) = \arg\max_{c=1,\ldots,C}\ P(c)\, P(x_1, \ldots, x_D \mid c)$

C) $c(\mathbf{x}) = \arg\max_{c=1,\ldots,C}\ P(c \mid x_1)\, P(x_2, \ldots, x_D \mid x_1, c)$

D) $c(\mathbf{x}) = \arg\max_{c=1,\ldots,C}\ P(x_1, c)\, P(x_2, \ldots, x_D \mid x_1, c)$

28 $\boxed{\text{B}}$ 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 asignada 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:

	$P(c \mid \mathbf{x})$			
$\mathbf{x}$	$c=1$	$c=2$	$P(\mathbf{x})$	$c(\mathbf{x})$
$\mathbf{x}_1$	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

A) $0/4 \leq \varepsilon < 1/4$.
B) $1/4 \leq \varepsilon < 2/4$. $\varepsilon = \frac{1}{3} \cdot 0 + \frac{1}{4} \cdot \frac{1}{4} + \frac{1}{4} \cdot \frac{3}{4} + \frac{1}{6} \cdot \frac{1}{2} = \frac{1}{3}$
C) $2/4 \leq \varepsilon < 3/4$.
D) $3/4 \leq \varepsilon \leq 4/4$.

29 $\boxed{\text{A}}$ Considérese la probabilidad de error del clasificador de Bayes, o error de Bayes, para el problema de clasificación descrito en la cuestión anterior. Dicho error, que denotamos como ε^*, es:
A) $0/4 \leq \varepsilon^* < 1/4$. $\varepsilon^* = \frac{1}{3} \cdot 0 + \frac{1}{4} \cdot \frac{1}{4} + \frac{1}{4} \cdot \frac{1}{4} + \frac{1}{6} \cdot \frac{1}{2} = \frac{5}{24} = 0.2083$
B) $1/4 \leq \varepsilon^* < 2/4$.
C) $2/4 \leq \varepsilon^* < 3/4$.
D) $3/4 \leq \varepsilon^* \leq 4/4$.

30 $\boxed{\text{C}}$ Dada la siguiente tabla de probabilidades:

B	0	0	1	1
C	0	1	0	1
$P(A=0 \mid B,C)$	0.222	0.298	0.234	0.118
$P(B,C)$	0.025	0.467	0.219	0.290

¿Cuál es el valor de $P(A=1, B=1 \mid C=0)$? $P(A=1, B=1 \mid C=0) = 0.689$

A) $P(A=1, B=1 \mid C=0) \leq 0.25$
B) $0.25 < \; P(A=1, B=1 \mid C=0) \leq 0.50$
C) $0.50 < \; P(A=1, B=1 \mid C=0) \leq 0.75$
D) $0.75 < \; P(A=1, B=1 \mid C=0) \leq 1.00$

31 $\boxed{\text{A}}$ 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. $P(C = \text{LLU} \mid S = \text{ACC}, L = \text{DIA}) = P(C = \text{LLU}, S = \text{ACC}, L = \text{DIA}) / P(S = \text{ACC}, L = \text{DIA})$
B) 0.03. $P(S = \text{ACC}, L = \text{DIA}) = \sum_c P(S = \text{ACC}, L = \text{DIA}, c) = 0.05$
C) 0.05. $P(C = \text{LLU} \mid S = \text{ACC}, L = \text{DIA}) = 0.03/0.05 = 0.60$
D) 0.02.

32 $\boxed{\text{D}}$ Sea $\mathbf{x}$ un objeto a clasificar en una clase de C posibles. Indica cuál de los siguientes clasificadores *no* es de error mínimo (o escoge la última opción si los tres son de error mínimo):
A) $c(\mathbf{x}) = \operatorname*{arg\,max}_{c=1,\ldots,C} \; p(c \mid \mathbf{x})^2$.
B) $c(\mathbf{x}) = \operatorname*{arg\,max}_{c=1,\ldots,C} \; \log p(\mathbf{x}, c)$.
C) $c(\mathbf{x}) = \operatorname*{arg\,max}_{c=1,\ldots,C} \; \sqrt{p(\mathbf{x}, c)} \,/\, p(\mathbf{x})$.
D) Los tres clasificadores anteriores son de error mínimo.

33 $\boxed{\text{C}}$ 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?

$\mathbf{x}$		$P(c \mid \mathbf{x})$			
x_1	x_2	$c{=}1$	$c{=}2$	$c{=}3$	$P(\mathbf{x})$
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) $0.4 \leq \varepsilon^* < 0.7$. $.2 \cdot .4 + .3 \cdot .2 + .2 \cdot .5 + .3 \cdot 2/3 = .44$
D) $0.7 \leq \varepsilon^*$.

34 $\boxed{\text{B}}$ En un problema de razonamiento probabilístico correspondiente a desplazamientos por carretera, con las variables aleatorias de interés: Climatología (C):{despejado (DES), nublado (NUB), lluvioso (LLU)}; Luminosidad (L):{dia (DIA), noche (NOC)}; Seguridad (S):{seguro (SEG), accidente (ACC)}. La probabilidad conjunta de las tres variables viene dada en la tabla:

	DIA			NOC		
$P(s,l,c)$	DES	NUB	LLU	DES	NUB	LLU
SEG	0.29	0.20	0.04	0.14	0.10	0.09
ACC	0.03	0.01	0.03	0.01	0.03	0.03

La probabilidad condicional $P(S = \text{ACC} \mid L = \text{NOC}, C = \text{DES})$ es:

A) 0.010
B) 0.067
C) 0.140
D) 0.150

35 $\boxed{\text{B}}$ En un problema de razonamiento probabilístico correspondiente a desplazamientos por carretera, con las variables aleatorias de interés: Climatología (C):{despejado (DES), nublado (NUB), lluvioso (LLU)}; Luminosidad (L):{dia (DIA), noche (NOC)}; Seguridad (S):{seguro (SEG), accidente (ACC)}. La probabilidad conjunta de las tres variables viene dada en la tabla:

	DIA			NOC		
$P(s,l,c)$	DES	NUB	LLU	DES	NUB	LLU
SEG	0.32	0.23	0.05	0.11	0.07	0.08
ACC	0.03	0.01	0.03	0.01	0.03	0.03

La probabilidad condicional $P(S = \text{ACC} \mid L = \text{NOC}, C = \text{NUB})$ es:

A) 0.140
B) 0.300
C) 0.030
D) 0.100

36 $\boxed{\text{C}}$ 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:

A) $\varepsilon < 0.25$.
B) $0.25 \leq \varepsilon < 0.50$.
C) $0.50 \leq \varepsilon < 0.75$.
D) $0.75 \leq \varepsilon$.

$\mathbf{x}$		$P(c \mid \mathbf{x})$				
x_1	x_2	$c{=}1$	$c{=}2$	$c{=}3$	$P(\mathbf{x})$	$c(\mathbf{x})$
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

$\varepsilon = 0.70$

37 $\boxed{\text{C}}$ 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

38 $\boxed{\text{C}}$ 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:

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	0.3	0.5	0	1
0	1	0.3	0.3	0.4	0.4	1
1	0	0.2	0.5	0.3	0.5	2
1	1	0.3	0.6	0.1	0.1	1

$\varepsilon = 0.60$

A) $\varepsilon < 0.25$.
B) $0.25 \leq \varepsilon < 0.50$.
C) $0.50 \leq \varepsilon < 0.75$.
D) $0.75 \leq \varepsilon$.

39 $\boxed{\text{C}}$ 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)	211	140	245	87	39	110	5	163

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

A) 0.317
B) 0.163
C) 0.652
D) 0.250

40 $\boxed{\text{A}}$ 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: $P = 0.23$

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

41 $\boxed{\text{D}}$ 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) $0.50 \leq \varepsilon^*$.

x		$P(c\mid\mathbf{x})$				$P(\mathbf{x})$
x_1	x_2	$c{=}1$	$c{=}2$	$c{=}3$	$c{=}4$	
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

$\varepsilon^* = 0.65$

42 $\boxed{\text{B}}$ Supóngase que tenemos dos cajas con 40 naranjas en la primera y 80 naranjas en la segunda. La primera caja contiene 26 naranjas Navelina y 14 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: $P = 0.46$

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

43 $\boxed{\text{D}}$ 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) $0.50 \leq \varepsilon^*$.

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

$\varepsilon^* = 0.61$

44 $\boxed{\text{B}}$ Dada la siguiente tabla de probabilidades 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
P(A,B,C)	0.035	0.089	0.085	0.054	0.215	0.161	0.165	0.196

¿Cuál es el valor de $P(A=1, B=1 \mid C=1)$? $P(A=1, B=1 \mid C=1) = 0.392$

A) $P(A=1, B=1 \mid C=1) \leq 0.25$
B) $0.25 < P(A=1, B=1 \mid C=1) \leq 0.50$
C) $0.50 < P(A=1, B=1 \mid C=1) \leq 0.75$
D) $0.75 < P(A=1, B=1 \mid C=1) \leq 1.00$

45 $\boxed{\text{A}}$ Dada la siguiente tabla de probabilidades 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
P(A,B,C)	0.093	0.100	0.133	0.163	0.157	0.150	0.117	0.087

¿Cuál es el valor de $P(A=1, B=1 \mid C=1)$? $P(A=1, B=1 \mid C=1) = 0.174$

A) $P(A=1, B=1 \mid C=1) \leq 0.25$
B) $0.25 < P(A=1, B=1 \mid C=1) \leq 0.50$
C) $0.50 < P(A=1, B=1 \mid C=1) \leq 0.75$
D) $0.75 < P(A=1, B=1 \mid C=1) \leq 1.00$

46 $\boxed{\text{A}}$ En un problema de razonamiento probabilístico correspondiente a diagnóstico de gripe, las variables aleatorias de interés son: Gripe (G):{positivo (POS), negativo (NEG)}; Ventilación (V):{alta (ALT), baja (BAJ)}; Actividad (A):{silencio (SIL), hablando (HAB), ejercicio (EJE)}. La probabilidad conjunta de las tres variables viene dada en la tabla siguiente: $P = 0.03$

	ALT			BAJ		
$P(g, v, a)$	SIL	HAB	EJE	SIL	HAB	EJE
POS	0.01	0.01	0.02	0.01	0.03	0.05
NEG	0.29	0.20	0.10	0.14	0.09	0.05

La probabilidad condicional $P(G = \text{POS} \mid V = \text{ALT}, A = \text{SIL})$ es:

A) $P \leq 0.25$
B) $0.25 < P \leq 0.50$
C) $0.50 < P \leq 0.75$
D) $0.75 < P \leq 1.0$

47 $\boxed{\text{D}}$ Sea $\mathbf{x}$ un objeto a clasificar en una clase de C posibles. Indica cuál de los siguientes clasificadores *no* es de error mínimo (o escoge la última opción si los tres son de error mínimo):

A) $c(\mathbf{x}) = \underset{c=1,\ldots,C}{\arg\min}\ -\log p(c \mid \mathbf{x})$
B) $c(\mathbf{x}) = \underset{c=1,\ldots,C}{\arg\max}\ e^{p(c|\mathbf{x})}$
C) $c(\mathbf{x}) = \underset{c=1,\ldots,C}{\arg\max}\ e^{p(\mathbf{x},c)} - e^{p(\mathbf{x})}$
D) Los tres clasificadores anteriores son de error mínimo.

48 $\boxed{\text{C}}$ 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:

A) $\varepsilon < 0.25$.
B) $0.25 \leq \varepsilon < 0.50$.
C) $0.50 \leq \varepsilon < 0.75$.
D) $0.75 \leq \varepsilon$.

$\mathbf{x}$		$P(c \mid \mathbf{x})$				
x_1	x_2	$c{=}1$	$c{=}2$	$c{=}3$	$P(\mathbf{x})$	$c(\mathbf{x})$
0	0	0.5	0.4	0.1	0.2	2
0	1	0.1	0.8	0.1	0.2	3
1	0	0.3	0.6	0.1	0.2	2
1	1	0.5	0.4	0.1	0.4	3

$\varepsilon = 0.74$

49 $\boxed{B}$ Supóngase que estamos aplicando el algoritmo Perceptrón, con factor de aprendizaje $\alpha = 1$ y margen $b = 0.1$, a un conjunto de 3 muestras bidimensionales de aprendizaje para un problema de 2 clases. Tras procesar las primeras 2 muestras se han obtenido los vectores de pesos $\mathbf{w}_1 = (0, 0, -2)^t$, $\mathbf{w}_2 = (0, 0, 2)^t$. A continuación, se procesa la última muestra $(\mathbf{x}_3, c_3)$ y se obtienen los vectores de pesos $\mathbf{w}_1 = (1, 1, -1)^t$, $\mathbf{w}_2 = (-1, -1, 1)^t$, ¿cuál de las siguientes es esa última muestra?

A) $((2, 3)^t, 1)$
B) $((1, 1)^t, 1)$
C) $((2, 1)^t, 2)$
D) $((2, 5)^t, 2)$

50 $\boxed{C}$ En un problema de razonamiento probabilístico correspondiente a diagnóstico de gripe, las variables aleatorias de interés son: Gripe (G):{positivo (POS), negativo (NEG)}; Ventilación (V):{alta (ALT), baja (BAJ)}; Actividad (A):{silencio (SIL), hablando (HAB), ejercicio (EJE)}. La probabilidad conjunta de las tres variables viene dada en la tabla siguiente:
$P = 0.56$

	ALT			BAJ		
$P(g, v, a)$	SIL	HAB	EJE	SIL	HAB	EJE
POS	0.01	0.02	0.02	0.01	0.03	0.05
NEG	0.29	0.19	0.10	0.14	0.10	0.04

La probabilidad condicional $P(G = \text{POS} \mid V = \text{BAJ}, A = \text{EJE})$ es:

A) $P \leq 0.25$
B) $0.25 < P \leq 0.50$
C) $0.50 < P \leq 0.75$
D) $0.75 < P \leq 1.0$

51 $\boxed{D}$ Sea $\mathbf{x}$ un objeto a clasificar en una clase de C posibles. Indica cuál de los siguientes clasificadores *no* es de error mínimo (o escoge la última opción si los tres son de error mínimo):

A) $c(\mathbf{x}) = \underset{c=1,\ldots,C}{\arg\max}\ \log p(\mathbf{x} \mid c) + \log p(c)$
B) $c(\mathbf{x}) = \underset{c=1,\ldots,C}{\arg\max}\ e^{p(c|\mathbf{x})} + e^{p(\mathbf{x})}$
C) $c(\mathbf{x}) = \underset{c=1,\ldots,C}{\arg\max}\ e^{p(\mathbf{x},c)} - e^{p(\mathbf{x})}$
D) Los tres clasificadores anteriores son de error mínimo.

52 $\boxed{C}$ 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:

A) $\varepsilon < 0.25$.
B) $0.25 \leq \varepsilon < 0.50$.
C) $0.50 \leq \varepsilon < 0.75$.
D) $0.75 \leq \varepsilon$.

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

$\varepsilon = 0.69$

53 $\boxed{A}$ Dada la siguiente tabla de probabilidades:

B	0	0	1	1
C	0	1	0	1
$P(A = 0 \mid B, C)$	0.921	0.900	0.378	0.273
$P(B, C)$	0.322	0.412	0.108	0.157

¿Cuál es el valor de $P(A = 1, B = 1 \mid C = 1)$? $P(A = 1, B = 1 \mid C = 1) = 0.201$

A) $P(A = 1, B = 1 \mid C = 1) \leq 0.25$
B) $0.25 < P(A = 1, B = 1 \mid C = 1) \leq 0.50$
C) $0.50 < P(A = 1, B = 1 \mid C = 1) \leq 0.75$
D) $0.75 < P(A = 1, B = 1 \mid C = 1) \leq 1.00$

2. Problemas

1. (Examen de SIN del 26 de Noviembre de 2012; tiempo estimado: 25 minutos)
Para diseñar un sistema de diagnóstico diferencial entre Gripe y Resfriado, se han elaborado histogramas de valores de temperatura corporal en pacientes con estas enfermedades. A partir de estos histogramas se han obtenido las siguientes distribuciones de temperaturas:

$t(^oC)$	36	37	38	39	40
$P(T = t \mid D = \text{GRIPE})$	0.05	0.10	0.20	0.30	0.35
$P(T = t \mid D = \text{RESFR})$	0.10	0.30	0.40	0.15	0.05

Sabiendo que la incidencia relativa de la gripe con respecto al resfriado es del 30 % (es decir, $P(D = \text{GRIPE}) = 0.3$), determínese:

a) La probabilidad a posteriori de que un paciente con 39 grados de fiebre tenga gripe.
b) El diagnóstico más probable para ese paciente y la probabilidad de que ese diagnóstico sea erróneo.
c) Las probabilidades de los diagnósticos GRIPE y RESFR $\forall t \in \{36, 37, 38, 39, 40\}$, así como el mínimo error global de diagnóstico ($P_\star(error)$) esperado para un sistema diseñado en base a las observaciones utilizadas.

Solución

a)

$$P(D = \text{GRIPE}) = 0.3; \qquad P(D = \text{RESFR}) = 1 - 0.3 = 0.7$$

$$P(D = \text{GRIPE} \mid T = 39) = \frac{P(D = \text{GRIPE})P(T = 39 \mid D = \text{GRIPE})}{P(T = 39)}$$

$$P(T = 39) = P(D = \text{GRIPE})P(T = 39 \mid D = \text{GRIPE}) + P(D = \text{RESFR})P(T = 39 \mid D = \text{RESFR}) =$$
$$0.3 \cdot 0.3 + 0.7 \cdot 0.15 = 0.195$$

$$P(D = \text{GRIPE} \mid T = 39) = \frac{0.3 \cdot 0.3}{0.195} = 0.462$$

b)

$$P(D = \text{RESFR} \mid T = 39) = \frac{0.7 \cdot 0.15}{0.195} = 0.538$$

Diagnóstico más probable:

$$d^\star(T = 39) = \operatorname*{arg\,max}_{d \in \{\text{GRIPE}, \text{RESFR}\}} P(D = d \mid T = 39) = \text{RESFR}$$

Probabilidad de que RESFR sea un diagnóstico erróneo para $t = 39$:

$$P_\star(\text{error} \mid T = 39) = 1 - \text{máx}(P(D = \text{GRIPE} \mid T = 39), P(D = \text{RESFR} \mid T = 39)) = 1 - \text{máx}(0.462, 0.538) = 0.462$$

c) Repitiendo los cálculos anteriores para $t \in \{36, 37, 38, 40\}$:

$t(^oC)$	36	37	38	39	40
$P(T = t)$	0.085	0.240	0.340	0.195	0.140
$P(D = \text{GRIPE} \mid T = t)$	0.176	0.125	0.176	0.462	0.750
$P(D = \text{RESFR} \mid T = t)$	0.824	0.875	0.824	0.538	0.250
$P_\star(\text{error} \mid t)$	0.176	0.125	0.176	0.462	0.250

$$P_\star(\text{error}) = \sum_{t=36}^{40} P_\star(\text{error} \mid T = t)P(T = t) =$$
$$0.176 \cdot 0.085 + 0.125 \cdot 0.240 + 0.176 \cdot 0.340 + 0.462 \cdot 0.195 + 0.250 \cdot 0.140 = .230$$

2. **Ejercicio Flores Iris, transpa #23 del Tema 1**

petal sizes in cm^2

$P(x \mid c)$	<1	1	2	3	4	5	6	7	8	9	10	>10
SETO	0.90	0.10	0	0	0	0	0	0	0	0	0	0
VERS	0	0	0	0.20	0.30	0.32	0.12	0.06	0	0	0	0
VIRG	0	0	0	0	0	0	0.08	0.12	0.24	0.14	0.20	0.22

Solution

Same prior probabilities

Assuming that the three classes have the same probability, calculate:

a) The conditional (posterior) probabilities $P(c \mid x)$, $c \in \{\text{SETO},\text{VERS},\text{VIRG}\}$, for a flower whose petal size is $x = 7\,\text{cm}^2$

$$P(\text{SETO} \mid 7) = \frac{P(7 \mid \text{SETO})P(\text{SETO})}{P(7)} = \frac{0 * 0.33}{P(7)} = 0$$

$$P(\text{VERS} \mid 7) = \frac{P(7 \mid \text{VERS})P(\text{VERS})}{P(7)} = \frac{0.06 * 0.33}{P(7)} = \frac{0.06 * 0.33}{0.06} = \frac{0.02}{0.06} = 0.33$$

$$P(\text{VIRG} \mid 7) = \frac{P(7 \mid \text{VIRG})P(\text{VIRG})}{P(7)} = \frac{0.12 * 0.33}{P(7)} = \frac{0.12 * 0.33}{0.06} = \frac{0.04}{0.06} = 0.67$$

$P(7) = P(7 \mid \text{SETO}) * P(\text{SETO}) + P(7 \mid \text{VERS}) * P(\text{VERS}) + P(7 \mid \text{VIRG}) * P(\text{VIRG}) = 0 + 0.06 * 0.33 + 0.12 * 0.33 = 0.0198 + 0.0396 = 0.06$

b) The decision of optimal classification for this flower and the probability of taking a wrong decision.
Best decision for this flower is class Virginica. The probability of taking a wrong decision is 0.33 ($1 - \text{máx}_{d \in \mathcal{D}} P(d \mid x) = 1 - 0.67 = 0.33$). $P(error \mid 7) = 0.33$.

c) The best decision and probability of error for petals $1, 2, \ldots, 10$ cm^2
For petal = 1cm

$$P(\text{SETO} \mid 1) = \frac{P(1 \mid \text{SETO})P(\text{SETO})}{P(1)} = \frac{0.10 * 0.33}{0.033} = \frac{0.033}{0.033} = 1$$

$$P(\text{VERS} \mid 1) = \frac{P(1 \mid \text{VERS})P(\text{VERS})}{P(1)} = \frac{0 * 0.33}{P(1)} = \frac{0}{0.033} = 0$$

$$P(\text{VIRG} \mid 1) = \frac{P(1 \mid \text{VIRG})P(\text{VIRG})}{P(1)} = \frac{0 * 0.33}{P(1)} = \frac{0}{0.033} = 0$$

$P(1) = P(1 \mid \text{SETO}) * P(\text{SETO}) + P(1 \mid \text{VERS}) * P(\text{VERS}) + P(1 \mid \text{VIRG}) * P(\text{VIRG}) = 0.10 * 0.33 + 0 * 0.33 + 0 * 0.33 = 0.033$

This case is clear. The best decision is the only one possible: Setosa. So the probability of error is 0. $P(error \mid 1) = 0$.

For petal = 2cm

$$P(\text{SETO} \mid 2) = \frac{P(2 \mid \text{SETO})P(\text{SETO})}{P(2)} = \frac{0 * 0.33}{0} = \frac{0}{0} = 0$$

$$P(\text{VERS} \mid 2) = \frac{P(2 \mid \text{VERS})P(\text{VERS})}{P(2)} = \frac{0 * 0.33}{P(2)} = \frac{0}{0} = 0$$

$$P(\text{VIRG} \mid 2) = \frac{P(2 \mid \text{VIRG})P(\text{VIRG})}{P(2)} = \frac{0 * 0.33}{P(2)} = \frac{0}{0} = 0$$

$P(2) = P(2 \mid \text{SETO}) * P(\text{SETO}) + P(2 \mid \text{VERS}) * P(\text{VERS}) + P(2 \mid \text{VIRG}) * P(\text{VIRG}) = 0 * 0.33 + 0 * 0.33 + 0 * 0.33 = 0$

This case is clear. There is no best decision possible so the probability of error is 1. $P(error \mid 2) = 1$.

For petal = 3cm

$$P(\text{SETO} \mid 3) = \frac{P(3 \mid \text{SETO})P(\text{SETO})}{P(3)} = \frac{0 * 0.33}{0.033} = \frac{0}{0.067} = 0$$

$$P(\textsc{vers} \mid 3) = \frac{P(3 \mid \textsc{vers})P(\textsc{vers})}{P(3)} = \frac{0.20 * 0.33}{P(3)} = \frac{0.067}{0.067} = 1$$

$$P(\textsc{virg} \mid 3) = \frac{P(3 \mid \textsc{virg})P(\textsc{virg})}{P(3)} = \frac{0 * 0.33}{P(3)} = \frac{0}{0.067} = 0$$

$P(3) = P(3 \mid \textsc{seto}) * P(\textsc{seto}) + P(3 \mid \textsc{vers}) * P(\textsc{vers}) + P(3 \mid \textsc{virg}) * P(\textsc{virg}) = 0*0.33+0.20*0.33+0*0.33 = 0.067$

This case is clear. The best decision is the only one possible: Versicolor. So the probability of error is 0. $P(error \mid 3) = 0$.

For petal = 4cm

$$P(\textsc{seto} \mid 4) = \frac{P(4 \mid \textsc{seto})P(\textsc{seto})}{P(4)} = \frac{0 * 0.33}{0.033} = \frac{0}{0.01} = 0$$

$$P(\textsc{vers} \mid 4) = \frac{P(4 \mid \textsc{vers})P(\textsc{vers})}{P(4)} = \frac{0.30 * 0.33}{P(4)} = \frac{0.01}{0.01} = 1$$

$$P(\textsc{virg} \mid 4) = \frac{P(4 \mid \textsc{virg})P(\textsc{virg})}{P(4)} = \frac{0 * 0.33}{P(4)} = \frac{0}{0.06} = 0$$

$P(4) = P(4 \mid \textsc{seto}) * P(\textsc{seto}) + P(4 \mid \textsc{vers}) * P(\textsc{vers}) + P(4 \mid \textsc{virg}) * P(\textsc{virg}) = 0*0.33+0.30*0.33+0*0.33 = 0.01$

This case is clear. The best decision is the only one possible: Versicolor. So the probability of error is 0. $P(error \mid 4) = 0$.

For petal = 5cm

$$P(\textsc{seto} \mid 5) = \frac{P(5 \mid \textsc{seto})P(\textsc{seto})}{P(5)} = \frac{0 * 0.33}{0.033} = \frac{0}{0.11} = 0$$

$$P(\textsc{vers} \mid 5) = \frac{P(5 \mid \textsc{vers})P(\textsc{vers})}{P(5)} = \frac{0.33 * 0.33}{P(5)} = \frac{0.11}{0.11} = 1$$

$$P(\textsc{virg} \mid 5) = \frac{P(5 \mid \textsc{virg})P(\textsc{virg})}{P(5)} = \frac{0 * 0.33}{P(5)} = \frac{0}{0.11} = 0$$

$P(5) = P(5 \mid \textsc{seto}) * P(\textsc{seto}) + P(5 \mid \textsc{vers}) * P(\textsc{vers}) + P(5 \mid \textsc{virg}) * P(\textsc{virg}) = 0*0.33+0.32*0.33+0*0.33 = 0.11$

This case is clear. The best decision is the only one possible: Versicolor. So the probability of error is 0. $P(error \mid 5) = 0$.

For petal = 6cm

$$P(\textsc{seto} \mid 6) = \frac{P(6 \mid \textsc{seto})P(\textsc{seto})}{P(6)} = \frac{0 * 0.33}{P(6)} = \frac{0}{0.066} = 0$$

$$P(\textsc{vers} \mid 6) = \frac{P(6 \mid \textsc{vers})P(\textsc{vers})}{P(6)} = \frac{0.12 * 0.33}{P(6)} = \frac{0.0396}{0.066} = 0.60$$

$$P(\textsc{virg} \mid 6) = \frac{P(6 \mid \textsc{virg})P(\textsc{virg})}{P(6)} = \frac{0.08 * 0.33}{P(6)} = \frac{0.0264}{0.066} = 0.40$$

$P(6) = P(6 \mid \textsc{seto}) * P(\textsc{seto}) + P(6 \mid \textsc{vers}) * P(\textsc{vers}) + P(6 \mid \textsc{virg}) * P(\textsc{virg}) = 0*0.33+0.12*0.33+0.08*0.33 = 0.0396 + 0.0264 = 0.066$

The best decision is Versicolor. So the probability of error is 0.40. $P(error \mid 6) = 0.40$.

For petal = 8cm

$$P(\textsc{seto} \mid 8) = \frac{P(8 \mid \textsc{seto})P(\textsc{seto})}{P(8)} = \frac{0 * 0.33}{P(8)} = \frac{0}{0.08} = 0$$

$$P(\textsc{vers} \mid 8) = \frac{P(8 \mid \textsc{vers})P(\textsc{vers})}{P(8)} = \frac{0 * 0.33}{P(8)} = \frac{0}{0.08} = 0$$

$$P(\text{VIRG} \mid 8) = \frac{P(8 \mid \text{VIRG})P(\text{VIRG})}{P(8)} = \frac{0.24 * 0.33}{P(8)} = \frac{0.08}{0.08} = 1$$

$P(8) = P(8 \mid \text{SETO}) * P(\text{SETO}) + P(8 \mid \text{VERS}) * P(\text{VERS}) + P(8 \mid \text{VIRG}) * P(\text{VIRG}) = 0 * 0.33 + 0 * 0.33 + 0.24 * 0.33 = 0.0792$

This case is clear. The best decision is the only one possible: Virginica. So the probability of error is 0. $P(error \mid 8) = 0$.

For petal = 9cm

$$P(\text{SETO} \mid 9) = \frac{P(9 \mid \text{SETO})P(\text{SETO})}{P(9)} = \frac{0 * 0.33}{P(9)} = \frac{0}{0.08} = 0$$

$$P(\text{VERS} \mid 9) = \frac{P(9 \mid \text{VERS})P(\text{VERS})}{P(9)} = \frac{0 * 0.33}{P(9)} = \frac{0}{0.08} = 0$$

$$P(\text{VIRG} \mid 9) = \frac{P(9 \mid \text{VIRG})P(\text{VIRG})}{P(9)} = \frac{0.14 * 0.33}{P(9)} = \frac{0.0462}{0.0462} = 1$$

$P(9) = P(9 \mid \text{SETO}) * P(\text{SETO}) + P(9 \mid \text{VERS}) * P(\text{VERS}) + P(9 \mid \text{VIRG}) * P(\text{VIRG}) = 0 * 0.33 + 0 * 0.33 + 0.14 * 0.33 = 0.0462$

This case is clear. The best decision is the only one possible: Virginica. So the probability of error is 0. $P(error \mid 9) = 0$.

For petal = 10cm

$$P(\text{SETO} \mid 10) = \frac{P(10 \mid \text{SETO})P(\text{SETO})}{P(10)} = \frac{0 * 0.33}{P(10)} = \frac{0}{0.08} = 0$$

$$P(\text{VERS} \mid 10) = \frac{P(10 \mid \text{VERS})P(\text{VERS})}{P(10)} = \frac{0 * 0.33}{P(10)} = \frac{0}{0.08} = 0.$$

$$P(\text{VIRG} \mid 10) = \frac{P(10 \mid \text{VIRG})P(\text{VIRG})}{P(10)} = \frac{0.20 * 0.33}{P(10)} = \frac{0.066}{0.066} = 1$$

$P(10) = P(10 \mid \text{SETO}) * P(\text{SETO}) + P(10 \mid \text{VERS}) * P(\text{VERS}) + P(10 \mid \text{VIRG}) * P(\text{VIRG}) = 0 * 0.33 + 0 * 0.33 + 0.20 * 0.33 = 0.066$

This case is clear. The best decision is the only one possible: Virginica. So the probability of error is 0. $P(error \mid 10) = 0$.

d) The minimum probability of error for any iris flower; that is, $P_\star(\text{error})$
$P_\star(\text{error}) = \sum_x P(error \mid x)P(x) = P(error \mid < 1)P(< 1) +$
$P(error \mid 1)P(1) + P(error \mid 2)P(2) + \ldots + P(error \mid 10)P(10) +$
$P(error \mid > 10)P(> 10) = 0 + 0 + 0 + \ldots + 0.40 * 0.066 + 0.33 * 0.06 + 0 + 0 + 0 + 0 = 0.0264 + 0.0198 = 0.0462 \approx 0.05 = 5\,\%$ error

Section e): Different prior probabilities

Repeat the same calculations assuming that the prior probabilities are:
$P(\text{Seto}) = 0.3,\ P(\text{Vers}) = 0.5,\ P(\text{Virg}) = 0.2$

a) The conditional (posterior) probabilities $P(c \mid x)$, $c \in \{\text{Seto}, \text{Vers}, \text{Virg}\}$, for a flower whose petal size is $x = 7\,\text{cm}^2$

$$P(\text{SETO} \mid 7) = \frac{P(7 \mid \text{SETO})P(\text{SETO})}{P(7)} = \frac{0 * 0.30}{P(7)} = 0$$

$$P(\text{VERS} \mid 7) = \frac{P(7 \mid \text{VERS})P(\text{VERS})}{P(7)} = \frac{0.06 * 0.50}{P(7)} = \frac{0.06 * 0.50}{0.054} = \frac{0.03}{0.054} = 0.56$$

$$P(\text{VIRG} \mid 7) = \frac{P(7 \mid \text{VIRG})P(\text{VIRG})}{P(7)} = \frac{0.12 * 0.20}{P(7)} = \frac{0.12 * 0.20}{0.054} = \frac{0.024}{0.054} = 0.44$$

$P(7) = P(7 \mid \text{SETO}) * P(\text{SETO}) + P(7 \mid \text{VERS}) * P(\text{VERS}) + P(7 \mid \text{VIRG}) * P(\text{VIRG}) = 0 + 0.06 * 0.50 + 0.12 * 0.20 = 0.03 + 0.024 = 0.054$

b) The decision of optimal classification for this flower and the probability of taking a wrong decision.
Best decision for this flower is class Versicolor. The probability of taking a wrong decision is 0.44 ($1 - \max_{d \in \mathcal{D}} P(d \mid x) = 1 - 0.56 = 0.44$). $P(error \mid 7) = 0.44$.

c) For the exercise in section c) and d), we only need to compute the values for petal size 6 and 7, which are the petal sizes which condition the value of the minimum average probability of error (minimum probability of error for any iris flower) as the values for the rest of petal sizes is 0.

$$P(\text{SETO} \mid 7) = \frac{P(7 \mid \text{SETO})P(\text{SETO})}{P(7)} = \frac{0 * 0.3}{P(7)} = 0$$

$$P(\text{VERS} \mid 7) = \frac{P(7 \mid \text{VERS})P(\text{VERS})}{P(7)} = \frac{0.06 * 0.5}{P(7)} = \frac{0.06 * 0.5}{0.054} = \frac{0.03}{0.054} = 0.56$$

$$P(\text{VIRG} \mid 7) = \frac{P(7 \mid \text{VIRG})P(\text{VIRG})}{P(7)} = \frac{0.12 * 0.2}{P(7)} = \frac{0.12 * 0.2}{0.054} = \frac{0.024}{0.054} = 0.44$$

$P(7) = P(7 \mid \text{SETO}) * P(\text{SETO}) + P(7 \mid \text{VERS}) * P(\text{VERS}) + P(7 \mid \text{VIRG}) * P(\text{VIRG}) = 0 + 0.06 * 0.50 + 0.12 * 0.20 = 0.03 + 0.024 = 0.054$

$$P(\text{SETO} \mid 6) = \frac{P(6 \mid \text{SETO})P(\text{SETO})}{P(6)} = \frac{0 * 0.30}{P(6)} = 0$$

$$P(\text{VERS} \mid 6) = \frac{P(6 \mid \text{VERS})P(\text{VERS})}{P(6)} = \frac{0.12 * 0.50}{P(6)} = \frac{0.06}{0.076} = 0.79$$

$$P(\text{VIRG} \mid 6) = \frac{P(6 \mid \text{VIRG})P(\text{VIRG})}{P(6)} = \frac{0.08 * 0.20}{P(6)} = \frac{0.016}{0.076} = 0.21$$

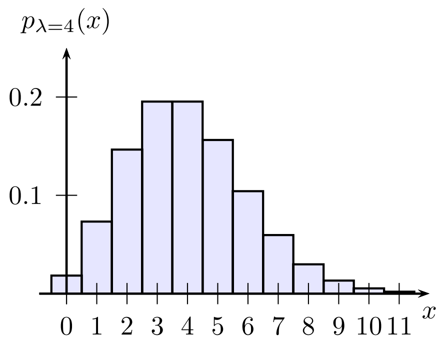
$P(6) = P(6 \mid \text{SETO}) * P(\text{SETO}) + P(6 \mid \text{VERS}) * P(\text{VERS}) + P(6 \mid \text{VIRG}) * P(\text{VIRG}) = 0 + 0.12 * 0.50 + 0.08 * 0.20 = 0.06 + 0.016 = 0.076$

d) The minimum probability of error for any iris flower; that is, $P_\star(\text{error})$
$P_\star(\text{error}) = \sum_x P(error \mid x)P(x) = P(error \mid < 1)P(< 1) +$
$P(error \mid 1)P(1) + P(error \mid 2)P(2) + \ldots + P(error \mid 10)P(10) +$
$P(error \mid > 10)P(> 10) = 0 + 0 + 0 + \ldots + 0.21 * 0.076 + 0.44 * 0.054 + 0 + 0 + 0 + 0 = 0.016 + 0.024 = 0.04 \rightarrow 4\,\%$ error

3. Sea $\lambda \in \mathbb{R}^+$. Decimos que una variable aleatoria $x \in \{0, 1, 2, \ldots\}$ es Poisson(λ) si su función de masa de probabilidad es:

$$p_\lambda(x) = \frac{\exp(-\lambda)\,\lambda^x}{x!}$$

La distribución de Poisson se emplea para modelizar la probabilidad de que un evento dado ocurra un cierto número de veces en un contexto prefijado. El parámetro λ puede interpretarse como la media de ocurrencias de dicho evento. Por ejemplo, x podría ser el número de llamadas telefónicas que recibimos en un día o el número de ocurrencias de una cierta palabra en un documento dado. La figura a la derecha muestra $p_{\lambda=4}(x)$ para todo $x \in \{0, 1, \ldots, 11\}$.



Sea un problema de clasificación en C clases para objetos representados mediante una característica de tipo contador, $x \in \{0, 1, 2, \ldots\}$. Para toda clase c, suponemos dadas:

- Su probabilidad a priori, $P(c)$.
- Su función de (masa de) probabilidad condicional, $P(x \mid c)$, la cual es Poisson(λ_c) con λ_c conocida.

Se pide:

a) (0.5 puntos) Sea el caso particular: $C = 2$, $P(c = 1) = P(c = 2) = \frac{1}{2}$, $\lambda_1 = 1$, $\lambda_2 = 2$ y $x = 2$. Determina la probabilidad *incondicional* de ocurrencia de $x = 2$, $P(x = 2)$.

b) (0.5 puntos) En el caso particular anterior, halla la probabilidad a posteriori $P(c = 2 \mid x = 2)$, así como la probabilidad de error si $x = 2$ se clasifica en la clase $c = 2$.

c) (0.5 puntos) Más generalmente, para cualquier número de clases C y cualesquiera probabilidades a priori, considera el caso en el que, dado un cierto $\tilde{\lambda} \in \mathbb{R}^+$, $\lambda_c = \tilde{\lambda}$ para todo c. En tal caso, existe una clase que no depende de x, c^*, en la que se puede clasificar todo x con mínima probabilidad de error. Determínala.

d) (0.5 puntos) En el caso general, prueba que el clasificador de Bayes para este problema puede expresarse como un clasificador basado en funciones discriminantes lineales como sigue (ln indica logaritmo natural):

$$c^*(x) = \arg\max_c \; g_c(x) \quad \text{con} \quad g_c(x) = w_c\, x + w_{c0}, \quad w_c = \ln \lambda_c \quad \text{y} \quad w_{c0} = \ln p(c) - \lambda_c$$

Solución:

a) $P(x = 2 \mid c = 1) = \frac{1}{2e} = 0.1839 \qquad P(x = 2 \mid c = 2) = \frac{2}{e^2} = 0.2707.$
$P(x = 2) = 0.5 \cdot 0.1839 + 0.5 \cdot 0.2707 = 0.2273.$

b) $P(c = 2 \mid x = 2) = \frac{P(c=2) \cdot P(x=2 \mid c=2)}{P(x=2)} = \frac{0.5 \cdot 0.2707}{0.2273} = 0.5955.$
$P(c \neq 2 \mid x = 2) = 1 - P(c = 2 \mid x = 2) = 0.4045.$

c) $c^*(x) = \arg\max_c P(c)\, P(x \mid c) = \arg\max_c P(c)\ \text{Poisson}(\lambda) = \arg\max_c P(c) = c^*.$

d)

$$\begin{aligned} c^*(x) &= \arg\max_c \; \ln P(c) + \ln P(x \mid c) \\ &= \arg\max_c \; \ln P(c) - \lambda_c + x \, \ln \lambda_c - \ln x! \\ &= \arg\max_c \; x \, \ln \lambda_c + (\ln P(c) - \lambda_c) \end{aligned}$$