Ejercicios Tema 5

Percepción

Curso 2019/2020

1. Sean A y B dos clases con la misma probabilidad a priori y f.d. condicional de clase de tipo Bernoulli:

$$p(\boldsymbol{x} \mid A) \sim Be_2(\boldsymbol{p}_A) \quad \text{con} \quad \boldsymbol{p}_A = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} \qquad\qquad p(\boldsymbol{x} \mid B) \sim Be_2(\boldsymbol{p}_B) \quad \text{con} \quad \boldsymbol{p}_B = \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix}$$

   a) Plantea el clasificador Bernoulli dados los parámetros anteriores

   b) Calcula el error global de este clasificador Bernoulli

2. Tenemos $N = 24$ vectores binarios bidimensionales aleatoriamente extraídos de $C = 2$ distribuciones de Bernoulli independientes:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
x_{n1}	1	0	0	1	1	1	1	1	0	1	1	0	1	1	0	0	0	1	0	0	1	0	0	0
x_{n2}	0	0	1	0	0	0	0	0	1	1	0	1	1	0	0	0	0	0	1	1	0	0	1	0
c_n	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2

   a) Calcula el clasificador Bernoulli más probable con respecto a estos datos

   b) Repite el cálculo anterior considerando únicamente el primer bit

   c) Compara los dos clasificadores Bernoulli calculados

3. Tenemos $N = 12$ vectores binarios tridimensionales aleatoriamente extraídos de $C = 2$ distribuciones Bernoulli independientes:

n	1	2	3	4	5	6	7	8	9	10	11	12
x_{n1}	1	1	1	1	1	1	1	1	0	1	1	0
x_{n2}	1	0	0	1	1	1	0	1	0	1	0	0
x_{n3}	1	0	1	0	1	0	0	0	0	0	0	0
c_n	1	1	1	1	1	1	2	2	2	2	2	2

   a) Calcula el clasificador Bernoulli más probable

   b) Clasifica la siguiente muestra de test $\boldsymbol{y} = (0\ 0\ 1)^t$

   c) Suaviza los prototipos Bernoulli con truncamiento simple ($\epsilon = 0.1$)

   d) Suaviza los prototipos Bernoulli con muestra ficticia

4. Se tiene un clasificador en dos clases basado en distribuciones Bernoulli bidimensionales, de forma que para la clase 1 se tiene $\boldsymbol{p_1} = (0.3\ 0.2)^t$, y para la clase 2 se tiene $\boldsymbol{p_2} = (0.6\ 0.8)^t$. Se pide clasificar la muestra $\boldsymbol{y} = (0\ 1)^t$ empleando $\arg\max_c P(c \mid \boldsymbol{y})$ sabiendo que la probabilidad condicional $p(\boldsymbol{y} \mid c) = \prod_d (p_{cd}\, y_d + (1 - p_{cd})(1 - y_d))$ y las probabilidades a priori son idénticas para ambas clases.

5. Tenemos $N = 12$ vectores binarios tridimensionales aleatoriamente extraídos de $C = 3$ distribuciones Bernoulli independientes:

n	1	2	3	4	5	6	7	8	9	10	11	12
x_{n1}	1	1	1	1	0	0	0	0	0	1	1	0
x_{n2}	0	0	0	0	1	1	0	1	1	1	1	1
x_{n3}	1	0	1	0	1	1	1	1	0	0	0	0
c_n	1	1	1	1	2	2	2	2	3	3	3	3

   a) Estima todos los parámetros del clasificador Bernoulli más probable

   b) Suaviza los prototipos Bernoulli con truncamiento simple ($\epsilon = \frac{1}{8}$) y clasifica la muestra $\boldsymbol{y} = (0\ 0\ 1)^t$

6. Tenemos $N = 12$ vectores 4-dimensionales etiquetados:

n	1	2	3	4	5	6	7	8	9	10	11	12
x_{n1}	1	0	1	1	0	1	0	0	0	0	0	0
x_{n2}	0	1	1	1	0	1	0	0	0	0	0	0
x_{n3}	0	0	0	0	0	0	1	1	0	1	0	1
x_{n4}	0	0	0	0	0	0	1	1	1	1	1	1
c_n	1	1	1	1	1	1	2	2	2	2	2	2

a) Calcula los parámetros del clasificador Bernoulli más probable respecto a estos datos, asumiendo que han sido aleatoriamente extraídos de $C = 2$ distribuciones Bernoulli independientes

b) Suaviza los parámetros Bernoulli del apartado anterior mediante truncamiento simple de $\epsilon = \frac{1}{3}$

7. Tenemos $N = 9$ vectores binarios bidimensionales aleatoriamente extraídos de $C = 2$ distribuciones Bernoulli independientes:

n	1	2	3	4	5	6	7	8	9
x_{n1}	1	0	0	1	1	1	1	1	0
x_{n2}	0	0	1	0	0	0	0	0	1
c_n	1	1	1	2	2	2	2	2	2

a) Calcula los parámetros del clasicador Bernoulli más probable respecto a estos datos

b) Calcula el error global

c) Suaviza los parámetros Bernoulli de ambas clases aplicando truncamiento simple con $\epsilon = \frac{1}{4}$

d) Clasifica el vector $\boldsymbol{y} = (1, 1)^t$ con el clasificador Bernoulli **suavizado** del apartado anterior

8. Sea A y B dos clases con igual probabilidad a priori y probabilidades condicionales de tipo multinomial:

$$p(\boldsymbol{x} \mid A) \sim Mult_2(x_+ = 5, \boldsymbol{p}_A) \;\text{ con }\; \boldsymbol{p}_A = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} \qquad p(\boldsymbol{x} \mid B) \sim Mult_2(x_+ = 5, \boldsymbol{p}_B) \;\text{ con }\; \boldsymbol{p}_B = \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix}$$

Para el correspondiente clasificador multinomial:

a) Calcula las funciones discriminantes

b) Calcula el error global

9. Tenemos $N = 20$ vectores de cuentas bidimensionales aleatoriamente extraídos de $C = 2$ distribuciones multinomiales independientes de longitud 4:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
x_{n1}	4	3	3	4	2	3	4	2	3	2	2	1	1	4	1	2	3	3	3	0
x_{n2}	0	1	1	0	2	1	0	2	1	2	2	3	3	0	3	2	1	1	1	4
c_n	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2

a) Calcula el clasificador multinomial más probable respecto a estos datos.

b) Calcula la probabilidad *a posteriori* para todo $\boldsymbol{x} = (x_1, x_2)^t \in \{0, 1, 2, 3, 4\}^2 \,:\, x_1 + x_2 = 4$.

10. Tenemos $N = 18$ vectores de cuentas 5-dimensionales aleatoriamente extraídos de $C = 3$ distribuciones multinomiales independientes:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
car	0	0	0	0	2	0	2	0	3	2	2	0	0	0	1	1	0	0
people	2	0	1	0	2	1	0	3	1	0	2	0	0	0	0	0	0	0
game	0	2	0	0	0	0	1	2	2	3	0	4	2	0	1	1	0	1
party	3	1	4	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
shell	0	0	0	0	0	0	3	0	0	0	0	0	1	1	0	0	1	0
c_n	1	1	1	1	1	1	2	2	2	2	2	2	3	3	3	3	3	3

a) Calcula el clasificador multinomial más probable respecto a estos datos

b) Clasifica la siguiente muestra de test $\boldsymbol{y} = (1\ 1\ 1\ 1\ 1)^t$

c) Suaviza los parámetros mediante Laplace con $\epsilon = 0.1$

d) Suaviza los parámetros con descuento absoluto de $\epsilon = 0.05$ y backing-off utilizando como distribución generalizada la distribución uniforme

e) Como el anterior pero con interpolación

11. Dados los parámetros de un clasificador multinomial $\hat{\boldsymbol{p}}_1 = (0.5\ 0.3\ 0.2\ 0.0\ 0.0)^t$ y $\hat{\boldsymbol{p}}_2 = (0.0\ 0.0\ 0.3\ 0.3\ 0.4)^t$ con probabilidades *a priori* idénticas, clasifica la muestra de test $\boldsymbol{y} = (1\ 1\ 1\ 1\ 1)^t$ tras aplicar los siguientes suavizados:

- Laplace con $\epsilon = 0.1$
- Descuento absoluto con $\epsilon = 0.05$ y *backing-off* utilizando como distribución generalizada la distribución uniforme

12. Tenemos $N = 12$ vectores de cuentas 5-dimensionales aleatoriamente extraídos de $C = 2$ distribuciones multinomiales independientes:

n	1	2	3	4	5	6	7	8	9	10	11	12
x_{n1}	1	2	1	1	2	1	0	0	0	0	0	0
x_{n2}	2	1	0	1	2	0	0	0	0	0	0	0
x_{n3}	0	0	0	0	0	0	1	1	0	1	0	1
x_{n4}	1	0	2	1	0	2	1	2	1	1	2	1
x_{n5}	0	0	0	0	0	0	1	3	1	1	1	1
c_n	1	1	1	1	1	1	2	2	2	2	2	2

a) Calcula los parámetros del clasificador multinomial más probable respecto a estos datos

b) Suaviza los parámetros multinomial mediante Laplace con $\epsilon = 0.2$

c) Suaviza los parámetros con descuento absoluto de $\epsilon = 0.1$ e interpolación usando como distribución generalizada, la distribución uniforme

d) Suaviza los parámetros con descuento absoluto de $\epsilon = 0.1$ e interpolación usando como distribución generalizada, la distribución $g = (0.1\ 0.2\ 0.4\ 0.2\ 0.1)$, donde g_i es la probabilidad asociada a la dimensión i-ésima de $\boldsymbol{p}_c$

e) Clasifica la muestra de test $\boldsymbol{y} = (1\ 1\ 1\ 1\ 1)^t$ con el clasificador multinomial resultante de aplicar los parámetros suavizados del apartado b)

13. Tenemos $N = 12$ vectores 4-dimensionales etiquetados:

n	1	2	3	4	5	6	7	8	9	10	11	12
x_{n1}	1	0	1	1	0	1	0	0	0	0	0	0
x_{n2}	0	1	1	1	0	1	0	0	0	0	0	0
x_{n3}	0	0	0	0	0	0	1	1	0	1	0	1
x_{n4}	0	0	0	0	0	0	1	1	1	1	1	1
c_n	1	1	1	1	1	1	2	2	2	2	2	2

a) Calcula los parámetros del clasificador multinomial más probable respecto a estos datos, asumiendo que han sido aleatoriamente extraídos de $C = 2$ distribuciones multinomiales independientes

b) Suaviza los parámetros multinomiales del apartado anterior con descuento absoluto de $\epsilon = 0.2$ e interpolación usando como distribución generalizada, la distribución uniforme

14. Tenemos $N = 6$ cadenas de longitud $x_+ = 2$ procedentes de un vocabulario $V = \{a, b, c\}$ aleatoriamente extraídas de $C = 2$ distribuciones multinomiales independientes, donde las muestras de la clase 1 son $\mathcal{X}_1 = \{aa, bb, aa\}$ y las muestras de la clase 2 son $\mathcal{X}_2 = \{ab, bc, ac\}$. Se pide

a) Calcula los parámetros del clasificador multinomial más probable respecto a estos datos

b) Calcula el error global (Nota: $0^0 = 1$)

c) Suaviza los parámetros multinomiales de ambas clases aplicando Laplace con $\epsilon = \frac{1}{3}$

d) Clasifica la cadena $y = cc$ con el clasificador multinomial **suavizado** del apartado anterior

15. Tenemos $N = 16$ vectores de cuentas 5-dimensionales aleatoriamente extraídos de $C = 2$ distribuciones multinomiales independientes:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
x_{n1}	1	2	1	1	2	1	1	0	0	0	0	0	0	0	0	0
x_{n2}	2	1	0	1	2	0	0	3	0	0	0	0	0	0	0	0
x_{n3}	0	0	0	0	0	0	0	0	1	1	0	1	0	1	1	1
x_{n4}	1	0	2	1	0	2	3	3	1	2	1	1	2	1	1	3
x_{n5}	0	0	0	0	0	0	0	0	1	3	1	1	1	1	3	1
c_n	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2

a) Calcula los parámetros del clasificador multinomial más probable respecto a estos datos

b) Suaviza los parámetros multinomial mediante Laplace con $\epsilon = 0.2$

c) Suaviza los parámetros con descuento absoluto de $\epsilon = 0.05$ e interpolación usando como distribución generalizada la distribución uniforme

d) Clasifica la muestra de test $\boldsymbol{y} = (1\ 1\ 1\ 1\ 1)^t$ con el clasificador multinomial resultante de aplicar los parámetros suavizados del apartado c).

16. Sea A y B dos clases con igual probabilidad *a priori* y probabilidades condicionales de clase gaussianas:

$$p(\boldsymbol{x} \mid A) \sim \mathcal{N}_2(\boldsymbol{\mu}_A, \Sigma_A) \qquad \text{y} \qquad p(\boldsymbol{x} \mid B) \sim \mathcal{N}_2(\boldsymbol{\mu}_B, \Sigma_B)$$

Para cada uno de los casos de abajo:

- Calcula funciones discriminantes para A y B
- Calcula la frontera de decisión para el clasificador

	$\boldsymbol{\mu}_A$	Σ_A	$\boldsymbol{\mu}_B$	Σ_B		$\boldsymbol{\mu}_A$	Σ_A	$\boldsymbol{\mu}_B$	Σ_B
1	$\begin{pmatrix}3\\0\end{pmatrix}$	$\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0\\3\end{pmatrix}$	$\begin{pmatrix}1&0\\0&1\end{pmatrix}$	**4**	$\begin{pmatrix}0\\0\end{pmatrix}$	$\begin{pmatrix}1/8&0\\0&1/4\end{pmatrix}$	$\begin{pmatrix}0\\2\end{pmatrix}$	$\begin{pmatrix}1/4&0\\0&1/2\end{pmatrix}$
2	$\begin{pmatrix}2\\7\end{pmatrix}$	$\begin{pmatrix}1&1\\1&2\end{pmatrix}$	$\begin{pmatrix}4\\9\end{pmatrix}$	$\begin{pmatrix}1&1\\1&2\end{pmatrix}$	**5**	$\begin{pmatrix}0\\0\end{pmatrix}$	$\begin{pmatrix}1/2&0\\0&1/8\end{pmatrix}$	$\begin{pmatrix}0\\2\end{pmatrix}$	$\begin{pmatrix}1/4&0\\0&1/2\end{pmatrix}$
3	$\begin{pmatrix}0\\0\end{pmatrix}$	$\begin{pmatrix}1&2\\2&5\end{pmatrix}$	$\begin{pmatrix}0\\4\end{pmatrix}$	$\begin{pmatrix}1&2\\2&5\end{pmatrix}$	**6**	$\begin{pmatrix}0\\0\end{pmatrix}$	$\begin{pmatrix}1/8&0\\0&1/2\end{pmatrix}$	$\begin{pmatrix}0\\2\end{pmatrix}$	$\begin{pmatrix}1/4&0\\0&1/2\end{pmatrix}$
					7	$\begin{pmatrix}0\\0\end{pmatrix}$	$\begin{pmatrix}1/4&0\\0&1/4\end{pmatrix}$	$\begin{pmatrix}0\\2\end{pmatrix}$	$\begin{pmatrix}1/4&0\\0&1/2\end{pmatrix}$

17. Tenemos $N = 6$ vectores reales bidimensionales extraídos aleatoriamente de $C = 2$ distribuciones gaussianas con matriz de covarianza común:

n	1	2	3	4	5	6
x_{n1}	3	3	6	6	9	9
x_{n2}	3	9	6	3	6	0
c_n	1	1	1	2	2	2

a) Calcula el clasificador gaussiano más probable respecto a estos datos

b) Calcula la frontera de decisión del clasificador anterior

c) Clasifica los puntos $\begin{pmatrix}3\\6\end{pmatrix}$ y $\begin{pmatrix}9\\3\end{pmatrix}$

18. Tenemos $N = 10$ valores reales extraídos aleatoriamente de $C = 2$ distribuciones gaussianas:

$$(1.0, A), (1.2, A), (1.3, A), (1.1, A), (0.9, A)$$
$$(1.3, B), (1.2, B), (1.4, B), (1.2, B), (1.3, B)$$

Se pide:

a) Calcula el clasificador gaussiano más probable respecto a estos datos

b) Calcula la frontera de decisión del clasificador anterior

c) Clasifica el punto $x = 1.15$ suponiendo que ambas clases son equiprobables

19. Sean A, B y C tres clases con igual probabilidad *a priori* y f.d. condicional de clase gaussianas gobernadas por sus parámetros

$$\boldsymbol{\mu}_A = \begin{pmatrix} 1 \\ \\ 1 \end{pmatrix} \quad \Sigma_A = \begin{pmatrix} \frac{1}{2} & 0 \\ \\ 0 & 1 \end{pmatrix} \qquad \boldsymbol{\mu}_B = \begin{pmatrix} \frac{1}{2} \\ \\ 1 \end{pmatrix} \quad \Sigma_B = \begin{pmatrix} \frac{1}{4} & 0 \\ \\ 0 & \frac{1}{2} \end{pmatrix} \qquad \boldsymbol{\mu}_C = \begin{pmatrix} 0 \\ \\ \frac{3}{4} \end{pmatrix} \quad \Sigma_C = \begin{pmatrix} 2 & 0 \\ \\ 0 & 1 \end{pmatrix}$$

Calcula $p(\boldsymbol{x} \mid A)$, $p(\boldsymbol{x} \mid B)$ y $p(\boldsymbol{x} \mid C)$ para $\boldsymbol{x} = \begin{pmatrix} 3/4 \\ 5/4 \end{pmatrix}$

20. Se tiene un clasificador en dos clases basado en distribuciones gaussianas unidimensionales, de forma que para la clase 1 se tiene $\mu_1 = 0$, $\sigma_1 = 2$, y para la clase 2 se tiene $\mu_2 = 1$, $\sigma_2 = 1$. Se pide clasificar el punto $y = 0.5$ empleando $\arg\max P(c|y)$ teniendo probabilidades *a priori* idénticas para ambas clases.

21. Se definen las clases A y B en un espacio bidimensional; cada clase está modelada respectivamente por una gaussiana, de forma que sus parámetros son:

$$\boldsymbol{\mu}_A = (\ 1 \quad 1\ )^t \quad \Sigma_A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \qquad \boldsymbol{\mu}_B = (\ 0 \quad -1\ )^t \quad \Sigma_B = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$$

Sabiendo que las probabilidades *a priori* de cada clase son $P(A) = 0.4$ y $P(B) = 0.6$, clasifica la muestra $\boldsymbol{x} = (\ 1 \quad -1\ )^t$.

22. Sea A, B y C tres clases con probabilidades *a priori* $p(A) = 1/2$ y $p(B) = p(C) = 1/4$, y f.d.p. condicionales de clase gaussianas $p(\boldsymbol{x}\,|\,A) \sim \mathcal{N}_2(\boldsymbol{\mu}_A, \Sigma_A)$, $p(\boldsymbol{x}\,|\,B) \sim \mathcal{N}_2(\boldsymbol{\mu}_B, \Sigma_B)$ y $p(\boldsymbol{x}\,|\,C) \sim \mathcal{N}_2(\boldsymbol{\mu}_C, \Sigma_C)$

$$\boldsymbol{\mu}_A = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Sigma_A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \qquad \boldsymbol{\mu}_B = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \quad \Sigma_B = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \qquad \boldsymbol{\mu}_C = \begin{pmatrix} -2 \\ -2 \end{pmatrix} \quad \Sigma_C = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$

a) Calcula funciones discriminantes para A, B y C

b) Calcula la frontera de decisión entre las clases A y B

c) Calcula la frontera de decisión entre las clases B y C

d) Clasifica la muestra $\boldsymbol{y} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

Nota: $\Sigma_A^{-1} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$ y $\Sigma_B^{-1} = \Sigma_C^{-1} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$

23. Sea A y B dos clases con probabilidades *a priori* $p(A) = 1/2$ y $p(B) = 1/2$, y f.d.p. condicionales de clase gaussianas $p(\boldsymbol{x}\,|\,A) \sim \mathcal{N}_2(\boldsymbol{\mu}_A, \Sigma_A)$ y $p(\boldsymbol{x}\,|\,B) \sim \mathcal{N}_2(\boldsymbol{\mu}_B, \Sigma_B)$

$$\boldsymbol{\mu}_A = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Sigma_A = \begin{pmatrix} 1/2 & 0 \\ 0 & 2 \end{pmatrix} \qquad \boldsymbol{\mu}_B = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \Sigma_B = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

a) Calcula funciones discriminantes para A y B

b) Calcula la frontera de decisión entre las clases A y B

c) Representa gráficamente la frontera y las regiones de decisión

d) Se desea suavizar las matrices de covarianza de ambas clases mediante *flat smoothing*. Calcula el valor de α necesario para que el clasificador resultante sea lineal

Nota: $\Sigma_A^{-1} = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix}$ y $\Sigma_B^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix}$

24. Sea A y B dos clases con probabilidades *a priori* $p(A) = 1/4$ y $p(B) = 3/4$, y f.d.p. condicionales de clase gaussianas $p(\boldsymbol{x}\,|\,A) \sim \mathcal{N}_2(\boldsymbol{\mu}_A, \Sigma_A)$ y $p(\boldsymbol{x}\,|\,B) \sim \mathcal{N}_2(\boldsymbol{\mu}_B, \Sigma_B)$

$$\boldsymbol{\mu}_A = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Sigma_A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \boldsymbol{\mu}_B = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \Sigma_B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$a)$ Calcula funciones discriminantes para A y B

$b)$ Calcula la frontera de decisión entre las clases A y B

$c)$ Representa gráficamente la frontera y las regiones de decisión

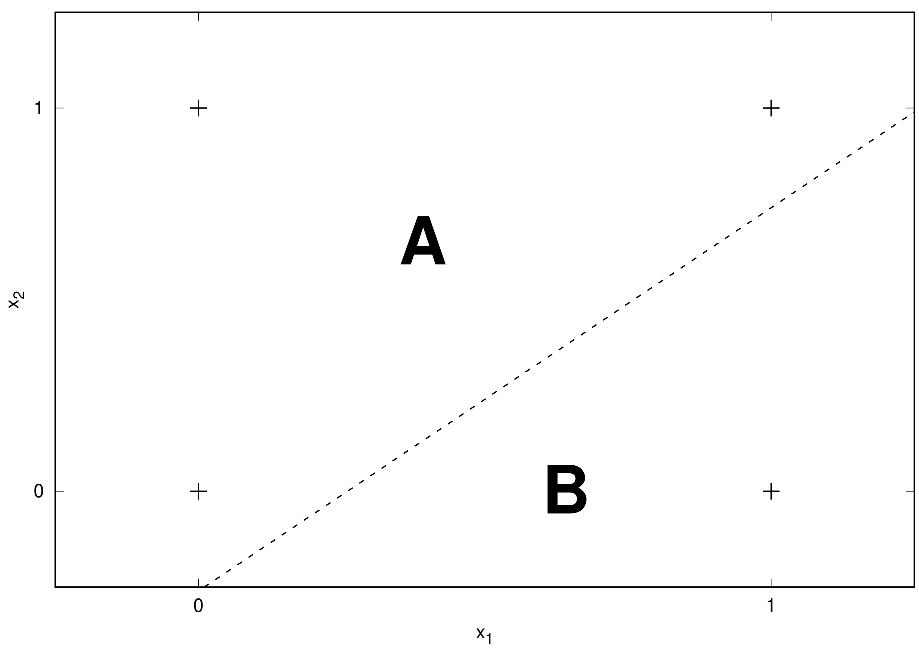
Soluciones

1. *a*)

$$g_A(\boldsymbol{x}) = \frac{1}{2}\left(\frac{1}{2}\right)^{x_1}\left(\frac{1}{2}\right)^{1-x_1}\left(\frac{1}{2}\right)^{x_2}\left(\frac{1}{2}\right)^{1-x_2} \qquad g_B(\boldsymbol{x}) = \frac{1}{2}\left(\frac{3}{4}\right)^{x_1}\left(\frac{1}{4}\right)^{1-x_1}\left(\frac{1}{4}\right)^{x_2}\left(\frac{3}{4}\right)^{1-x_2}$$

Si tuvieramos que calcular la frontera de decisión entre la clase A y B, es más sencillo hacerlo para los clasificadores equivalentes usando $\log_2$:

$$\begin{aligned}
\log_2 g_A(\boldsymbol{x}) &= \log_2 g_B(\boldsymbol{x})\\
-1 - x_1 - (1-x_1) - x_2 - (1-x_2) &= -1 + x_1 \log_2 \frac{3}{4} + (1-x_1)\log_2 \frac{1}{4} + x_2 \log_2 \frac{1}{4} + (1-x_2)\log_2 \frac{3}{4}\\
-3 &= -1 + x_1(\log_2 3 - 2) - 2(1-x_1) - 2\,x_2 + (1-x_2)(\log_2 3 - 2)\\
-3 &= -1 + x_1 \log_2 3 - 2 - 2\,x_2 + \log_2 3 - 2 - x_2 \log_2 3 + 2x_2\\
-3 &= -5 + \log_2 3 + x_1 \log_2 3 - x_2 \log_2 3\\
x_2 &= \frac{-2 + \log_2 3}{\log_2 3} + x_1 = -0.26 + x_1
\end{aligned}$$



b) En general, el error global $p(e)$ se calcula como

$$p(e) = \sum_{\boldsymbol{x}} p(\boldsymbol{x}, e) = \sum_{\boldsymbol{x}} p(\boldsymbol{x})\, p(e \mid \boldsymbol{x}) = \sum_{\boldsymbol{x}} p(\boldsymbol{x}) \left(1 - \max_c p(c \mid \boldsymbol{x})\right).$$

Aplicamos Bayes y simplificamos la expresión teniendo en cuenta que sólo tenemos dos clases

$$\begin{aligned}
p(e) &= \sum_{\boldsymbol{x}} p(\boldsymbol{x}) \min_c p(c \mid \boldsymbol{x}) = \sum_{\boldsymbol{x}} p(\boldsymbol{x}) \min_c \frac{p(c)\, p(\boldsymbol{x} \mid c)}{p(\boldsymbol{x})} = \sum_{\boldsymbol{x}} \min_c p(c)\, p(\boldsymbol{x} \mid c)\\
&= \sum_{\boldsymbol{x}} \min_c p(c) \prod_d p_{cd}^{x_d} (1 - p_{cd})^{(1-x_d)}\\
&= \sum_{\boldsymbol{x}} \min\left(\frac{1}{2} \cdot \frac{1}{2}^{x_1}\left(1 - \frac{1}{2}\right)^{(1-x_1)} \cdot \frac{1}{2}^{x_2}\left(1 - \frac{1}{2}\right)^{(1-x_2)}, \frac{1}{2} \cdot \frac{3}{4}^{x_1}\left(1 - \frac{3}{4}\right)^{(1-x_1)} \cdot \frac{1}{4}^{x_2}\left(1 - \frac{1}{4}\right)^{(1-x_2)}\right)\\
&= 0.34375
\end{aligned}$$

donde $\boldsymbol{x} \in \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$

2. *a*)

$$g_1(\boldsymbol{x}) = \frac{1}{2}\left(\frac{2}{3}\right)^{x_1}\left(\frac{1}{3}\right)^{1-x_1}\left(\frac{1}{3}\right)^{x_2}\left(\frac{2}{3}\right)^{1-x_2} \qquad g_2(\boldsymbol{x}) = \frac{1}{2}\left(\frac{1}{3}\right)^{x_1}\left(\frac{2}{3}\right)^{1-x_1}\left(\frac{1}{3}\right)^{x_2}\left(\frac{2}{3}\right)^{1-x_2}$$

b)

$$g_1(\boldsymbol{x}) = \frac{1}{2}\left(\frac{2}{3}\right)^{x_1}\left(\frac{1}{3}\right)^{1-x_1} \qquad g_2(\boldsymbol{x}) = \frac{1}{2}\left(\frac{1}{3}\right)^{x_1}\left(\frac{2}{3}\right)^{1-x_1}$$

c) Son equivalentes

3. *a*)

$$g_1(\boldsymbol{x}) = \frac{1}{2}(1)^{x_1}(0)^{1-x_1}\left(\frac{2}{3}\right)^{x_2}\left(\frac{1}{3}\right)^{1-x_2}\left(\frac{1}{2}\right)^{x_3}\left(\frac{1}{2}\right)^{1-x_3}$$

$$g_2(\boldsymbol{x}) = \frac{1}{2}\left(\frac{2}{3}\right)^{x_1}\left(\frac{1}{3}\right)^{1-x_1}\left(\frac{1}{3}\right)^{x_2}\left(\frac{2}{3}\right)^{1-x_2}(0)^{x_3}(1)^{1-x_3}$$

b) Ambas F.D. dan valor 0, no es posible clasificar.

c)

$$\tilde{\boldsymbol{p}}_1 = \left(\frac{9}{10}, \frac{2}{3}, \frac{1}{2}\right)^t \qquad \tilde{\boldsymbol{p}}_2 = \left(\frac{2}{3}, \frac{1}{3}, \frac{1}{10}\right)^t$$

d)

$$\tilde{\boldsymbol{p}}_1 = \left(\frac{7}{8}, \frac{5}{8}, \frac{1}{2}\right)^t \qquad \tilde{\boldsymbol{p}}_2 = \left(\frac{5}{8}, \frac{3}{8}, \frac{1}{8}\right)^t$$

4. (Examen Recuperación Primer Parcial 2013)

$$\hat{c}(\boldsymbol{y}) = \arg\max_c P(c \mid \boldsymbol{y}) \approx \arg\max_c p(\boldsymbol{y} \mid c)\, p(c)$$

Para la clase 1:

$$\begin{aligned} p(\boldsymbol{y} = (0\ 1)^t \mid c = 1) &= (p_{11}\, y_1 + (1 - p_{11})(1 - y_1)) \cdot (p_{12}\, y_2 + (1 - p_{12})(1 - y_2)) \\ &= (0.3 \cdot 0 + (1 - 0.3) \cdot (1 - 0)) \cdot (0.2 \cdot 1 + (1 - 0.2) \cdot (1 - 1)) \\ &= 0.14 \end{aligned}$$

Para la clase 2:

$$\begin{aligned} p(\boldsymbol{y} = (0\ 1)^t \mid c = 2) &= (p_{21}\, y_1 + (1 - p_{21})(1 - y_1)) \cdot (p_{22}\, y_2 + (1 - p_{22})(1 - y_2)) \\ &= (0.6 \cdot 0 + (1 - 0.6)(1 - 0)) \cdot (0.8 \cdot 1 + (1 - 0.8) \cdot (1 - 1)) \\ &= 0.32 \end{aligned}$$

Dado que $p(c = 1) = p(c = 2) = 0.5$, la muestra $\boldsymbol{y}$ se clasifica en la clase 2.

5. (Examen Recuperación Segundo Parcial 2013)

   a)

$$p(1) = p(2) = p(3) = \frac{1}{3}$$

$$\hat{\boldsymbol{p}}_1 = \frac{1}{4}\begin{pmatrix} 1+1+1+1 \\ 0+0+0+0 \\ 1+0+1+0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1/2 \end{pmatrix}$$

$$\hat{\boldsymbol{p}}_2 = \frac{1}{4}\begin{pmatrix} 0+0+0+0 \\ 1+1+0+1 \\ 1+1+1+1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3/4 \\ 1 \end{pmatrix}$$

$$\hat{\boldsymbol{p}}_3 = \frac{1}{4}\begin{pmatrix} 0+1+1+0 \\ 1+1+1+1 \\ 0+0+0+0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1 \\ 0 \end{pmatrix}$$

   b)

$$\hat{\boldsymbol{p}}_1 = \begin{pmatrix} 1 \\ 0 \\ 1/2 \end{pmatrix} \rightarrow \hat{\boldsymbol{p}}_1 = \begin{pmatrix} 7/8 \\ 1/8 \\ 1/2 \end{pmatrix}$$

$$\hat{\boldsymbol{p}}_2 = \begin{pmatrix} 0 \\ 3/4 \\ 1 \end{pmatrix} \rightarrow \hat{\boldsymbol{p}}_2 = \begin{pmatrix} 1/8 \\ 3/4 \\ 7/8 \end{pmatrix}$$

$$\hat{\boldsymbol{p}}_3 = \begin{pmatrix} 1/2 \\ 1 \\ 0 \end{pmatrix} \rightarrow \hat{\boldsymbol{p}}_3 = \begin{pmatrix} 1/2 \\ 7/8 \\ 1/8 \end{pmatrix}$$

   En nuestro caso, dado que las priors son idénticas, la regla de clasificación se reduce a:

$$p(\boldsymbol{y} = (0\ 0\ 1) \mid c = 1) = (7/8)^0(1-7/8)^{(1-0)}(1/8)^0(1-1/8)^{(1-0)}(1/2)^1(1-1/2)^{(1-1)} = 1/8 \cdot 7/8 \cdot 1/2 = \frac{7}{128}$$
$$p(\boldsymbol{y} = (0\ 0\ 1) \mid c = 2) = (1/8)^0(1-1/8)^{(1-0)}(3/4)^0(1-3/4)^{(1-0)}(7/8)^1(1-7/8)^{(1-1)} = 7/8 \cdot 1/4 \cdot 7/8 = \frac{49}{256}$$
$$p(\boldsymbol{y} = (0\ 0\ 1) \mid c = 3) = (1/2)^0(1-1/2)^{(1-0)}(7/8)^0(1-7/8)^{(1-0)}(1/8)^1(1-1/8)^{(1-1)} = 1/2 \cdot 1/8 \cdot 1/8 = \frac{1}{128}$$

   La muestra $\boldsymbol{y}$ se clasifica en la clase 2.

6. (Examen Recuperación Segundo Parcial 2014)

   a)

$$p(1) = p(2) = \frac{6}{12} = 0.5$$

$$\hat{\boldsymbol{p}}_1 = \frac{1}{6}\begin{pmatrix} 1+0+1+1+0+1 \\ 0+1+1+1+0+1 \\ 0+0+0+0+0+0 \\ 0+0+0+0+0+0 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 4 \\ 4 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2/3 \\ 2/3 \\ 0.0 \\ 0.0 \end{pmatrix}$$

$$\hat{\boldsymbol{p}}_2 = \frac{1}{6}\begin{pmatrix} 0+0+0+0+0+0 \\ 0+0+0+0+0+0 \\ 1+1+0+1+0+1 \\ 1+1+1+1+1+1 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 0 \\ 0 \\ 4 \\ 6 \end{pmatrix} = \begin{pmatrix} 0.0 \\ 0.0 \\ 2/3 \\ 1.0 \end{pmatrix}$$

b)

$$\hat{\boldsymbol{p}}_1 = \begin{pmatrix} 2/3 \\ 2/3 \\ 0.0 \\ 0.0 \end{pmatrix} \rightarrow \hat{\boldsymbol{p}}_1 = \begin{pmatrix} 2/3 \\ 2/3 \\ 1/3 \\ 1/3 \end{pmatrix}$$

$$\hat{\boldsymbol{p}}_2 = \begin{pmatrix} 0.0 \\ 0.0 \\ 2/3 \\ 1.0 \end{pmatrix} \rightarrow \hat{\boldsymbol{p}}_2 = \begin{pmatrix} 1/3 \\ 1/3 \\ 2/3 \\ 2/3 \end{pmatrix}$$

7. (Examen Recuperación Segundo Parcial 2015)

a)

$$p(1) = \frac{3}{9} = \frac{1}{3} \quad p(2) = \frac{6}{9} = \frac{2}{3}$$

$$\hat{\boldsymbol{p}}_1 = \frac{1}{3}\begin{pmatrix} 1+0+0 \\ 0+0+1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}$$

$$\hat{\boldsymbol{p}}_2 = \frac{1}{6}\begin{pmatrix} 1+1+1+1+1+0 \\ 0+0+0+0+0+1 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 5 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{5}{6} \\ \frac{1}{6} \end{pmatrix}$$

b) En general, el error global $p(e)$ se calcula como

$$p(e) = \sum_{\boldsymbol{x}} p(\boldsymbol{x}, e) = \sum_{\boldsymbol{x}} p(\boldsymbol{x})\, p(e \mid \boldsymbol{x}) = \sum_{\boldsymbol{x}} p(\boldsymbol{x}) \left(1 - \max_c p(c \mid \boldsymbol{x})\right).$$

Aplicamos Bayes y simplificamos la expresión teniendo en cuenta que sólo tenemos dos clases

$$p(e) = \sum_{\boldsymbol{x}} p(\boldsymbol{x}) \min_c p(c \mid \boldsymbol{x}) = \sum_{\boldsymbol{x}} p(\boldsymbol{x}) \min_c \frac{p(c)\, p(\boldsymbol{x} \mid c)}{p(\boldsymbol{x})} = \sum_{\boldsymbol{x}} \min_c p(c)\, p(\boldsymbol{x} \mid c) = \sum_{\boldsymbol{x}} \min_c p(c) \prod_d p_{cd}^{x_d} (1 - p_{cd})^{(1-x_d)}$$

$$= \sum_{\boldsymbol{x}} \min\left(\frac{1}{3} \cdot \frac{1}{3}^{x_1}\left(1 - \frac{1}{3}\right)^{(1-x_1)} \cdot \frac{1}{3}^{x_2}\left(1 - \frac{1}{3}\right)^{(1-x_2)}, \frac{2}{3} \cdot \frac{5}{6}^{x_1}\left(1 - \frac{5}{6}\right)^{(1-x_1)} \cdot \frac{1}{6}^{x_2}\left(1 - \frac{1}{6}\right)^{(1-x_2)}\right)$$

$\boldsymbol{x}$		$p(c)\, p(\boldsymbol{x} \mid c)$		
x_1	x_2	$c = 1$	$c = 2$	$\min_c$
0	0	$\frac{1}{3} \cdot \frac{2}{3} \cdot \frac{2}{3}$	$\frac{2}{3} \cdot \frac{1}{6} \cdot \frac{5}{6}$	$\frac{10}{108}$
0	1	$\frac{1}{3} \cdot \frac{2}{3} \cdot \frac{1}{3}$	$\frac{2}{3} \cdot \frac{1}{6} \cdot \frac{1}{6}$	$\frac{2}{108}$
1	0	$\frac{1}{3} \cdot \frac{1}{3} \cdot \frac{2}{3}$	$\frac{2}{3} \cdot \frac{5}{6} \cdot \frac{5}{6}$	$\frac{2}{27}$
1	1	$\frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3}$	$\frac{2}{3} \cdot \frac{5}{6} \cdot \frac{1}{6}$	$\frac{1}{27}$

Sumando el mínimo de $p(c)\, p(\boldsymbol{x} \mid c)$ para cada $\boldsymbol{x}$:

$$p(e) = \frac{10}{108} + \frac{2}{108} + \frac{2}{27} + \frac{1}{27} = \frac{2}{9}$$

c)

$$\tilde{\boldsymbol{p}}_1 = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}$$

$$\tilde{\boldsymbol{p}}_2 = \begin{pmatrix} \frac{3}{4} \\ \frac{1}{4} \end{pmatrix}$$

d) Aplicamos la regla de clasificación de Bayes:

$$\hat{c}(\boldsymbol{y}) = \arg\max_c p(c)\, p(\boldsymbol{y} \mid c) = p(c) \prod_d p_{cd}^{y_d} \, (1 - p_{cd})^{(1-y_d)}$$

$$p(c=1)\, p(\boldsymbol{y} = (1\ 1)^t \mid c=1) = \frac{1}{3} \cdot \frac{1}{3}^{y_1} \left(1 - \frac{1}{3}\right)^{(1-y_1)} \cdot \frac{1}{3}^{y_2} \left(1 - \frac{1}{3}\right)^{(1-y_2)} = \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{27}$$

$$p(c=2)\, p(\boldsymbol{y} = (1\ 1)^t \mid c=2) = \frac{2}{3} \cdot \frac{3}{4}^{y_1} \left(1 - \frac{3}{4}\right)^{(1-y_1)} \cdot \frac{1}{4}^{y_2} \left(1 - \frac{1}{4}\right)^{(1-y_2)} = \frac{2}{3} \cdot \frac{3}{4} \cdot \frac{1}{4} = \frac{1}{8}$$

La muestra $\boldsymbol{y}$ se clasifica en la clase 2.

8. *a*)

$$g_A(\boldsymbol{x}) = \left(\frac{1}{2}\right)^{x_1} \left(\frac{1}{2}\right)^{x_2} \qquad g_B(\boldsymbol{x}) = \left(\frac{3}{4}\right)^{x_1} \left(\frac{1}{4}\right)^{x_2}$$

b) $p(\text{err}) = \frac{568}{2048} \approx 0.2773$

9. *a*)

$$\hat{\boldsymbol{p}}_1 = \begin{pmatrix} \frac{3}{4} \\ \frac{1}{4} \end{pmatrix} \qquad \hat{\boldsymbol{p}}_2 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$

b)

x^t	(0,4)	(1,3)	(2,2)	(3,1)	(4,0)
$p(1 \mid \boldsymbol{x})$	$\frac{1}{17}$	$\frac{12}{76}$	$\frac{54}{150}$	$\frac{108}{172}$	$\frac{81}{97}$
$p(2 \mid \boldsymbol{x})$	$\frac{16}{17}$	$\frac{64}{76}$	$\frac{96}{150}$	$\frac{64}{172}$	$\frac{16}{97}$

10. *a*)

$$g_1(\boldsymbol{x}) = \left(\frac{1}{10}\right)^{x_1} \left(\frac{3}{10}\right)^{x_2} \left(\frac{1}{10}\right)^{x_3} \left(\frac{1}{2}\right)^{x_4} (0)^{x_5}$$

$$g_2(\boldsymbol{x}) = \left(\frac{3}{10}\right)^{x_1} \left(\frac{1}{5}\right)^{x_2} \left(\frac{2}{5}\right)^{x_3} (0)^{x_4} \left(\frac{1}{10}\right)^{x_5}$$

$$g_3(\boldsymbol{x}) = \left(\frac{1}{5}\right)^{x_1} (0)^{x_2} \left(\frac{1}{2}\right)^{x_3} (0)^{x_4} \left(\frac{3}{10}\right)^{x_5}$$

b) Todas las F.D. dan valor 0, no es posible clasificar.

c)

$$\tilde{\boldsymbol{p}}_1 = \left(\frac{2}{15}, \frac{4}{15}, \frac{2}{15}, \frac{2}{5}, \frac{1}{15}\right)^t \qquad \tilde{\boldsymbol{p}}_2 = \left(\frac{4}{15}, \frac{1}{5}, \frac{1}{3}, \frac{1}{15}, \frac{2}{15}\right)^t \qquad \tilde{\boldsymbol{p}}_3 = \left(\frac{1}{5}, \frac{1}{15}, \frac{2}{5}, \frac{1}{15}, \frac{4}{15}\right)^t$$

d)

$$\tilde{\boldsymbol{p}}_1 = \left(\frac{1}{20}, \frac{1}{4}, \frac{1}{20}, \frac{9}{20}, \frac{1}{5}\right)^t \qquad \tilde{\boldsymbol{p}}_2 = \left(\frac{1}{4}, \frac{3}{20}, \frac{7}{20}, \frac{1}{5}, \frac{1}{20}\right)^t \qquad \tilde{\boldsymbol{p}}_3 = \left(\frac{3}{20}, \frac{3}{40}, \frac{9}{20}, \frac{3}{40}, \frac{1}{4}\right)^t$$

e)

$$\tilde{\boldsymbol{p}}_1 = \left(\frac{9}{100}, \frac{29}{100}, \frac{9}{100}, \frac{49}{100}, \frac{4}{100}\right)^t \qquad \tilde{\boldsymbol{p}}_2 = \left(\frac{29}{100}, \frac{19}{100}, \frac{39}{100}, \frac{4}{100}, \frac{9}{100}\right)^t \qquad \tilde{\boldsymbol{p}}_3 = \left(\frac{18}{100}, \frac{3}{100}, \frac{48}{100}, \frac{3}{100}, \frac{28}{100}\right)^t$$

11. (Segundo Parcial Junio 2013)

   a)

$$\hat{\boldsymbol{p}}_1 = \begin{pmatrix} 0.5 \\ 0.3 \\ 0.2 \\ 0.0 \\ 0.0 \end{pmatrix} \rightarrow \hat{\boldsymbol{p}}_1 = \frac{1}{1.5}\begin{pmatrix} 0.5+0.1 \\ 0.3+0.1 \\ 0.2+0.1 \\ 0.0+0.1 \\ 0.0+0.1 \end{pmatrix} = \frac{1}{1.5}\begin{pmatrix} 0.6 \\ 0.4 \\ 0.3 \\ 0.1 \\ 0.1 \end{pmatrix}$$

$$\hat{\boldsymbol{p}}_2 = \begin{pmatrix} 0.0 \\ 0.0 \\ 0.3 \\ 0.3 \\ 0.4 \end{pmatrix} \rightarrow \hat{\boldsymbol{p}}_2 = \frac{1}{1.5}\begin{pmatrix} 0.0+0.1 \\ 0.0+0.1 \\ 0.3+0.1 \\ 0.3+0.1 \\ 0.4+0.1 \end{pmatrix} = \frac{1}{1.5}\begin{pmatrix} 0.1 \\ 0.1 \\ 0.4 \\ 0.4 \\ 0.5 \end{pmatrix}$$

En nuestro caso, dado que las priors son idénticas, la regla de clasificación se reduce a:

$$p(\boldsymbol{y} = (1\ 1\ 1\ 1\ 1) \mid c = 1) = \frac{0.6}{1.5} \cdot \frac{0.4}{1.5} \cdot \frac{0.3}{1.5} \cdot \frac{0.1}{1.5} \cdot \frac{0.1}{1.5} = 0.00009$$
$$p(\boldsymbol{y} = (1\ 1\ 1\ 1\ 1) \mid c = 2) = \frac{0.1}{1.5} \cdot \frac{0.1}{1.5} \cdot \frac{0.4}{1.5} \cdot \frac{0.4}{1.5} \cdot \frac{0.5}{1.5} = 0.0001$$

La muestra $\boldsymbol{y}$ se clasifica en la clase 2.

   b)

$$\hat{\boldsymbol{p}}_1 = \begin{pmatrix} 0.5 \\ 0.3 \\ 0.2 \\ 0.0 \\ 0.0 \end{pmatrix} \rightarrow \hat{\boldsymbol{p}}_1 = \begin{pmatrix} 0.5-0.05 & \\ 0.3-0.05 & \\ 0.2-0.05 & \\ 0.0 & +\frac{1}{2}\cdot 0.15 \\ 0.0 & +\frac{1}{2}\cdot 0.15 \end{pmatrix} = \begin{pmatrix} 0.45 \\ 0.25 \\ 0.15 \\ 0.075 \\ 0.075 \end{pmatrix}$$

$$\hat{\boldsymbol{p}}_2 = \begin{pmatrix} 0.0 \\ 0.0 \\ 0.3 \\ 0.3 \\ 0.4 \end{pmatrix} \rightarrow \hat{\boldsymbol{p}}_2 = \begin{pmatrix} 0.0 & +\frac{1}{2}\cdot 0.15 \\ 0.0 & +\frac{1}{2}\cdot 0.15 \\ 0.3-0.05 & \\ 0.3-0.05 & \\ 0.4-0.05 & \end{pmatrix} = \begin{pmatrix} 0.075 \\ 0.075 \\ 0.25 \\ 0.25 \\ 0.35 \end{pmatrix}$$

$$p(\boldsymbol{y} = (1\ 1\ 1\ 1\ 1) \mid c = 1) = 0.45 \cdot 0.25 \cdot 0.15 \cdot 0.075 \cdot 0.075 = 0.00009$$
$$p(\boldsymbol{y} = (1\ 1\ 1\ 1\ 1) \mid c = 2) = 0.075 \cdot 0.075 \cdot 0.25 \cdot 0.25 \cdot 0.35 = 0.0001$$

La muestra $\boldsymbol{y}$ se clasifica en la clase 2.

12. (Segundo Parcial Junio 2014)

   a)

$$p(1) = p(2) = \frac{6}{12} = 0.5$$

$$\hat{\boldsymbol{p}}_1 = \frac{1}{20}\begin{pmatrix} 1+2+1+1+2+1 \\ 2+1+0+1+2+0 \\ 0+0+0+0+0+0 \\ 1+0+2+1+0+2 \\ 0+0+0+0+0+0 \end{pmatrix} = \frac{1}{20}\begin{pmatrix} 8 \\ 6 \\ 0 \\ 6 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.4 \\ 0.3 \\ 0.0 \\ 0.3 \\ 0.0 \end{pmatrix}$$

$$\hat{\boldsymbol{p}}_2 = \frac{1}{20}\begin{pmatrix} 0+0+0+0+0+0 \\ 0+0+0+0+0+0 \\ 1+1+0+1+0+1 \\ 1+2+1+1+2+1 \\ 1+3+1+1+1+1 \end{pmatrix} = \frac{1}{20}\begin{pmatrix} 0 \\ 0 \\ 4 \\ 8 \\ 8 \end{pmatrix} = \begin{pmatrix} 0.0 \\ 0.0 \\ 0.2 \\ 0.4 \\ 0.4 \end{pmatrix}$$

b)

$$\hat{\boldsymbol{p}}_1 = \begin{pmatrix} 0.4 \\ 0.3 \\ 0.0 \\ 0.3 \\ 0.0 \end{pmatrix} \rightarrow \hat{\boldsymbol{p}}_1 = \frac{1}{2.0} \begin{pmatrix} 0.4 + 0.2 \\ 0.3 + 0.2 \\ 0.0 + 0.2 \\ 0.3 + 0.2 \\ 0.0 + 0.2 \end{pmatrix} = \begin{pmatrix} 0.30 \\ 0.25 \\ 0.10 \\ 0.25 \\ 0.10 \end{pmatrix}$$

$$\hat{\boldsymbol{p}}_2 = \begin{pmatrix} 0.0 \\ 0.0 \\ 0.2 \\ 0.4 \\ 0.4 \end{pmatrix} \rightarrow \hat{\boldsymbol{p}}_2 = \frac{1}{2.0} \begin{pmatrix} 0.0 + 0.2 \\ 0.0 + 0.2 \\ 0.2 + 0.2 \\ 0.4 + 0.2 \\ 0.4 + 0.2 \end{pmatrix} = \begin{pmatrix} 0.10 \\ 0.10 \\ 0.20 \\ 0.30 \\ 0.30 \end{pmatrix}$$

c)

$$\hat{\boldsymbol{p}}_1 = \begin{pmatrix} 0.4 \\ 0.3 \\ 0.0 \\ 0.3 \\ 0.0 \end{pmatrix} \rightarrow \hat{\boldsymbol{p}}_1 = \begin{pmatrix} 0.4 - 0.1 + \frac{1}{5} \cdot 0.3 \\ 0.3 - 0.1 + \frac{1}{5} \cdot 0.3 \\ 0.0 \qquad + \frac{1}{5} \cdot 0.3 \\ 0.3 - 0.1 + \frac{1}{5} \cdot 0.3 \\ 0.0 \qquad + \frac{1}{5} \cdot 0.3 \end{pmatrix} = \begin{pmatrix} 0.36 \\ 0.26 \\ 0.06 \\ 0.26 \\ 0.06 \end{pmatrix}$$

$$\hat{\boldsymbol{p}}_2 = \begin{pmatrix} 0.0 \\ 0.0 \\ 0.2 \\ 0.4 \\ 0.4 \end{pmatrix} \rightarrow \hat{\boldsymbol{p}}_2 = \begin{pmatrix} 0.0 \qquad + \frac{1}{5} \cdot 0.3 \\ 0.0 \qquad + \frac{1}{5} \cdot 0.3 \\ 0.2 - 0.1 + \frac{1}{5} \cdot 0.3 \\ 0.4 - 0.1 + \frac{1}{5} \cdot 0.3 \\ 0.4 - 0.1 + \frac{1}{5} \cdot 0.3 \end{pmatrix} = \begin{pmatrix} 0.06 \\ 0.06 \\ 0.16 \\ 0.36 \\ 0.36 \end{pmatrix}$$

d)

$$\hat{\boldsymbol{p}}_1 = \begin{pmatrix} 0.4 \\ 0.3 \\ 0.0 \\ 0.3 \\ 0.0 \end{pmatrix} \rightarrow \hat{\boldsymbol{p}}_1 = \begin{pmatrix} 0.4 - 0.1 + 0.1 \cdot 0.3 \\ 0.3 - 0.1 + 0.2 \cdot 0.3 \\ 0.0 \qquad + 0.4 \cdot 0.3 \\ 0.3 - 0.1 + 0.2 \cdot 0.3 \\ 0.0 \qquad + 0.1 \cdot 0.3 \end{pmatrix} = \begin{pmatrix} 0.33 \\ 0.26 \\ 0.12 \\ 0.26 \\ 0.03 \end{pmatrix}$$

$$\hat{\boldsymbol{p}}_2 = \begin{pmatrix} 0.0 \\ 0.0 \\ 0.2 \\ 0.4 \\ 0.4 \end{pmatrix} \rightarrow \hat{\boldsymbol{p}}_2 = \begin{pmatrix} 0.0 \qquad + 0.1 \cdot 0.3 \\ 0.0 \qquad + 0.2 \cdot 0.3 \\ 0.2 - 0.1 + 0.4 \cdot 0.3 \\ 0.4 - 0.1 + 0.2 \cdot 0.3 \\ 0.4 - 0.1 + 0.1 \cdot 0.3 \end{pmatrix} = \begin{pmatrix} 0.03 \\ 0.06 \\ 0.22 \\ 0.36 \\ 0.33 \end{pmatrix}$$

e) En nuestro caso, dado que las priors son idénticas, la regla de clasificación se reduce a:

$$p(\boldsymbol{y} = (1\ 1\ 1\ 1\ 1) \mid c = 1) = 0.30 \cdot 0.25 \cdot 0.10 \cdot 0.25 \cdot 0.10 = 0.0001875$$
$$p(\boldsymbol{y} = (1\ 1\ 1\ 1\ 1) \mid c = 2) = 0.10 \cdot 0.10 \cdot 0.20 \cdot 0.30 \cdot 0.30 = 0.00018$$

La muestra $\boldsymbol{y}$ se clasifica en la clase 1.

13. (Recuperación Segundo Parcial Junio 2014)

a)

$$
\begin{aligned}
p(1) &= p(2) = \frac{6}{12} = 0.5 \\
\hat{\boldsymbol{p}}_1 &= \frac{1}{8}\begin{pmatrix} 1+0+1+1+0+1 \\ 0+1+1+1+0+1 \\ 0+0+0+0+0+0 \\ 0+0+0+0+0+0 \end{pmatrix} = \frac{1}{8}\begin{pmatrix} 4 \\ 4 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.5 \\ 0.0 \\ 0.0 \end{pmatrix} \\
\hat{\boldsymbol{p}}_2 &= \frac{1}{10}\begin{pmatrix} 0+0+0+0+0+0 \\ 0+0+0+0+0+0 \\ 1+1+0+1+0+1 \\ 1+1+1+1+1+1 \end{pmatrix} = \frac{1}{10}\begin{pmatrix} 0 \\ 0 \\ 4 \\ 6 \end{pmatrix} = \begin{pmatrix} 0.0 \\ 0.0 \\ 0.4 \\ 0.6 \end{pmatrix}
\end{aligned}
$$

b)

$$
\begin{aligned}
\hat{\boldsymbol{p}}_1 &= \begin{pmatrix} 0.5 \\ 0.5 \\ 0.0 \\ 0.0 \end{pmatrix} \to \hat{\boldsymbol{p}}_1 = \begin{pmatrix} 0.5-0.2+\frac{1}{4}\cdot 0.4 \\ 0.5-0.2+\frac{1}{4}\cdot 0.4 \\ 0.0 \quad\quad +\frac{1}{4}\cdot 0.4 \\ 0.0 \quad\quad +\frac{1}{4}\cdot 0.4 \end{pmatrix} = \begin{pmatrix} 0.4 \\ 0.4 \\ 0.1 \\ 0.1 \end{pmatrix} \\
\hat{\boldsymbol{p}}_2 &= \begin{pmatrix} 0.0 \\ 0.0 \\ 0.4 \\ 0.6 \end{pmatrix} \to \hat{\boldsymbol{p}}_2 = \begin{pmatrix} 0.0 \quad\quad +\frac{1}{4}\cdot 0.4 \\ 0.0 \quad\quad +\frac{1}{4}\cdot 0.4 \\ 0.4-0.2+\frac{1}{4}\cdot 0.4 \\ 0.6-0.2+\frac{1}{4}\cdot 0.4 \end{pmatrix} = \begin{pmatrix} 0.1 \\ 0.1 \\ 0.3 \\ 0.5 \end{pmatrix}
\end{aligned}
$$

14. (Segundo Parcial Junio 2015)

a) Representando nuestros datos como vectores de contadores tenemos

	aa	bb	aa	ab	bc	ac
x_a	2	0	2	1	0	1
x_b	0	2	0	1	1	0
x_c	0	0	0	0	1	1
c_n	1	1	1	2	2	2

La estimación de los parámetros del clasificador multinomial es

$$
\begin{aligned}
p(1) &= p(2) = \frac{3}{6} = \frac{1}{2} \\
\hat{\boldsymbol{p}}_1 &= \frac{1}{6}\begin{pmatrix} 2+0+2 \\ 0+2+0 \\ 0+0+0 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 4 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \\ 0 \end{pmatrix} \\
\hat{\boldsymbol{p}}_2 &= \frac{1}{6}\begin{pmatrix} 1+0+1 \\ 1+1+0 \\ 0+1+1 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}
\end{aligned}
$$

b) En general, el error global $p(e)$ se calcula como

$$
p(e) = \sum_{\boldsymbol{x}} p(\boldsymbol{x}, e) = \sum_{\boldsymbol{x}} p(\boldsymbol{x})\, p(e \mid \boldsymbol{x}) = \sum_{\boldsymbol{x}} p(\boldsymbol{x}) \, \Big(1 - \max_c p(c \mid \boldsymbol{x})\Big).
$$

Como sólo tenemos dos clases con igual prior y aplicando Bayes

$$
\begin{aligned}
p(e) &= \sum_{\boldsymbol{x}} p(\boldsymbol{x}) \ \min_c p(c \mid \boldsymbol{x}) = \sum_{\boldsymbol{x}} p(\boldsymbol{x}) \ \min_c \frac{p(c)\, p(\boldsymbol{x} \mid c)}{p(\boldsymbol{x})} = \sum_{\boldsymbol{x}} \min_c p(c)\, p(\boldsymbol{x} \mid c) \\
&= \frac{1}{2} \sum_{\boldsymbol{x}} \min_c p(\boldsymbol{x} \mid c) = \frac{1}{2} \sum_{\boldsymbol{x}} \frac{x_+!}{x_1!\, x_2!\, x_3!} \min \left(p_{1a}^{x_a} \cdot p_{1b}^{x_b} \cdot p_{1c}^{x_c},\, p_{2a}^{x_a} \cdot p_{2b}^{x_b} \cdot p_{2c}^{x_c} \right) \\
&= \frac{1}{2} \sum_{\boldsymbol{x}} \frac{x_+!}{x_1!\, x_2!\, x_3!} \min \left(\left(\frac{2}{3}\right)^{x_a} \cdot \left(\frac{1}{3}\right)^{x_b} \cdot 0^{x_c},\, \left(\frac{1}{3}\right)^{x_a} \cdot \left(\frac{1}{3}\right)^{x_b} \cdot \left(\frac{1}{3}\right)^{x_c} \right)
\end{aligned}
$$

$\boldsymbol{x}$				$\prod_d p_{cd}^{x_d}$		
x_a	x_b	x_c	$\frac{x_+!}{x_1!\, x_2!\, x_3!}$	$c=1$	$c=2$	$\min_c$
0	0	2	1	0	$\frac{1}{9}$	0
0	2	0	1	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
2	0	0	1	$\frac{4}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
0	1	1	2	0	$\frac{1}{9}$	0
1	0	1	2	0	$\frac{1}{9}$	0
1	1	0	2	$\frac{2}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

Sumando para cada $\boldsymbol{x}$:

$$
p(e) = \frac{1}{2} \left(1 \cdot 0 + 1 \cdot \frac{1}{9} + 1 \cdot \frac{1}{9} + 2 \cdot 0 + 2 \cdot 0 + 2 \cdot \frac{1}{9} \right) = \frac{2}{9}
$$

c)

$$
\hat{\boldsymbol{p}}_1 = \frac{1}{2} \begin{pmatrix} \frac{2}{3} + \frac{1}{3} \\ \frac{1}{3} + \frac{1}{3} \\ 0 + \frac{1}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{3} \\ \frac{1}{6} \end{pmatrix}
$$

$$
\hat{\boldsymbol{p}}_2 = \frac{1}{2} \begin{pmatrix} \frac{1}{3} + \frac{1}{3} \\ \frac{1}{3} + \frac{1}{3} \\ \frac{1}{3} + \frac{1}{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}
$$

d) En nuestro caso, dado que las priors son idénticas, la regla de clasificación se reduce a:

$$
\hat{c}(\boldsymbol{y}) = \arg\max_c p(\boldsymbol{y} \mid c) = \arg\max_c \prod_d p_{cd}^{y_d}
$$

$$
\begin{aligned}
p(\boldsymbol{y} = (0\ 0\ 2)^t \mid c = 1) &= \left(\frac{1}{2}\right)^0 \cdot \left(\frac{1}{3}\right)^0 \cdot \left(\frac{1}{6}\right)^2 = \frac{1}{36} \\
p(\boldsymbol{y} = (0\ 0\ 2)^t \mid c = 2) &= \left(\frac{1}{3}\right)^0 \cdot \left(\frac{1}{3}\right)^0 \cdot \left(\frac{1}{3}\right)^2 = \frac{1}{9}
\end{aligned}
$$

La muestra $\boldsymbol{y}$ se clasifica en la clase 2.

15. (Recuperación Segundo Parcial Junio 2016)

a)

$$
\begin{aligned}
p(1) &= p(2) = \frac{8}{16} = 0.5 \\
\hat{\boldsymbol{p}}_1 &= \frac{1}{20}\begin{pmatrix} 1+2+1+1+2+1+1+0 \\ 2+1+0+1+2+0+0+3 \\ 0+0+0+0+0+0+0+0 \\ 1+0+2+1+0+2+3+3 \\ 0+0+0+0+0+0+0+0 \end{pmatrix} = \frac{1}{30}\begin{pmatrix} 9 \\ 9 \\ 0 \\ 12 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.3 \\ 0.3 \\ 0.0 \\ 0.4 \\ 0.0 \end{pmatrix} \\
\hat{\boldsymbol{p}}_2 &= \frac{1}{20}\begin{pmatrix} 0+0+0+0+0+0+0+0 \\ 0+0+0+0+0+0+0+0 \\ 1+1+0+1+0+1+1+1 \\ 1+2+1+1+2+1+1+3 \\ 1+3+1+1+1+1+3+1 \end{pmatrix} = \frac{1}{30}\begin{pmatrix} 0 \\ 0 \\ 6 \\ 12 \\ 12 \end{pmatrix} = \begin{pmatrix} 0.0 \\ 0.0 \\ 0.2 \\ 0.4 \\ 0.4 \end{pmatrix}
\end{aligned}
$$

b)

$$
\begin{aligned}
\hat{\boldsymbol{p}}_1 &= \begin{pmatrix} 0.3 \\ 0.3 \\ 0.0 \\ 0.4 \\ 0.0 \end{pmatrix} \rightarrow \hat{\boldsymbol{p}}_1 = \frac{1}{2.0}\begin{pmatrix} 0.3+0.2 \\ 0.3+0.2 \\ 0.0+0.2 \\ 0.4+0.2 \\ 0.0+0.2 \end{pmatrix} = \begin{pmatrix} 0.25 \\ 0.25 \\ 0.10 \\ 0.30 \\ 0.10 \end{pmatrix} \\
\hat{\boldsymbol{p}}_2 &= \begin{pmatrix} 0.0 \\ 0.0 \\ 0.2 \\ 0.4 \\ 0.4 \end{pmatrix} \rightarrow \hat{\boldsymbol{p}}_2 = \frac{1}{2.0}\begin{pmatrix} 0.0+0.2 \\ 0.0+0.2 \\ 0.2+0.2 \\ 0.4+0.2 \\ 0.4+0.2 \end{pmatrix} = \begin{pmatrix} 0.10 \\ 0.10 \\ 0.20 \\ 0.30 \\ 0.30 \end{pmatrix}
\end{aligned}
$$

c)

$$
\begin{aligned}
\hat{\boldsymbol{p}}_1 &= \begin{pmatrix} 0.4 \\ 0.3 \\ 0.0 \\ 0.3 \\ 0.0 \end{pmatrix} \rightarrow \hat{\boldsymbol{p}}_1 = \begin{pmatrix} 0.4-0.05+\frac{1}{5}\cdot 0.15 \\ 0.3-0.05+\frac{1}{5}\cdot 0.15 \\ 0.0 \qquad\quad +\frac{1}{5}\cdot 0.15 \\ 0.3-0.05+\frac{1}{5}\cdot 0.15 \\ 0.0 \qquad\quad +\frac{1}{5}\cdot 0.15 \end{pmatrix} = \begin{pmatrix} 0.38 \\ 0.28 \\ 0.03 \\ 0.28 \\ 0.03 \end{pmatrix} \\
\hat{\boldsymbol{p}}_2 &= \begin{pmatrix} 0.0 \\ 0.0 \\ 0.2 \\ 0.4 \\ 0.4 \end{pmatrix} \rightarrow \hat{\boldsymbol{p}}_2 = \begin{pmatrix} 0.0 \qquad\quad +\frac{1}{5}\cdot 0.15 \\ 0.0 \qquad\quad +\frac{1}{5}\cdot 0.15 \\ 0.2-0.05+\frac{1}{5}\cdot 0.15 \\ 0.4-0.05+\frac{1}{5}\cdot 0.15 \\ 0.4-0.05+\frac{1}{5}\cdot 0.15 \end{pmatrix} = \begin{pmatrix} 0.03 \\ 0.03 \\ 0.18 \\ 0.38 \\ 0.38 \end{pmatrix}
\end{aligned}
$$

d) En nuestro caso, dado que las priors son idénticas, la regla de clasificación se reduce a:

$$
\begin{aligned}
p(\boldsymbol{y} = (1\ 1\ 1\ 1\ 1) \mid c = 1) = 0.38 \cdot 0.28 \cdot 0.03 \cdot 0.28 \cdot 0.03 = 2.7 \cdot 10^{-5} \\
p(\boldsymbol{y} = (1\ 1\ 1\ 1\ 1) \mid c = 2) = 0.03 \cdot 0.03 \cdot 0.18 \cdot 0.38 \cdot 0.38 = 2.3 \cdot 10^{-5}
\end{aligned}
$$

La muestra $\boldsymbol{y}$ se clasifica en la clase 1.

16.
- Caso 1:
$$-g_A(\boldsymbol{x}) = x_1^2 + x_2^2 - 6x_1 + 9$$
$$-g_B(\boldsymbol{x}) = x_1^2 + x_2^2 - 6x_1 + 9$$
Frontera: $x_1 = x_2$
- Caso 2:
$$-g_A(\boldsymbol{x}) = 2x_1^2 + x_2^2 - 2x_1x_2 + 6x_1 - 10x_2 + 29$$
$$-g_B(\boldsymbol{x}) = 2x_1^2 + x_2^2 - 2x_1x_2 - 10x_2 + 41$$
Frontera: $x_1 = 3$
- Caso 3:
$$-g_A(\boldsymbol{x}) = 5x_1^2 + x_2^2 - 4x_1x_2$$
$$-g_B(\boldsymbol{x}) = 5x_1^2 + x_2^2 - 4x_1x_2 + 16x_1 - 8x_2 + 16$$
Frontera: $x_2 = 2x_1 + 2$
- Caso 4:
$$-g_A(\boldsymbol{x}) = -5\log 2 + 8x_1^2 + 4x_2^2$$
$$-g_B(\boldsymbol{x}) = -3\log 2 + 8 + 4x_1^2 + 2x_2^2 - 8x_2$$
Frontera: $2x_1^2 + x_2^2 + 4x_2 = \log 2 + 4$
- Caso 5:
$$-g_A(\boldsymbol{x}) = -4\log 2 + 2x_1^2 + 8x_2^2$$
$$-g_B(\boldsymbol{x}) = -3\log 2 + 8 + 4x_1^2 + 2x_2^2 - 8x_2$$
Frontera: $2x_1^2 - 6x_2^2 - 8x_2 + \log 2 + 8 = 0$
- Caso 6:
$$-g_A(\boldsymbol{x}) = -4\log 2 + 8x_1^2 + 2x_2^2$$
$$-g_B(\boldsymbol{x}) = -3\log 2 + 8 + 4x_1^2 + 2x_2^2 - 8x_2$$
Frontera: $4x_1^2 + 8x_2 = \log 2 + 8$
- Caso 7:
$$-g_A(\boldsymbol{x}) = -4\log 2 + 4x_1^2 + 4x_2^2$$
$$-g_B(\boldsymbol{x}) = -3\log 2 + 8 + 4x_1^2 + 2x_2^2 - 8x_2$$
Frontera: $2x_2^2 + 8x_2 = \log 2 + 8$

17. *a*) $P(1) = P(2) = \frac{1}{2}$, $\boldsymbol{\mu}_1 = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$, $\boldsymbol{\mu}_2 = \begin{pmatrix} 8 \\ 3 \end{pmatrix}$, $\Sigma_1 = \Sigma_2 = \begin{pmatrix} 2 & 0 \\ 0 & 6 \end{pmatrix}$
$g_1(\boldsymbol{x}) = \frac{1}{2}x_1^2 - \frac{1}{6}x_2^2 - 4x_1 - 2x_2 + 14$, $g_2(\boldsymbol{x}) = \frac{1}{2}x_1^2 - \frac{1}{6}x_2^2 - 8x_1 - x_2 + \frac{67}{2}$
b) $8x_1 - 2x_2 - 39 = 0$
c) Clases 1 y 2, respectivamente

18. *a*) $\mu_A = 1.1$, $\mu_B = 1.28$, $\sigma_A^2 = 0.02$, $\sigma_B^2 = 0.0056$,
$$g_A(x) = P(A)(0.02)^{-\frac{1}{2}}\exp(-25(x-1.1)^2) \qquad g_B(x) = P(B)(0.0056)^{-\frac{1}{2}}\exp\left(-\frac{(x-1.28)^2}{0.0112}\right)$$
b) $64.286x^2 - 173.57x + 176.536 + \log P(A) + \log P(B) - \frac{1}{2}\log 0.02 - \frac{1}{2}\log 0.0056 = 0$
c) Clase A

19. $p(\boldsymbol{x}|A) \approx 0.205$, $p(\boldsymbol{x}|B) \approx 0.373$, $p(\boldsymbol{x}|C) \approx 0.086$

20. (Examen Primer Parcial 2013)
Para clase 1: $p(0.5|1) = \frac{1}{2\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{0.5-0}{2}\right)^2} = \frac{1}{2}\frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{32}}$
Para clase 2: $p(0.5|2) = \frac{1}{1\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{0.5-1}{1}\right)^2} = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{8}}$
Tomando log en clase 1: $\log p(0.5|1) = \log\frac{1}{2} + \log\frac{1}{\sqrt{2\pi}} - \frac{1}{32}$
Tomando log en clase 2: $\log p(0.5|2) = \log\frac{1}{\sqrt{2\pi}} - \frac{1}{8}$
Llamando $K = \log\frac{1}{\sqrt{2\pi}}$, tenemos para clase 1: $\log p(0.5|1) = -\log 2 - \frac{1}{32} + K = -0.69 - 0.03 + K = -0.72 + K$
Y para clase 2: $\log p(0.5|2) = -\frac{1}{8} + K = -0.13 + K = -0.13 + K$
Por tanto, al tener idénticas probabilidades a priori, se clasifica en clase 2

21. (Examen Recuperación Segundo Parcial 2013)

Usando clasificación por máxima verosimilitud y regla de Bayes, el problema se reduce a calcular

$$c^* = \underset{c \in \{A,B\}}{\arg\max}\, P(c)p(\boldsymbol{x}|c)$$

Por tanto:

$$p(\boldsymbol{x}|A) = \frac{1}{|\Sigma_A|^{\frac{1}{2}} 2\pi} \cdot e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_A)^t \Sigma_A^{-1} (\boldsymbol{x}-\boldsymbol{\mu}_A)} = \frac{1}{2\sqrt{2}\pi} \cdot e^{-\frac{1}{2}\begin{pmatrix} 0 & -2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 0 \\ -2 \end{pmatrix}} = \frac{1}{2\sqrt{2}\pi} \cdot e^{-1}$$

$$p(\boldsymbol{x}|B) = \frac{1}{|\Sigma_B|^{\frac{1}{2}} 2\pi} \cdot e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_B)^t \Sigma_B^{-1} (\boldsymbol{x}-\boldsymbol{\mu}_B)} \frac{1}{2\sqrt{3}\pi} \cdot e^{-\frac{1}{2}\begin{pmatrix} 1 & 0 \end{pmatrix}\begin{pmatrix} \frac{1}{3} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}} = \frac{1}{2\sqrt{3}\pi} \cdot e^{-\frac{1}{6}}$$

$$P(A|\boldsymbol{x}) = P(A)p(\boldsymbol{x}|A) = 0.4 \cdot \frac{1}{2\sqrt{2}\pi} \cdot e^{-1} \approx 0.0166$$

$$P(B|\boldsymbol{x}) = P(B)p(\boldsymbol{x}|B) = 0.6 \cdot \frac{1}{2\sqrt{3}\pi} \cdot e^{-\frac{1}{6}} \approx 0.0467$$

Por tanto, se clasifica en la clase **B**.

22. (Examen Segundo Parcial 2014)

a) En general, la función discriminante de un clasificador basado en la regla de Bayes se define como

$$g_c(\boldsymbol{x}) = p(c) \cdot p(\boldsymbol{x}|c)$$

o de forma equivalente

$$g_c(\boldsymbol{x}) = \log p(c) + \log p(\boldsymbol{x}|c)$$

siendo

$$p(\boldsymbol{x} \mid c) \sim \mathcal{N}_2(\boldsymbol{\mu}_c, \Sigma_c) = (2\pi)^{-1} \cdot \det \Sigma_c{}^{-\frac{1}{2}} \cdot \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu_c})^t \Sigma_c^{-1} (\boldsymbol{x} - \boldsymbol{\mu_c})\right)$$

aplicamos logaritmo neperiano y operamos para simplificar la expresión resultante eliminando la constante $(2\pi)^{-1}$

$$g_c(\boldsymbol{x}) = -\frac{1}{2}\boldsymbol{x}^t \Sigma_c^{-1} \boldsymbol{x} + \boldsymbol{\mu}_c^t \Sigma_c^{-1} \boldsymbol{x} + \left(\log p(c) - \frac{1}{2} \log \det \Sigma_c - \frac{1}{2}\boldsymbol{\mu}_c^t \Sigma_c^{-1} \boldsymbol{\mu}_c\right)$$

Las funciones discriminantes de las clases serían:

$$\begin{aligned} g_A(\boldsymbol{x}) =& -\frac{1}{2}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^t \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix}^t \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &+ \left(\log \frac{1}{2} - \frac{1}{2} \log \det \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 0 \\ 0 \end{pmatrix}^t \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 0 \end{pmatrix}\right) \\ =& -\frac{1}{4}x_1^2 - \frac{1}{4}x_2^2 + \log \frac{1}{2} - \frac{1}{2} \log 4 \end{aligned} \tag{1}$$

$$\begin{aligned} g_B(\boldsymbol{x}) =& -\frac{1}{2}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^t \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 2 \\ 2 \end{pmatrix}^t \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &+ \left(\log \frac{1}{4} - \frac{1}{2} \log \det \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 2 \\ 2 \end{pmatrix}^t \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 2 \\ 2 \end{pmatrix}\right) \\ =& -x_1^2 - x_1x_2 - \frac{1}{2}x_2^2 + 6x_1 + 4x_2 + \log \frac{1}{4} - 10 \end{aligned} \tag{2}$$

$$
\begin{aligned}
g_C(\boldsymbol{x}) = &-\frac{1}{2}\begin{pmatrix}x_1\\x_2\end{pmatrix}^t\begin{pmatrix}1 & -1\\-1 & 2\end{pmatrix}^{-1}\begin{pmatrix}x_1\\x_2\end{pmatrix}+\begin{pmatrix}-2\\-2\end{pmatrix}^t\begin{pmatrix}1 & -1\\-1 & 2\end{pmatrix}^{-1}\begin{pmatrix}x_1\\x_2\end{pmatrix}\\
&+\left(\log\frac{1}{4}-\frac{1}{2}\log\det\begin{pmatrix}1 & -1\\-1 & 2\end{pmatrix}-\frac{1}{2}\begin{pmatrix}-2\\-2\end{pmatrix}^t\begin{pmatrix}1 & -1\\-1 & 2\end{pmatrix}^{-1}\begin{pmatrix}-2\\-2\end{pmatrix}\right)\\
= &-x_1^2-x_1x_2-\frac{1}{2}x_2^2-6x_1-4x_2+\log\frac{1}{4}-10
\end{aligned}
$$

b) La frontera de decisión entre las clases A y B se obtiene igualando sus respectivas funciones discriminantes

$$g_A(\boldsymbol{x}) = g_B(\boldsymbol{x})$$

Por tanto,

$$-\frac{1}{4}x_1^2-\frac{1}{4}x_2^2+\log\frac{1}{2}-\frac{1}{2}\log 4 = -x_1^2-x_1x_2-\frac{1}{2}x_2^2+6x_1+4x_2+\log\frac{1}{4}-10$$

que resulta en una curva cuadrática

$$\frac{3}{4}x_1^2+x_1x_2+\frac{1}{4}x_2^2-6x_1-4x_2+10=0$$

c) Al igual que en el apartado b) igualamos las funciones discriminantes, pero en este caso las de las clase B y C

$$-x_1^2-x_1x_2-\frac{1}{2}x_2^2+6x_1+4x_2+\log\frac{1}{4}-10 = -x_1^2-x_1x_2-\frac{1}{2}x_2^2-6x_1-4x_2+\log\frac{1}{4}-10$$

que como esperábamos definen una frontera de decisión lineal

$$3x_1+2x_2=0 \rightarrow x_2=-\frac{3}{2}x_1$$

dado que la matriz de covarianza es común a ambas clases.

d) Aplicamos la regla de clasificación a la muestra $\boldsymbol{y}$

$$\hat{c}\,(\boldsymbol{y}) = \arg\max_c g_c(\boldsymbol{y})$$

Esto es

$$
\begin{aligned}
g_A(\boldsymbol{y}) &= -\frac{1}{4}1^2-\frac{1}{4}1^2+\log\frac{1}{2}-\frac{1}{2}\log 4 = -1.89\\
g_B(\boldsymbol{y}) &= -1^2-1-\frac{1}{2}1^2+6+4+\log\frac{1}{4}-10 = -3.89\\
g_C(\boldsymbol{y}) &= -1^2-1-\frac{1}{2}1^2-6-4+\log\frac{1}{4}-10 = -23.89
\end{aligned}
$$

Por tanto, $\hat{c}\,(\boldsymbol{y}) = A$.

23. (Examen Segundo Parcial 2016)

a) En general, la función discriminante de un clasificador basado en la regla de Bayes se define como

$$g_c(\boldsymbol{x}) = p(c)\cdot p(\boldsymbol{x}|c)$$

o de forma equivalente

$$g_c(\boldsymbol{x}) = \log p(c) + \log p(\boldsymbol{x}|c)$$

siendo

$$p(\boldsymbol{x}\,|\,c) \sim \mathcal{N}_2(\boldsymbol{\mu}_c,\Sigma_c) = (2\pi)^{-1}\cdot\det{\Sigma_c}^{-\frac{1}{2}}\cdot\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu_c})^t\Sigma_c^{-1}(\boldsymbol{x}-\boldsymbol{\mu_c})\right)$$

aplicamos logaritmo neperiano y operamos para simplificar la expresión resultante eliminando la constante $(2\pi)^{-1}$

$$g_c(\boldsymbol{x}) = -\frac{1}{2}\boldsymbol{x}^t\Sigma_c^{-1}\boldsymbol{x}+\boldsymbol{\mu}_c^t\Sigma_c^{-1}\boldsymbol{x}+\left(\log p(c)-\frac{1}{2}\log\det\Sigma_c-\frac{1}{2}\boldsymbol{\mu}_c^t\Sigma_c^{-1}\boldsymbol{\mu}_c\right)$$

Las funciones discriminantes de las clases serían:

$$
\begin{aligned}
g_A(\boldsymbol{x}) =& -\frac{1}{2}\begin{pmatrix}x_1\\x_2\end{pmatrix}^t \begin{pmatrix}1/2 & 0\\0 & 2\end{pmatrix}^{-1} \begin{pmatrix}x_1\\x_2\end{pmatrix} + \begin{pmatrix}0\\0\end{pmatrix}^t \begin{pmatrix}1/2 & 0\\0 & 2\end{pmatrix}^{-1} \begin{pmatrix}x_1\\x_2\end{pmatrix}\\
&+ \left(\log\frac{1}{2} - \frac{1}{2}\log\det\begin{pmatrix}1/2 & 0\\0 & 2\end{pmatrix} - \frac{1}{2}\begin{pmatrix}0\\0\end{pmatrix}^t \begin{pmatrix}1/2 & 0\\0 & 2\end{pmatrix}^{-1}\begin{pmatrix}0\\0\end{pmatrix}\right)\\
=& -x_1^2 - \frac{1}{4}x_2^2 - \log 2
\end{aligned}
$$

$$
\begin{aligned}
g_B(\boldsymbol{x}) =& -\frac{1}{2}\begin{pmatrix}x_1\\x_2\end{pmatrix}^t \begin{pmatrix}1 & 0\\0 & 2\end{pmatrix}^{-1} \begin{pmatrix}x_1\\x_2\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}^t \begin{pmatrix}1 & 0\\0 & 2\end{pmatrix}^{-1} \begin{pmatrix}x_1\\x_2\end{pmatrix}\\
&+ \left(\log\frac{1}{2} - \frac{1}{2}\log\det\begin{pmatrix}1 & 0\\0 & 2\end{pmatrix} - \frac{1}{2}\begin{pmatrix}0\\1\end{pmatrix}^t \begin{pmatrix}1 & 0\\0 & 2\end{pmatrix}^{-1}\begin{pmatrix}0\\1\end{pmatrix}\right)\\
=& -\frac{1}{2}x_1^2 - \frac{1}{4}x_2^2 + \frac{1}{2}x_2 - \log 2 - \frac{1}{2}\log 2 - \frac{1}{4}
\end{aligned}
$$

b) La frontera de decisión entre las clases A y B se obtiene igualando sus respectivas funciones discriminantes

$$g_A(\boldsymbol{x}) = g_B(\boldsymbol{x})$$

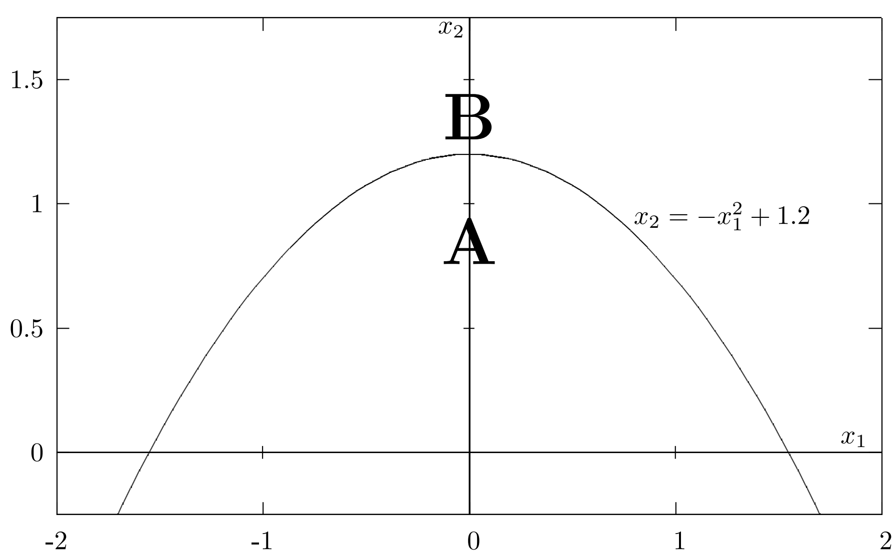
Por tanto,

$$-x_1^2 - \frac{1}{4}x_2^2 - \log 2 = -\frac{1}{2}x_1^2 - \frac{1}{4}x_2^2 + \frac{1}{2}x_2 - \log 2 - \frac{1}{2}\log 2 - \frac{1}{4}$$

que resulta en una parábola vertical

$$x_2 = -x_1^2 + \log 2 + \frac{1}{2} \approx -x_1^2 + 1.2$$

c)



d) El suavizado por *flat smoothing* de un matriz de covarianza $\hat{\Sigma}_c$ se calcula como una combinación lineal de dicha matriz y la matriz de identidad I, de forma que la matriz de covarianzas resultante es:

$$\tilde{\Sigma}_c = \alpha\,\hat{\Sigma}_c + (1-\alpha)\,I \quad \forall c$$

En nuestro caso

$$
\begin{aligned}
\tilde{\Sigma}_A &= \alpha\begin{pmatrix}\frac{1}{2} & 0\\0 & 2\end{pmatrix} + (1-\alpha)\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix} = \begin{pmatrix}\frac{\alpha}{2}+1 & 0\\0 & \alpha+1\end{pmatrix}\\
\tilde{\Sigma}_B &= \alpha\begin{pmatrix}1 & 0\\0 & 2\end{pmatrix} + (1-\alpha)\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix} = \begin{pmatrix}1 & 0\\0 & \alpha+1\end{pmatrix}
\end{aligned}
$$

Por otra parte, sabemos que un clasificador Gaussiano es lineal si la matriz de covarianzas es la misma para todas las clases. Si igualamos $\tilde{\Sigma}_A$ y $\tilde{\Sigma}_B$ para los elementos no nulos, tenemos

$$\begin{aligned} \frac{\alpha}{2} + 1 &= 1 \\ \alpha + 1 &= \alpha + 1 \end{aligned}$$

Si resolvemos este sistema de ecuaciones, vemos que la solución es $\alpha = 0$, es decir, obviamente si $\tilde{\Sigma}_A = \tilde{\Sigma}_B = I$.

24. (Examen Recuperación Segundo Parcial 2017)

a) En el caso de matrices de covarianza común, la función discriminante se define como:

$$g_c(\boldsymbol{x}) = \boldsymbol{w}_c^t\,\boldsymbol{x} + b_c$$

donde

$$\boldsymbol{w}_c = \Sigma^{-1}\boldsymbol{\mu}_c \qquad b_c = \log P(c) - \frac{1}{2}\boldsymbol{\mu}_c^t\Sigma^{-1}\boldsymbol{\mu}_c$$

Las funciones discriminantes de las clases serían:

$$\begin{aligned} g_A(\boldsymbol{x}) &= \boldsymbol{w}_A^t\,\boldsymbol{x} + b_A \\ \boldsymbol{w}_A &= \Sigma^{-1}\boldsymbol{\mu}_A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ b_A &= \log P(A) - \frac{1}{2}\boldsymbol{\mu}_A^t\Sigma^{-1}\boldsymbol{\mu}_A = \log\frac{1}{4} - \frac{1}{2}\begin{pmatrix} 0 \\ 0 \end{pmatrix}^t \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = -\log 4 \\ g_A(\boldsymbol{x}) &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}^t \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \log 4 = -\log 4 \end{aligned}$$

$$\begin{aligned} g_B(\boldsymbol{x}) &= \boldsymbol{w}_B^t\,\boldsymbol{x} + b_B \\ \boldsymbol{w}_B &= \Sigma^{-1}\boldsymbol{\mu}_B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ b_B &= \log P(B) - \frac{1}{2}\boldsymbol{\mu}_B^t\Sigma^{-1}\boldsymbol{\mu}_B = \log\frac{3}{4} - \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}^t \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \log\frac{3}{4} - 1 \\ g_B(\boldsymbol{x}) &= \begin{pmatrix} 1 \\ 1 \end{pmatrix}^t \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \log\frac{3}{4} - 1 = x_1 + x_2 + \log 3 - \log 4 - 1 \end{aligned}$$

b) La frontera de decisión entre las clases A y B se obtiene igualando sus respectivas funciones discriminantes

$$g_A(\boldsymbol{x}) = g_B(\boldsymbol{x})$$

Por tanto,

$$\begin{aligned} -\log 4 &= x_1 + x_2 + \log 3 - \log 4 - 1 \\ 0 &= x_1 + x_2 + \log 3 - 1 \\ x_2 &= -x_1 - \log 3 + 1 = -x_1 - 0.1 \end{aligned}$$

que es una frontera lineal.

$c)$