1-a

$p(\text{method 1 "+"} \mid \text{no disease}) = 0.2$ $\quad p(\text{method 1 "−"} \mid \text{no disease}) = 0.8$

$p(\text{method 1 "−"} \mid \text{has disease}) = 0.15$ $\quad p(\text{method 1 "+"} \mid \text{has disease}) = 0.85$

$$p(\text{has disease} \mid \text{method 1 "+"}) = \frac{p(\text{method 1 "+"} \mid \text{has disease}) \times p(\text{has disease})}{p(\text{method 1 "+"})}$$

$$= \frac{0.85 \times 0.0001}{0.85 \times 0.0001 + 0.2 \times 0.9999} = 4.25 \times 10^{-4}$$

$$p_1(\text{no disease} \mid \text{method 1 "+"}) = \frac{p(\text{method 1 "+"} \mid \text{no disease}) \times p(\text{no disease})}{p(\text{method 1 "+"})}$$

$$= \frac{0.2 \times 0.9999}{0.85 \times 0.0001 + 0.2 \times 0.9999} = 0.99958$$

Therefore, the person should have no disease.

1-b

$p(\text{method 1 "+"} \mid \text{has disease}) = 85\%$

$p(\text{method 1 "+"} \mid \text{no disease}) = 20\%$

Therefore, by ML approach, the person should have disease.

1-C

p(has disease | method 1 "+" method 2 "+")

$$= \frac{p(\text{method 1 "+", method 2 "+"} \mid \text{has disease}) \times p(\text{has disease})}{p(\text{method 1 "+", method 2 "+"})}$$

p(method 1 "+", method 2 "+" | has disease)

$= p$(method 1 "+" | has disease) $\times$ p(method 2 "+" | has disease)

$= 0.85 \times 0.96$

p(method 2 "+" | no disease) $= 0.07$ p(method 2 "+" | has disease) $= 0.96$

p(method 2 "−" | ~~has~~ no disease) $= 0.93$ p(method 2 "−" | has disease) $= 0.04$

p(has disease) $= 0.0001$

p(method 1 "+", method 2 "+") $= p$(method 1 "+") $\times$ p(method 2 "+")

p(method 1 "+") $= 0.2 \times 0.99 + 0.85 \times 0.0001 = 0.198085$

p(method 2 "+") $= 0.07 \times 0.99 + 0.96 \times 0.0001 = 0.069396$

p(has disease | method 1 "+" method 2 "+")

$$= \frac{0.85 \times 0.96 \times 0.0001}{0.198085 \times 0.069396} = 6 \times 10^{-3}$$

Therefore, the person only has 6×10^{-3} chance of having disease.

2-a Scaled dataset

X1	X2	label
0.68	0.79	+
1.00	1.00	+
0.00	0.00	+
0.24	0.06	−
0.46	0.52	−
0.63	0.81	−

2-b after scaling $\begin{bmatrix} 3.9 \\ 4 \end{bmatrix}$: $\vec{x} = \begin{bmatrix} 1.02 \\ 0.10 \end{bmatrix}$

~~distances:~~

			distance $\|\|x - x_i\|\|_2$
+	0.68	0.79	0.75
+	1.00	1.00	0.90
+	0.00	0.00	1.02
−	0.24	0.06	0.78
−	0.46	0.52	$\boxed{0.70}$
−	0.63	0.81	0.81

k=1, therefore, the predicted label is "−"

3-a QDA

$$u_+ = \begin{bmatrix} 1.833 \\ 3.200 \end{bmatrix} \quad \Sigma_+ = \begin{bmatrix} 2.536 & 3.933 \\ 3.933 & 6.140 \end{bmatrix} \quad \phi_+ = 0.5$$

$$u_- = \begin{bmatrix} 1.500 \\ 2.533 \end{bmatrix} \quad \Sigma_- = \begin{bmatrix} 0.420 & 0.990 \\ 0.990 & 2.442 \end{bmatrix} \quad \phi_- = 0.5$$

3-b $\vec{x} = \begin{pmatrix} 3.1 \\ 1.7 \end{pmatrix}$

$p(+|x) \Rightarrow p(x|+) \cdot p(+)$

$$= -\frac{1}{2}\log|\Sigma_+| - \frac{1}{2}(x-u_+)^T(\Sigma_+)^{-1}(x-u_+) + \log\phi_+$$

$$= -147.672$$

$|\Sigma_+| = 0.103 \quad (\Sigma_+)^{-1} = \begin{bmatrix} 59.612 & -38.184 \\ -38.184 & 24.621 \end{bmatrix}$

$p(-|x) \Rightarrow p(x|-) \cdot p(-)$

$$= -\frac{1}{2}\log|\Sigma_-| - \frac{1}{2}(x-u_-)^T(\Sigma_-)^{-1}(x-u_-) + \log\phi_-$$

$$= -98.95726$$

$|\Sigma_-| = 0.046 \quad (\Sigma_-)^{-1} = \begin{bmatrix} 53.087 & -21.522 \\ -21.522 & 9.130 \end{bmatrix}$

Therefore, the class of $\vec{x}$ should be "$-$"

3-C

pros: we can find a linear decision boundary, It simplifies a lot of work.

cons: If the actual distributions of data are different, it would generate significant bias on Σ.

4-a

~~4. (a)~~ add-m smoothing : $m=0.2$

$p(x_1=Low|+) = \frac{2+0.2}{3+0.2\times3} = 0.611$

$p(x_2=Yes|+) = \frac{0+0.2}{3+0.2\times2} = 0.059$

$p(x_3=Green|+) = \frac{2+0.2}{3+0.2\times2} = 0.647$

$p(x_1=Low|-) = \frac{1+0.2}{4+0.2\times3} = 0.261$

$p(x_2=Yes|-) = \frac{3+0.2}{4+0.2\times2} = 0.727$

$p(x_3=Green|-) = \frac{3+0.2}{4+0.2\times2} = 0.727$

4-b

$p(x|+) = \frac{22}{36} \times \frac{2}{34} \times \frac{22}{34} = 0.023$

$p(x|-) = \frac{12}{46} \times \frac{32}{44} \times \frac{32}{44} = 0.138$

4-c $p(x|-) > p(x|+)$, the ML label for x should be "—"

4-d

$p(+|x) = \frac{p(x|+) \times p(+)}{p(x)} = \frac{22\times2\times22\times3}{36\times34\times34\times7} \times \frac{1}{p(x)}$

$= 9.97\times10^{-3} \times \frac{1}{p(x)}$

$p(-|x) = \frac{p(x|-)\times p(-)}{p(x)} = \frac{12\times32\times32\times4}{46\times44\times44\times7} \times \frac{1}{p(x)}$

$= \cancel{0.079}\ 7.89\times10^{-2} \times \frac{1}{p(x)}$

$\boxed{\text{MAP Label: "—"}}$