Homework II

Madalena Yang 110206
Madalena Mota 110355

1. a)

D	y_1	y_2	y_3	y_4	y_5	y_6
x_1	0,52	0,80	0	1	1	N
x_2	0,53	0,92	0	0	0	N
x_3	0,42	0,48	0	1	1	N
x_4	0,49	0,58	1	0	1	P
x_5	0,62	0,31	0	0	1	P
x_6	0,44	0,38	0	0	1	P

(i) $\{y_1, y_2\}$ → 2D Gaussian $m = 2$ dimensions

→ for class P:
($n = 3$ samples)

D	y_1	y_2	y_6
x_4	0,49	0,58	P
x_5	0,62	0,31	P
x_6	0,44	0,38	P

- $$\mu_P = \frac{1}{3}\left(\begin{bmatrix}0,49\\0,58\end{bmatrix} + \begin{bmatrix}0,62\\0,31\end{bmatrix} + \begin{bmatrix}0,44\\0,38\end{bmatrix}\right) =$$

$$= \frac{1}{3}\begin{bmatrix}0,49+0,62+0,44\\0,58+0,31+0,38\end{bmatrix} = \begin{bmatrix}0,51(6)\\0,42(3)\end{bmatrix} \simeq \begin{bmatrix}0,5167\\0,423\end{bmatrix}$$

- $$\Sigma_P = \begin{bmatrix}\overset{\sigma_1^2}{\overset{\shortparallel}{cov(y_1,y_1)}} & cov(y_1,y_2)\\ cov(y_2,y_1) & \underset{\sigma_2^2}{\underset{\shortparallel}{cov(y_2,y_2)}}\end{bmatrix}$$

$$\sigma_1^2 = \frac{1}{3-1}\left[(0,49-0,5167)^2 + (0,62-0,5167)^2 + (0,44-0,5167)^2\right] \simeq$$

$$\simeq 0,00863$$

$$\sigma_2^2 = \frac{1}{2}\left[(0,58-0,423)^2 + (0,31-0,423)^2 + (0,38-0,423)^2\right] \simeq 0,0196$$

$$cov(y_1,y_2) = cov(y_2,y_1) = \frac{1}{2}\left[(0,49-0,5167)(0,58-0,423) + (0,62-0,5167)(0,31-0,423) + (0,44-0,5167)(0,38-0,423)\right]$$

$$\simeq -0,00628$$

$$\Sigma_P = \begin{bmatrix}0,00863 & -0,00628\\ -0,00628 & 0,0196\end{bmatrix}$$

$$|\Sigma_P| = 0,00863 \times 0,0196 - (-0,00628)^2 \simeq 0,00013$$

- $$\Sigma_P^{-1} = \frac{1}{0,00013}\begin{bmatrix}0,0196 & 0,00628\\ 0,00628 & 0,00863\end{bmatrix}$$

PRIORS:

$$P(y_6 = P) = \frac{3}{6} = \frac{1}{2}$$

$$P(y_6 = N) = \frac{3}{6} = \frac{1}{2}$$

$$P(A|B) = \frac{\overbrace{P(B|A)}^{\text{likelihood}}\ \overbrace{P(A)}^{\text{PRIOR}}}{P(B)}$$ → posterior

Gaussian/Normal Distribution:

μ mean
σ standard deviation
σ^2 variance

$$N(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma}\exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)$$ 1 dimension

$$N(x|\mu,\Sigma) = \frac{1}{(2\pi)^{m/2}\sqrt{|\Sigma|}}\exp\left[-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right]$$ m dimensions

μ - m-dimensional mean vector
Σ - m×m covariance matrix
$|\Sigma|$ - determinant of Σ

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$$ n - nº of observations

$$\Sigma = \begin{bmatrix}\overset{\sigma_1^2}{\overset{\shortparallel}{cov(y_1,y_1)}} & cov(y_1,y_2)\\ cov(y_2,y_1) & \underset{\sigma_2^2}{\underset{\shortparallel}{cov(y_2,y_2)}}\end{bmatrix}$$

$$cov(y_1,y_2) = \frac{\sum_{i=1}^{n}(x_{1i}-\bar{y}_1)(x_{2i}-\bar{y}_2)}{n-1}$$ (samples)

$$\Sigma = \begin{bmatrix}a & b\\ c & d\end{bmatrix}$$

$$|\Sigma| = ad - cb$$

$$\Sigma^{-1} = \frac{1}{|\Sigma|}\begin{bmatrix}d & -b\\ -c & a\end{bmatrix}$$

→ for class N :
(n = 3 samples)

D	y_1	y_2	y_6
x_1	0,52	0,80	N
x_2	0,53	0,92	N
x_3	0,42	0,48	N

• $\mu_N = \frac{1}{3}\begin{bmatrix} 0,52+0,53+0,42 \\ 0,80+0,92+0,42 \end{bmatrix} = \begin{bmatrix} 0,49 \\ 0,7(3) \end{bmatrix} \approx \begin{bmatrix} 0,49 \\ 0,73 \end{bmatrix}$

$\sigma_1^2 = \frac{1}{2}\left[(0,52-0,49)^2+(0,53-0,49)^2+(0,42-0,49)^2\right] = 0,0037$

$\sigma_2^2 = \frac{1}{2}\left[(0,80-0,73)^2+(0,92-0,73)^2+(0,48-0,73)^2\right] = 0,05175$

$cov(y_1,y_2) = cov(y_2,y_1) = \frac{1}{2}\left[(0,52-0,49)(0,80-0,73)+(0,53-0,49)(0,92-0,73)+(0,42-0,49)(0,48-0,73)\right] = 0,0136$

$\Sigma_N = \begin{bmatrix} 0,0037 & 0,0136 \\ 0,0136 & 0,05175 \end{bmatrix}$ • $\Sigma_N^{-1} = \frac{1}{6,515\times10^{-6}}\begin{bmatrix} 0,05175 & -0,0136 \\ -0,0136 & 0,0037 \end{bmatrix}$

$|\Sigma_N| = 0,0037 \times 0,05175 - 0,0136^2 = 6,515\times10^{-6}$

For $c \in (N,P)$ and a test vector $z = [y_1, y_2]^T$:

$$N(z\,|\,\mu_c,\Sigma_c) = \frac{1}{(2\pi)^{n/2}\sqrt{|\Sigma_c|}} \exp\left[-\frac{1}{2}(z-\mu_c)^T \Sigma_c^{-1}(z-\mu_c)\right]$$

• $N(z\,|\,\mu_P,\Sigma_P) = \frac{1}{2\pi\sqrt{0,00013}} \exp\left(-\frac{1}{2}\begin{bmatrix} y_1-0,5167 \\ y_2-0,423 \end{bmatrix}^T \frac{1}{0,00013}\begin{bmatrix} 0,0196 & 0,00628 \\ 0,00628 & 0,00863 \end{bmatrix}\begin{bmatrix} y_1-0,5167 \\ y_2-0,423 \end{bmatrix}\right)$

• $N(z\,|\,\mu_N,\Sigma_N) = \frac{1}{2\pi\sqrt{6,515\times10^{-6}}} \exp\left(-\frac{1}{2}\begin{bmatrix} y_1-0,49 \\ y_2-0,73 \end{bmatrix}^T \frac{1}{6,515\times10^{-6}}\begin{bmatrix} 0,05175 & -0,0136 \\ -0,0136 & 0,0037 \end{bmatrix}\begin{bmatrix} y_1-0,49 \\ y_2-0,73 \end{bmatrix}\right)$

(ii) $\{y_3, y_4\}$ → Joint Distribution

Let's calculate the empirical joint probabilities for each combination of (y_3, y_4) values within each class:

Since y_3 and y_4 are discrete, we will use a frequency based approach.

→ for class P: (n=3 samples)

D	y_3	y_4	y_6
x_4	1	0	P
x_5	0	0	P
x_6	0	0	P

• $P(y_3=0, y_4=0 \,|\, y_6=P) = \frac{2}{3}$

• $P(y_3=0, y_4=1 \,|\, y_6=P) = 0$

• $P(y_3=1, y_4=0 \,|\, y_6=P) = \frac{1}{3}$

• $P(y_3=1, y_4=1 \,|\, y_6=P) = 0$

→ for class N: ($n=3$ samples)

D	y_3	y_4	y_6
x_1	0	1	N
x_2	0	0	N
x_3	0	1	N

- $P(y_3=0, y_4=0 \mid y_6=N) = \frac{1}{3}$
- $P(y_3=0, y_4=1 \mid y_6=N) = \frac{2}{3}$
- $P(y_3=1, y_4=0 \mid y_6=N) = 0$
- $P(y_3=1, y_4=1 \mid y_6=N) = 0$

(iii) $\{y_5\}$ → Bernoulli Distribution

D	y_5	y_6
x_1	1	N
x_2	0	N
x_3	1	N
x_4	1	P
x_5	1	P
x_6	1	P

- $P(y_5=1 \mid y_6=P) = \frac{3}{3} = 1$
- $P(y_5=0 \mid y_6=P) = 0$
- $P(y_5=1 \mid y_6=N) = \frac{2}{3}$
- $P(y_5=0 \mid y_6=N) = \frac{1}{3}$

b)(i) $x_7 = [0{,}45 \quad 0{,}80 \quad 0 \quad 0 \quad 1]^T$ the sets are independent

$$P(y_6=P \mid x_7) = \frac{\overset{a)}{P(y_1=0{,}45, y_2=0{,}8 \mid y_6=P)}\ \overset{b)}{P(y_3=0, y_4=0 \mid y_6=P)}\ \overset{c)}{P(y_5=1 \mid y_6=P)}\ \overset{d)}{P(y_6=P)}}{P(x_7)}$$

a) $$P(y_1=0{,}45, y_2=0{,}8 \mid y_6=P) = \frac{1}{2\pi\sqrt{0{,}0013}} \exp\left(-\frac{1}{2}\begin{bmatrix}0{,}45-0{,}5167\\0{,}8-0{,}423\end{bmatrix}^T \frac{1}{0{,}0013}\begin{bmatrix}0{,}0196 & 0{,}00628\\0{,}00628 & 0{,}00863\end{bmatrix}\begin{bmatrix}0{,}45-0{,}5167\\0{,}80-0{,}423\end{bmatrix}\right) \overset{calc}{\simeq} 0{,}301$$

b) $P(y_3=0, y_4=0 \mid y_6=P) = \frac{2}{3}$ c) $P(y_5=1 \mid y_6=P) = 1$ d) $P(y_6=P) = \frac{1}{2}$

$$\Rightarrow P(y_6=P \mid x_7) = \frac{0{,}301 \times \frac{2}{3} \times 1 \times \frac{1}{2}}{\alpha} = 0{,}100(3) \simeq 0{,}1003$$

α → we ignore the denominator as it does not alter decision

$$P(y_6=N \mid x_7) = \frac{\overset{a)}{P(y_1=0{,}45, y_2=0{,}8 \mid y_6=N)}\ \overset{b)}{P(y_3=0, y_4=0 \mid y_6=N)}\ \overset{c)}{P(y_5=1 \mid y_6=N)}\ \overset{d)}{P(y_6=N)}}{P(x_7)}$$

a) $$P(y_1=0{,}45, y_2=0{,}8 \mid y_6=N) = \frac{1}{2\pi\sqrt{6{,}515\times10^{-6}}} \exp\left(-\frac{1}{2}\begin{bmatrix}0{,}45-0{,}49\\0{,}8-0{,}73\end{bmatrix}^T \frac{1}{6{,}515\times10^{-6}}\begin{bmatrix}0{,}05175 & -0{,}0136\\-0{,}0136 & 0{,}0037\end{bmatrix}\begin{bmatrix}0{,}45-0{,}49\\0{,}8-0{,}73\end{bmatrix}\right) \simeq 1{,}2\times10^{-4}$$

b) $P(y_3=0, y_4=0 \mid y_6=N) = \frac{1}{3}$ c) $P(y_5=1 \mid y_6=N) = \frac{2}{3}$ d) $P(y_6=N) = \frac{1}{2}$

$$\Rightarrow P(y_6=N \mid x_7) = \frac{1{,}2\times10^{-4} \times \frac{1}{3} \times \frac{2}{3} \times \frac{1}{2}}{\alpha} \simeq \frac{1{,}33\times10^{-5}}{\alpha}$$

Conclusion x_7: $P(y_6=P \mid x_7) > P(y_6=N \mid x_7)$ ⇒ x_7 is classified as P

(ii) $x_8 = [0{,}50 \quad 0{,}70 \quad 0 \quad 1 \quad 1]^T$

$$P(y_6 = P \mid x_8) = \frac{\overset{a)}{P(y_1=0{,}5,\, y_2=0{,}7 \mid y_6=P)}\ \overset{b)}{P(y_3=0,\, y_4=1 \mid y_6=P)}\ \overset{c)}{P(y_5=1 \mid y_6=P)}\ \overset{d)}{P(y_6=P)}}{P(x_8)}$$

a) $$P(y_1=0{,}5, y_2=0{,}7 \mid y_6=P) = \frac{1}{2\pi\sqrt{0{,}00013}} \exp\left(-\frac{1}{2}\begin{bmatrix}0{,}5-0{,}5167\\0{,}7-0{,}423\end{bmatrix}^T \frac{1}{0{,}00013}\begin{bmatrix}0{,}0196 & 0{,}00628\\0{,}00628 & 0{,}00863\end{bmatrix}\begin{bmatrix}0{,}5-0{,}5167\\0{,}7-0{,}423\end{bmatrix}\right) \approx$$

$\overset{calc}{\approx} 1{,}34$

b) $P(y_3=0, y_4=1 \mid y_6=P) = 0 \quad \Rightarrow P(y_6 = P \mid x_8) = 0$

$$P(y_6 = N \mid x_8) = \frac{\overset{a)}{P(y_1=0{,}5, y_2=0{,}7 \mid y_6=N)}\ \overset{b)}{P(y_3=0, y_4=1 \mid y_6=N)}\ \overset{c)}{P(y_5=1 \mid y_6=N)}\ \overset{d)}{P(y_6=N)}}{P(x_8)}$$

a) $$P(y_1=0{,}5, y_2=0{,}7 \mid y_6=N) = \frac{1}{2\pi\sqrt{6{,}515\times10^{-6}}} \exp\left(-\frac{1}{2}\begin{bmatrix}0{,}5-0{,}49\\0{,}7-0{,}73\end{bmatrix}^T \frac{1}{6{,}515\times10^{-6}}\begin{bmatrix}0{,}05175 & -0{,}0136\\-0{,}0136 & 0{,}0037\end{bmatrix}\begin{bmatrix}0{,}5-0{,}49\\0{,}7-0{,}73\end{bmatrix}\right) \approx$$

$\overset{calc}{\approx} 17{,}35$

b) $P(y_3=0, y_4=1 \mid y_6=N) = \frac{2}{3}$ c) $P(y_5=1 \mid y_6=N) = \frac{2}{3}$ d) $P(y_6=N) = \frac{1}{2}$

$$\Rightarrow P(y_6 = N \mid x_8) = \frac{17{,}35 \times \frac{2}{3} \times \frac{2}{3} \times \frac{1}{2}}{\alpha} \approx \frac{3{,}86}{\alpha}$$

Conclusion x_8: $P(y_6 = N \mid x_8) > P(y_6 = P \mid x_8) \Rightarrow x_8$ is classified as N

2.

D	y_3	y_4	y_5	y_6	$\hat{z}$ (predicted)	
x_1	0	1	1	N	N	✓
x_2	0	0	0	N	P	✗
x_3	0	1	1	N	N	✓
x_4	1	0	1	P	P	✓
x_5	0	0	1	P	P	✓
x_6	0	0	1	P	P	✓
x_7	0	0	1	P	P	✓
x_8	0	1	1	N	N	✓

$$ACC = \frac{TP + TN}{ALL}$$

classify x_1:

$H(x_i,x_j)$	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1	2	0	2	1	1	1	0
		N					N

$\hat{z}_1 = \text{mode}(N, N) = N$

classify x_2:

$H(x_i,x_j)$	x_1	x_3	x_4	x_5	x_6	x_7	x_8
x_2	2	2	2	1	1	1	2
				P	P	P	

$\hat{z}_2 = \text{mode}(P, P, P) = P$

classify x_3:

$H(x_i,x_j)$	x_1	x_2	x_4	x_5	x_6	x_7	x_8
x_3	0	2	2	1	1	1	0
	N						N

$\hat{z}_3 = \text{mode}(N, N) = N$

classify x_4:

$H(x_i, x_j)$	x_1	x_2	x_3	x_5	x_6	x_7	x_8
x_4	2	2	2	1	1	1	2
				P	P	P	

$\hat{z}_4 = \text{mode}(P, P, P) = P$

classify x_5:

$H(x_i, x_j)$	x_1	x_2	x_3	x_4	x_6	x_7	x_8
x_5	1	1	1	1	0	0	1
					P	P	

$\hat{z}_5 = \text{mode}(P, P) = P$

classify x_6:

$H(x_i, x_j)$	x_1	x_2	x_3	x_4	x_5	x_7	x_8
x_6	1	1	1	1	0	0	1
					P	P	

$\hat{z}_6 = \text{mode}(P, P) = P$

classify x_7:

$H(x_i, x_j)$	x_1	x_2	x_3	x_4	x_5	x_6	x_8
x_7	1	1	1	1	0	0	1
					P	P	

$\hat{z}_7 = \text{mode}(P, P) = P$

classify x_8:

$H(x_i, x_j)$	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_8	0	2	0	2	1	1	1
	N		N				

$\hat{z}_8 = \text{mode}(N, N) = N$

$$Acc = \frac{TP + TN}{All} = \frac{7}{8} = 87{,}5\%$$

3. • $p \in (1/2; 1]$

• $P(\theta = 0) = p$

• $\forall x \in X,\ P(X = x \mid \theta = 0) = P(X = x \mid \theta = 1)$, $P(X = x) > 0$

a) $\theta_{Bayes} = \arg\max_\theta P(x \mid \theta)\, P(\theta) = \arg\max_\theta P(\theta)$

Since $P(X = x \mid \theta = 0) = P(X = x \mid \theta = 1)$ $\quad p \in (1/2; 1] \Rightarrow p > 1/2 \Rightarrow p > 1 - p$
$(p \le 1)$

$P(\theta = 0) = p$ so MAP classifier always predict $\theta_{Bayes} = 0$

$E_{Bayes} = P(\theta \neq \theta_{Bayes}) = P(\theta \neq 0) = P(\theta = 1) = 1 - p$

b) $E_{1NN} = P(\theta = \theta_{1NN}) = 2 \cdot P(\theta = 0 \mid X = x)\, P(\theta = 1 \mid X = x)$

$$P(\theta = 0 \mid X = x) = \frac{P(X = x \mid \theta = 0)\, P(\theta = 0)}{P(X = x)} \qquad P(\theta = 1 \mid X = x) = \frac{P(X = x \mid \theta = 1)\, P(\theta = 1)}{P(X = x)}$$

$P(X = x) = P(X = x \mid \theta = 0)\, P(\theta = 0) + P(X = x \mid \theta = 1)\, P(\theta = 1) =$

$P(X = x \mid \theta = 0) = P(X = x \mid \theta = 1)$

$= P(X = x \mid \theta)\, P(\theta = 0) + P(X = x \mid \theta)\, P(\theta = 1) =$

$= P(X = x \mid \theta)\,(p + 1 - p) =$

$= P(X = x \mid \theta)$

$$P(\theta=0 \mid x=u) = \frac{\cancel{P(x=u \mid \theta)}\, P(\theta=0)}{\cancel{P(x=u \mid \theta)}} = P(\theta=0) = p$$

$$P(\theta=1 \mid x=u) = \frac{\cancel{P(x=u \mid \theta)}\, P(\theta=1)}{\cancel{P(x=u \mid \theta)}} = P(\theta=1) = 1-p$$

$$\epsilon_{1NN} = P(\theta \neq \theta_{1NN}) = 2 \overbrace{P(\theta=0 \mid x=u)}^{p} \overbrace{P(\theta=1 \mid x=u)}^{1-p} =$$

$$= 2p(1-p)$$

e) $\epsilon_{1NN} \overset{\text{from b)}}{=} 2p(1-p) = 2p\, \epsilon_{Bayes} = 2(1-\epsilon_{Bayes})\, \epsilon_{Bayes} =$

from a) $\epsilon_{Bayes} = 1-p \Leftrightarrow$ $\quad = 2\,\epsilon_{Bayes}(1-\epsilon_{Bayes})$

$\Leftrightarrow p = 1 - \epsilon_{Bayes}$