Eléments de théorie bayésienne de la décision

Exercice 1

On considère le cas de deux classes dans $\mathbb{R}$ définies par les fonctions de densité de probabilité conditionnelle données ci-dessous, où x ne peut prendre que trois valeurs (0, 1 ou 2).

x	$p(x\|\omega_0)$	$p(x\|\omega_1)$
0	0.3	0.1
1	0.4	0.1
2	0.3	0.8

Les probabilités *a priori* des classes sont $P(\omega_0) = 0.6$ et $P(\omega_1) = 0.4$. Les coûts de décision sont donnés par :

$$\lambda_{00} = \lambda_{11} = 0 \qquad \lambda_{01} = 3 \qquad \lambda_{10} = 5$$

1. On choisit la règle de décision du maximum de probabilité *a posteriori*. Énoncer celle-ci. Indiquer les décisions prises lorsqu'on observe 0, 1 ou 2.
2. On choisit la règle de décision de Bayes. Énoncer celle-ci. Indiquer les décisions prises lorsqu'on observe 0, 1 ou 2.

Exercice 2

Soient deux classes ω_0 et ω_1 dans $\mathbb{R}^2$ gouvernées chacune par une loi normale. Les paramètres respectifs de ces lois sont :

$$\boldsymbol{m}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \boldsymbol{m}_1 = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \qquad \boldsymbol{\Sigma}_0 = \boldsymbol{\Sigma}_1 = \begin{pmatrix} 1 & 0.25 \\ 0.25 & 1 \end{pmatrix}$$

Les deux classes sont supposées équiprobables.

1. Formuler la règle de décision de Bayes dans le cas où les coûts de classification valent :

$$\lambda_{00} = \lambda_{11} = 0 \qquad \lambda_{01} = 2\,\lambda_{10}$$

2. Calculer la probabilité d'erreur de la règle de Bayes.
3. Donner l'expression de la règle de décision pour des coût 0/1. Rappeler comment cette règle est autrement nommée, ainsi que le critère qu'elle minimise.
4. Calculer la probabilité d'erreur de la règle de décision établie à la question 3.
5. Comparer les probabilités d'erreur obtenues aux questions 2. et 4. Conclure.

Exercice 3

Soient deux classes ω_0 et ω_1 dans $\mathbb{R}^2$ gouvernées chacune par une loi normale. Les paramètres respectifs de ces lois sont :

$$\boldsymbol{m}_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \boldsymbol{m}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \boldsymbol{\Sigma}_0 = \boldsymbol{\Sigma}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Les deux classes sont supposées équiprobables.
On choisit la règle de décision du maximum de probabilité *a posteriori*.

1. Formuler la règle de décision correspondante et la représenter dans $\mathbb{R}^2$.

2. On suppose que les vraies valeurs de $\boldsymbol{m}_0$ et $\boldsymbol{m}_1$ sont inconnues. Les matrices de variance-covariance $\boldsymbol{\Sigma}_0$ et $\boldsymbol{\Sigma}_1$ demeurent en revanche celles fournies en début d'exercice. On dispose de l'ensemble d'apprentissage suivant pour estimer $\boldsymbol{m}_0$ et $\boldsymbol{m}_1$.

$$\text{classe } \omega_0 \;:\; \left\{ \begin{pmatrix} -0.03 \\ -1.02 \end{pmatrix}, \begin{pmatrix} 0.08 \\ -0.76 \end{pmatrix}, \begin{pmatrix} -0.49 \\ -1.61 \end{pmatrix}, \begin{pmatrix} 0.14 \\ 0.06 \end{pmatrix} \right\}$$
$$\text{classe } \omega_1 \;:\; \left\{ \begin{pmatrix} 1.56 \\ 2.7 \end{pmatrix}, \begin{pmatrix} 0.60 \\ 0.76 \end{pmatrix}, \begin{pmatrix} 2.77 \\ 0.77 \end{pmatrix}, \begin{pmatrix} 0.04 \\ 0.89 \end{pmatrix} \right\}$$

Calculer les estimées $\hat{\boldsymbol{m}}_0$ et $\hat{\boldsymbol{m}}_1$ de $\boldsymbol{m}_0$ et $\boldsymbol{m}_1$, respectivement.

3. Formuler la règle de décision correspondante reposant sur $\{\hat{\boldsymbol{m}}_0, \hat{\boldsymbol{m}}_1, \boldsymbol{\Sigma}_0, \boldsymbol{\Sigma}_1\}$ et la représenter dans $\mathbb{R}^2$.

4. Comparer graphiquement les deux règles.

5. On estime la probabilité d'erreur par la méthode de resubstitution (ensemble d'apprentissage et de test identiques). Calculer la valeur estimée de la probabilité d'erreur.

Exercice 4

On suppose que les individus $\boldsymbol{x}$ à classer dans ω_0 ou ω_1 sont des éléments de $\mathbb{R}^2$, distribués selon des lois normales $\mathcal{N}(\mu_0, \boldsymbol{\Sigma}_0)$ et $\mathcal{N}(\mu_1, \boldsymbol{\Sigma}_1)$. On suppose que les coûts λ_{01} et λ_{10} de mauvaise classification valent 1, et que les coûts de bonne classification λ_{00} et λ_{11} valent 0. On suppose par ailleurs que les classes ω_0 et ω_1 sont équiprobables.

1. On considère tout d'abord que $\boldsymbol{\Sigma}_0 = \boldsymbol{\Sigma}_1 = \sigma^2 \boldsymbol{I}$, où σ est un réel et $\boldsymbol{I}$ désigne la matrice identité dans $\mathbb{R}^2$. Donner la règle de Bayes.

2. On suppose à présent que $\boldsymbol{\Sigma}_0 = \sigma_0^2 \boldsymbol{I}$ et $\boldsymbol{\Sigma}_1 = \sigma_1^2 \boldsymbol{I}$. Énoncer la règle de Bayes.

3. On retient le modèle obtenue en 1. Estimer $\{\mu_0, \mu_1, \sigma^2\}$ à partir des données d'apprentissage suivantes. Énoncer la règle de décision correspondante.

classe	ω_0	ω_0	ω_0	ω_0	ω_1	ω_1	ω_1	ω_1
x_1	3.58	-1.35	0.72	0.71	1.87	3.41	2.67	2.71
x_2	2.77	3.03	-0.06	-0.20	4.49	4.42	1.79	4.63

4. On retient à présent le modèle 2. Estimer $\{\mu_0, \mu_1, \sigma_0^2, \sigma_1^2\}$ à partir des données ci-dessus. Énoncer la règle de décision correspondante.

5. Tracer les séparateurs correspondants après avoir repéré les positions des moyennes.

Exercice 5

Soit une source émettant un signal $x(k)$ constitué d'une séquence de symboles $+a$ et $-a$ au travers d'un canal bruité. Le signal reçu par le récepteur s'écrit $y(k) = x(k) + b(k)$, avec $b(k)$ un bruit additif indépendant de $x(k)$. A l'instant k, l'échantillon $x(k)$ prend les valeurs $+a$ ou $-a$ avec les probabilités respectives p et $1 - p$. Le bruit $b(k)$ est i.i.d. et distribué selon une loi normale centrée de variance σ_b^2.

La règle de décodage adoptée par le récepteur consiste à considérer que le symbole $+a$ a été émis si $y(k) > \eta$, sinon $-a$, où η désigne un seuil à déterminer.

1. Déterminer la règle de décision au sens du maximum de probabilité *a posteriori*.

2. Calculer la probabilité d'erreur du récepteur en fonction de a, σ_b^2 et p.

3. Etudier le cas $p = \frac{1}{2}$ et montrer que la probabilité d'erreur du décodeur s'exprime en fonction de a et σ. Interpréter les résultats.

On suppose à présent que le décodeur dispose de N observations $y(k_i)$ du signal reçu pour déterminer le symbole émis.

1. Répondre aux mêmes questions que précédemment dans l'hypothèse où les N observations $y(k_i)$ sont indépendantes.

2. Interpréter les résultats lorsque N tend vers l'infini.