Contrôle de Connaissances

Si 221 : UE Bases de la Reconnaissance des Formes

Novembre 2012

Autorisés : polycopié, notes de cours et calculette

Exercice 1

***note** : La question 5 de cet exercice est indépendante des questions 3 et 4*

Soient deux distributions mono-dimensionnelles 'triangle' de même paramètre δ, correspondant respectivement à deux classes ω_1 et ω_2, de moyennes m_1 et m_2 . On suppose que $m_2 > m_1$. On a :

$$\forall i = 1,2$$

$$p(x|\omega_i) = \frac{\delta - |x - m_i|}{\delta^2} \quad \text{si} \quad |x - m_i| \prec \delta$$

$$p(x|\omega_i) = 0 \quad \text{sinon}$$

1) On introduit la quantité d:

$d=(m_2-m_1)/\delta$

Que représente cette quantité ?
on suppose par la suite que $1 < d < 2$. Pourquoi ?

2) On suppose que les probabilités a priori des 2 classes sont égales.
Quel est le point frontière suivant la règle de décision bayésienne ?

3) Donner les expressions des probabilités d'erreur de type I et II, ainsi que du risque bayésien C* en fonction de m_1, m_2 et δ .

4) Application : calculer C* avec $m_1=1$, $m_2=2.5$, $\delta=1$

5) On ne suppose plus les probabilités a priori égales.
Donner l'expression du point frontière entre les 2 classes en fonction de m_1, m_2 et des probabilités a priori.

Exercice 2: Chaîne de Markov binaire et modèle de durée

Ce problème a pour but d' établir quelques compléments de cours.
On considère un système avec deux états notés 0 et 1
(attention au léger changement de notation par rapport au cours).
Le système démarre, à l'instant $t = 0$, sur l'état initial $X_0 = 1$,
et son évolution est supposée régie par une chaîne de Markov *stationnaire*
de matrice de transition:

$$A = \begin{array}{c} {\scriptstyle 0} \\ {\scriptstyle 1} \end{array} \begin{bmatrix} 1 & 0 \\ 1-\alpha & \alpha \end{bmatrix} \qquad \text{avec } 0 \leq \alpha \leq 1$$

Identifier les différents éléments de matrice a_{ij} $(i, j = 0, 1)$
et répondre aux questions suivantes :

Q. 1 Si le système se trouve dans l'état $X_t = 0$ à un certain instant $t \geq 1$, qu'arrive-t-il par la suite ? Un tel état est appelé absorbant.

Q. 2 Que représente α ?

Q. 3 Indiquer graphiquement le type de trajectoire suivie par le système. Préciser les cas $\alpha = 0$ et $\alpha = 1$.

Q. 4 En partant de la relation

$$P(X = x) = \sum_y P(X = x \mid Y = y)\ P(Y = y)$$

valable pour tout couple de variables aléatoires discrètes X, Y à nombre de valeurs fini, montrer que l'on a

$$P(X_t = 1) = \alpha\ P(X_{t-1} = 1) \qquad \forall t \geq 1$$

On suppose maintenant $\alpha < 1$ et l'on s'intéresse à la *durée de vie* du système dans l'état 1 (c'est-a-dire le nombre d'instants où l'état du système prend la valeur 1), que l'on représente par une variable aléatoire D **entière**
(i.e. à valeurs dans $\mathbb{N}$).

Q. 5 Ecrire $P(D > t)$ en fonction de $P(X_t = \ldots)$ $\forall t \in \mathbb{N}$.

Q. 6 On admettra que pour toute variable aléatoire **entière** D **d'espérance finie**, cette espérance peut s'écrire

$$\mathbf{E}[D] = \sum_{t=0}^{+\infty} P(D > t)$$

(si vous pouvez le re-démontrer, c'est très bien !). En admettant également que la condition $\alpha < 1$ implique $\mathbf{E}[D]$ finie, calculer cette espérance et retrouver ainsi la formule que l'on a vue dans le cours.

Exercice 3

1) Redonnez le déroulement de l'algorithme des k-moyennes

2) Si l'algorithme converge de deux manières différentes (selon deux initialisations différentes), comment choisir la solution la plus "intéressante" ?

3) Montrez quel critère optimal vérifie l'algorithme des k-moyennes : pour cela on adopte la norme L2. Quel serait le résultat pour la norme L1 ?