

Apprentissage Automatique Numérique

M1 2021/2022

TD 1

Exercice 1 - Détection de fraude

L'apprentissage automatique peut-être utilisé pour détecter les fraudes : l'exercice suivant en est une illustration très simple. On dispose d'un dé à 6 faces, parfaitement équilibré. On confie ce dé à des individus en leur demandant de procéder à un certain nombre de lancers et de faire part de leurs résultats. La population est composée de personnes honnêtes (H) qui font exactement ce qu'on leur demande, mais aussi d'un certain nombre de tricheurs (T) qui chaque fois qu'on leur demande de lancer une fois le dé, le lancent deux fois et annoncent le plus grand des nombres obtenus. Ainsi, si l'on demande à un tricheur de lancer 5 fois le dé, il pourra obtenir la suite de résultats 2, 2, 5, 2, 4, 1, 5, 4, 6, 6 et annoncer 2,5,4,5,6.

1. Calculez $p(i|H)$ et $p(i|T)$ pour i allant de 1 à 6.
2. Calculez $p(25456|H)$ et $p(25456|T)$
3. On suppose que la population contient 28% de tricheurs.

Que doit-on décider sur l'honnêteté d'un individu qui annonce 25456 si l'on suit

- (a) la règle majoritaire ?
- (b) la règle du maximum de vraisemblance ?
- (c) la règle de décision de Bayes ?

Exercice 2

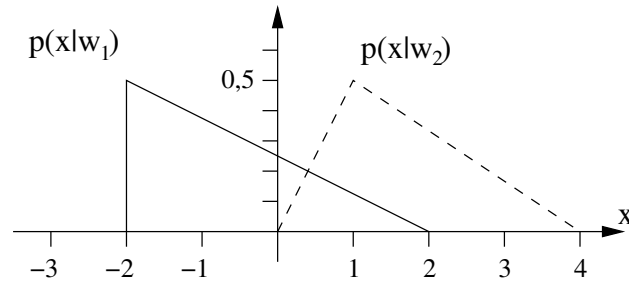
On considère un classifieur Bayésien pour deux classes ω_1 et ω_2 dans l'espace vectoriel de dimension 1. Les densités conditionnelles des deux classes sont données par :

$$p(x|\omega_1) = \begin{cases} \frac{1}{6}x + \frac{1}{3} & \text{si } x \in [-2, 0] \\ -\frac{1}{12}x + \frac{1}{3} & \text{si } x \in [0, 4] \\ 0 & \text{sinon} \end{cases} \quad p(x|\omega_2) = \begin{cases} \frac{1}{8}x - \frac{1}{4} & \text{si } x \in [2, 6] \\ 0 & \text{sinon} \end{cases}$$

- a) Soit $P(\omega_1) = P(\omega_2) = \frac{1}{2}$.
Déterminer la règle de décision de ce classifieur Bayésien et calculer la probabilité d'erreur.
- b) Refaire les mêmes calculs avec $P(\omega_1) = \frac{3}{4}$ et $P(\omega_2) = \frac{1}{4}$. Comparer et conclure.

Exercice 3

On considère deux classes ω_1 et ω_2 dans l'espace vectoriel de dimension 1, l'observation x variant dans l'intervalle $[0, 5]$. Les densités conditionnelles des deux classes sont données par la figure suivante (attention, les échelles des abscisses et des ordonnées sont différentes) :



Soit $P(\omega_1) = P(\omega_2) = \frac{1}{2}$.

Déterminer **par calcul** la règle de décision de ce classifieur Bayésien et calculer la probabilité d'erreur.