

TP 1 Apprentissage supervisé 2A

Eléments de correction - B. Gelein

Septembre 2019

Exercice 1 : Règle et risque de Bayes en discrimination binaire

1. Les trois cas sont des cas de discrimination binaire d'où la règle de Bayes f^* :

$$f^*(x) : \mathbb{R} \rightarrow \{0; 1\}$$

$$x \mapsto f^*(x) = \begin{cases} 1 & \text{si } P(Y = 1|X = x) > 0.5 \\ 0 & \text{sinon} \end{cases}$$

Cas 1 :

- Pour $X_i \leq 0$:
 $P(Y_i = 1|X_i = x_i) = P(U_i \leq 0.1) = 0.1$ d'où $f^*(x_i) = 0$
- Pour $X_i > 0$:
 $P(Y_i = 1|X_i = x_i) = P(U_i > 0.2) = 0.8$ d'où $f^*(x_i) = 1$

On a donc la **règle de Bayes** : $f^*(x) = \begin{cases} 1 & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$

Risque de Bayes :

$$\begin{aligned} R_p^* &= P(Y \neq f^*(X)) \\ &= P(Y = 1 \cap f^*(X) = 0) + P(Y = 0 \cap f^*(X) = 1) \\ &= P(Y = 1 \cap X \leq 0) + P(Y = 0 \cap X > 0) \\ &= P(Y = 1|X \leq 0)P(X \leq 0) + P(Y = 0|X > 0)P(X > 0) \\ &= 0.1 \times 0.5 + 0.2 \times 0.5 = 0.15 \end{aligned}$$

Cas 2 : Règle de Bayes : $f^*(x) = \begin{cases} 1 & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$

Risque de Bayes :

$$\begin{aligned} R_p^* &= P(Y \neq f^*(X)) \\ &= 0.2 \times 0.5 + 0.4 \times 0.5 = 0.3 \end{aligned}$$

Cas 3 : La loi jointe de (X, Y) n'étant pas connue, il est impossible de définir la règle de Bayes et donc impossible de calculer le risque de Bayes.

2. Le risque de Bayes constitue un indicateur de complexité. On classe seulement les cas 1 et 2. Le cas 2 est plus complexe que le cas 1.