Arnaud Ly FTML

Exercice 1.

a) Le risque empirique pour EST_1 =

6	1
8	2

$$\frac{1}{17} \times (1 \times 8 + x \times 1) = \frac{8+x}{17}$$

Le risque empirique pour EST_2 :

4	3
1	9

$$\frac{1}{17} \times (1 \times 1 + x \times 3) = \frac{1+3x}{17}$$

$$\frac{8}{17} + \frac{x}{17} \geqslant \frac{1}{17} + \frac{3x}{17} \iff \frac{7}{17} \geqslant 2 \times \frac{x}{17}$$

$$\iff 7 \geqslant 2 \times x \iff \frac{7}{2} \geqslant x$$

b) On suppose $x = 2$

Pour (X_1, X_2)

$(1,2) = 4T$ et $1F$

$(2,2) = 1T$ et $3F$

$(1,1) = 1T$ et $2F$

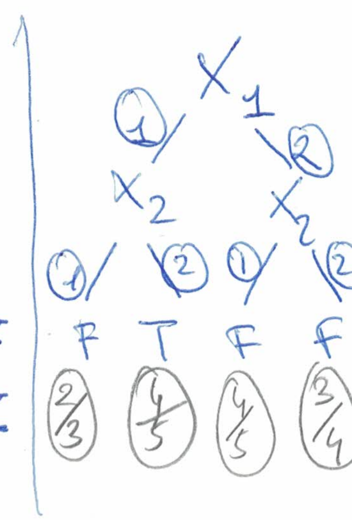
$(2,1) = 1T$ et $4F$

d'où $(1,2) = T$

$(2,2) = F$

$(1,1) = F$

$(2,1) = F$



On obtient ainsi l'estimateur 2 :

Pour $x=2$

d'où $\frac{1}{17} \times (1 + 3x) = \frac{7}{17}$

c) Un autre décision renverra des résultats optimaux.

–

–

d)

$Y=T$	$X=1$	$X=2$
X_1	5	2
X_2	2	5

$Y=F$	$X=1$	$X=2$
X_1	3	7
X_2	6	4

$$P(Y=T \mid X_1=1 \text{ et } X_2=1) = P(Y=T) \times P(X_1=1 \mid Y=T) \times P(X_2=1 \mid Y=T)$$

$$= 7/17 \times 5/7 \times 2/7 = 10/119$$

$$\approx 0{,}08$$

$$P(Y=F \mid X_1=1 \text{ et } X_2=1) = 10/17 \times \frac{3}{10} \times \frac{6}{10} = \frac{9}{85} \approx \boxed{0{,}10}$$

$P(Y=T \mid X_1=1 \text{ et } X_2=2) = \frac{7}{17} \times \frac{5}{7} \times \frac{5}{7} = \frac{25}{119} \approx \boxed{0{,}21}$

$P(Y=F \mid X_1=1 \text{ et } X_2=2) = \frac{10}{17} \times \frac{3}{10} \times \frac{4}{10} = \frac{6}{85} \approx 0{,}07$

$P(Y=T \mid X_1=2 \text{ et } X_2=1) = \frac{7}{17} \times \frac{2}{7} \times \frac{2}{7} = \frac{4}{119} \approx 0{,}03$

$P(Y=F \mid X_1=2 \text{ et } X_2=1) = \frac{10}{17} \times \frac{7}{10} \times \frac{6}{10} = \frac{21}{85} \approx \boxed{0{,}24}$

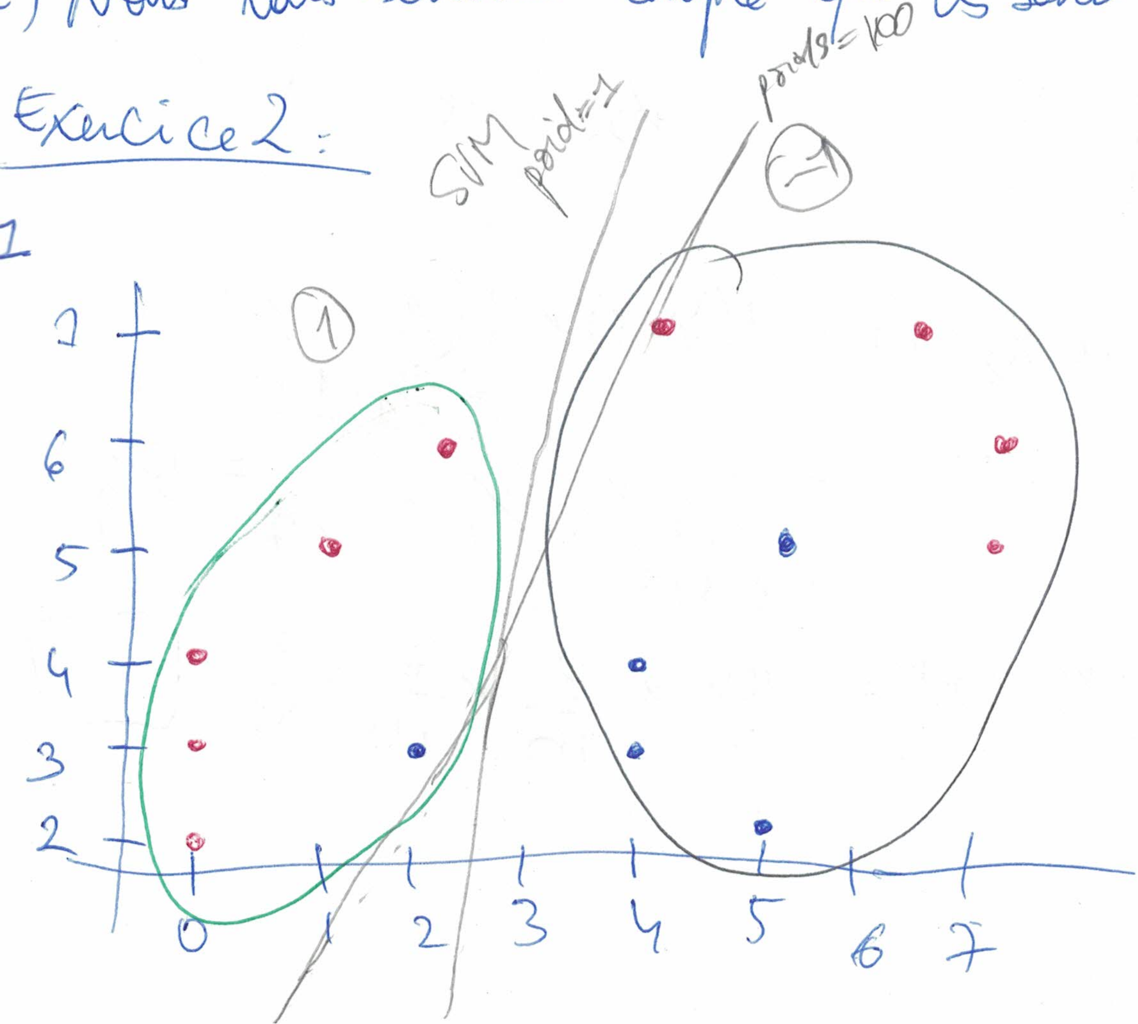
$P(Y=T \mid X_2=2 \text{ et } X_1=2) = \frac{7}{17} \times \frac{2}{7} \times \frac{5}{7} = \frac{10}{119} \approx 0{,}08$

$P(Y=F \mid X_2=2 \text{ et } X_1=2) = \frac{10}{17} \times \frac{7}{10} \times \frac{4}{10} = \frac{14}{85} \approx \boxed{0{,}16}$

2) Nous nous rendons compte qu'ils sont identiques

Exercice 2 :

1



On peut utiliser un Kmeans ou CHA

b.) /

c) Oui il est possible de faire passer un hyperplan pour séparer les 2 clusters

d) poids = 1
poids

Exercice 3 :

a) Nous pouvons utiliser le test du χ^2.

On calcul les effectifs croisée espéré :

$$E_{ij} = \frac{1}{m} \sum Card(\{X=i\}) \times Card(\{Y=j\})$$

Ensuite les effectifs croisés observés :

$$O_{i,j} = Card(\{X=i \text{ et } Y=j\})$$

Pour finalement obtenir l'écart relatif entre les 2 variables

$$\text{Ecart relatif} = \sum_{i,j} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

plus celui-ci est faible et plus ça signifie qu'elles sont indépendants.

b) Il faudrait chercher la dimension de VC.
Celle-ci correspond au nombre maximum d'éléments pulverisable.

c) ?