

## Exercice 1:

a) Risque empirique:

$$\text{EST-1} : (8 + x) \quad 1/7$$

$$\text{EST-2} : (1 + 3x) \quad 1/7$$

$$1 + 3x \geq 8 + x \Leftrightarrow x \geq 3/2$$

Pour  $x \geq \frac{3}{2}$  EST-2 a un plus grand risque.

b)	$x_1$	$x_2$	$Y$	Risque empirique pour $x=2$		Meilleur estimateur
				pour T	pour F	
	1	1	1T 2F	2	2	T
	1	2	4T 1F	1	8	T
	2	1	4T 4F	4	2	F
	2	2	1T 3F	3	2	F

d)	$Y=T$	$X=1$	$X=2$	$Y=F$	$X=1$	$X=2$
	$x_1$	5	2	$x_1$	3	7
	$x_2$	2	5	$x_2$	6	4

$$e) \quad P(Y=T | X=1) = P(Y=T) \cdot P(X_1=1 | Y=T) \times P(X_2=1 | Y=T)$$

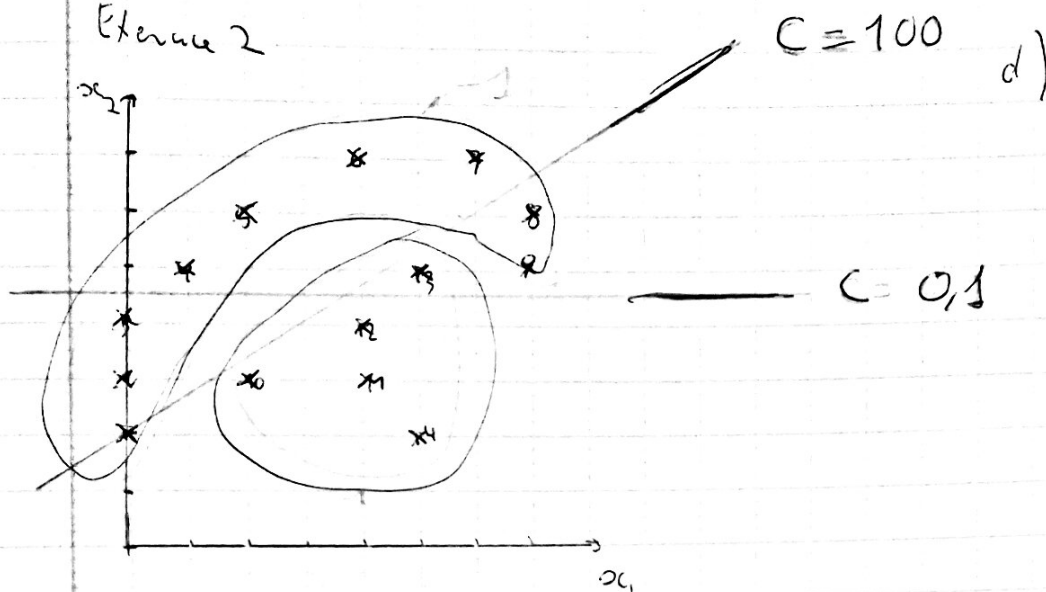
$$= 7/17 \times 5/7 \times 2/7$$

$$\approx 0,084$$

$x_1$	$x_2$	$Y=T$	$Y=F$	Estimateur bayésien
1	1	0,084	0,106	F
1	2	0,21	0,07	T
2	1	0,037	0,247	F
2	2	0,084	0,165	F

Les estimateurs sont semblables.

Exercice 2



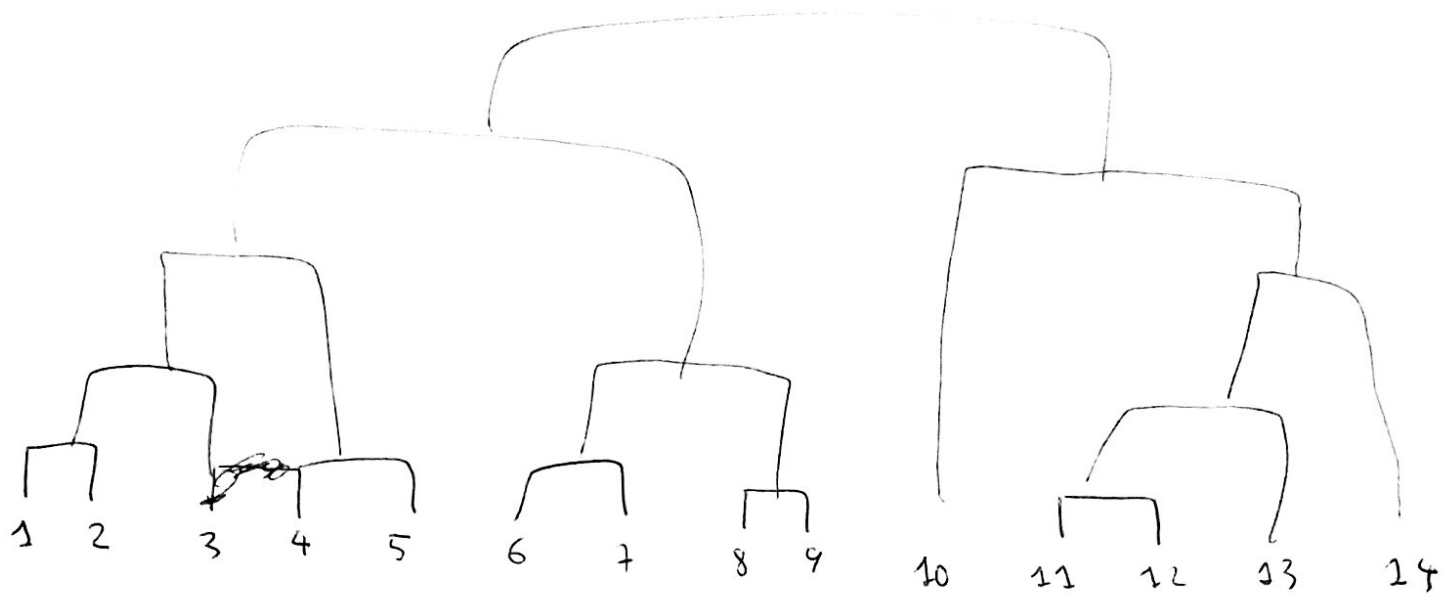
a) Une classification hiérarchique ascendante permettrait de distinguer ces 2 classes (contrairement à un K-means).

c) Impossible de les séparer avec un SVM linéaire.

e)  $f(x_1, x_2) = (x_1, x_2 - 10 \times x_2)$

b) Mesure de dissimilitude inter-classes : distance entre les individus les + éloignés.

d)



### Exercice 3

- a) Avec le test d'indépendance du  $\chi^2$  chi-deux.
- c) Le risque, c'est lorsque les probabilités sont connues.  
L'ambiguïté, c'est lorsque elles ne le sont pas.