Exercice 1 :

a) Est_1 : Faux positif : 8

Faux negatif : 1

Donc le risque empirique est de $\frac{8+x}{17}$

est_2 : Faux positif : 1

Faux negatif : 3

Donc le risque empirique est de $\frac{1+3x}{17}$

Ainsi : Si $x > \frac{7}{2}$ (3,5) est_1 est avantage.

Sinon ($x < \frac{7}{2}$) est_2 est avantage.

~~b)~~ Matrice de Confusion :

est_1 :

Real \ Predict	True	False
True	6	1
False	8	2

est_2 :

Real \ Predict	True	False
True	4	3
False	1	9

b) $p = 2$

est_1 a pour risque enpirique $\frac{8+x}{17} = \frac{9}{17}$

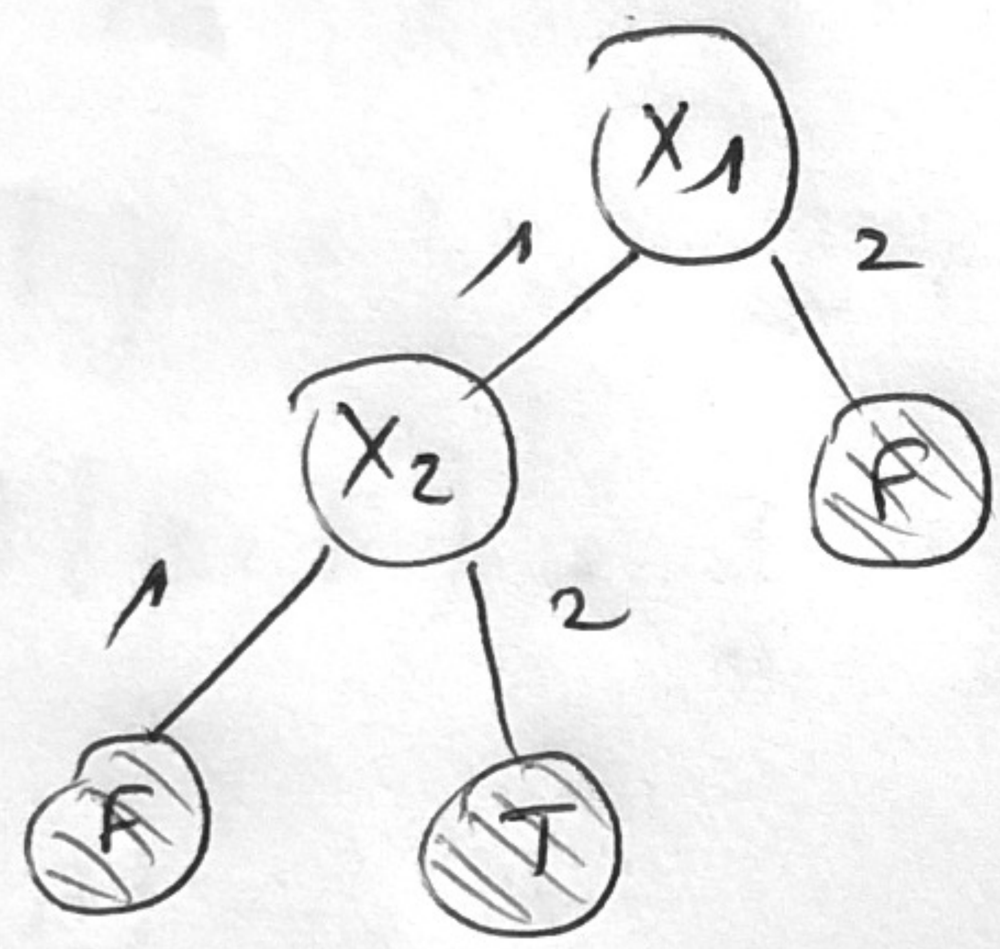
est_2 a pour risque enpirique $\frac{1+3x}{17} = \frac{7}{17}$

<u>Y = True</u>

$x_1 \backslash x_2$	1	2
1	1	4
2	1	1

<u>Y = False</u>

$x_1 \backslash x_2$	1	2
1	2	1
2	4	3



d)

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

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

• est_1 a pour risque empirique : $\frac{\;}{17}$: $\frac{\;}{17}$ $\frac{\;}{17}$

• est_2 a pour risque empirique : $\frac{1+3x}{17}$: $\frac{1+3\times 2}{17} = \frac{7}{17}$

e) • $P((1,1) \mid y=T) = P(X_1=1 \mid y=T) \times P(X_2=1 \mid y=T) \times P(y=T)$

$$= \frac{5}{7} \times \frac{2}{\cancel{7}} \times \frac{\cancel{7}}{17} = \frac{5\times 2}{7\times 17} = \frac{10}{7\times 17}$$

• $P((1,1) \mid y=F) : \frac{3}{\cancel{10}} \times \frac{6}{10} \times \frac{\cancel{10}}{17} = \frac{9}{5\times 17}$

• $P((1,2) \mid y=T) = \frac{5}{7} \times \frac{5}{\cancel{7}} \times \frac{\cancel{7}}{17} = \frac{25}{7\times 17}$

• $P((1,2) \mid y=F) : \frac{3}{10} \times \frac{4}{\cancel{10}} \times \frac{\cancel{10}}{17} = \frac{6}{5\times 17}$

• $P((2,1) \mid y=T) = \frac{2}{7} \times \frac{2}{\cancel{7}} \times \frac{\cancel{7}}{17} = \frac{4}{7\times 17}$

$P((2,1) \mid y=F) = \frac{7}{10} \times \frac{6}{\cancel{10}} \times \frac{\cancel{10}}{17} = \frac{7\times 3}{5\times 17}$

$P((2,2) \mid y=T) = \frac{2}{7} \times \frac{5}{\cancel{7}} \times \frac{\cancel{7}}{17} = \frac{10}{7\times 17}$

$P((2,2) \mid y=F) = \frac{7}{10} \times \frac{4}{\cancel{10}} \times \frac{\cancel{10}}{17} = \frac{14}{5\times 17}$

$\Rightarrow (1,1) \rightarrow F$

$(1,2) \rightarrow T$

$(2,1) \rightarrow F$

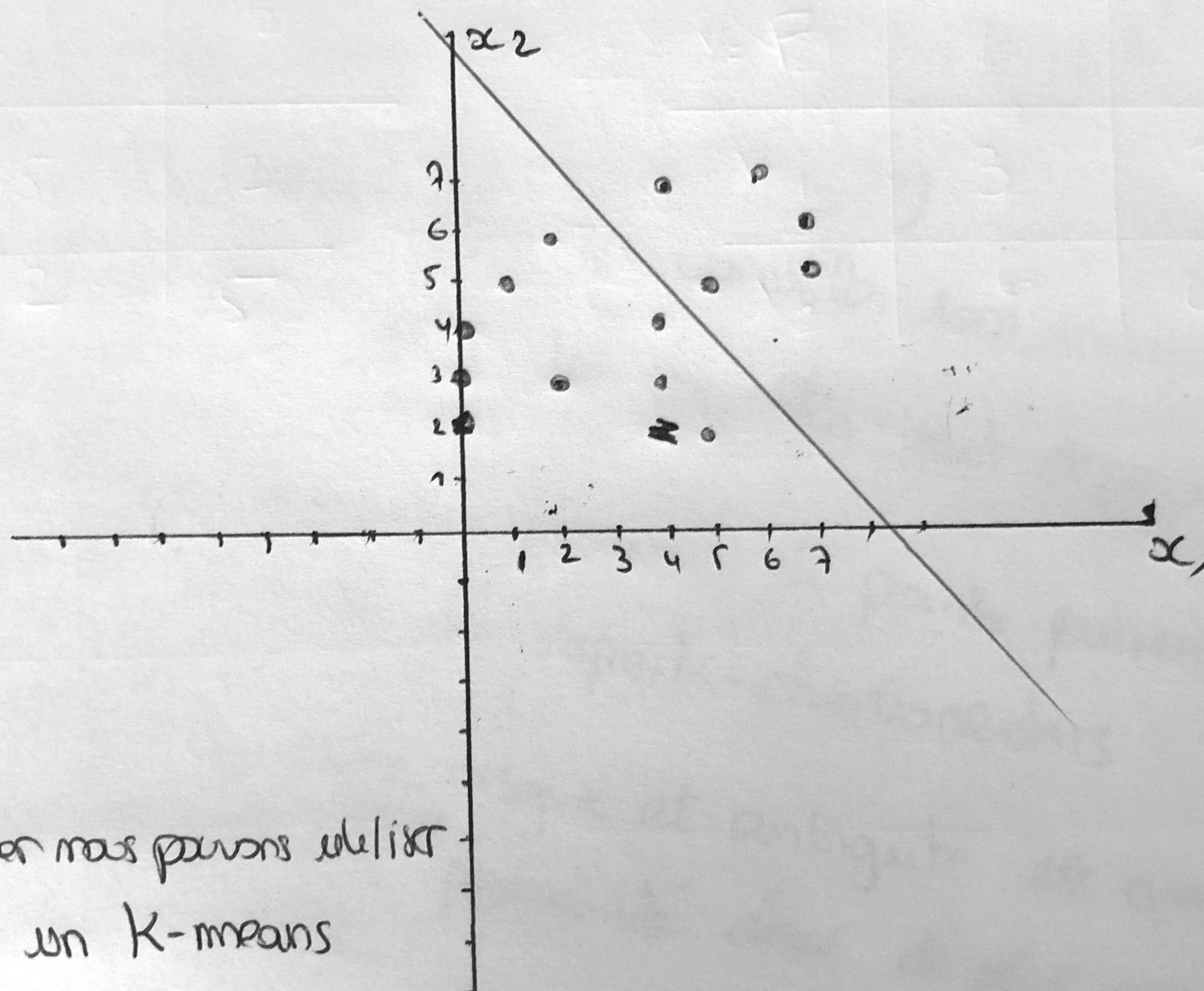
$(2,2) \rightarrow F$

L'arbre construit au petit 6 et l'estimateur bayesien naif predisent la même chose.

on calcule l'ecart ...

Exercice 2 :

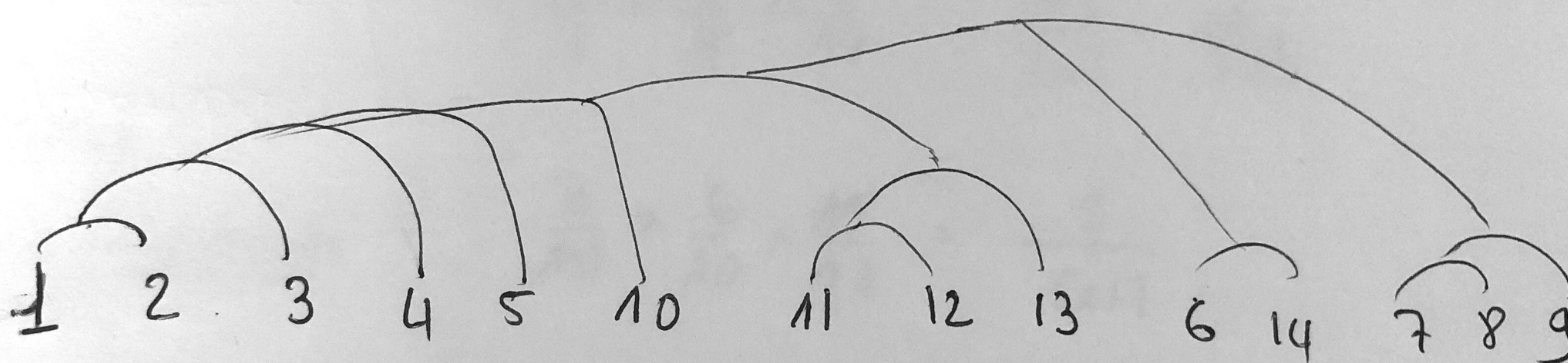
a)



→ Pour les distinguer nous pouvons utiliser un CHA ou un K-means

... que soit Il y'a absence de ...
... des données très ...
n'a jamais ...

b)



Pour évaluer le coût de l'intégration j'ai utilisé la distance.

Exercice 3 :

a) Pour ~~d~~determiner l'independence de deux variable à partir d'un tableau Croisé d'effectifs :

(1) on commence par calculer les effectifs croisés esperés :

$$E_{ij} = \frac{1}{n} \sum N(X=i)\, N(Y=j)$$

N etant le nb d'evenements

(2) On calcule les effectifs ~~espere~~ observé

$$O_{ij} = N(X=i \text{ et } Y=j)$$

(3) On calcule l'ecart relatif entre O_{ij} et E_{ij} en utilisant la methode des moindres carrés :

$$T = \sum_{ij} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

plus T est faible plus les variables sont independantes.
plus T est elevé plus les variables sont dependantes

b) Pour trouver le nombre maximum de points pulverisables, nous pouvons utiliser la methode de Vapnik-Chervonenkis

c) ~~La~~ difference entre risque et ambiguité est que un risque n'est pas censé avoir une probabilité elevé de plus un Risque est considéré Comme une erreur.
Or, une ambiguité signifie que soit Il y'a absence de sens dans les données (des données très contradictoire) soit un cas de figure n'a jamais été abordé.