

Examen de Scoring Appliqué à la Détection du Risque

Durée : 60 minutes. L'utilisation d'un poste informatique, de logiciels de statistique, du cours et d'internet sont autorisés.

Exercice 1.

mydata est un jeu de données de R indiquant le flux, l'emprunt et la classe de dix clients bancaires :

```
mydata=data.frame(flow=c(10,8,6,7,4,7,5,7,4,3),loan=c(20,22,51,78,34,45,51,43,65,70),  
class=c(1,1,1,1,1,2,2,2,2,2))
```

On souhaite estimer la classe de M. Li dont le flux vaut 5 et l'emprunt 52.

1. On considère les modèles Gaussiens : hétéroscédastique et homoscédastique.
 - (a) Quel est le meilleur modèle selon le critère BIC ?
 - (b) A quelle classe et avec quelle probabilité le meilleur modèle selon BIC affecte-t-il M. Li ?
 2. On considère les modèle de Régression Logistique.
 - (a) A quelle classe et avec quelle probabilité affecte-t-il M. Li ?
 - (b) Quel est le taux de mal classés parmi les données de mydata ?
 3. On suppose : (i) les flux de chaque classe sont distribués selon une loi exponentielle et (ii) en moyenne les flux de Classe 1 sont 1,25 fois plus grands que ceux de classe 2.

A quelle classe et avec quelle probabilité ce modèle de flux affecte-t-il M. Li ?
-