Question 17

On va justifier le choix du classifieur suivant :

$$\begin{cases} 1 & \text{si } X^T \hat{\Sigma}^{-1}(\hat{\mu}_+ - \hat{\mu}_-) > \frac{1}{2}\hat{\mu}_+^T \hat{\Sigma}^{-1}\hat{\mu}_+ - \frac{1}{2}\hat{\mu}_- + \\ -1 & \text{sinon} \quad \log(1 - 1 - m/n) - \log(m/n) \end{cases}$$

Pour mieux asseoir notre justification, il faut expliquer quelques termes suivant :

- $X^T \hat{\Sigma}^{-1}(\hat{\mu}_+ - \hat{\mu}_-)$: la fonction discriminante linéaire.

Elle calcule la distance de Mahalanobis entre X et les moyennes estimées des deux classes $\hat{\mu}_+$ et $\hat{\mu}_-$. $\hat{\Sigma}^{-1}$ est l'inverse de la matrice de covariance commune estimée. $\hat{\Sigma}^{-1}$ ajuste l'influence de chaque caractéristique ou variables lors de la classification.

- $\frac{1}{2}\hat{\mu}_+^T \Sigma^{-1}\hat{\mu}_+$: représente la distance de Mahalanobis de la moyenne de la classe positive par rapport à l'origine dans l'espace redimensionné ~~de~~ par $\hat{\Sigma}^{-1}$.
- $\frac{1}{2}\hat{\mu}_-^T \Sigma^{-1}\hat{\mu}_-$: idem mais pour une classe négative.
- $\log(m/n) = \log(\pi_+)$: algorithme naturel de la probabilité a priori de la classe positive.
- $\log(1 - m/n) = \log(\pi_-)$: idem mais pour une classe négative.

Maintenant que nous savons les différents termes, on peut justifier la pertinence du classifieur.

...