

## TD Analyse Discriminante Linéaire

Ce TD utilise le package MASS

1.

On va examiner l'analyse discriminante linéaire sur un exemple historique : les iris de Fischer

On commence par charger le package MASS

2.

Chargement des données iris, ici la commande `data(iris)` fait l'affaire car les données sont incluses dans le package.

remarquer que les 3 espèces d'Iris sont bien documentées. On va examiner s'il est possible de les reconnaître

3.

On lance directement la commande `lda`, plus tard on travaillera dans le premier plan principal (Regardez la documentation avant de lancer la commande, commentez en classe).

4.

A ce stade, les calculs ont été faits avec R

La question : ces calculs sont-ils significatifs ?

Avant de répondre, examinons quelques résultats donnés par R.

Vous devriez retrouver dans le résultat :

- les proportions de chacune des espèces d'iris,
- le décompte de chacune des espèces d'iris,
- le nombre total de fleurs dans la population.

Examiner les centres de gravités de chacune des espèces numériquement, on voit que ces centres ont l'air d'être bien séparés vous comprenez ici comment une analyse en composante principale permettra plus tard de dessiner "intelligemment" les individus de notre échantillon dans le premier plan principal (alors, qu'en dimension 4 c'est plus difficile à représenter).

Retrouver les coefficients qui sont employés par R pour séparer les classes. Ces coefficients ont été calculés avec la méthode présentée en cours, nous ne les recalculons pas dans le TD.

5.

Pour voir le résultat, gardez les valeurs numériques de iris dans un tableau, et effectuez la prévision des classes d'après le modèle généré, en utilisant la fonction R `predict`.

Examiner à quoi ressemble le résultat rendu par R.

6.

Donnez la matrice de confusion ...

**7.**

Remarquez que les probabilités telles que présentées dans les slides sont aussi calculées par R, retrouvez les.

**8.**

Vérifiez que ces probabilités calculées permettent de déduire les prévisions ...

**9.**

Maintenant, en classe, prenez des valeurs arbitraires et effectuez des prévisions

**10.**

Maintenant, vous faites une ACP, puis une segmentation hiérarchique et examinez la matrice de confusion obtenue.

**11.**

Si vous avez le temps, vous faites la même chose sur les deux exemples suivants du TD précédent :

- Les vins italiens,
- Les véhicules.