

UTILISATION DES STATISTIQUES ÉLÉMENTAIRES

Projet 4 : Le problème de la classification

Mots-clé : Statistiques descriptives, comparaison de moyennes, méthodes de classification.

Description du projet

I Echantillonnage et mesures

Comme chacun sait, il existe plusieurs espèces de pins (pin pignon, pin noir d'Autriche, etc...). La question est de savoir s'il est possible de reconnaître ces diverses espèces sur des critères morphométriques et d'apprécier sur cette base la « distance » qui les séparent les unes des autres.

Pour cela on reconnaîtra tout d'abord les individus des espèces de pin du Parc (Pin d'Alep, Pin des Canaries, Pin Pignon...). Pour chacune de ces espèces on choisira des individus sur lesquels on mesurera la longueur des aiguilles, et la longueur et diamètres des cônes («pommes de pin»). A l'occasion d'une sortie de week end, il serait bon de prélever aussi quelques individus de Pin sylvestre et de Pin Mésogéen.

II Analyse des données (quelques pistes seulement)

Partie 1

- Calculer pour chaque individu et chaque espèce présumée les statistiques élémentaires (moyennes et variances, histogrammes).
- Comparer les moyennes entre individus. Comparez les moyennes entre espèces.
- Conclure sur la pertinence du critère longueur des feuilles.

Partie 2

- On dispose donc pour chaque individu de deux mesures : longueurs des feuilles, volume du cône femelle. Reporter dans un diagramme de dispersion les coordonnées des individus. Peut-on identifier visuellement les espèces ? A partir de ce nombre appliquer la méthode des moyennes mobiles.
- A partir de la matrice des distances euclidiennes, construire une classification hiérarchique ascendante et comparer ce résultat avec le précédent.

III Conclure votre travail sur l'intérêt des mesures biométriques en matière de classification.