Université de Nantes

Licence Professionnelle "métrologie chimique et nucléaire"

Module de statistiques

Méthodes factorielles - Exercices

Exercices : Classification – Analyse des classes

Deux classifications ont été réalisées à partir du jeu de données 1-ACP voitures-LPro.xlsx, l'une avec toute les variables du jeu de données initiale, l'autre en ne prenant pas en compte la variable poids.





1- S'il n'y avait pas la légende au-dessus de chaque dendrogramme, serait-il possible d'identifier de façon certaine le diagramme "avec une variable en moins" ?

2- Si on réalise une partition en 2 classes, les classes sont-elles sensiblement équivalentes (taille comparable, individus sensiblement les mêmes) selon que l'on prend en compte 6 ou 5 variables ?

3- Décrire sommairement les 2 classes formées, dans les deux cas de graphiques.

4- On réalise maintenant une partition en trois classes. Les classes sont-elles équivalentes entre les dendrogrammes 5 variables et 6 variables ?

*La suite de l'exercice concerne maintenant une partition en 3 classes, réalisées sur les 24 individus, décrits par les 6 variables.*

5- Partant du tableau Excel 1-ACP voitures-LPro.xlsx, réalisez-en trois duplicatas, (un par classe). Pour chaque variable, calculez la moyenne, l'écart type et le coefficient de variation (exemple ci –dessous).



6- Partant du dendrogramme "à 6 variables" identifiez par surlignage sur chacun des duplicatas les véhicules appartenant à la classe considérée (un duplicata = une classe).

Sur chaque duplicata, supprimez les valeurs des variables pour les individus n'appartenant pas à la classe considérée, cela restreint aux seuls véhicules de la classe le calcul de la moyenne et de l'écart-type.

7- A quelles variations vous attendez-vous concernant la moyenne ?

8- A quelles variations vous attendez-vous concernant l'écart-type ?

9- A quelles variations vous attendez-vous concernant le coefficient de variation ?

10- Vérifiez les variations pour les trois classes.

11- A partir des 3 duplicatas, choisissez un ou des représentant(s) pour chaque classe. Il est rappelé que le centre de gravité est un représentant idéal de chaque classe, mais qu'il n'y a pas toujours un individu exactement au centre de gravité.

Le dendrogramme indique qu'une des classes pourrait aisément être scindée en deux sous-groupes.

12- Réalisez deux duplicatas du tableau de la classe en question, identifiez les individus appartenant au sous-groupe considéré puis supprimez les valeurs des autres individus. Qu'observez-vous ?

13- En vous appuyant sur le diagramme des individus (remis en séance) construisez les classes en reliant les individus appartenant à la même classe.