

USING GENERALIZED DISTANCES IN CLASSIFICATION OF GROUPS⁽¹⁾

P. Dagnelie et A. Merckx

Faculté des Sciences agronomiques
B-5030 Gembloux (Belgique)

pierre@dagnelie.be

ABSTRACT

The Mahalanobis generalized distance can advantageously be used to achieve the hierarchical clustering of groups of individuals. With a set of nearly 12,000 biometrical data, comprising populations of 14 different species of clover, we tried four methods to cluster those populations, in order to compare their results and to see whether the numerical classification obtained agrees with the botanical taxonomy. One of those methods is a conventional hierarchical clustering technique, based upon the Euclidean distances between the means of the populations, while the three other methods make use, with an increasing degree of complexity, of the generalized distances. These methods gave obviously better results.

RÉSUMÉ

La distance généralisée de Mahalanobis peut être utilisée avantageusement pour réaliser une classification hiérarchisée de groupes d'individus. Au moyen d'un jeu de près de 12.000 données biométriques relatives à des populations de 14 espèces différentes de trèfle, nous avons testé quatre méthodes de classification de ces populations, afin de comparer leurs résultats et de voir dans quelle mesure la classification numérique obtenue correspond à la taxonomie botanique. Une de ces méthodes consiste en une classification hiérarchisée classique au moyen des distances euclidiennes entre les moyennes des populations, tandis que les trois autres techniques font intervenir les distances généralisées, à différents degrés de complexité. Ces méthodes ont manifestement donné de meilleurs résultats.

⁽¹⁾ *Biometrical Journal* **33** (6), 683-695, 1991.