

Classification de variables avec la procédure Varclus de SAS : applications à la BPE (base permanente des équipements) et aux indicateurs européens de développement durable

Brigitte GELEIN¹ et Olivier SAUTORY²

Si les méthodes de classification les plus connues s'appliquent à la réalisation de typologies d'individus, la classification de variables s'avère également utile. En effet, l'émergence de bases de données toujours plus volumineuses confère à la réduction des dimensions une place cruciale. Or les méthodes factorielles (ACP, AFC et ACM) ne répondent pas à tous les besoins. Par la création de groupes de variables corrélées, la classification de variables offre la possibilité de construire des représentants de classes (variables synthétiques) - ou encore de choisir parmi les variables initiales celles qui sont les plus représentatives de leur classe (parangons). Ce nouvel ensemble plus restreint de variables sera plus facilement gérable et interprétable.

Certains domaines s'attachent davantage à la classification de variables qu'à celle d'individus. C'est le cas notamment de l'analyse sensorielle (création de groupes de descripteurs) ou encore de la médecine (élaboration de syndromes à partir d'un ensemble de symptômes). Cette utilisation de la classification de variables existe également dans l'économique, le social et l'environnement - thématiques dans lesquelles nous avons retenu deux exemples d'applications. Le premier porte sur la création de gammes d'équipements et le second sur l'identification de groupes de facteurs de développement durable.

On exposera, dans une première partie, l'approche divisive utilisée par la procédure Varclus de SAS. La mise en œuvre pratique de cette procédure sera ensuite décrite au travers des résultats obtenus par nos travaux sur la BPE (Base permanente des équipements) et sur les indicateurs européens de développement durable.

¹ Ensai - Tél. : 02 99 05 32 65 - bgelein@ensai.fr

² Cepe, Insee - Tél. : 01 41 17 57 46 - sautory@ensae.fr