

Une introduction à l'analyse discriminante avec SPSS pour Windows

Dominique DESBOIS

INRA-ESR Nancy et SCEES

251 rue de Vaugirard, 75732 Paris Cedex 15.

Fax : +33 1 49 55 85 00

Mél : desbois@jouy.inra.fr

RÉSUMÉ : cette note initie l'utilisateur débutant à la mise en œuvre de l'**analyse discriminante** par l'intermédiaire de la procédure **DISCRIM** du logiciel **SPSS pour Windows**. Cette mise en œuvre concerne le classement d'individus caractérisés par des variables quantitatives, les affectant à des groupes a priori au moyen de **fonctions discriminantes**. Les sorties du logiciel sont commentées par la présentation du formulaire de l'analyse discriminante associé à chacun des résultats obtenus.

MOTS-CLÉS : analyse discriminante linéaire, analyse factorielle discriminante, fonctions linéaires discriminantes, logiciel **SPSS 11.0.1**, mise en œuvre.

I) Problématiques de discrimination

Dans beaucoup de domaines, les professionnels sont amenés à prévoir les comportements sur la base de certains critères : c'est le cas par exemple d'un médecin établissant un diagnostic pour prescrire un traitement, ou d'un banquier accordant un crédit à un particulier ou une entreprise.

Une stratégie très empirique mais suffisamment éprouvée consiste à comparer ces caractéristiques aux relevés effectués sur un ensemble d'observations où le comportement des individus qu'ils soient malades ou emprunteurs, entreprises ou particuliers, est connu. Diagnostic ou prévision s'effectuent alors sur la base des corrélations observées entre les critères retenus et la situation des individus. Cette stratégie est d'ailleurs appliquée de façon plus ou moins implicite par la plupart des décideurs sur la base de leurs expériences antérieures dans des circonstances similaires.

Dans certains cas, les enjeux financiers ou humains sont suffisamment importants pour qu'on envisage de formaliser ce processus décisionnel en proposant une aide au diagnostic ou au jugement de l'expert. Par exemple dans le secteur bancaire, les techniques de *credit-scoring* cherchent à distinguer les bons des mauvais payeurs sur la base d'un ensemble de variables appelés **descripteurs** comme l'âge, le revenu annuel et le profil des transactions bancaires. Disposant de comptabilités d'entreprises, on peut également chercher à minimiser le risque de défaillance ou à prévenir les défaillances des emprunteurs, en classant selon une échelle de risque fixée a priori les agents économiques sur la base des descripteurs que sont les ratios financiers (cf. [Bardos, 2001]). **L'affectation à une classe** de risque décide alors de leur accès aux ressources financières ou de leur éligibilité à des mesures d'aide publique.

Consistant à prévoir l'appartenance du sujet aux groupes d'une partition a priori, la plupart de ces applications fondent leurs prédictions sur une technique statistique multidimensionnelle, l'**analyse discriminante linéaire**, introduite par **Ronald Fisher**.