

## *L'analyse factorielle des correspondances*

*Notes de cours et exemples.*

*par: Claire Durand, professeur, département de sociologie*

### *Un peu d'histoire...*

Jean-Paul Benzécri a élaboré l'analyse des correspondances et s'en est ensuite fait le promoteur. Dans "L'analyse des données", publié par Dunod en 1973 pour la première fois, il énonce les cinq grands principes sur lesquels l'analyse des données devrait reposer selon lui. Il est intéressant de voir que le premier principe s'énonce ainsi :

*"Statistique n'est pas probabilité. Sous le nom de statistique mathématique, des auteurs (qui, je vous le dis en français, n'écrivent guère dans notre langue...) ont édifié une pompeuse discipline, riche en hypothèses qui ne sont jamais satisfaites dans la pratique. Ce n'est pas de ces auteurs qu'il faut attendre la solution à nos problèmes typologiques"*  
Benzécri (1984, t.2 p. 3).

Le deuxième principe énoncé stipule que "Le modèle doit suivre les données". Le troisième spécifie qu'"Il faut traiter simultanément des informations concernant le plus grand nombre possible de dimensions". Enfin, les deux derniers principes ont trait à l'essentielle utilisation de l'ordinateur pour traiter les données. Ainsi, Benzécri, qui est un mathématicien, s'oppose aux bases – probabilistes – sur lesquelles reposent les procédures statistiques habituellement utilisées dans le monde anglo-saxon et les méthodes de type hypothético-déductif. Pour lui, l'analyse doit viser à élaborer des typologies et elle ne doit pas préjuger des résultats que la méthode fera apparaître. Ce sont les données qui servent de guide et non la théorie ou l'idéologie.

Cibois (1981) et Van Meter et coll. (1994) nous informent que Benzécri agissait comme un maître avec des disciples qu'il chargeait de répandre la *bonne parole*. Pour ce faire, un vocabulaire propre, un langage d'initiés a été créé. Comme il est normal de s'y attendre de la part d'un mathématicien, les premiers textes présentant l'analyse des correspondances comprenaient de nombreuses formules et équations mathématiques, certes utiles à qui pouvait les comprendre, mais non essentielles et difficiles à comprendre au chercheur en sciences sociales qui voulait simplement faire une utilisation intelligente de la méthode. Par ailleurs, dès les débuts, Benzécri utilise le terme générique d'"analyse des données" pour parler de sa méthode. On en viendra par la suite à parler de l'Ecole française d'analyse des données lorsque l'on se référera à l'analyse des correspondances. Ce type d'analyse prendra un certain temps avant de se répandre hors du cercle français.

On s'étonne que l'école anglo-saxonne ne s'y intéresse pas pendant un certain temps. C'est avec Greenacre (1984), lui-même élève de Benzécri, que la méthode fait officiellement son apparition dans le monde anglo-saxon. Deux raisons sont invoquées pour le peu d'intérêt pour la méthode dans le monde anglo-saxon au début : d'une part, la non disponibilité de textes en anglais

puisque les Français ne publiaient et ne faisaient de conférences qu'en français, et d'autre part, le fait que l'analyse factorielle et le multidimensional scaling remplissaient des fonctions équivalentes à l'analyse des correspondances. Toutefois, comme on le verra plus loin, il s'est développé dans l'Ecole française une vision intégrée de l'analyse, combinant l'analyse des correspondances et l'analyse de classification (voir Roux, 1994), ce qui justifie mieux l'adoption du terme "Analyse des données".

Benzécri proposait également une vision philosophique de sa méthode. Elle devait permettre de découvrir le sens profond des choses, la nature première, sans qu'il y ait intervention des "préjugés" du chercheur, de ses a priori théoriques ou autres. Cibois (1981) affirme que Benzécri croit en l'existence de facteurs "naturels" préexistants; il pense que sa méthode lui permettra de s'opposer aux idéologues, marxistes entre autres. Il demeure que la méthode a une valeur, indépendamment des positions de son auteur. Cibois (1981) la décrit comme wébérienne, en ce sens qu'elle permet de rechercher des "types idéaux" – via la combinaison des méthodes, entre autres – contrairement aux méthodes privilégiées par l'Ecole anglo-saxonne, plus durkheimiennes et qui visent la validation de modèles et la vérification d'hypothèses.

Van Meter et coll. (1994), référant à Rouanet (1988), identifient trois périodes dans l'évolution de l'Ecole française d'analyse des données, la période d'émergence (1963-1973) pendant laquelle les premiers textes et les premières thèses sont publiées, la période d'isolation (1973-1981) où l'analyse a un grand succès en France, entre autres grâce à l'utilisation que Bourdieu (1979) en fait et qui aide à populariser la méthode. Enfin, à partir de 1981, la période de "reluctant admittance" où les grands logiciels américains polyvalents (SPSS, BMDP, SAS) font leur apparition en France et à l'inverse, la méthode française se répand hors de France et finit par être intégrée dans les logiciels polyvalents (d'abord dans BMDP puis dans SPSS). Van Meter et coll. (1994) notent que la méthode se répand en dépit de l'opposition des statisticiens.

### *Qu'est-ce que l'analyse des correspondances ?*

C'est d'abord une procédure mathématique permettant de décrire en les plaçant sur des plans géométriques, l'ensemble des relations entre des variables de type catégoriel (nominales ou ordinales). Son originalité réside entre autres dans la possibilité de projeter sur ce plan des variables qui ne déterminent pas le plan, ce qui permet d'une certaine manière de conduire une analyse similaire à la régression multiple. Elle permet aussi de superposer le plan des variables et le plan des cas (appelés *objets* en analyse des correspondances). L'analyse des correspondances s'apparente fortement à l'analyse en composantes principales (sans rotation) mais se distingue de cette dernière par le fait qu'elle est basée sur des écarts pondérés. Cette particularité donne la même importance relative aux écarts à l'indépendance quel que soit le nombre de cas -- c'est-à-dire que l'on décompose la structure des écarts. L'information relative au nombre de cas est tenue en compte par la "masse", qui détermine le centre (appelé centroïde) du plan. En analyse des correspondances, ce sont uniquement les écarts à l'indépendance des observations qui nous intéressent. Il faut souligner qu'un des principaux avantages de l'analyse des correspondances est sa capacité à analyser des données de type catégoriel, sans qu'il y ait de postulats à respecter quant aux caractéristiques des distributions (normales, multi-normales,

homoscédasticité, etc.)(Van Meter et coll. 1994).

Toutefois, comme l'analyse des correspondances permet de positionner sur un plan à  $x$  dimensions l'ensemble des cas en fonction de la similitude de leurs réponses aux questions, ces cas peuvent être catégorisés par leur positionnement sur chacun des axes. A partir de ces positionnements, la pratique est de poursuivre le processus par une analyse de classification qui peut alors suggérer des regroupements des individus. Le but de l'opération est d'en arriver à des typologies. Toutefois, il s'agit d'une deuxième analyse que nous n'aborderons pas dans ce texte.

### ***Exemples d'utilisation***

L'exemple le plus connu, celui qui a sans doute aidé à faire connaître cette procédure d'analyse en France même et en dehors de la France, c'est l'utilisation qu'en a faite Bourdieu dans *La Distinction* (1979:293-301) où il met en relation le capital économique et culturel d'une part et les goûts et connaissances artistiques d'autre part. On rencontre aussi des analyses des correspondances dans les travaux de Jean Lojkine, dans le livre de Doise et coll. (1992) sur les représentations sociales. Pour un texte plus récent, voir Doucet, Durand et Smith (2008) portant sur les écarts salariaux selon le genre chez les professeurs d'université.

On utilise l'analyse des correspondances lorsque l'on veut représenter l'ensemble des relations entre des variables de type nominal ou ordinal. Les variables les plus fréquemment utilisées dans ce contexte sont des variables socio-démographiques, des variables de préférences ou de comportement et certaines variables d'attitudes dont les échelles de réponse sont de type Likert. Avec ce dernier type de variables, l'analyse des correspondances est théoriquement plus appropriée et souvent plus révélatrice (Blasius et Thissen, 1998) que l'analyse factorielle habituellement utilisée (dont l'utilisation devrait théoriquement être réservée aux variables métriques). Toutefois, l'usage nord-américain privilégie dans ce cas l'utilisation de l'analyse factorielle que Cibois appelle *l'analyse factorielle des psychologues*. On voit comment les traditions disciplinaires et culturelles s'entremêlent pour privilégier certaines méthodes. Il demeure que, lorsque les mesures d'attitudes de type Likert sont de fait des approximations de mesure métriques, les deux méthodes donnent habituellement des résultats similaires (à tout le moins pour ce qui est du nombre de dimensions significatives et de l'appartenance des variables à chaque dimension). De même, une dichotomisation des variables nominales ou ordinales soumises ensuite à l'analyse factorielle donne des résultats qui peuvent être presque identiques à ceux de l'analyse des correspondances (Durand, 1993). En conclusion, le choix d'une méthode d'analyse "factorielle" (en composantes principales, en facteurs communs ou "analyse factorielle des psychologues", analyse factorielle des correspondances ou analyse des données) pour explorer les données dépend plutôt de l'aptitude de chaque méthode à donner des résultats interprétables et utiles compte tenu de la forme des données et des buts de l'analyse.

### **Références:**

Benzécri, Jean-Paul (1984) *L'analyse des données*, Paris: Dunod,t.2, Partie A, p. 3-17.

Bourdieu (1979) *La Distinction*, Paris : Editions de Minuit, 670p.

Cibois, P. (1981). *Analyse des données et sociologie. L'Année sociologique* (31), p.333-348.

Doise, W., Clemence, A. Et F. Lorenzi-Cioldi (1992). *Représentations sociales et analyses de données*. Grenoble: PUG.

Doucet, C., Durand, C. and M. Smith (2008). Who gets Market Supplements? Gender Differences within a Large Canadian University, *Canadian Review of Higher Education*, 38(1), 67-103.

Roux, M. (1994) *Classification des données d'enquête* in Grangé, D. Et Lebart, L., ed. *Traitement statistique des enquêtes*. Paris Dunod, p. 91-112.

Van Meter, K., Schiltz, M.-A., Cibois, P et Mounier, L. (1994). Correspondence analysis: A history and French sociological perspective in Greenacre, M. et Blasius, J. Ed. *Correspondence analysis in the social sciences*. London: Academic Press, p. 128-137.

### ***Le petit vocabulaire de l'analyse factorielle des correspondances***

- Analyse en facteurs communs (de Thurstone), analyse factorielle des psychologues, analyse factorielle.
- Analyse en composantes principales
- Analyse des correspondances, analyse factorielle des correspondances, analyse des données (Attention le Que sais-je? (Cibois, 1981) intitulé *L'Analyse factorielle* porte en fait presque exclusivement sur l'analyse des correspondances, en relation avec l'analyse en composantes principales et ne porte pas du tout sur l'analyse factorielle en facteurs communs tel que le terme est compris habituellement).
- inertie  $\cong$  variance
- dimension ou axe  $\cong$  facteur
- contribution  $\cong$  saturation
- masse ou poids  $\cong$  fréquences
- objet (object)  $\cong$  cas (dans une matrice de données)
- Object scores : position d'un cas sur une dimension (donnant sa proximité -- ou son éloignement -- avec les diverses catégories de réponses).
- coordonnées des catégories: position des catégories sur le plan lorsque les variables sont de type nominal simple, ordinal ou métrique.
- quantification d'une catégorie : position d'une catégorie sur une dimension = moyenne des scores d'objets appartenant à cette catégorie. Position des catégories sur le plan lorsque les variables sont de type nominal multiple.
- optimal scaling
- *conjonction et opposition* : Voir Cibois (1981,1990). Deux catégories sont dites en conjonction lorsque l'angle entre les droites partant de l'origine à chacune des catégories est inférieur à  $90^\circ$  ; lorsque ces droites sont en opposition (angle  $>90^\circ$ ), cela signifie que les catégories s'excluent

mutuellement (les objets qui sont dans une catégorie ne sont pas en même temps dans l'autre) ; les catégories sont indépendantes quand le fait de savoir qu'un objet appartient à une catégorie ne nous permet pas de tirer quelque déduction que ce soit sur son appartenance à une autre catégorie.

### *L'analyse des correspondances avec SPSS*

Plusieurs logiciels existent pour faire l'analyse des correspondances, certains même gratuits (voir entre autres sur les sites de Philippe Cibois et de Ludovic Lebart (probablement le plus développé). D'autres logiciels français dont SPAD permettent de produire des analyses et des graphiques sophistiqués. Nous présentons ici le module de SPSS puisqu'il est inclus dans la plupart des licences des Universités. Ce n'est pas le meilleur mais il permet d'avoir une idée des données.

SPSS comprend un module complémentaire, appelé "Categories" qui permet de faire de l'analyse des correspondances. Ce module comprend un certain nombre de procédures. Les trois procédures qui apparaissent les plus pertinentes pour l'utilisation traditionnelle de l'AFC sont les suivantes : ANACOR (pour Analyse des Correspondances), HOMALS (pour Homogeneity analysis using Alternating least squares) et PRINCALS (pour Principal component analysis using Alternating least squares). Le module OVERALS permet de faire une analyse entre deux ensemble de variables.

#### **ANACOR:**

Le module ANACOR (CORRESPONDENCE ANALYSIS dans le sous-menu DATA REDUCTION de STATISTICS) de SPSS permet de procéder à une analyse des correspondances "traditionnelle", c'est-à-dire une **analyse entre deux variables ayant un nombre plus ou moins grand de catégories**. Cette restriction à deux variables peut apparaître gênante. Toutefois, il faut se rappeler qu'il est possible de combiner les catégories de deux variables pour en faire une seule lorsque cela est pertinent (les variables de sexe et d'âge peuvent être combinées en une variable ayant des catégories de femmes et d'hommes de divers groupes d'âge). L'avantage d'ANACOR réside surtout selon SPSS (Categories , p. B-26) dans la qualité des informations qu'elle produit sur l'adéquation de la représentation (proportion d'inertie expliquée par chaque dimension, contribution de chaque rang et de chaque colonne à l'inertie, etc.

Par défaut, la procédure produit l'ensemble des informations et graphiques nécessaires pour une extraction sur deux dimensions. Notons que le nombre de dimensions maximales pour expliquer l'ensemble des écarts à l'indépendance est égal à une dimension de moins que le nombre de catégories de la variable ayant le moins de catégories : pour un tableau croisé d'une variable à 5 catégories avec une autre à 10 catégories, le nombre de dimensions nécessaires pour expliquer l'ensemble des écarts à l'indépendance est de 4, c'est-à-dire 5-1. L'analyse des correspondances "classique" est donc particulièrement utile pour le croisement de deux variables dont une aurait un nombre important de catégories (les départements d'une université d'une part et les taux d'abandon regroupés d'autre part, par exemple).

Pour effectuer l'analyse des correspondances classique (ANACOR), il est très simple d'utiliser le tableau qui apparaît dans SPSS pour Windows et d'y inscrire la variable "colonne" (à expliquer) et la variable "ligne" (variable explicative). Toutefois, par défaut, la procédure ne donne pas la représentation graphique conjointe des deux variables (SIC) et il faut donc cocher cet item (*Plot joint*) dans les options.

La commande telle qu'enregistrée dans le fichier de syntaxe (via PASTE) devrait avoir cette apparence :

```
ANACOR
  TABLE=age(1 5) BY ancien(1 5)
  /DIMENSION=2
  /NORMALIZATION CANONICAL
  /PRINT TABLE SCORES CONTRIBUTIONS
  /PLOT ROWS COLUMNS JOINT NDIM(ALL,MAX).
```

**Note : version 9 : nouvelle procédure : correspondance, plus d'options...**

## HOMALS

La procédure HOMALS permet d'effectuer l'analyse des correspondances multiples, c'est-à-dire portant sur plusieurs variables de type multi-nominal. L'expression multi-nominal réfère ici non pas au fait qu'il y aurait plusieurs catégories de type nominal mais au fait que chaque catégorie peut déterminer plus d'une dimension. On utilise HOMALS lorsque l'on ne présume pas d'un ordonnancement dans les catégories d'aucune des variables, ni du fait qu'une variable devrait se retrouver sur une seule dimension.

Pour procéder à cette analyse dans SPSS, il suffit d'accéder au menu Optimal scaling du menu DATA REDUCTION de STATISTICS.. Le tableau qui apparaît offre un choix quant au type de variables et quant au nombre d'ensembles de variables ("sets"). Pour un ensemble de variables (plusieurs ensembles nous amènent à la corrélation canonique non linéaire, procédure OVERALS), si on indique que toutes les variables doivent être analysées au niveau nominal multiple, le logiciel nous amène automatiquement dans la procédure HOMALS. Il suffit alors de donner à la procédure les variables que l'on veut utiliser pour l'analyse en indiquant la catégorie maximum. *La procédure présume que la première catégorie est le 1 et cette information n'est pas modifiable, ce qui entraîne que si la variable est sur une échelle de 0 à 10, on devra ajouter 1 à la variable pour créer une nouvelle échelle de 1 à 11, sinon toutes les valeurs 0 seront considérées comme des valeurs manquantes.*

Les options prévues par défaut sont appropriées et il n'y a pas lieu de les modifier. Elles donnent les informations et graphiques de base et il n'est pas possible d'obtenir plus. A noter que par défaut, la procédure présente deux dimensions. Il peut être intéressant d'en prévoir trois ou même quatre selon les situations. Il est alors préférable d'éditer la syntaxe pour les graphiques de telle sorte de ne pas se retrouver avec un graphique en trois dimensions (que l'on peut faire "rouler" dans l'éditeur de graphique, mais qui est habituellement strictement non imprimable et non interprétable). Toutefois, il semble que l'on doit faire rouler la commande pour chaque graphique que l'on demande parce qu'il est impossible de demander plusieurs graphiques. Le nombre maximum de dimensions d'un tel type d'analyse est égal au nombre de catégories au total auquel on soustrait le nombre de variables, ce qui équivaut en fait au nombre de degrés de liberté i.e l'addition du nombre de catégories moins un pour chaque variable.

```
HOMALS
/VARIABLES=age(5) domain(6) sexe(2) statut(3) ancien(5)
/ANALYSIS=age domain sexe statut ancien
/DIMENSION=6
/PRINT FREQ EIGEN DISCRIM QUANT
/PLOT QUANT OBJECT DISCRIM NDIM(ALL,MAX)
/MAXITER = 100
/CONVERGENCE = .00001 .
```

## PRINCALS

La procédure PRINCALS permet d'effectuer une analyse en composantes principales non linéaires lorsque les variables dont on veut analyser les relations sont de plusieurs types, c'est à dire soit ordinales, soit nominales simples (single-nominal), soit nominales multiples (multiple-nominal), soit numériques. Il est important de bien comprendre les conséquences de l'attribution d'un type plutôt qu'un autre à l'analyse des variables. Il faut d'abord comprendre qu'il s'agit non pas du type de variable mais de la manière dont celles-ci sont analysées.

*Lorsqu'une variable de type Likert est analysée au niveau ordinal*, l'analyse préservera l'ordonnement entre les catégories (très, assez, peu, pas du tout) sans toutefois maintenir une distance égale entre les catégories alors que si elle est analysée comme variable métrique, à la fois l'ordre et la distance entre les catégories seront préservées et la ligne droite joignant les catégories passera par le centre du graphique. Par contre, si la même variable est analysée au niveau simple nominal, l'ordonnement n'est pas respecté, c'est-à-dire que la catégorie "très" pourrait se retrouver entre la catégorie "peu" et la catégorie "assez" ou même près de la catégorie "pas du tout" auquel cas on peut penser la dimension distingue d'une part les personnes ayant des positions tranchées, dans un sens ou dans l'autre, et d'autre part, les gens ayant une position plus neutre. En analysant au niveau simple nominal toutefois, la variable se maintient sur une seule dimension et la droite qui joint les points passe aussi par le centre du graphique. Si on analysait cette même variable au niveau multiple nominal, non seulement l'ordonnement ne serait pas préservé mais de plus certaines catégories pourraient être fortement reliées à une dimension alors que d'autres catégories seraient liées à une autre dimension. On peut aisément se retrouver avec une ligne en  $\cup$  joignant les catégories, une première dimension allant du négatif au positif et une deuxième dimension distinguant les positions extrêmes des positions centrales.

Un exemple de cette situation survient lorsque l'on demande le type de diplôme et le domaine du diplôme à une population de professionnels au Québec. Les professionnels plus âgés ont parfois complété ce qui s'appelait un bac. ès arts et qui est l'équivalent de 15 ans de scolarité (un an de moins qu'un baccalauréat universitaire). Ces personnes inscrivent donc, à diplôme "bac" et à domaine "arts". Les autres diplômés de baccalauréat, de maîtrise ou de doctorat inscrivent comme domaine de diplôme, les domaines habituels (sciences pures, administration, sciences humaines, etc.). Lorsque l'on permet à cette variable (domaine du diplôme) d'être analysée comme nominale multiple, conjointement avec les variables de diplôme, d'âge et de sexe, les divers domaines de diplomation, à l'exception des Arts, se retrouvent sur un axe avec le sexe, les sciences étant plus près de la catégorie Homme et les sciences humaines et lettres plus près de la catégorie femme. Toutefois, le domaine Arts se retrouve sur une deuxième dimension fortement liée à l'âge et en conjonction avec les catégories d'âge plus élevées.

Pour effectuer une analyse de type PRINCALS, il faut choisir non seulement les variables à analyser mais aussi décider du niveau où chaque variable sera analysée. Lorsqu'aucune variable

n'est analysée au niveau nominal multiple, chaque valeur propre doit être plus grande que  $1/nb$  de variables pour apparaître importante. Par contre, lorsqu'une des variables est analysée au niveau multiple nominal, les règles quant au nombre de facteurs apparaissent plus floues. A noter d'ailleurs que lorsqu'une variable est analysée à ce niveau, le calcul des saturations de la variable elle-même sur les dimensions est impossible. Elle apparaît alors au centre du plan, ce qui porte à confusion.

```
PRINCALS
/VARIABLES=age(5) ancien(5) domain(6) sexe(3) statut(3)
/ANALYSIS=age(ORDI) ancien(ORDI) domain(MNOM) sexe(SNOM) statut(SNOM)
/DIMENSION=2
/PRINT FREQ EIGEN LOADINGS QUANT
/PLOT QUANT OBJECT LOADINGS NDIM(ALL,MAX)
/MAXITER = 100
/CONVERGENCE = .00001 .
```

### *Comment faire...*

Les étapes;

1) d'abord définir le but de l'analyse et les variables qui seront analysées

- pour deux ou trois variables avec de multiples catégories, privilégier ANACOR
- pour plusieurs variables n'ayant aucun ordonnancement (de type nominal), privilégier HOMALS
- pour des variables de plusieurs types (nominal, ordinal, métrique), PRINCALS de façon automatique.

2) Attention avant de procéder à l'analyse, s'assurer d'un certain nombre de détails techniques à régler pour obtenir des graphiques lisibles:

- a) dans le "page setup" de l'imprimante, voir les options et s'assurer de demander "full page height"
- b) dans les options de graphique, noter "cycle through patterns" plutôt que "cycle through colors" de façon à pouvoir distinguer les diverses variables.

3) procéder à l'analyse en pensant, dans le cas des variables qui peuvent être analysées à divers niveaux, à vérifier que le niveau choisi est adéquat. Dans ce cas, lorsque l'on analyse la variable au niveau nominal multiple, on obtient des résultats similaires à une analyse au niveau nominal simple ou ordinal selon le cas.

## Références:

sur Internet : [Http://www.soc.surrey.ac.uk/sru/sru7.html](http://www.soc.surrey.ac.uk/sru/sru7.html)

- \* Cibois, P. L'analyse des données en sociologie. PUF, 1984.  
plus particulièrement les pages 56-73 pour une explication détaillée et simplifiée de l'algorithme.
- Cibois, P. (1981). Analyse des données et sociologie. L'Année sociologique (31), p.333-348.  
(excellent article portant sur les liens entre Ecoles d'analyse des données (française vs anglo-saxonne) et positions idéologiques et méthodologiques.
- Doise, W., Clemence, A. Et F. Lorenzi-Cioldi (1992). Représentations sociales et analyses de données. Grenoble: PUG. chap 3 p. 71-83 et chap 8, p. 171-185.
- Greenacre, M. et Blasius, J. (1994). Correspondence analysis in the social sciences. London: Academic Press, 370.p.  
Rassemble des chapitres de divers auteurs dont Van Meter et coll.. Très intéressant. Facile à lire et à comprendre.
- Michel, F. (1994). Intensité de liaison et masse d'information des tableaux de contingence: Deux problèmes de mesure en analyse des données. *Bulletin de méthodologie sociologique*, (44), p. 73-92.
- Van Meter, K., Schiltz, M.-A., Cibois, P et Mounier, L. (1994). Correspondence analysis: A history and French sociological perspective in Greenacre, M. et Blasius, J. Ed. *Correspondence analysis in the social sciences*. London: Academic Press, p. 128-137.
- \* Volle, Michel (1978). L'analyse des données. *Economie et Statistique* (96), 1-23

## Applications:

- Cibois, P. (1989). Eclairer le vocabulaire des questions ouvertes par les questions fermées: le tableau lexical des questions. *Bulletin de méthodologie sociologique*, (26), p. 12-23.
- \* Durand, C. (1993). L'aspiration à la mobilité en emploi chez les professionnels, diverses formes, différents déterminants. Thèse de doctorat. Université de Montréal. p. 147-169.  
+ ajout listing
- \* AU - Langlois A; Razin E  
TI - SELF-EMPLOYMENT AMONG FRENCH-CANADIANS - THE ROLE OF THE REGIONAL MILIEU  
SO - Ethnic & Racial Studies 1995 Jul;18(3):581-604