

REVUE DE STATISTIQUE APPLIQUÉE

ANDRÉ VIDAL

L'analyse de groupes à variables multiples ou analyse typologique

Revue de statistique appliquée, tome 3, n° 4 (1955), p. 87-94

http://www.numdam.org/item?id=RSA_1955__3_4_87_0

© Société française de statistique, 1955, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

L'ANALYSE DE GROUPES A VARIABLES MULTIPLES OU ANALYSE TYPOLOGIQUE

par

André VIDAL

Le développement des connaissances scientifiques et techniques s'est fait moins par l'application indéfinie des mêmes méthodes à des lignes de recherche initiales, que par la découverte de nouveaux domaines susceptibles de se prêter à une étude systématique et de nouvelles méthodes permettant de progresser dans ces domaines ou de renouveler d'anciennes lignes qui paraissaient épuisées.

Dans le domaine de la statistique, les méthodes d'étude de données qualitatives, par exemple, commencent à se faire jour à côté des procédés plus classiques applicables à des faits quantitatifs.

Le but de cet article est de présenter une nouvelle technique d'étude statistique (fort ancienne à bien des égards) qui est de nature à compléter ou à renouveler nos moyens d'approche dans des domaines très divers.

Il s'agit de la découverte des types; ou, si l'on préfère, de la découverte de la manière dont, dans une population d'objets caractérisés par plusieurs variables, les objets se groupent en familles analogues au point de vue de toutes les variables considérées simultanément; ou, si l'on préfère encore, la découverte des zones de condensation dans un espace à n dimensions.

L'ANALYSE DE GROUPES A VARIABLES MULTIPLES OU ANALYSE TYPOLOGIQUE

Soit une population d'objets, dont chacun est caractérisé par des valeurs dans, par exemple, cinq variables. Il s'agit, si l'on veut, de divers échantillons de pétrole. Chaque échantillon a une teneur a dans le constituant x_1 , b dans le constituant x_2 , etc...

On peut représenter cet échantillon par une ligne brisée, sur un graphique comprenant cinq ordonnées, comme dans la figure 1.

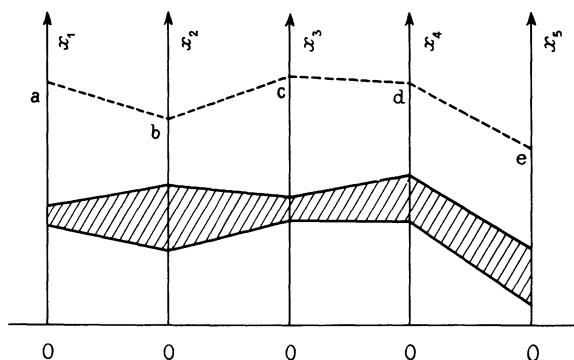


Fig. 1

Il peut arriver que les formes et les emplacements des différents profils, tels que a b c d e, soient très différents d'un échantillon à l'autre. Il peut arriver aussi qu'un nombre plus ou moins important d'échantillons aient des profils voisins et semblablement placés, qui passent tous dans un canal plus étroit, tel que celui qui a été hachuré sur la Fig. 1. Ce groupement est défini, si l'on veut, par l'axe du canal et par ses épaisseurs aux différentes variables. Il constitue ce qu'on appelle un **type**. Nous appelons "intervalles de tolérance" les épaisseurs du canal aux différentes variables.

Si les "intervalles de tolérance" pour chaque variable sont grands, c'est-à-dire si le canal hachuré est très large, il est normal qu'il y passe beaucoup d'échantillons. Si, au contraire, il passe beaucoup d'échantillons à travers un canal très étroit, il est hautement probable que cela n'est pas dû au hasard. La "significativité" de l'existence d'un groupement est donc fonction directe de l'effectif du groupe et fonction inverse de la largeur du canal.

Nous ne connaissons encore aucune méthode pour chiffrer cette significativité. Ce problème est actuellement entre les mains de mathématiciens, amis de cette Revue, et qui espèrent arriver à une solution commode de ce problème fort difficile. Mais ce point n'est pour nous que secondaire, ou plus exactement second. Le problème que nous souhaitons examiner ici est celui de la **découverte même des types**.

Etant donné une population d'objets caractérisés sur différentes variables, si on reporte le profil de chaque objet sur la même figure 1, on aboutit à un réseau inextricable. Fixant a priori des intervalles de tolérance et un effectif minimum, le problème consiste à se demander s'il y a dans ce réseau des canaux de l'épaisseur ainsi définie, **quels que puissent être leur forme et leur emplacement**, et contenant un nombre de profils supérieur ou égal à l'effectif admis comme minimum. Pour le lecteur qui préfère "voir" les choses dans l'espace, chaque objet est représenté par un point dans un espace à n dimensions, n étant le nombre de variables. Un type est une sorte de grumeau dans ce nuage de points. On décide qu'il y aura grumeau quand on trouvera plus de m points dans une zone de "diamètres" déterminés. Bien entendu, il y a lieu de ne rien présupposer en ce qui concerne l'emplacement du grumeau, car il peut être n'importe où. Le problème consiste à trouver une méthode pour détecter les grumeaux, tous les grumeaux répondant aux seules conditions tolérances et effectif, comme le radar détecte l'avion dans l'espace qu'il balaye.

Sur le plan mathématique, la question est pratiquement insoluble par les voies habituelles au-dessus d'un tout petit nombre de variables. Au vrai, elle est traditionnellement considérée comme assortie de difficultés telles que beaucoup de statisticiens, même fort avertis, en ont oublié jusqu'à l'existence. Notre chance a été de trouver une combinaison qui s'est révélée valable, de considérations mathématiques et de "tours de mains" dans le maniement de la carte perforée classique. ⁽¹⁾

Appliqué à différents problèmes de statistique industrielle et de recherche psychologique, ce procédé nous a donné l'occasion de mesurer clairement l'intérêt méthodologique de premier ordre attaché à cette voie d'approche. Nous en reparlerons plus loin.

LIAISONS ET FRONTIÈRES AVEC LA STATISTIQUE CLASSIQUE

Les rapports de l'analyse typologique avec certaines autres techniques statistiques habituelles se laissent assez bien figurer par le tableau ci-dessous. On y voit que l'analyse typologique vient combler une lacune ancienne dans le domaine des statistiques multi-dimensionnelles.

Partant de données quantitatives multidimensionnelles, le statisticien peut avoir à découvrir de nouvelles variables, encore inconnues; ou bien il peut avoir

⁽¹⁾ Il n'est pas dans nos intentions de publier, au moins pour le moment, cette méthode qui d'ailleurs n'est pas à la portée de chacun, notamment parce qu'elle utilise une trieuse d'un modèle particulièrement évolué.

- N.d.l.R. Un service spécialisé de l'Organisation dirigée par M. Vidal est à la disposition des statisticiens pour effectuer de tels travaux.

	Méthode de découverte	Technique de prédiction
Domaine Quantitatif	Analyse Factorielle	Equations de régression Corrélations canoniques
Domaine Qualitatif	Analyse Typologique	Fonctions discriminantes.

à prédire des valeurs pour d'autres variables, dont l'existence est déjà connue. Cette distinction correspond aux deux colonnes du tableau. D'autre part, la variable à découvrir ou à prédire peut être d'ordre quantitatif, ou bien consister simplement en données qualitatives, comme l'existence ou la non-existence d'un caractère, l'appartenance ou la non-appartenance à un groupe. Cette opposition correspond aux deux lignes du tableau.

On aperçoit ainsi que l'analyse typologique est la démarche inverse des fonctions discriminantes. Celles-ci supposent les types connus et permettent de rapporter les nouvelles observations à l'un de ces types connus, de façon à rendre minimum le nombre d'erreurs de classification. L'analyse typologique se pose le problème logiquement précédent, qui est celui de la détermination de ces types.

On aperçoit également que l'analyse typologique a ceci de commun avec l'analyse factorielle qu'elle est aussi une méthode inductive, mais ces deux méthodes semblent explorer deux "versants" opposés des phénomènes. L'analyse factorielle s'intéresse à la définition des dimensions de l'espace représentatif du domaine étudié. Elle s'intéresse ainsi à la projection des vecteurs représentant les observations sur certaines variables primaires. **Elle nous dit comment se répartissent ou se groupent les scores dans les différentes variables, mais non comment se répartissent ou se groupent les faits ou les objets étudiés eux-mêmes.** Il y a là une optique entièrement différente. L'analyse typologique considère les observations toutes dimensions réunies et peut ainsi mettre en lumière des faits importants que rien ne décèle sur le "versant variables" des mêmes données.

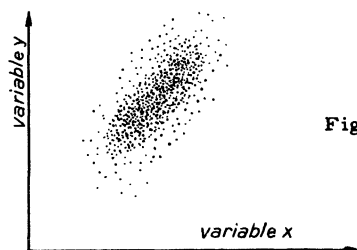


Fig. 2

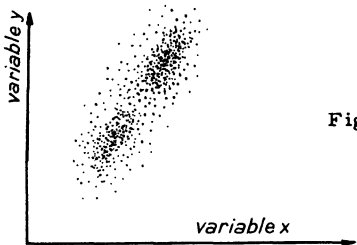


Fig. 3

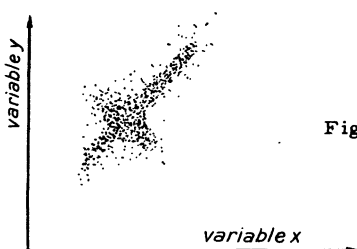


Fig. 4

Par exemple, sur deux variables, il n'est pas du tout indifférent que la population s'organise comme sur la figure 2 ou comme sur la figure 3.

Dans le phénomène représenté par la figure 3, beaucoup d'objets ont un score faible dans les deux variables, beaucoup ont un score élevé dans les deux variables, très peu sont caractérisés par un couple de scores de valeur moyenne. Ce fait, dont l'intérêt pratique peut être fondamental, ne peut pas nous apparaître par toute méthode axée sur les variables. Plus encore, le calcul des corrélations, par exemple, peut nous conduire à des conclusions erronées dans le détail, (quoique valables en gros) si nous ignorons la structure typologique de la population. Par exemple, deux répartitions à deux variables aussi différentes que les répartitions des figures 2 et 4 peuvent donner le même coefficient de corrélation entre x et y .

Dans le cas de la figure 4, il existe dans la population deux types d'objets, fortement groupés et qui réagissent de manière opposée à l'égard des variables x et y . Dans le cas de la figure 2, les objets paraissent se comporter à peu près au hasard par rapport aux deux variables. De toute évidence, il s'agit là de deux phénomènes complètement différents, que rien ne nous permettra de déceler, et de distinguer, aussi longtemps que nous nous bornerons à une étude des variables. D'une manière générale, les corrélations caractérisant le nuage d'ensemble sont, ou peuvent

être, sans rapport avec la structure interne du nuage et sa décomposition possible en groupes plus ou moins nettement distincts. Ce fait fondamental doit être rappelé, car il est souvent méconnu, même de statisticiens avertis.

Dans le tableau présenté plus haut, il correspond à la situation opposée de l'analyse typologique et des équations de régression ou, plus généralement, des corrélations canoniques.

On peut remarquer en passant que le phénomène "corrélations", s'il n'a certainement rien à voir avec le phénomène "existence ou non existence de types" a vraisemblablement quelque chose à voir avec la notion de "significativité" de l'existence d'un type. Par exemple, pour reprendre nos comparaisons concrètes, une population d'objets à trois variables fortement corrélées entre elles donne un nuage en forme de boudin. Il peut ou non y avoir des grumeaux dans ce boudin, tout comme dans la masse sphérique que l'on aurait si les variables n'étaient pas corrélées. Cependant, il doit être moins surprenant (significativité) de trouver des grumeaux dans le boudin que dans la sphère. Mais laissons ce sujet, où l'étude mathématique en cours est très susceptible de nous démontrer que cette "vraisemblance" ne correspond à rien.

EFFORTS ANTÉRIEURS VERS LA DÉCOUVERTE DES TYPES.

L'originalité de la méthode apparaîtra mieux si on la compare aux procédés proposés jusqu'ici pour la découverte des types. On peut écarter dès l'abord toutes les recherches qui portaient de types **a priori** (comme ce fut le cas de beaucoup de typologies du caractère). Il est bien évident que le fait que des individus sélectionnés comme introvertis ou extravertis, par exemple, réagissent différemment à certains tests n'est nullement la preuve de l'existence du type introverti ou extraverti. La réalité d'un type implique beaucoup plus que des différences individuelles sur une dimension.

On peut écarter également la découverte intuitive des types, propre à l'anthropologie ou au botaniste, car ces méthodes ne sont plus applicables dans des domaines où les types sont moins clairement marqués.

Par contre, une approche intéressante avait été fournie par M. DELAPORTE (Biotypologie, Juillet 1946). Sa méthode consistait à projeter le nuage à n dimensions sur les $\frac{n(n-1)}{2}$ plans passant par les combinaisons des n axes deux à deux. Il put ainsi mettre en évidence l'existence de différents types de fonte, dans des échantillons caractérisés sur 18 variables. Mais cette méthode ne va pas directement au fait essentiel du groupement. Elle repose uniquement sur le fait que les groupes ont quelque chance de se différencier sur certaines dimensions. Il est très facile d'imaginer des genres de groupement que rien ne trahirait sur des projections en deux dimensions et que cette méthode serait incapable de déceler.

D'autres méthodes ont été proposées, sous la dénomination d'analyse en "clusters". Notre ami R.B. CATTELL en propose quatre dans un article intitulé : "A note on correlation clusters and cluster search methods" (Psychometrika, 1944, 9, 169-184). La première revient à la comparaison deux à deux des profils correspondant à chaque observation et au regroupement des observations qui ont toutes un minimum de similarité entre elles. Elle est évidemment impraticable dès que le nombre d'observations dépasse quelques dizaines.

Il en est de même de la seconde qui nécessite le réarrangement de la table des corrélations entre profils, pour faire apparaître des groupes homogènes le long de la diagonale. La troisième est la méthode de Tryon, qui corrèle les colonnes de corrélations interprofils elles-mêmes. Bien que parfois intéressante, cette méthode ne simplifie en rien le travail de recherche des groupes, puisqu'elle doit partir d'une table de corrélations de second ordre, aussi importante que la table des corrélations originales.

La Q-technique de Stevenson part de la même table de corrélations entre profils, mais la traite en analyse factorielle. Il y a évidemment quelques chances pour que les facteurs passent au centre des clusters, définissant ainsi des types. Mais il peut en être autrement et, de toute manière, on ne saurait découvrir plus

de types qu'on ne découvrira de facteurs, quel que soit le nombre de types effectivement présents dans le nuage de points.

Les différentes méthodes qui partent de la table de corrélations entre profils ont d'ailleurs en commun deux graves inconvénients inhérents à ce point de départ même. Dans la table entre profils, les lignes et les colonnes représentent les objets et les coefficients représentent les corrélations entre deux objets sur l'ensemble des scores. Elle nécessite donc le calcul de $\frac{n(n-1)}{2}$ coefficients de corrélations, n étant le nombre d'objets. Or la recherche de types impose un nombre d'objets suffisant pour que l'échantillon soit représentatif. Un échantillon d'une centaine peut être insuffisant, et il correspond déjà au calcul de 4.950 coefficients... D'autre part, pour que les coefficients de corrélation soient significatifs, il faut que l'on dispose d'un nombre élevé de variables, de l'ordre de plusieurs dizaines au minimum. Or les problèmes concrets ont leurs variables propres et il est tout aussi difficile d'en multiplier le nombre que de le réduire.

La dernière méthode proposée par CATTELL, celle de la délimitation approximative des groupes, est sans doute la plus rapide, mais comporte encore la nécessité de comparer chaque observation à toutes les autres, ce qui en limite irrémédiablement la valeur pratique.

Jamais à notre connaissance une méthode pratique n'a été proposée pour attaquer directement le problème typologique.

Toutes les recherches de types ont toujours cherché à mettre en évidence une dimension privilégiée sur laquelle se différenciaient les groupes. Seule la méthode proposée (ou une autre ayant les mêmes caractéristiques) peut permettre de typer directement des milliers d'observations sur plus d'une dizaine de dimensions.

POSITION MÉTHODOLOGIQUE DE L'ANALYSE DES GROUPES.

Nous pensons très réellement que cette méthode est de nature à déclencher une manière de petite révolution en statistique appliquée. Cela ne se fera pas très vite, car même dans les domaines essentiellement évolutifs, et la statistique appliquée en est un, l'homme a une tendance bien affirmée à considérer la nouveauté d'hier comme classique et la nouveauté de demain comme dangereuse. En fait, la mise en évidence de types d'objets, c'est-à-dire de zones de condensation dans le nuage d'ensemble est un progrès de toute première importance dans la compréhension du phénomène à l'étude. Elle permet en effet de conclure à l'existence, entre les objets relevant d'un même type, de liens de parenté profonde **qui ne sont convenablement décrits par aucune des variables utilisées.** D'où résulte la nécessité et l'intérêt d'une investigation destinée à trouver la nouvelle variable convenable, c'est-à-dire à comprendre ce qui différencie véritablement deux types l'un de l'autre. Il y a les plus grandes chances pour qu'un progrès dans la voie de cette compréhension accroisse considérablement l'intérêt théorique et pratique des résultats d'ensemble de la recherche. Autant d'exemples que l'on voudra viennent à l'esprit, dans les domaines les plus divers.

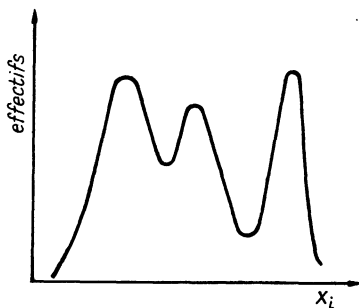


Fig. 5

S'il se trouve que les objets étudiés se différencient fondamentalement d'une certaine façon et que le chercheur ait eu la chance, ou le flair, de prendre la caractéristique correspondante parmi ses variables d'étude, le phénomène lui apparaîtra immédiatement car l'une des projections du nuage aura la forme nettement multi-modale (fig. 5) et il ne lui restera plus qu'à reprendre l'ensemble de son étude en séparant les différentes populations. S'il n'a pas eu cette chance, le sous-nuage constituant le type sera orienté de manière quelconque par rapport au faisceau d'axes et il n'a aucune chance de le voir autrement qu'au terme d'une recherche systématique.

Sur le plan théorique, la recherche des groupes à variables multiples représente donc un effort pour atteindre, si l'on peut dire, l'essence des phénomènes. Le choix des variables a toujours quelque chose d'accidentel. En manifestant une parenté de comportement à l'égard des variables utilisées, les objets appartenant à un même groupe suggèrent quelque chose comme une identité de nature. Et, bien sûr, cette identité de nature n'est peut-être que groupement sur une variable provisoirement inconnue, mais cela n'enlève rien à son intérêt, au contraire.

Sur le plan pratique, il est tout à fait certain que les jugements formulés sur la simple connaissance des aspects mesurés par les variables risquent de paraître fort sommaires, si on les reconsidère à la lumière des résultats d'une analyse des condensations. Dans les mêmes conditions, les décisions faisant suite à ces jugements, notamment en statistique industrielle ou commerciale, risquent de paraître passablement aventurées. Par exemple, (voir fig. 2 et 4) on peut décider d'agir sur une variable dans le but d'améliorer les scores dans une deuxième variable corrélée avec la première, sans se douter que tous les membres de l'un des deux types vont réagir exactement à l'inverse des membres de l'autre, et de la population d'ensemble qui a donné lieu à la corrélation utilisée. Plus généralement, l'existence de types ayant été constatée, tant que le chercheur ne connaît pas l'aspect qui les caractérise, il n'a aucun espoir d'agir sur eux de manière rationnelle. A plus forte raison s'il ignore leur existence.

Il y a longtemps que la statistique, sur le versant "variables", s'efforce de juger, de prévoir ou de décider, toutes variables prises en considération simultanément (pronostics ou jugements par indices pondérés et corrélations multiples; analyses de la variance et de la covariance, etc.) Mais l'existence non-décélée de types dans une population donnée prouve précisément que ces procédés ont été insuffisants, car ils ont laissé dans l'ombre le fait réel, peut-être fondamental, qu'est cette existence même.

Différents cas sont d'ailleurs possibles. Il peut arriver (sans qu'on le sache, bien entendu, au départ) que la totalité du phénomène, y compris l'existence des types, soit descriptible intégralement dans la dimensionnalité utilisée, c'est-à-dire avec les variables retenues. L'existence de types implique alors la probabilité de quelque effet inconnu, de nature discontinue, dans le jeu combiné de plusieurs de ces variables. Ce cas est probablement assez fréquent lorsque plusieurs variables correspondent à des caractéristiques **causales**. Par exemple, certaines valeurs de variables décrivant la manière dont a été réglée une opération de chimie peuvent aboutir, dans la réalité, à une certaine variante du produit fini, - d'autres valeurs à une autre variante du même produit. Une fois ces deux types identifiés, il sera vraisemblablement possible de définir la projection selon laquelle ils se distinguent tout à fait bien l'un de l'autre, c'est-à-dire la combinaison des anciennes variables qui décrit parfaitement la dualité constatée, en même temps qu'elle en suggère aux techniciens une explication correcte.

Dans d'autres cas, notamment si plusieurs variables sont **descriptives** (et non causales), l'existence de types entraîne une forte présomption pour qu'il existe **ailleurs** une nouvelle variable, non encore identifiée et qui conditionne les différences entre les deux types. Aucune étude statistique conduite sur le versant "variables" ne saurait donner une telle indication, sauf bien entendu si elle a abouti, grâce à une projection ou à une factorisation "heureuse", à la découverte de types. De même aucune combinaison des variables utilisées ne saurait se substituer à cette variable inconnue, car il s'agit ici de changer la dimensionnalité de l'étude. Par exemple, si la dualité constatée dans le produit fabriqué se trouve dépendre au premier chef de la provenance d'une des matières premières, et que cette caractéristique ne figure pas dans les variables d'étude, il n'y a aucune chance pour qu'elle apparaisse par une combinaison de ces variables. Certes, l'analyse typologique n'entraîne pas automatiquement l'identification de cette nouvelle variable, mais elle en suggère l'existence et donne le moyen de la rechercher, **par une revue systématique de tout ce que les objets appartenant à un de ces types peuvent avoir de commun, que n'aient pas les objets d'un autre type.**

Il convient de signaler, en terminant, l'erreur qui consisterait à conclure que les résultats d'une étude des condensations sont de nature à déconsidérer les résultats d'une étude des variables effectuée sur le même matériel. Les deux lignes de recherche sont inoffensives l'une pour l'autre. C'est seulement au stade

des conclusions, sur le plan de la connaissance ou sur le plan de l'action, que le chercheur risque d'être victime de celle des deux qu'il aura négligée ...

Un dernier aspect, à la fois théorique et pratique, de la démarche typologique mérite d'être signalé. L'existence d'un groupe permet dans une certaine mesure de raisonner sur une portion connue de l'échantillon **comme si** tous les objets appartenant au groupe étaient identiques. D'où la possibilité de substituer au groupe l'un de ses représentants centraux (ou quelques-uns) pour toute étude ou action qui serait impraticable sans cette simplification. Qu'il s'agisse de choses aussi différentes que de calculer le prix de revient d'une commande de fabrication représentative de son type, d'habiller des conscrits ou d'essayer une drogue sur un malade représentatif d'un syndrome, il est certain qu'il y a là une perspective méthodologique, à la fois théorique et pratique, d'un extrême intérêt.

CONCLUSIONS - QUELQUES CAS CONCRETS.

Le champ ouvert à ce mode de recherche est en recouvrement à peu près complet avec l'ensemble du domaine de la statistique appliquée. Il cesse de convenir dans les seuls cas où le phénomène apparaît comme une covariation continue de variables, et où la notion d'objet ou de fait scorable n'a pas de sens. En biologie, normale ou pathologique, en génétique, les lignes d'application sont évidentes. Il en est de même en économie, en démographie, et plus généralement pour toutes les sciences de l'homme.

Quand l'être vivant est concerné, aucune recherche ne peut prétendre maîtriser plus qu'une proportion infime des variables réellement en jeu. La tendance générale est alors - d'innombrables études en témoignent - d'essayer de rattacher l'individu, le cas, l'observation, le processus ... à des "types", seul moyen de mettre de l'ordre dans une multiplicité inabordable par d'autres voies. Et certes, cette ligne est justifiée, mais à la condition que le type résulte d'une élaboration statistique et non d'une définition a priori ou d'une synthèse intuitive, comme c'est en général le cas.

En psychologie appliquée, l'analyse des groupes à variables multiples est de nature à faire progresser sensiblement la technique, spécialement dans les études de postes et dans l'exploitation des profils d'aptitudes. La même méthode permettra de donner aux études de Job Rating la cohérence qui leur fait assez souvent défaut.

En statistique industrielle, la plupart des études de qualité concernant les matières premières, ainsi que les produits semi-finis et finis relèvent à quelque degré de cette méthode qui ouvre en particulier d'intéressantes perspectives d'extension pour les techniques devenues classiques du Quality Control. Elle peut avoir un intérêt du tout premier ordre pour l'étude des opérations de fabrication, notamment dans le cas où celles-ci revêtent un caractère discontinu, et se laissent décrire en populations d'échantillons, comme c'est le cas de presque toute l'industrie chimique.

En statistique commerciale, technique dont on commence à deviner actuellement le très grand avenir, la méthode apporte une aide de première importance à l'étude des ordres-clients, des populations d'intermédiaires ou de consommateurs, des gammes d'articles du point de vue de leur comportement à l'égard du marché, etc ...

Il sera nécessaire de conduire un nombre important d'applications avant d'asseoir définitivement la méthodologie de cette technique et d'en apprécier l'exacte étendue.

A ce jour, nous avons pu procéder à quatre essais, intéressants par leur diversité même.

Une fabrication industrielle aboutit à un produit, de qualité variable, caractérisé sur six variables. L'analyse typologique permet une classification objective des différentes qualités du produit, clef possible de toute une série de recherches, "en amont" comme "en aval", sur la ligne qui part de la matière première et va jusqu'à l'utilisation du produit fini. Les objets sont ici des échantillons de produit, les variables les résultats des analyses.

Une autre usine livre un produit caractérisé par trois dimensions, qui sont spécifiées par le client. L'analyse aboutit à une normalisation de fait, définissant comment les commandes clients se groupent autour de formats que l'on peut prendre désormais comme standards. Bel exemple pour ceux que préoccupent les questions de normalisation. Cette fois, les objets sont les commandes clients, les variables les dimensions demandées.

Une autre usine conduit des opérations de fabrication qui se caractérisent par des incidents de différentes natures (par exemple : rebut de tel genre). L'analyse typologique classe ces opérations par familles analogues au point de vue de tous les incidents, ouvrant une ligne de recherche intéressante sur les techniques de fabrication considérées dans leur ensemble, inabordable par toute autre voie. Dans ce cas, les objets sont des opérations de fabrication, les variables des événements caractéristiques du succès ou de l'insuccès de ces opérations.

Dans le cadre de notre propre Département Recherches (en psychologie) les profils de quatre cents malades mentaux au test M.M.P.I. (7 variables principales) ont été traités par la même méthode. Dans la mesure (c'est une large mesure) où chaque variable est significative d'un trouble mental bien défini, c'est une véritable "nosographie objective" qui est constituée de cette manière.

Nous espérons avoir bientôt d'autres cas à présenter (plus en détail) aux lecteurs de cette Revue, mais nous ne saurions cette fois abuser davantage de son aimable hospitalité.