

S. RÉGNIER

Classification et analyse des expressions plastiques non-figuratives de malades mentaux

Mathématiques et sciences humaines, tome 82 (1983), p. 45-65

http://www.numdam.org/item?id=MSH_1983__82__45_0

© Centre d'analyse et de mathématiques sociales de l'EHESS, 1983, tous droits réservés.

L'accès aux archives de la revue « Mathématiques et sciences humaines » (<http://msh.revues.org/>) 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/>

Math. Sci. hum. (21^e année, n°82, 1983, p.45-65)

CLASSIFICATION ET ANALYSE DES EXPRESSIONS
PLASTIQUES NON-FIGURATIVES DE MALADES MENTAUX *

S. REGNIER

INTRODUCTION

En vue de mettre en lumière l'influence hypothétique du trouble psychique sur les expressions plastiques non-figuratives de malades mentaux, le C.I.D.E.P.** nous a fourni :

- a) Les descriptions formelles de 200 gouaches non-figuratives.
- b) Des classifications en 7 classes effectuées sans critère imposé par des examinateurs bénévoles : médecins, peintres, etc. sur ces mêmes 200 tableaux.
- c) Le diagnostic psychiatrique de l'auteur de chaque tableau.

Notre objectif principal était de détecter une influence du trouble psychique dans les structures des oeuvres. Au préalable, il semblait bon de comparer les descriptions formelles systématiques fournies par le C.I.D.E.P. avec le contenu perceptible des oeuvres tel qu'il pouvait ressortir des partitions b).

* Note 22 / Janvier 1966

** Centre International de Documentation concernant les Expressions Plastiques.
Fondation Singer-Polignac.

Nous avons étudié :

1/. D'abord la cohérence des 27 classifications b/ et extrait une sorte de partition moyenne définie comme un centre des moyennes distances entre les 27 points donnés.

2/. Ensuite, la distance ou corrélation (χ^2) entre cette "partition centrale" et les partitions "formelles" qui se trouvent définies par les critères de la description formelle a/ : présence ou absence d'une certaine couleur, d'un certain procédé, etc.

3/. Enfin, les corrélations entre ces partitions formelles et la partition définie par les diagnostics c/ : névroses, psychoses non-schizophréniques ou schizophrénie.

Ces dernières liaisons se révèlent moins nombreuses et moins fortes que les précédentes. Quoiqu'ils rendent bien compte du contenu perceptible des oeuvres tel qu'il ressort de b/ les critères formels a/ sont presque tous indépendants du contenu pathologique éventuel défini par c/.

4/. Un point de vue nouveau pour l'étude et la comparaison de partitions a été élaboré à partir du point 1/. L'ensemble $P(E)$ de partitions de E (fini) peut être représenté dans $R^{E \times E}$ par une partie du cube $(0,1)^{E \times E}$ (matrices de 0 et de 1).

Des techniques linéaires plus variées que les techniques du χ^2 ont pu alors être employées.

- . Pondération d'un système de partitions C.
- . Construction d'un centre de moyenne distance, d'abord dans le cube (centre de gravité) puis sur P (projection).
- . Calcul inverse par projection (moindres carrés) d'une pondération dont le centre de gravité soit le plus proche possible d'une partition donnée D.
- . Appliquée à un nouvel objet, cette pondération définit une mesure de ressemblance optima par rapport à la partition donnée D; d'où résulte une technique nouvelle d'analyse discriminante rattachant chaque nouvel objet à une classe de D.

I Perception : Etude intrinsèque des classifications subjectives

Chaque examinateur a fourni une partition des 200 tableaux en 7 classes et donné des titres à ces 7 classes. L'analyse de ces noms de classe révèlent des références plus souvent intellectuelles chez les peintres professionnels, plus souvent affectives ou perceptives chez les autres examinateurs, mais toujours très variables de l'un à l'autre.

On s'est facilement convaincu que les noms de classes résultent de la confrontation examinateur-tableau au fil même de la formation des classes (les 200 gouaches étaient présentées en même temps par diapositifs sur une table lumineuse). Ce sont des justifications a posteriori. L'examineur forme progressivement une relation d'équivalence entre tableaux, basée sur leurs ressemblances deux à deux, et donne à chaque classe une définition qui peut varier progressivement avec son contenu. On a voulu tester la cohérence de ces partitions et définir une sorte de moyenne des 27 partitions observées, abstraction faite des désignations des 7 classes.

Il est impossible d'additionner des partitions comme des nombres réels pour former $\text{moy}(x) = \frac{1}{n} \sum x_i$; mais on peut définir entre partitions des distances assez naturelles. Les relations d'équivalence associées à 2 partitions X, Y de E sont 2 parties X', Y' de E x E, et une distance D naturelle est donnée par le cardinal de la différence symétrique de X' et Y'. C'est :

$D^2 = \text{nombre de paires } i, j \in E \times E \text{ classées ensemble par X et pas par Y}$
ou par Y et pas par X.

Ici X et Y représentent 2 examinateurs et D^2 prend un sens très concret. C'est le nombre de concessions que doit faire X pour accepter la partition de Y, chaque concession consistant :

- soit à réunir 2 tableaux que X avait dissociés
- soit à séparer 2 tableaux que X avait associés.

Cette distance étant définie, on peut considérer un centre des moyennes distances des 27 partitions s_k

X tel que $\sum_{k=1}^{27} D^2(X, s_k)$ soit minimum

Ce problème peut avoir plusieurs solutions. Nous appellerons chacune une partition centrale.

C'est une partition sur laquelle les 27 examinateurs pourraient se mettre d'accord au prix d'un nombre minimum de concessions mutuelles.

C'est donc une bonne notion de partition moyenne. Sur le plan calculatoire, la distance D ci-dessus présente de grands avantages.

Chaque partition de

E = les tableaux, peut être représentée par une matrice s ($n \times n$) où :

$s_{ij} = 1$ si i et j sont dans la même classe.

0 sinon

L'ensemble $P(E)$ des partitions de E est ainsi représenté dans

$R^{n \times n}$ = ensemble des matrices $n \times n$, par une partie du cube

$(0,1)^{n \times n}$ = ensemble des matrices booléennes = matrices d'une relation quelconque.

La distance D^2 entre partitions est la restriction à $P(E)$ de la distance Euclidienne de $R^{n \times n}$: $D^2(X, Y) = \sum \sum (x_{ij} - y_{ij})^2$

Considérons alors les 27 matrices s^k $k = 1, 2 \dots 27$. leur centre de moyenne distance dans $R^{n \times n}$ est simplement le centre de gravité

$$m = \frac{1}{27} \sum_{k=1}^{27} s^k$$

et toute autre matrice X vérifie l'identité de Koenig

$$H(X) = \sum_{k=1}^{27} D^2(X, s_k) = 27 D^2(X, m) + H(m)$$

Ainsi $H(X)$ et $D^2(X, m)$ atteindront leurs valeurs minima sur le même ensemble de partitions centrales X .

La matrice m résume toute l'information utile pour ce problème. Nous l'avons appelée Jugement moyen. Chaque coordonnée

$$m_{ij} = \frac{1}{27} \sum_{k=1}^{27} s_{ij}^k$$

représente la proportion d'examineurs qui ont groupé ensemble les tableaux i et j . C'est donc une mesure naturelle de la ressemblance des 2 tableaux, sur le plan de la perception. Dans la littérature classificatoire, une telle matrice s'appelle en général une SIMILARITE. Les 27 distances $D^2(m, s^k)$ permette de classer les examineurs dans un ordre intéressant. Les peintres professionnels se sont révélés plus "excentriques" que les médecins : plus éloignés de m .

Le problème :

$$\begin{aligned} \text{Min } D^2(X, m) \\ X \in P \end{aligned}$$

se simplifie en remarquant que $P(E)$ et le cube $(0,1)^n \times n$ sont inscrits dans une sphère de centre H :

$$h_{ij} = 1/2 \quad \forall ij, \text{ et de rayon } n/4,$$

si bien que X doit maximiser le produit scalaire Euclidien :

$$H X \cdot H M = \sum (x_{ij} - \frac{1}{2}) (m_{ij} - \frac{1}{2})$$

ou encore la forme linéaire homogène :

$$L(X) = \sum_{i,j} x_{ij} t_{ij} \quad \left(\text{si } t_{ij} = m_{ij} - \frac{1}{2} \right)$$

Ce problème linéaire relève alors en principe de la programmation linéaire.

Un algorithme heuristique beaucoup moins lourd peut se construire en enchaînant des transferts de tableau d'une classe dans une autre de façon à augmenter $L(X)$ le plus possible. Pour plus de détail on pourra se reporter à (2).

L'algorithme des transferts permet naturellement de borner le nombre de classes. On a donc établi une partition centrale en 7 classes, meilleure que toute autre, en 7 classes ou moins.

A partir du jugement moyen m_{ij} , on peut considérer une relation

$$R : m_{ij} > S \quad (S = 1/2 \quad 2/3 \text{ par exemple})$$

en définissant des paires de tableaux "fortement" associés, relativement au seuil S .

Sa fermeture transitive est une relation d'équivalence $\bar{R}$. Deux objets i et j seront R - équivalents s'il existe une chaîne d'objets $i = i_1, i_2 \dots i_n = j$ le long de laquelle 2 objets consécutifs sont toujours fortement associés. On démontre que pour $S = 1/2$, la partition associée est moins fine que la partition centrale.

Test de cohésion.

Les 27 partitions peuvent être considérées comme 27 éléments aléatoires indépendants de même loi L sur $P(E)$ et L se décompose en un produit

$$L(s) = L_1(n_i) \times L_2(s / n_i)$$

ou $L_1(n_i)$ est la loi de répartition marginale du système d'effectifs des 7 classes

$$n_1, n_2 \qquad n_7$$

(une partition de l'entier $n = 200$)

et $L_2(s / n_i)$ est la loi conditionnelle de s quand le système d'effectif est fixé.

Les 27 points donnés s^k permettent en principe d'estimer la loi L . L'information contenue serait nulle si nous avions affaire à 27 examinateurs aveugles ou tricheurs, qui auraient classé les tableaux sans les regarder!

$L_1(n_i)$ pourrait être une loi quelconque, mais

$L_2(s / n_i)$ serait une loi invariante par permutation des objets

$$L_2(s / n_i) = \frac{n_1! \ n_2! \ \dots \ n_7!}{n!}$$

loi uniforme sur l'ensemble des partitions s d'effectifs n_i .

Cette formule exprime une hypothèse composite qu'il conviendrait de tester; indépendamment de la formulation physique un peu caricaturale qui nous y conduit.

La construction d'un test rigoureux serait très difficile, nous nous sommes contentés jusqu'ici d'examiner l'hypothèse simple H_0 où les :

$$\left\{ \begin{array}{l} a^k = \text{numéro de la classe attribué à l'objet } i \text{ par l'examineur } k, \\ \text{seraient des variables aléatoires uniforme indépendantes sur } \{1, 2 \dots 7\} \end{array} \right\}$$

$$L_1(n_1, n_2 \dots n_7) = n! / 7^n n_1! n_2! \dots n_7!$$

$s_{ij}^k = 0$, ou 1 avec la probabilité

$$E(s_{ij}^k) = \Pr \left[a_i^k = a_j^k \right] = \frac{1}{7} \quad \text{si } i \neq j$$

chaque m_{ij} aurait la même espérance mathématique et pour variance :

$$\frac{1}{27} - \frac{1}{7} \left(1 - \frac{1}{7} \right)$$

On a constaté que les $\frac{200 \times 199}{2} m_{ij}^2$ avaient une moyenne empirique de : 0.095 qui autorise à rejeter cette hypothèse H_0 très particulière.

II - Création et perception

Les 200 gouaches utilisées, comme les autres tableaux archivés, sont soumis par le C I D E P à une analyse formelle systématique très détaillée (voir annexe (1)). On en a retenu 52 items bivalents particulièrement pertinents dans le cas non figuratifs. Chaque tableau est décrit par la présence de certains items : couleur noire, trait cerné, surface entièrement couverte etc ...

A - Comparaison d'un Item avec une partition

Une première question intéressante était d'évaluer, pour chaque examinateur, ou pour la partition centrale X, dans quelle mesure un item particulier intervient dans la formation des 7 classes.

Un tableau croisé tel que le suivant :

Surface non limitée

classe n°	1	2	3	4	5	6	7	Total
item présent	9	14	8	4	30	3	5	73
item absent	14	11	16	32	5	28	21	127
total	23	25	24	36	35	31	26	200

$$\chi^2 = 63,83$$

indique la fréquence de l'item dans chacune des 7 classes. Lorsque la fréquence relative s'écarte beaucoup de la fréquence générale $73 / 200$, on peut présumer que cet item a joué un rôle dans la formation de la classe considérée : par exemple ici, cet item est fortement présent en classe 5 et absent en classe 6. Pour reconnaître si ces écarts sont significatifs dans leur ensemble, on peut former le chi 2 habituel

$$\chi^2 = n \sum_{ij} (n_{ij} - n_i n_j / n)^2 / n_i n_j$$

et le comparer aux tables du chi 2 avec $(7 - 1)(2 - 1)$ degrés de liberté. En effet cette distribution s'appliquera si pour chaque tableau, la probabilité d'appartenir à chacune des 7 classes est la même dans les 2 lignes, indépendamment de la présence ou l'absence de l'item considéré.

Cette statistique est de plus invariante par permutation des colonnes, et indépendante de la numérotation des 7 classes. La page montre tous les items classés par χ^2 décroissant, relativement à la partition centrale en 7 classes.

B - Comparaison globale

B 1 construction à partir des items formels d'une similarité moyenne admettant X pour partition centrale.

La procédure

$T : m \rightarrow X$

tel que $D^2(X, m)$ soit minimum, qui définit la partition centrale X n'est pas biunivoque.

Les points m qui conduisent au même point X forme un cône convexe $T'(X)$, ensemble infini de demi-droites issues de H = centre du cube dans $(0, 1)^{n \times n}$ de $R^{n \times n}$. Il est limité par les hyperplans médiateurs des couples de points X, Y , pour tout Y pris dans $P(E)$.

On s'est demandé si un point G de $T'(X)$ pouvait être une fonction simple des 52 partitions définies par les items formels, de façon à voir si, globalement, ces derniers fournissent la même information que les 27 partitions des examinateurs

Chaque item définit une partition de E en 2 classes :

$$C^k \in P(E)$$

et toute moyenne pondérée des points C^k sera un nouveau jugement moyen G analogue à m . Mais ici il n'y a aucune raison d'attribuer aux 52 items des poids égaux. (Alors que les 27 examinateurs étaient naturellement sur le même pied). Nous prendrons donc les 52 poids p_k comme des inconnues.

Sous la condition a) $\sum p_k = 1$, le point $G = \sum p_k C^k$ parcourt la variété affine A engendrée par les C^k dans $R^n \times n$ et seulement la fermeture convexe Q des C^k si $p_k \geq 0$ pour tout k .

Comme $T(G)$ ne dépend que de la demi droite HG on peut remplacer a) par :

$$b) : \sum p_k > 0$$

l'équation en G : $T(G) = X$, équivalente à :

$$G \quad T'(X)$$

ne sera possible que si A coupe le cône convexe $T'(X)$.

Dans une première étape du calcul, il est plus simple de laisser G parcourir toute la variété linéaire $A' \supset A$, engendrée par H et les C_k , en posant :

$$HG = \sum p_k H C_k \quad \text{sans restriction sur les poids.}$$

L'équation $T(G) = X$ définit alors un système d'inéquations concernant les poids p_k .

Dans la perspective de l'algorithme des transferts défini ci-dessus, $T(G)$ est une fonction multivoque définissant 1 ou plusieurs "maxima locaux".

Si X comporte p classes, non vides, X pris dans $T(G)$ implique $n(p-1)$ inéquations relatives au transfert de chacun des n objets dans une autre classe que la sienne

+ n inéquation de transfert dans une classe vide (pour former une $p+1^{\text{e}}$ classe de 1 objet)

+ $p \frac{(p-1)}{2}$ inéquations relatives aux réunions de classes.

Dans la perspective de la programmation linéaire rigoureuse on aura autant d'inéquations que le polyèdre engendré par $P(E)$ comporte de sommets adjacents à X .

Dans les 2 cas le nombre d'inéquations dépasse de beaucoup le nombre d'inconnues (52), et l'équation

C) $x \in T(G)$ risque d'être presque toujours impossible.

Pour obtenir quand même un résultat utile, il est naturel de remplacer

C) par le problème suivant :

Soient $a_i(G) \leq b_i$ les inégalités de C). On considère alors le programme linéaire

$$d) \quad \begin{cases} G \in A' \text{ (ou } A) \\ y_i^+ - y_i^- = a_i(G) - b_i \\ y_i^+ \text{ et } y_i^- \geq 0 \\ \text{Min } z = \sum y_i^+ \end{cases}$$

Ce critère ne laissera positives que les variables y_i^+ correspondant à des $a_i(G) > b_i$.

La solution va définir un point G tel que la somme des écarts $a_i(G) - b_i$ relatifs aux inégalités i non vérifiées soit minima.

Le critère sera nul lorsque le problème C) est possible. Dans le cas contraire, il est bon que z ait un sens économique utile. Pour cela, dans la perspective des transferts on impose la condition normalisante :

$$L(X) = H G \cdot X = 1 \quad \text{c'est-à-dire}$$

$$\sum p_k (H C^k \cdot X) = 1$$

De cette façon les divers écarts y_i^+ ont une signification analogue.

Une condition de cette espèce est d'ailleurs nécessaire pour exclure la solution triviale de d) :

$$\begin{aligned} p_k &= 0 \quad \forall \quad k. & (G = I) \\ z &= 0 \end{aligned}$$

qui correspond au fait que pour $G = H$ $T(G) = P(E)$.

Toutes les partitions sont centrales, à la même distance $n^2/4$ de H .

B.2 - Construction d'une similarité moyenne la plus voisine d'un point donné.

Les calculs ci-dessus, assez lourds, n'ont pas encore été programmés. Une procédure plus simple consiste à chercher, dans A ou A' , le point G le plus voisin de la partition X donnée. Cette idée de voisinage mérite au départ une analyse. La procédure $T : G \rightarrow T(G) = X$ réalise le minimum de $D^2(X, G)$ et le maximum de $H X \cdot H G$ par rapport à X , et ne dépend que de la demi droite $H G$.

Par rapport à cette procédure : $\alpha = (\overrightarrow{H X}, \overrightarrow{H G})$ définit une bonne distance classificatoire entre les points X et G , définis pour tout X et G dans $\mathbb{R}^2$.

Nous appelons pondération optima le système de poids p_k qui définit :

$$\text{tel que } \begin{cases} H G = \sum p_k H G^k \\ \cos(H X, H G) \text{ soit maximum :} \end{cases}$$

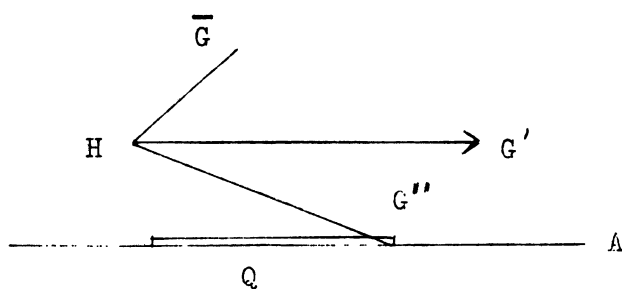
$$\cos(H X, H G) \geq \cos(H X, H Y) \text{ pour tout } Y \in A'$$

La solution quand $G \in A'$, sans restriction sur les poids, est donnée par

$$\overline{G} = \text{projection de } X \text{ sur } A'$$

et se calcule par la méthode des moindres carrés.

Quand on impose $G \in A$, soit $\sum p_k = 1$, la solution G' se déduit de $\overline{G}$ par des considérations géométriques simples : les 2 demi droites $H \overline{G}$ et $H G'$ sont identiques si la somme des poids de $\overline{G}$ est positive. Dans le cas contraire $H \overline{G}$ est parallèle à la projection de $H \overline{G}$ sur A .



Par contre la détermination d'un optimum G'' dans le polyèdre convexe Q engendré par les C_k relèverait de la programmation linéaire, sur une sphère $(H, 1)$, et serait sensiblement plus compliquée, excepté lorsque la demi-droite $H \overline{G}$ traverse Q .

Résultats

Au contraire de la précédente, cette construction s'applique même si X est un point quelconque, et non pas une partition. On l'a donc appliqué à la fois à :

X = partition centrale

m = jugement moyen

les résultats sont :

	$\cos^2 \alpha$	Distance $D^2(X, G)$	Somme des poids
pour X	0.514	4853	- 0.39
pour m	0.8455	8964	- 0.43

$$\text{Distance } D^2(G, G') = 35,02$$

G est ici la projection de X ou m sur A' (engendré par H et les C^k) Les distances $D^2(X, G)$ doivent être comparées à leur borne supérieure :

$D^2 = n^2 = 40.000$ pour la diagonale du cube, elles sont donc petites, et la distance 35 entre les 2 points G, G' , projection de X et m , est très petite. Ceci dit, nous ne savons pas interpréter statistiquement tous ces résultats. La question reste ouverte. On peut noter que si $H X$ (ou $H m$) était distribué uniformément en direction dans l'espace R^m des matrices symétriques, avec

$$m = \frac{n(n-1)}{2}, \cot^2 \alpha : \frac{52}{m-52} \text{ serait un } F \text{ de Snédécour avec :}$$

52 et $m - 52$ degrés de liberté, et $\cos \alpha$ serait distribué comme un coefficient de corrélation multiple.

Ici $n = 200$, $m = 19.900$, ce point de vue attribue au $\cos \alpha$ des seuils de signification excessivement petits.

Une hypothèse nulle un peu plus réaliste serait :

H_0 : Les s_k sont des variables aléatoires indépendantes de loi uniforme sur $P(E)$

On établit facilement que la moyenne de cette loi, dans $R^{n \times n}$,

$$\text{est une matrice } X \text{ où } X_{ii} = 1 \quad X_{ij} = C_n = \frac{P_n - 1}{P_n}, \text{ en notant } P_n$$

le cardinal de $P(E)$ = nombre de partitions de n objets

$(P_n = \sum_{m=1}^n P_n^m, \text{ où } P_n^m \text{ désigne le nombre de partitions en } m \text{ classes non vides}$
 $= \text{nombre de Stirling de } 2^{\text{è}} \text{ espèce})$

La variance de cette loi peut être définie par une matrice carrée V à m lignes et colonnes, où :

$$\left\{ \begin{array}{l} \text{pour } i < j \text{ et } k < 1 \quad (2 \text{ paires d'objets}) \\ V_{ij}^{kl} = C_n (C_{n-1} - C_n) \text{ sauf sur la diagonale} \\ V_{ij}^{ij} = 0 \end{array} \right.$$

Alors la moyenne m de k ($=27$) partitions s_k est répartie à peu près normalement avec la même moyenne $\bar{X}$ et pour variance $V / \sqrt{k}$.

On pourrait alors préciser la loi de : α = angle de H m avec le sous espace A' de dimension 52, (non aléatoire).

II - Création et trouble psychique

Sur une partie des 200 tableaux, le C I D E P nous a fourni une partition D en 3 diagnostics :

N = Névrose	(51 auteurs)
P = Psychose non schizophrénique	(27 auteurs)
S = Psychose schizophrénique	(100 auteurs)

Sur cet échantillon restreint, de 178 oeuvres, on peut refaire les mêmes études que ci-dessus :

A - Comparaison d'un item formel avec D

Les χ^2 ont été calculés. Deux seulement sont significatifs, et les tableaux croisés correspondant indiquent que les couleurs "noir" et "gris" étaient employées le plus par les névrosés, le moins dans le groupe P. - voir p. 64.

B - Comparaison globale

On a également projeté D sur la variété affine A', et obtenu :

$\cos^2 \alpha$	$D^2 (D, G)$	somme des poids
0.049	7532,5	- 0,131

Ces résultats concordent avec ceux de A : le système d'items formels recouvre très mal la partition D, beaucoup moins que la partition X centrale des examinateurs, et que leur jugement moyen.

à $(3 - 1)(7 - 1) = 12$ degrés de liberté, significatif.

H D, A' forme presque un angle droit, alors que

$$A', H m = 23^\circ, 1$$

$$A', H X = 44^\circ, 1$$

H D est plus "loin" de A' que H X. Il n'est donc pas surprenant de constater que les partitions X et D sont très indépendantes.

Le χ^2 de leur tableau croisé vaut 6,958, avec $(7 - 1)(3 - 1)$ degrés de liberté.

C - Analyse discriminante d'un nouveau tableau

A partir d'un système de poids p_k considéré comme optimum, la description formelle d'un nouveau tableau t définit 178 (ou 200) mesures de ressemblance entre t et les tableaux $i = 1, 2 \dots 178$ (ou 200) déjà utilisés

$$s_{it} = \sum_{k=1}^{52} p_k^* c_{it}^k$$

C_{it}^k étant la variable indicatrice = 1 si i et t sont dans la même classe pour l'item k

$$= 0 \quad \text{sinon}$$

On peut en déduire les "attractions" de t vers les diverses classes d'une partition quelconque D (ou X)

$$A_t^c = \sum_{i \in c} s_{it} - \frac{1}{2}$$

Ces attractions jouent un rôle naturel dans l'algorithme des transferts : si la partition D (ou X) est centrale relativement à la similarité

$$G = \sum p_k H G^k,$$

pour les n premiers objets, l'extension X' de X aux $n + 1$ objets, obtenue en ajoutant 1 classe contenant t tout seul, n'est plus centrale en général.

Le transfert le plus avantageux de l'objet t est défini par le max. de Λ_t^C et le gain correspondant sur la forme linéaire

$$L(Y) = \sum X_{ij} s_{ij} - \frac{1}{2}$$

est précisément

D'une façon purement heuristique, nous sommes conduits à poser la règle suivante :

$$R \quad \begin{cases} \text{Si max. } A_t^G = 0 \quad \text{le nouveau tableau } t \text{ est "unique en son genre"} \\ \text{Sinon, la classe } C(t) \text{ correspondante est celle ou il s'int\grave{e}gre le mieux.} \end{cases}$$

Nous avons appliqué cette procédure

a) A la partition D avec les poids p_k donnant G le plus voisin de D

$$b) \quad \begin{array}{ccccccc} & & X & & p_1 & & G \\ & & & & | & & \\ & & & & \vdots & & \\ & & & & x & & \end{array}$$

avec 50 nouveaux tableaux t. Le premier cas se prête à un contrôle : la

procédure R attribue à chaque tableau t une sorte de diagnostic automatique selon la classe d'intégration :

$$C(t) = N \quad P \text{ ou } S;$$

On peut comparer ce résultat au diagnostic réel, en espérant de fréquentes coïncidences. Les résultats ont été très mauvais, la règle R intègre presque tous les nouveaux tableaux dans la classe la plus nombreuse de $D : S = 100$ tableaux de schizophrènes, (contre 51 et 27) alors que leur répartition réelle est : 7 N, 6 P, 34 S. Cet échec n'était pas absolument surprenant, compte tenu des corrélations faibles entre la partition D et les items formels, pris individuellement ou globalement. De plus, on a constaté que G, projection de D sur A', n'appartient pas à T' (D). C'est-à-dire que la procédure R appliquée aux 178 tableaux de base ne respecterait pas non plus les diagnostics.

C'est en raison de ce résultat que nous avons défini la procédure B₁. Il est cependant probable que A' ne coupe pas le cône T' (D).

Dans le cas b) les résultats semblent à priori moins choquants : sur 50 nouveaux tableaux, la procédure R définit :

0	tableaux "d'un nouveau genre" et intègre	
0	" dans la classe 1 dont l'effectif de base est 23	
0	" 2	" 25
0	" 3	" 24
14	" 4	" 36
22	" 5	" 35
14	" 6	" 31
0	" 7	" 26
total	50	200

Mais il est difficile d'effectuer un contrôle précis de la valeur des résultats : La partition X est d'origine subjective, on ne sait pas a priori à quelle classe rattacher un nouveau tableau, à moins de demander à chaque examinateur son avis et A' en tirer à nouveau un jugement moyen etc...

Comparaison de la procédure R et des techniques classiques d'analyses discriminantes

Le problème abordé par ces techniques est le suivant :

1^{ère} forme : Etant donné g lois de probabilité L sur un même ensemble F estimer si une observation x dans F obéit à l'une ou l'autre. La solution du maximum de vraisemblance affectera à la population l où $p(x / L)$ est maximum, p désignant soit la probabilité ponctuelle quand L , soit la densité en x quand L , quand les g lois sont continues.

2^{ème} forme : Les g lois de répartitions ne sont pas exactement connues mais estimées à partir de g échantillons "de base" quelconque. A partir de ces estimations on peut opérer comme dans le cas 1. Ce problème est particulièrement bien étudié lorsque les g populations hypothétiques sont normales avec la même matrice de variance inconnue dans un espace $F = R^n$. Voir références [1] et [2]
Les surfaces discriminantes, séparant les régions où telle densité est plus grande que les autres, sont alors des hyperplans. Notre problème est de la 2^{ème} forme mais notre ensemble F est fini :

les descriptions formelles de tableaux sont des points de l'ensemble $F = (0, 1)^{52}$, de cardinal 2^{52} .

Chaque population de malades définit une loi de répartition inconnue sur cet ensemble fini. Et chacune des trois lois est estimée à l'aide d'un échantillon : 51, 27, ou 100 points. On peut évidemment considérer les lois empiriques définies par chaque échantillon comme la meilleure estimation de la loi réelle.

Mais ces lois empiriques ont des supports très petits dans $(0, 1)^{52}$.

Un nouveau tableau tombera en général en dehors des 3 supports, avec une description formelle distincte des 178 descriptions de base. Les 3 vraisemblances seront nulles et le critère du maximum de vraisemblance est inopérant.

Sans hypothèse supplémentaire, cette méthode ne classe donc que des tableaux "formellement identiques" à l'un de ceux qui définissent les 3 échantillons. Il semble d'autre part difficile de formuler une hypothèse supplémentaire efficace et raisonnable.

a) On peut penser à l'hypothèse de normalité :

$(0,1)^{52}$ peut se représenter comme un cube de R^{52}

Il est évidemment possible, mais très déformant, de traiter les 3 distributions sur ce cube comme des lois normales.

b) L'hypothèse d'indépendance

Pour chaque population, les distributions de 2 items pourraient être indépendantes. Les lois ne dépendraient plus que des 52 paramètres :

p_k = probabilité de présence de l'item k. Leur estimation sera différente de 0 et 1 pour presque tous les items disons m sur 52, et le support de la loi occupe alors 2^m points sur les 2^{52} . Le critère du maximum de vraisemblance est alors opérant.

Mais cette hypothèse est certainement erronée dans le cas présent : dans une même population pour certaine combinaison d'items, tels que noir - rouge, la probabilité conditionnelle de présence de l'un va être affecté par la présence de l'autre.

De ce point de vue, l'hypothèse de normalité est beaucoup moins restrictive puisqu'elle introduit, en plus des paramètres (p_k) = point moyen, d'où résultent les variances $V_k = p_k (1 - p_k) = s_k^2$, les 52.51/2.

coefficients de corrélation $R_{kl} = \text{covariance}(k, l) / s_k s_l$

A la différence de ces deux procédures plus directes, la procédure R n'introduit aucune hypothèse arbitraire. C'est pourquoi elle nous semble d'un certain intérêt théorique. Il faudra étudier ses propriétés statistiques dans les modèles stochastiques les plus utiles :

1) Probabilité des erreurs de classement

2) Test de validité de la méthode :

L'analyse discriminante classique comporte un test indiquant si les g échantillons sont significativement différents. Dans la négative la procédure n'est plus valable. De même dans la technique R, on peut présumer que lorsque G est trop "loin" du cône T' (D) les items formels sont trop indépendants de la partition D, et que la procédure est a priori non valable.

CONCLUSIONS

L'étude intrinsèque des partitions fondées sur la perception ne jouait qu'un rôle préliminaire. Les deux comparaisons :

création - perception

création - trouble psychique

nous conduisent au résultat suivant :

Les critères sélectionnés, pris individuellement ou globalement, rendent beaucoup mieux compte des classifications d'origine subjective que de la partition selon les diagnostics psychiatriques.

Si ce résultat devait se confirmer cela pourrait signifier :

- que notre codage trop "formel" des tableaux est inapte à transmettre leur "contenu pathologique" éventuel. (Un expert pourrait expliquer que "certain signe caractéristique" de la schizophrénie n'a pas été codé)
- ou même que cette notion de contenu pathologique ne recouvre rien dans les oeuvres non-figuratives, et que la manière d'un peintre (interné) dépend très peu de son cas, beaucoup moins que de certains autres caractères socio-culturels.

On pourrait invoker le style de Van Gogh, plus marqué sans doute par ses origines flamandes que par sa maladie.

Ces extrapolations tentantes doivent être limitées du fait que les peintres de notre étude, sauf exception, ne peignent que sur l'incitation du milieu au cours de leur internement. Leur activité est plus comparable à celle des enfants dans les ateliers de peinture spontanée qu'à celle des peintres professionnels.

Pour terminer, nous rappellerons que ce travail suscite plusieurs réserves

- 1) La description clinique est sommaire, des analyses plus fines, par présence et absence de symptômes, pourraient "entrer en corrélation" avec les analyses formelles des oeuvres. Il serait bon d'ajouter une description sociologique du peintre.

- 2) L'échantillon de 200 tableaux est petit par rapport au nombre de paramètres à estimer dans cette étude.
- 3) De plus, mis à part la technique du χ^2 pour la comparaison de deux partitions, les procédures employées sont à notre connaissance nouvelle, et nous ignorons leurs seuils de signification dans les modèles stochastiques adéquats. Il serait particulièrement utile de les mettre à l'épreuve sur des problèmes de classification moins ... complexes.
-

Tri croisé : NOIR / Diagnostic (p. 12 A)

		Diagnostics			
$\chi^2 = 11.59$		N	P	S	Total
Couleur noire	présente	<u>37</u>	10	48	95
	absente	14	<u>17</u>	52	83
	total	51	27	100	178

Le χ^2 calculé est significatif. La couleur noire est donc plus employée par les névrosés, moins pour les psychoses non-schizophréniques. Ceux-ci l'emploie cependant plus que les autres dans leurs oeuvres figuratives, voir (5)

Références bibliographiques

1. Cl. WIART - Automatisation Documentaire et Expressions Plastiques
Projet d'une analyse formelle et sémantique des oeuvres picturales.
- L'Encéphale n°5, 1964, pp. 604 à 613 -
2. S. REGNIER - Sur quelques aspects mathématiques des problèmes de classification automatique
- I C C, Bulletin Vol.4 NO = 5 -
3. Cl. WIART - Expressions picturales et psycho-pathologie.
Essai d'analyse et d'automatique documentaires. Principes, méthode codification.
- DOIN Paris 1966 -
4. Cl. WIART - Expression picturale non-figurative. L'oeuvre, son créateur, son spectateur. (A partir de 200 gouaches de malades mentaux présentés à 27 "examineurs").
- A paraître -
5. J. UNAL - La couleur dans les productions picturales de malades mentaux (étude statistique)
- (Thèse déposée) clinique des maladies mentales et de l'Encéphale PARIS.
6. T.W. ANDERSON - An Introduction to Multivariate Statistical Analysis.
- (1958) WILEY -
7. M. COOLEY and R. LOHES - Multivariate Procedures for the Behavioral Sciences.
- (1962) WILEY -