

M. JAMBU

Sur l'interprétation mutuelle d'une classification hiérarchique et d'une analyse des correspondances

Revue de statistique appliquée, tome 24, n° 2 (1976), p. 45-73

http://www.numdam.org/item?id=RSA_1976__24_2_45_0

© Société française de statistique, 1976, 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/>

SUR L'INTERPRÉTATION MUTUELLE D'UNE CLASSIFICATION HIÉRARCHIQUE ET D'UNE ANALYSE DES CORRESPONDANCES (1)

M. JAMBU (*)



1 – Introduction.

Les praticiens de la classification automatique sont souvent déconcertés par les difficultés d'interprétation des représentations arborescentes des données et par l'instabilité des classes constituées par les méthodes de classification. Ces mêmes praticiens sont parfois déconcertés par la représentation spatiale des données que leur offre l'analyse des correspondances, dans la mesure où cette méthode ne s'appuie pas sur la détermination ou la constitution effective de classes. Au lieu d'opposer ces deux techniques, on propose ici de les utiliser simultanément, et de les assortir de quelques calculs et représentations supplémentaires qui, à défaut de tests de signification très difficiles à mettre en oeuvre en classification, aideront le taxinomiste tant dans l'interprétation de classes que dans l'interprétation des facteurs d'une analyse des correspondances. Pour illustrer ces calculs, on a choisi de présenter un exemple réel portant sur la comparaison internationale des budgets-temps.

2 – PRESENTATION DE L'EXEMPLE.

L'objet principal d'une étude portant sur les budgets-temps est de déterminer une typologie tant des activités exercées par la population que de la population observée, en ne tenant compte que du temps passé à exercer ces activités. Les tableaux de données illustrant un budget-temps se présentent alors ainsi.

• Soit I la population étudiée et J l'ensemble des activités exerçables par la population I (exerçables pendant un temps fixé à l'avance, par exemple, une semaine, un mois, ...). La case d'indice (i, j) du tableau des données $I \times J$ contient le nombre d'heures (ou de minutes) passées à exercer l'activité j de J par le sujet i de I pendant le temps fixé de l'enquête. L'exemple qui nous sert d'illustration est un cas particulier de ce cas général. L'ensemble I est constitué de 28 classes de sujets dont les variables de base sont le sexe, le pays, l'activité professionnelle, le mariage. L'ensemble J est constitué de

(*) Attaché de recherches au CNRS, laboratoire de statistique de l'Université Pierre et Marie Curie. (Paris VI)

(1) Article remis en janvier 1975, révisé en janvier 1976.

10 classes d'activités regroupées pour les besoins de la comparaison des budgets-temps ⁽¹⁾. La case d'indice (i, j) contient le nombre d'heures que la classe i des sujets de l'enquête a passé à exercer l'activité j de J pendant le temps fixé de l'enquête.

Composition de l'ensemble I	Composition de l'ensemble J
HAUS – hommes actifs dès U.S.A.	*** 1 – Travail professionnel
FAUS – femmes actives ” ”	
FNAU – ” non actives ”	*** 2 – Occupations dûes ou
HMUS – hommes mariés des ”	ou liées au travail
FMUS – femmes mariés ”	professionnel
HCUS – hommes célibataires des U.S.A.	(transports. . .)
HAWE – hommes actifs des pays de l'Ouest	*** 3 – Travail ménager
FAWE – femmes actives ” ” ”	
FNAW – ” non actives ” ”	*** 4 – Occupations liées aux
HMWE – hommes mariés des ” ”	enfants
FMWE – femmes mariées des ” ”	
HCWE – hommes célib. ” ” ”	*** 5 – Les courses ou
HAES – hommes actifs des pays de l'Est	emplettes ménagères
FAES – femmes actives ” ” ”	*** 6 – La toilette, les soins
FNAE – femmes non actives ” ”	personnels
HMES – hommes mariés des ” ”	*** 7 – Les repas
FMES – femmes mariées ” ” ”	
HCES – hommes célib. ” ” ”	*** 8 – Le sommeil
HAYO – hommes actifs de Yougoslavie	
FAYO – femmes actives ” ”	*** 9 – La télévision
FNAY – ” non actives ” ”	
HMYO – hommes mariés de Yougoslavie	** 10 – Les autres loisirs
FMYO – femmes mariées ” ”	
HCYO – homme célib. de ” ”	
FCUS – femmes célib. des U.S.A.	
FCWE – ” ” des pays de l'Ouest	
FCES – ” ” ” ” de l'Est	
FCYO – ” ” de Yougoslavie	

(1) Dans cette note, aucune critique ne sera formulée sur la façon dont on est parvenu à la construction du tableau des données. Le tableau des données présenté sert uniquement d'illustration aux calculs utiles à l'interprétation des données, non de justification du choix des ensembles I et J – Pour une étude approfondie d'une telle donnée il eût fallu que l'ensemble I soit la population observée sans regroupements "a priori" et que l'ensemble J décrive un ensemble d'activités plus large que celles utilisées ici.

TABLEAU DES DONNEES

	***1	***2	***3	***4	***5	***6	***7	***8	***9	**10
HAUS	610.	140.	60.	10.	120.	95.	115.	760.	175.	315.
FAUS	475.	90.	250.	30.	140.	120.	100.	775.	115.	305.
FNAU	10.	0.	495.	110.	170.	110.	130.	785.	160.	430.
HMUS	615.	141.	65.	10.	115.	90.	115.	765.	180.	305.
FMUS	179.	29.	421.	87.	161.	112.	119.	776.	143.	373.
HCUS	585.	115.	50.	0.	150.	105.	100.	760.	150.	385.
FCUS	482.	94.	196.	18.	141.	130.	96.	775.	132.	336.
HAWE	652.	100.	95.	7.	57.	85.	150.	807.	115.	330.
FAWE	510.	70.	307.	30.	80.	95.	142.	815.	87.	262.
FNAW	20.	7.	567.	87.	112.	90.	180.	842.	125.	367.
HMWE	655.	97.	97.	10.	52.	85.	152.	807.	122.	320.
FMWE	168.	22.	529.	69.	102.	83.	174.	825.	119.	392.
HCWE	642.	105.	72.	0.	62.	77.	140.	812.	100.	387.
FCWE	389.	34.	262.	14.	92.	97.	147.	848.	84.	392.
HAYO	650.	140.	120.	15.	85.	90.	105.	760.	70.	365.
FAYO	560.	105.	375.	45.	90.	90.	95.	745.	60.	235.
FNAY	10.	10.	710.	55.	145.	85.	130.	815.	60.	380.
HMYO	650.	145.	112.	15.	85.	90.	105.	760.	80.	357.
FMYO	260.	52.	576.	59.	116.	85.	117.	775.	65.	295.
HCYO	615.	125.	95.	0.	115.	90.	85.	760.	40.	475.
FCYO	433.	89.	318.	23.	112.	96.	102.	774.	45.	409.
HAES	650.	142.	122.	22.	76.	94.	100.	764.	96.	334.
FAES	578.	106.	338.	42.	106.	94.	92.	752.	64.	228.
FNAE	24.	8.	594.	72.	158.	92.	128.	840.	86.	398.
HMES	652.	133.	134.	22.	68.	94.	102.	762.	122.	310.
FMES	434.	77.	431.	60.	117.	88.	105.	770.	73.	229.
HCES	627.	148.	68.	0.	88.	92.	86.	770.	58.	463.
FCES	433.	86.	296.	21.	128.	102.	94.	758.	58.	379.

3 – POSITION DU PROBLEME

Pour analyser un tel tableau de données, nous utiliserons l'analyse des correspondances (qui permet de connaître les facteurs principaux associés à ce tableau) et la classification automatique sur les deux ensembles en correspondance (qui permet de constituer des classes de sujets et des classes d'activités). Il reste à enrichir les résultats obtenus par ces deux méthodes par différents calculs dont le but est de répondre aux questions suivantes :

- Quelles sont les activités et les classes de sujets qui contribuent le plus à la formation des axes factoriels ?
- Sur quels axes factoriels les classes obtenues par la classification automatique se séparent-elles ?
- Quelle contribution les axes factoriels apportent-ils à la formation de ces classes (classes de I et de J) ?
- Quels sont les éléments de J qui caractérisent le plus les classes obtenues sur I (et réciproquement) ?
- A quelle classe obtenue sur J correspond une classe obtenue sur I ?

La méthode de classification automatique employée est celle associée à l'analyse des correspondances (cf. § 5)

4 – L'ANALYSE DES CORRESPONDANCES. RAPPEL DES NOTATIONS ET DES RESULTATS PRINCIPAUX.

Soient I et J deux ensembles finis et $k_{ij}, i \in I, j \in J$ le tableau des données correspondant. (Dans l'ensemble présente, k_{ij} est le nombre d'heures passées par le "sujet" i de I à exercer l'activité j de J, pendant le temps de l'enquête).

$$k_i = \sum_{j \in J} k_{ij} \quad ; \quad k_j = \sum_{i \in I} k_{ij} \quad ; \quad k = \sum_{i \in I} k_i = \sum_{j \in J} k_j$$

$$f_{ij} = \frac{k_{ij}}{k} \quad \quad f_i^j = \frac{f_{ij}}{f_j} \quad \quad f_j^i = \frac{f_{ij}}{f_i}$$

avec $f_j \neq 0$ avec $f_i \neq 0$

$$N_J(I) = \left\{ \begin{array}{l} i \in I, j^e \text{ coordonnée de } i \text{ dans } R_J = \frac{f_{ij}}{f_i}, \\ \text{masse affectée au point } i = f_i \end{array} \right\} \subset R_J$$

$$N_I(J) = \left\{ \begin{array}{l} j \in J, i^e \text{ coordonnée de } j \text{ dans } R_I = \frac{f_{ij}}{f_j} \\ \text{masse affectée au point } j = f_j \end{array} \right\} \subset R_I$$

Profil du point i sur J : $f_j^i = \left\{ \frac{f_{ij}}{f_i}, j \in J \right\}$

Profil du point j sur I : $f_i^j = \left\{ \frac{f_{ij}}{f_j}, i \in I \right\}$

- Moments d'inertie des nuages $N_J(I)$ et $N_I(J)$

$$\begin{aligned} M^2(N_J(I)) &\stackrel{\text{def}}{=} \sum_{i \in I} m_i \|i - g\|^2 \\ &= \sum_{i \in I} f_i \|f_j^i - f_j\|_{f_j}^2 && \begin{array}{l} g \text{ centre de gravité du nuage} \\ N_J(I), m_i \text{ masse associée au} \\ \text{point } i \end{array} \\ &= \sum_{\substack{i \in I \\ j \in J}} (f_{ij} - f_i \cdot f_j)^2 / f_i \cdot f_j \end{aligned}$$

$$M^2(N_J(I)) = M^2(N_I(J)) = \|f_{IJ} - f_I \cdot f_J\|_{f_I, f_J}^2 = \text{Lien}_2(I, J) \quad (\text{cf. § 5})$$

- Distance distributionnelle entre i et i' de I.

$$d^2(i, i') = \|f_j^i - f_j^{i'}\|_{f_j}^2$$

- Facteurs et axes factoriels

soient

$$\{F_\alpha(i); i \in I, \alpha \in A\}, \quad \{G_\alpha(j); j \in J, \alpha \in A\}$$

les facteurs associés aux ensembles I et J

$$\text{Card } A = \inf(\text{Card } J - 1, \text{Card } I - 1)$$

les facteurs sont de moyenne nulle, de variance λ_α , et sont non corrélés deux à deux.

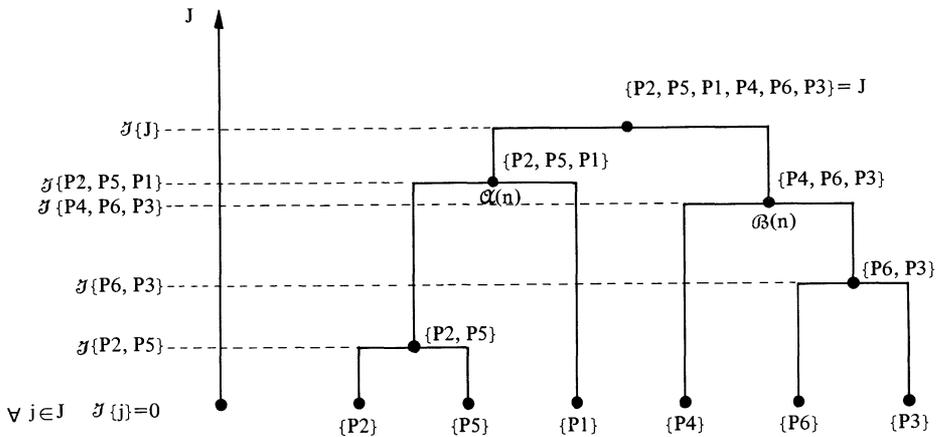
• Moments d'inertie des Nuages $N_J(I)$ et $N_I(J)$ dans la direction de l'axe et inertie totale.

$$\sum_{i \in I} f_i F_\alpha^2(i) = \sum_{j \in J} f_j G_\alpha^2(j) = \lambda_\alpha$$

$$\sum_{\alpha \in A} \lambda_\alpha = \sum_{i \in I} f_i \cdot \sum_{j \in J} \left(\frac{f_{ij}}{f_i} - f_j \right)^2 / f_j = \text{Lien}_2(I, J) \quad (\text{cf. } \S 5)$$

5 – LA CLASSIFICATION AUTOMATIQUE – RAPPEL DES NOTATIONS ET DES RESULTATS PRINCIPAUX.

5.1. Notations générales pour les classifications ascendantes hiérarchiques.



Représentation d'une hiérarchie de parties, totale indicée A(J) sur J

Soit J fini, $A \subset \mathcal{P}(J) - \emptyset$; A ordonné par la relation \subset on note :

Som.	(A)	l'ensemble des éléments maximaux de A
Ter	(A)	'' '' minimaux de A
Nod	(A)	'' '' éléments non minimaux de A
Pred	(a, A)	'' '' prédécesseurs de l'élément a dans A
Suc	(a, A)	'' '' successeurs '' '' a '' A
Suci	(a, A)	'' '' '' immédiats de a '' A

$\mathcal{J}(a)$ est l'indice de diamètre (ou d'agrégation, ou de stratification) de l'élément a de A .

$$\mathcal{J}(\{j\}) = 0 \quad \forall j \in J$$

5.2. La méthode de classification choisie.

Soit I et J deux ensembles finis et $\{k_{ij}\}$ le tableau des données correspondant. On calcule $\{f_{ij} = k_{ij} / k\}$ (cf. § 4).

On considère les nuages $N_J(I)$ et $N_I(J)$ définis au § précédent ;

$$\begin{aligned} \text{on a : } \text{Lien}_2(I, J) &= \sum_{i \in I, j \in J} (f_{ij} - f_i f_j)^2 / f_i \cdot f_j = \sum_{i \in I, j \in J} f_i f_j \left[\left(\frac{f_{ij}}{f_i f_j} \right)^2 - 1 \right] \\ &= \| f_{IJ} - f_I \cdot f_J \|^2_{f_I f_J} \\ &= M^2(N_I(J)) = M^2(N_J(I)) \end{aligned}$$

Soit Q une partition de I , et K une partition de J .

Soit $N_J(Q)$ le nuage associé à la partition Q , et $N_I(K)$ le nuage correspondant à la partition K , q une classe de Q , k une classe de K

$$N_J(Q) = \{q \in Q, j^e \text{ coordonnée de } q \text{ dans } R_J = \frac{f_{qj}}{f_q}, \text{ masse affectée au point } q = f_q\}$$

$$N_I(K) = \{k \in K, i^e \text{ coordonnée de } k \text{ dans } R_I = \frac{f_{ik}}{f_k}, \text{ masse affectée au point } k = f_k\}$$

$$N_J(Q) = \{f_q^j, f_q\} \subset R_J ; N_I(K) = \{f_i^k, f_k\} \subset R_I$$

On définit

$$\begin{aligned} \bullet \text{Lien}_2(Q, J) &= \| f_{Q \cdot J} - f_Q \cdot f_J \|^2_{f_Q \cdot f_J} \\ &= \sum_{\substack{q \in Q \\ j \in J}} (f_{qj} - f_q \cdot f_j)^2 / f_q \cdot f_j \quad \left(f_q = \sum_{j \in J} f_j \right) \\ &= M^2(N_J(Q)) = M^2(N_Q(J)) \end{aligned}$$

La quantité $\text{Lien}_2(Q, J)$ matérialise la dépendance globale entre deux partitions Q de I et J . Pour construire une hiérarchie indicée totale sur I , par exemple, on considère les quantités $\{\text{Lien}_2(Q, J)\}$. On procède par agrégation binaire sur $\mathfrak{R}(I)$. Pour construire une classification sur J , on définira de même $\text{Lien}_2(I, K)$ et on procédera par agrégation binaire sur $\mathfrak{R}(J)$

Prenons l'exemple de l'édification d'une classification sur I . Supposons l'étape de rang $(n-1)$ réalisée. Le sommet de la hiérarchie est la partition \mathcal{Q}_{n-1} de I . On cherche à agréger deux parties de I , (soit deux classes s et t

de Q_{n-1}) qui maximisera la quantité $\text{Lien}_2(Q_n, J)$ ($\text{Lien}_2(Q_{n-1}, J)$ est fixé pour cette étape ; Q_n est la partition constituée de Q_{n-1} , à laquelle on a retiré s et t et à laquelle on ajoute $s \cup t$).

On a donc les résultats suivants :

$$\begin{aligned}
 \underset{\text{indice de stratification}}{\mathcal{J}(n)} &= \mathcal{J}(s \cup t) = \text{Lien}_2(Q_{n-1}, J) - \text{Lien}_2(Q_n, J) \\
 &= M^2(N_J(Q_{n-1})) - M^2(N_J(Q_n)) \\
 &= \frac{f_t \cdot f_s}{f_t + f_s} \|f_J^s - f_J^t\|_{f_t}^2 \\
 &= \frac{f_t \cdot f_s}{f_t + f_s} \sum_{j \in J} \left(\frac{f_{sj}}{f_s} - \frac{f_{tj}}{f_t} \right)^2 / f_j \\
 &= M^2(s \cup t) - M^2(s) - M^2(t)
 \end{aligned}$$



De ces propriétés on obtient les propriétés suivantes, en sommant les indices de stratification des classes :

soient I et $A(I)$ la hiérarchie de parties sur I
 soient J et $B(J)$ " " " " " J

$$M^2(N_J(I)) = M^2(N_I(J)) = \sum_{n \in \text{Nod}(A(I))} \mathcal{J}(n) = \sum_{n \in \text{Nod}(B(J))} \mathcal{J}(n) = \text{Lien}_2(I, J)$$

- Taux d'inertie associé au Nœud n de la hiérarchie $A(I)$

$$\Gamma_A(n) \stackrel{\text{def}}{=} \frac{\mathcal{J}(n)}{\text{Lien}_2(I, J)}$$

- Taux d'inertie associé à une partition Q de I , construite sur la hiérarchie $A(I)$

$$\Gamma_I(Q) \stackrel{\text{def}}{=} \frac{M^2(N_J(Q))}{M^2(N_J(I))} = \frac{\text{Lien}_2(Q, J)}{\text{Lien}_2(I, J)} \leq 1$$

$$[\text{Lien}_2(Q, J) \leq \text{Lien}(I, J)]$$

Soit $A(Q)$ la hiérarchie indicée sur I , dont le sommet est la partition Q .
 On a alors :

$$\sum_{n \in \text{Nod}(A(Q))} \mathcal{J}(n) = \text{Lien}_2(I, J) - \text{Lien}_2(Q_n, J)$$

donc si $X = \sum_{n \in \text{Nod}(A(I))} \mathcal{J}(n)$ et $X' = \sum_{n \in \text{Nod}(A(Q))} \mathcal{J}(n)$

$$\text{Lien}_2(Q_n, J) = X - X'$$

donc
$$\Gamma_I(Q) = \frac{X - X'}{X'}$$

On pourra comparer les taux d'inertie associés aux facteurs et les taux d'inertie associés à une classification.

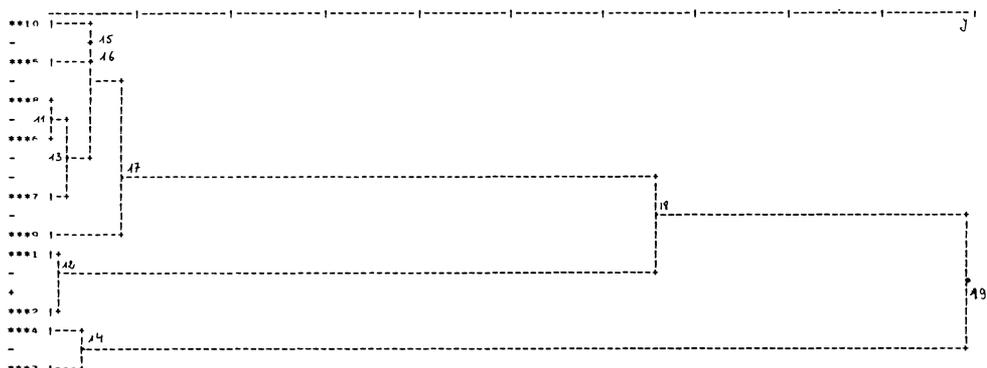
5.3. L'exemple.

Les tableaux suivants donnent la description des hiérarchies constituées sur I et sur J.

A chaque Nœud n est associé son indice $\mathcal{J}(n)$, son taux d'inertie $\Gamma(n)$, de successeurs immédiats $\mathcal{A}(n)$ et $\mathcal{B}(n)$, le nombre d'éléments de la classe $P(n)$ les éléments de J qui composent le nœud n . Les nœuds sont numérotés de $\text{Card } J + 1$ à $2 \cdot \text{Card } J - 1$ pour une hiérarchie binaire totale indiquée construire sur un ensemble de $\text{Card } J$ éléments.

5.3.1. Hiérarchie construite sur l'ensemble J de l'activités.

NŒUD	TAUX	INDICE	AINE	BENJAMIN	POIDS	DESCRIPTION DE LA HIÉRARCHIE
NŒ11	0.007	0.001	8	6	2	***8 ***6
NŒ12	0.007	0.001	1	2	2	***1 ***2
NŒ13	0.014	0.002	11	7	3	***8 ***6 ***7
NŒ14	0.021	0.003	4	3	2	***4 ***3
NŒ15	0.021	0.003	10	5	2	***10 ***5
NŒ16	0.028	0.004	15	13	5	***10 ***5 ***8 ***6 ***7
NŒ17	0.042	0.006	16	9	6	***10 ***5 ***8 ***6 ***7 ***9
NŒ18	0.350	0.050	17	12	8	***10 ***5 ***8 ***6 ***7 ***9 ***1 ***2
NŒ19	0.520	0.075	18	14	10	***10 ***5 ***8 ***6 ***7 ***9 ***1 ***2 ***4 ***3



*L'interprétation**. Le 1^{er} niveau sépare les activités traditionnellement réservées aux femmes qui restent au foyer (temps passé à faire les courses et à faire le ménage, temps passé à s'occuper des enfants) de toutes les autres activités. Pour les activités restantes, le 2nd niveau sépare les activités professionnelles habituellement exercées par les hommes des autres activités. Ces activités peuvent être exercées par des célibataires non incluses dans une structure familiale traditionnelle. Le 3^e niveau sépare la télévision des activités restantes. Ce qui signifie qu'elle constitue un point de différenciation pour les groupes de sujets retenus. Le 4^e niveau sépare, d'une part, certains types d'activités

(*) On laisse le soin au lecteur de confirmer ou d'infirmer l'interprétation que nous suggérons ici en reprenant tous les graphiques et calculs proposés dans cette note. Ceux-ci ont été intégralement reproduits à partir des "sorties" d'ordinateurs afin de permettre au lecteur de reprendre tous les éléments qui ont servi de base à l'interprétation.

externes (courses, loisirs) de certaines activités internes (sommeil, repas, toilette) et, d'autre part, une différence entre le travail ménager et l'occupation due aux enfants.

5.3.2. Hiérarchie construite sur l'ensemble I des groupes de sujets enquêtés.

L'interprétation. On observe deux niveaux principaux (cf. figure 4)

- le 1^{er} niveau sépare les femmes dites inactives ou mariées de toutes les autres catégories de population, quel que soit le pays où a été effectuée l'enquête. Le 1^{er} niveau est un niveau de structure sociale liée au ménage, à la famille. C'est la séparation des tâches dans les sociétés : les femmes mariées, ou ayant des enfants à élever à la maison d'un coté ; les hommes, les célibataires et les femmes actives.

- le 2nd niveau sépare les hommes quel que soit leur pays d'appartenance des femmes actives ou célibataires. On peut raisonnablement croire que ces deux groupes vivent professionnellement comme les hommes mais consacrent une partie de leur temps à des activités différentes en dehors de leur travail professionnel

- le 3^e niveau sépare les données en pays (N = 52, 53, 51) : à quelques exceptions près, les pays de l'Ouest des pays de l'Est. Cette séparation se fait à l'intérieur des classes constituées au niveau 1 et 2.

- le 4^e niveau sépare les femmes célibataires des femmes actives.

- le 5^e niveau sépare les hommes ou femmes des pays européens de l'Ouest des hommes ou femmes des Etats-Unis.

A la suite de ces deux interprétations, on souhaiterait pouvoir confirmer la validité des classes constituées et les expliquer. Par quelles activités sont caractérisées les classes obtenues par la hiérarchie sur I : c'est ce à quoi on essaye de répondre en illustrant les représentations spatiales ou arborescentes de calculs propres à affiner notre interprétation.

29 0.000 4 1 2 HMUS HAUS
30 0.000 11 8 2 HMWE HAME
31 0.000 18 15 2 HMYO HAYO
32 0.000 23 16 2 FAFS FAYO
33 0.000 21 28 2 FCYO FCFS
34 0.000 22 25 2 HAES HMFS
35 0.000 20 27 2 HCYO HCES
36 0.000 7 2 2 FCUS FAUS
37 0.000 13 30 3 HCWE HMWE HAWF
38 0.000 31 34 4 HMYO HAYO HAES HMFS
39 0.000 6 29 3 HCUS HMUS HAUS
40 0.001 17 24 2 FNAV FNAE
41 0.001 12 10 2 FWE FNAW
42 0.001 32 9 3 FAES FAYO FAME
43 0.001 14 33 3 FCWE FCYO FCES
44 0.001 5 3 2 FMUS FNAU
45 0.001 42 26 4 FAES FAYO FAME FMES
46 0.001 38 37 7 HMYO HAYO HAES HMES HCWE HMWE HAME
47 0.002 41 44 4 FMWE FNAW FMUS FNAU
48 0.002 36 43 5 FCUS FAUS FCWE FCYO FCES
49 0.002 19 40 3 FMYO FNAV FNAE
50 0.002 46 35 9 HMYO HAYO HAFS HMFS HCWE HMWE HAME HCYO HCES
51 0.003 49 47 7 FMYO FNAV FNAE FMWE FNAW FMUS FNAU
52 0.003 50 39 12 HMYO HAYO HAES HMES HCWE HMWE HAME HCYO HCES HCUS HMUS HAUS
53 0.004 48 45 9 FCUS FAUS FCWE FCYO FCES FAES FAYO FAME FMES
54 0.022 53 52 21 FCUS FAUS FCWE FCYO FCFS FAFS FAYO FAME FMES HMYO HAYO HAES HMES HCWE HMWE HAME HCYO HCES HCUS HMUS HAUS
55 0.095 51 54 28 FMYO FNAV FNAE FMWE FNAW FMUS FNAU FCUS FAUS FCWE FCYO FCES FAFS FAYO FAME FMES HMYO HAYO HAES HMES HCWE HMWE HAME HCYO HCES HCUS HMUS HAUS

6 – LES AIDES A L'INTERPRETATION.

On rappelle en premier lieu le calcul de contributions usités pour l'interprétation d'une analyse factorielle des correspondances, puis on donnera différents calculs qui mettent en relation les classes d'une hiérarchie sur J avec les classes d'une hiérarchie sur I, puis les classes de chaque hiérarchie aux axes factoriels extraits par l'analyse des correspondances.

6.1. Le calcul des contributions en analyse factorielle.

6.1.1. Les formules.

Définition de la distance d'un point $i \in N_J(I)$ au centre de gravité du Nuage $N_J(I)$

$$\rho^2(i) = \|f_J^i - f_J\|_{f_J}^2 = \sum_{j \in J} \left(\frac{f_{ij}}{f_i} - f_j \right)^2 / f_j$$

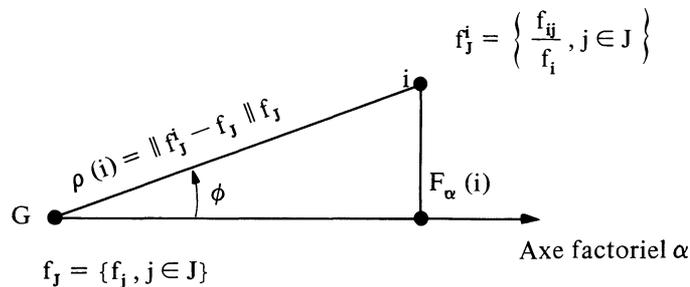
On a :

$$\bullet M^2(N_J(I)) = \sum_{\alpha \in A} \lambda_\alpha = \sum_{i \in I} f_i \cdot \rho^2(i) \quad [F1]$$

$$\lambda_\alpha = \sum_{i \in I} f_i \cdot F_\alpha^2(i) \quad [F2]$$

$$\rho^2(i) = \sum_{\alpha \in A} F_\alpha^2(i) \quad [F3]$$

- $f_i \cdot \rho^2(i)$ = contribution absolue du point i à $M^2(N_J(I))$ ou variance du point i
- $f_i F_\alpha^2(i)$ = contribution absolue du point i au moment d'inertie λ_α
- $f_i F_\alpha^2(i) / \lambda_\alpha$ = contribution relative du point i au moment d'inertie λ_α
- $F_\alpha^2(i) / \rho^2(i)$ = contribution du point i à cause de la distance au sens du χ^2 entre i et le centre du nuage $N_J(I)$ – soit $\rho^2(i)$



$$\cos^2 \theta = \frac{F_\alpha^2(i)}{\rho^2(i)} = \text{contribution relative du facteur } \alpha \text{ au point } i$$

On a alors : $\sum_{\alpha \in A} F_{\alpha}^2(i) / \rho^2(i) = 1$

On a les formules analogues pour les points j du Nuage N₁ (J)

6.1.2. Applications à l'exemple (figures 1, 2, 3).

$\lambda_1 = 0.118$; $\lambda_2 = 0.010$; $\lambda_3 = 0.007$; $\lambda_4 = 0.004$; $\lambda_5 = 0.002$; $\lambda_6 = 0.001$

I	1#F	COR	CTR	2#F	COR	CTR	3#F	COR	CTR	4#F	COR	CTR	5#F	COR	CTR	6#F	COR	CTR	7#F	COR	CTR	
1#AUS	227	774	23	130	119	44	99	69	531	3#F	0	111	-32	8	231	-50	14	101	7	0	101	
2#AUS	40	124	0	70	32	1	47	168	121	72	140	431	-44	172	471	-31	74	42	-16	23	52	
3#AUS	-50	844	102	222	124	141	34	3	6	45	8	261	84	18	152	-33	3	48	6	0	6	
4#AUS	223	764	23	124	106	50	114	10	70	28	4	71	-43	4	241	59	24	141	10	1	19	
5#AUS	-72	824	42	137	109	41	54	18	171	65	18	171	65	24	441	-41	10	71	0	0	6	
6#AUS	113	724	40	140	151	43	-7	0	0	71	34	421	-41	40	441	-2	0	0	25	5	114	
7#AUS	103	344	4	141	214	21	24	20	31	74	144	471	-44	144	441	-24	21	26	-35	42	225	
8#AWE	312	647	94	114	1	1	20	20	31	-121	130	121	-4	0	11	-14	2	91	0	0	6	
9#AWE	8	3	0	-01	491	27	65	108	24	-78	294	531	-31	47	221	-34	48	91	-4	2	71	
10#AWE	-402	444	104	44	14	23	31	2	51	-92	23	721	44	3	251	26	2	271	-13	1	44	
11#AWE	306	424	24	70	4	1	10	22	14	-130	144	1431	0	0	0	-17	3	131	4	0	31	
12#AWE	-434	424	40	44	4	4	4	0	0	-04	44	421	23	3	121	31	5	401	10	0	14	
13#AWE	344	477	44	44	10	4	-44	17	121	-104	94	941	4	0	0	-3	0	1	14	2	44	
14#AWE	-34	49	0	47	70	7	-89	244	431	-113	411	1041	-44	110	731	-44	97	1231	-5	1	71	
15#AWE	40	432	24	-44	24	4	-41	18	101	11	1	44	21	441	4	1	31	0	0	0	0	
16#AWE	44	4	0	-214	431	142	84	132	401	31	20	410	21	4	41	-4	0	2	4	0	2	
17#AWE	-44	452	144	-41	13	22	-93	17	471	0	0	0	-72	11	114	57	4	1331	0	0	0	
18#AWE	114	454	30	-44	14	4	-24	4	3	14	2	21	44	14	401	-19	4	14	-4	0	4	
19#AWE	-404	444	40	-144	124	74	32	2	3	11	1	11	-13	1	41	24	4	321	4	0	4	
20#AWE	304	444	28	-4	0	0	-194	284	2071	52	20	221	26	4	141	-14	2	111	-24	4	1071	
21#AWE	-32	140	0	-47	2	14	-12	702	431	24	23	31	0	0	0	1	0	0	-3	1	3	
22#AWE	302	444	28	-33	12	4	21	4	2	11	1	11	52	24	441	14	4	111	-24	4	44	
23#AWE	24	24	0	-140	717	101	90	163	444	44	60	241	14	0	0	-24	12	241	11	2	21	
24#AWE	-411	442	113	44	3	4	1	44	11	241	19	3	31	-14	1	4	0	1	4	0	3	
25#AWE	240	447	24	-19	4	1	71	55	24	17	1	11	45	13	271	22	5	14	-20	5	91	
26#AWE	-147	414	4	-14	343	44	102	154	44	41	24	141	4	1	11	-21	7	201	19	5	631	
27#AWE	344	402	34	0	0	0	-14	165	142	32	4	41	4	16	421	24	4	26	-15	2	44	
28#AWE	-14	19	0	-42	124	4	-43	601	481	44	144	191	-30	47	211	-16	19	111	-6	3	91	
I	1000			1000			1000			1000			1000			1000			1000			499

Tableau des facteurs, corrélations et contributions des points de I

I	J	1#F	COR	CTR	2#F	COR	CTR	3#F	COR	CTR	4#F	COR	CTR	5#F	COR	CTR	6#F	COR	CTR	7#F	COR	CTR	
1	1	444	453	372	-44	30	167	3#F	6	401	-1	0	0	3	0	11	-15	1	54	9	0	74	
2	1	414	434	401	-44	25	24	3#F	1	120	44	1201	44	4	501	110	40	44	-15	1	44		
3	1	444	444	444	-144	49	243	19	1	7	6	0	11	-17	1	24	0	1	531	1	0	11	
4	1	444	404	74	-44	4	4	243	74	124	44	4	241	280	44	644	-72	7	45	0	0	6	
5	1	444	-171	442	11	104	144	44	2	11	194	444	407	-44	41	114	-23	7	31	30	11	204	
6	1	444	-4	1	44	147	4	14	13	11	70	334	451	-26	51	171	-47	244	144	-53	200	574	
7	1	444	-44	142	4	40	41	14	26	91	-174	447	243	-4	1	11	4	1	31	0	2	22	
8	1	444	-14	214	1	7	44	21	-4	43	41	-24	447	451	-4	74	211	-44	32	171	-4	15	
9	1	444	24	4	0	4	22	344	240	351	332	-7	0	11	-24	4	211	44	15	911	1	0	
10	1	444	-12	4	0	44	274	121	-147	473	442	0	0	0	34	35	44	4	2	11	4	0	
I		1000			1000			1000			1000			1000			1000			1000			1000

Tableau des facteurs, corrélations et contributions des points j de J

- les colonnes notées COR contiennent les contributions relatives du facteur α au point i (ou j) considéré. ($F_{\alpha}^2(i) / \rho^2(i)$), pour chaque valeur de α .
- les colonnes notées CTR contiennent les contributions relatives du point i (ou j) au moment d'inertie λ_{α} ; ($f_i F_{\alpha}^2(i) / \lambda_{\alpha}$), pour chaque valeur de α .
- les colonnes notées 1 ≠ F ; 2 ≠ F ; 3 ≠ F ; 4 ≠ F ; 5 ≠ F ; 6 ≠ F ; 7 ≠ F contiennent les valeurs des facteurs {F α (i) ; $\alpha = 1,7$ }

On a

$$QLT(i)^* = \sum_{\alpha=1,6} F_{\alpha}^2(i) / \rho^2(i) \in [0.9, 1] \quad \forall i \in I$$

$$QLT(j) = \sum_{\alpha=1,6} G_{\alpha}^2(j) / \rho^2(j) \in [0.9, 1] \quad \forall j \in I$$

(*) QLT (i) représente la qualité d'explication du point i (ou j) dans l'espace factoriel retenu.

Les points i et j sont donc bien expliqués par les 6 premiers facteurs.

Les facteurs sont sous la dépendance des éléments qui leur apportent les plus fortes contributions. Le tableau suivant résume pour chaque facteur les résultats obtenus en donnant les éléments dont les contributions relatives ($f_i F_\alpha^2(i) / \lambda_\alpha$ et $f_j G_\alpha^2(j) / \lambda_\alpha$) sont fortes, affectés de leur signe (facteur > 0 ou < 0).

- 1^{er} facteur • FNAW-, FMWE-, FNAY-, FMYO-; FNAE-, FNAU-, ***3-, ***4-
 - HAUS+, HMUS+, HCUS+, HAWÉ+, HMWE+, HCWE+, HAYO+, HMYO+, HCYO+, HAES+, HMES+, HCES+, ***1+, ***2+.
- 2^e facteur • HAUS+, FNAU+, HMUS+, FMUS+, HCUS+, ***5+, ***9+, **10+
 - FAYO-, FMYO-, FAES-, FMES-, FAWÉ-, ***1-, ***2-, ***3-
- 3^e facteur • HAUS+, HMUS+, FAWÉ+, FAYO+, FAES+, FMES+, HMES+, ***4+, ***9+
 - FCWE-, FNAY-, HCYO-, FCYO-, FNAE-, HCES-, FCES-, **10-
- 4^e facteur • FCUS+, FAUS+, avec $F^2(i) / \rho^2(i)$ élevé, ***5+, ***6+
 - FCWE-, FAWÉ-, HMWE-, HAWÉ-, ***7-, ***8-, avec $F_\alpha^2(i) / \rho^2(i)$ élevé.

D'une façon très résumée, on dira que le premier facteur sépare ceux qui ont une activité professionnelle de ceux qui n'en exercent pas — c'est la séparation des tâches liée aux structures sociales (opposition travail — soins donnés aux enfants et courses).

Le deuxième facteur sépare les pays de l'Ouest des pays de l'Est concrétisé par un temps passé aux loisirs de toutes sortes plus important chez les premiers que chez les seconds.

Corollairement, on note un temps plus important consacré aux travaux (professionnels ou ménagers) dans les pays de l'Est que dans les pays de l'Ouest.

Le troisième facteur sépare les célibataires des non-célibataires, séparation due aux types de loisirs. Les premiers se consacrent à des loisirs externes à leur habitat ; les seconds se consacrent à des loisirs familiaux (occupations dues aux enfants, télévision).

Le quatrième facteur, d'importance très faible, fait apparaître une séparation à l'intérieur des pays occidentaux entre les Etats-Unis et les pays européens. Cette séparation est due à une importance plus grande accordée aux repas et au sommeil dans les pays européens au détriment (par rapport aux Etats-Unis) du temps passé aux courses et aux soins personnels. Cette séparation est nette chez les femmes des deux groupes cités (cf. figures n° 1, 2 et 3).

Remarque : Dans un budget-temps, par construction du tableau des données on a $f_i = \text{cte}$. $COR_\alpha(j)$ représente le carré du coefficient de corrélation entre le facteur de rang α et la j^{e} activité (cf. interprétation du 4^e facteur).

6.2. La mise en éléments supplémentaires des classes d'une hiérarchie indicée.

6.2.1. Formules.

On sait qu'on peut adjoindre des éléments supplémentaires de masse nulle à une analyse des correspondances en ne connaissant que son vecteur de description (ou son profil). Par extension, on peut illustrer les résultats en projetant dans l'espace factoriel des classes d'éléments en calculant les centres de gravité de ces classes dans l'espace factoriel.

Soit une correspondance $I \times J$ et $\{f_{ij}, i \in I, j \in J\}$ le tableau des fréquences correspondant. Soit s une ligne supplémentaire de fréquence marginale f_s et de profil $\left\{ \frac{f_{sj}}{f_s}, j \in J \right\}$. On calcule ses facteurs (ou ses coordonnées dans l'espace factoriel) de la façon suivante :

$$F_\alpha (s) = \sum_{j \in J} \frac{1}{\sqrt{\lambda_\alpha}} \cdot \frac{f_{sj}}{f_s} \cdot G_\alpha (j) \quad \forall \alpha \in A$$

avec
$$f_{sj} = \sum_{i \in s} f_{ij} \quad \text{et} \quad f_s = \sum_{i \in s} f_i$$

On fait de même pour une colonne supplémentaire t et on aura :

$$G_\alpha (t) = \sum_{i \in I} \frac{1}{\sqrt{\lambda_\alpha}} \cdot \frac{f_{it}}{f_t} \cdot F_\alpha (i) \quad \alpha \in A$$

On peut donc représenter les classes d'une hiérarchie indicée dans l'espace factoriel en adoptant les mêmes formulations.

Pour une classe n de A (I), et s et t , ses successeurs immédiats de la hiérarchie indicée par l'application \mathcal{J} , on a :

$$f_n = f_s + f_t ; f_{nj} = f_{sj} + f_{tj} \quad \forall j \in J$$

On adopte les formules analogues à celles utilisées pour la mise en éléments supplémentaires de lignes. On a donc :

$$F_\alpha (n) = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{j \in J} \frac{f_{nj}}{f_n} \cdot G_\alpha (j) \quad \forall \alpha \in A$$

ce qui entraîne :

$$F_\alpha (n) = \frac{f_s \cdot F_\alpha (s) + f_t F_\alpha (t)}{f_s + f_t}$$

De même pour une classe m de B (J), avec v et w , ses successeurs immédiats dans la hiérarchie, on aura :

$$G_\alpha (m) = \frac{1}{\sqrt{\lambda_\alpha}} \sum_{i \in I} \frac{f_{im}}{f_m} \cdot F_\alpha (i) \quad \forall \alpha \in A$$

et

$$G_\alpha(m) = \frac{f_v \cdot G_\alpha(v) + f_w G_\alpha(w)}{f_v + f_w} \quad \forall \alpha \in A$$

6.2.2. Applications à l'exemple. (figures 1, 2, 3)

A partir des deux hiérarchies A (I) et B (J) on calcule pour toutes les classes de A (I) (mises en lignes supplémentaires) et toutes les classes de B (J) (mises en colonnes supplémentaires) leurs coordonnées dans l'espace factoriel correspondant par les formules ci-dessus citées.

On calcule de plus (cf. § 6.1) les contributions absolues de chaque classe q de A (I) au moment d'inertie λ_α pour chaque α retenu dans l'analyse (soit $f_q \cdot F_\alpha^2(q)$) et les contributions relatives du facteur α à la classe q de A (I), (soit $F_\alpha^2(q) / \rho^2(q)$).

Représentation des classes k de B (J) dans l'espace factoriel de dimension 6

		JSPUJ						JSPUJ						JSPUJ								
		14F	CDM	CTR1	24F	CDM	CTR1	34F	CDM	CTR1	44F	CDM	CTR1	54F	CDM	CTR1	64F	CDM	CTR1			
24J	20J	222	722	671	127	112	106	106	79	122	32	7	17	-33	8	48	54	21	240	9	1	24
30J	30J	311	441	441	10	5	21	40	14	171	-125	190	206	-2	0	0	-15	2	221	5	0	9
31J	31J	311	441	441	-42	14	121	-33	11	121	13	2	31	44	14	44	14	2	151	-2	0	4
32J	32J	16	4	0	-202	401	271	88	149	45	44	37	32	12	2	4	-14	4	20	7	1	18
33J	33J	-26	24	0	-64	170	201	-104	490	129	37	0	22	-14	14	10	-7	4	8	-4	2	11
34J	34J	264	634	441	-24	4	4	44	23	23	1	0	0	44	21	42	19	4	29	-19	4	14
35J	35J	259	744	441	4	0	0	-174	224	347	42	12	29	44	10	41	4	0	23	4	0	2
36J	36J	72	44	4	13	17	34	44	14	73	250	90		-54	170	154	-24	43	46	-24	36	247
37J	37J	119	447	421	24	4	4	11	1	21	-120	124	363	0	0	0	-11	1	14	4	1	41
38J	38J	40	441	111	-34	14	14	4	0	1	7	1	2	44	20	164	16	3	45	-11	1	10
39J	39J	224	744	441	154	140	144	44	34	74	44	14	41	-43	14	122	34	4	151	14	2	110
40J	40J	-24	477	241	-24	1	4	-79	14	49	9	0	1	-45	5	91	32	2	41	2	0	2
41J	41J	-119	444	144	44	12	23	14	1	4	-44	22	144	24	4	45	1	24	4	4	0	7
42J	42J	14	4	0	-144	704	271	40	144	104	3	0	0	-1	0	0	-21	14	40	3	0	4
43J	43J	-24	44	1	-51	44	4	-102	740	171	-12	12	4	-24	44	44	-22	41	47	-5	2	14
44J	44J	-47	440	144	174	114	210	44	7	22	40	13	41	44	14	140	-37	4	114	4	0	12
45J	45J	-32	40	1	-144	744	144	44	204	144	12	4	4	0	0	0	-21	14	40	7	1	32
46J	46J	-10	444	241	-4	1	2	4	1	3	-47	23	132	24	4	43	4	0	4	-2	0	11
47J	47J	-44	444	101	114	4	144	32	0	22	-17	1	11	47	4	4	4	0	4	1	0	0
48J	48J	11	24	1	7	10	1	-44	414	40	22	47	20	-40	207	174	-24	121	132	-13	36	174
49J	49J	-31	474	241	-44	14	44	-44	4	34	10	0	3	-54	4	40	10	3	112	3	0	4
50J	50J	114	472	241	-4	1	2	-33	11	44	-27	4	44	24	4	144	4	0	4	0	0	3
51J	51J	-24	444	441	34	4	4	0	0	0	-5	0	2	12	0	20	10	0	30	2	0	4
52J	52J	317	444	144	20	4	33	-7	1	4	-4	1	4	10	1	24	12	1	70	3	0	14
53J	53J	-7	1	0	-44	744	141	12	22	7	17	44	23	-22	74	57	-23	90	211	-4	4	37
54J	54J	174	454	200	-12	4	12	0	0	1	2	0	1	-4	0	7	-2	0	10	0	0	1

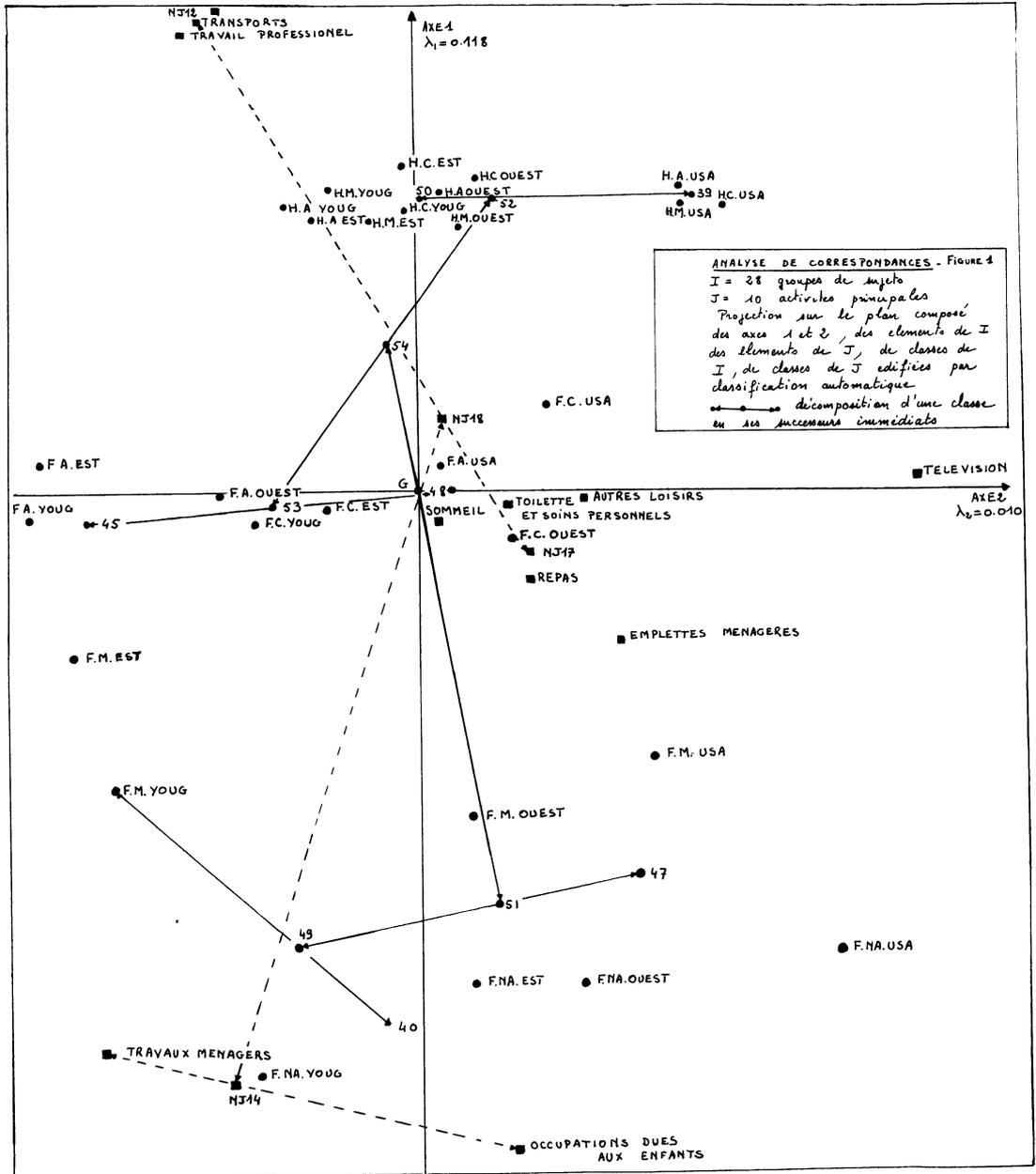
Les classes de A (I) mises en éléments supplémentaires portent des numéros allant de Card I + 1 à 2. Card I - 1 (cf. § 5.3.2). Les classes de B (J) mises en éléments supplémentaires portent des numéros allant de Card J + 1 à 2. Card J - 1 (cf. § 5.3.2). Les classes sont situées au barycentre de leurs successeurs

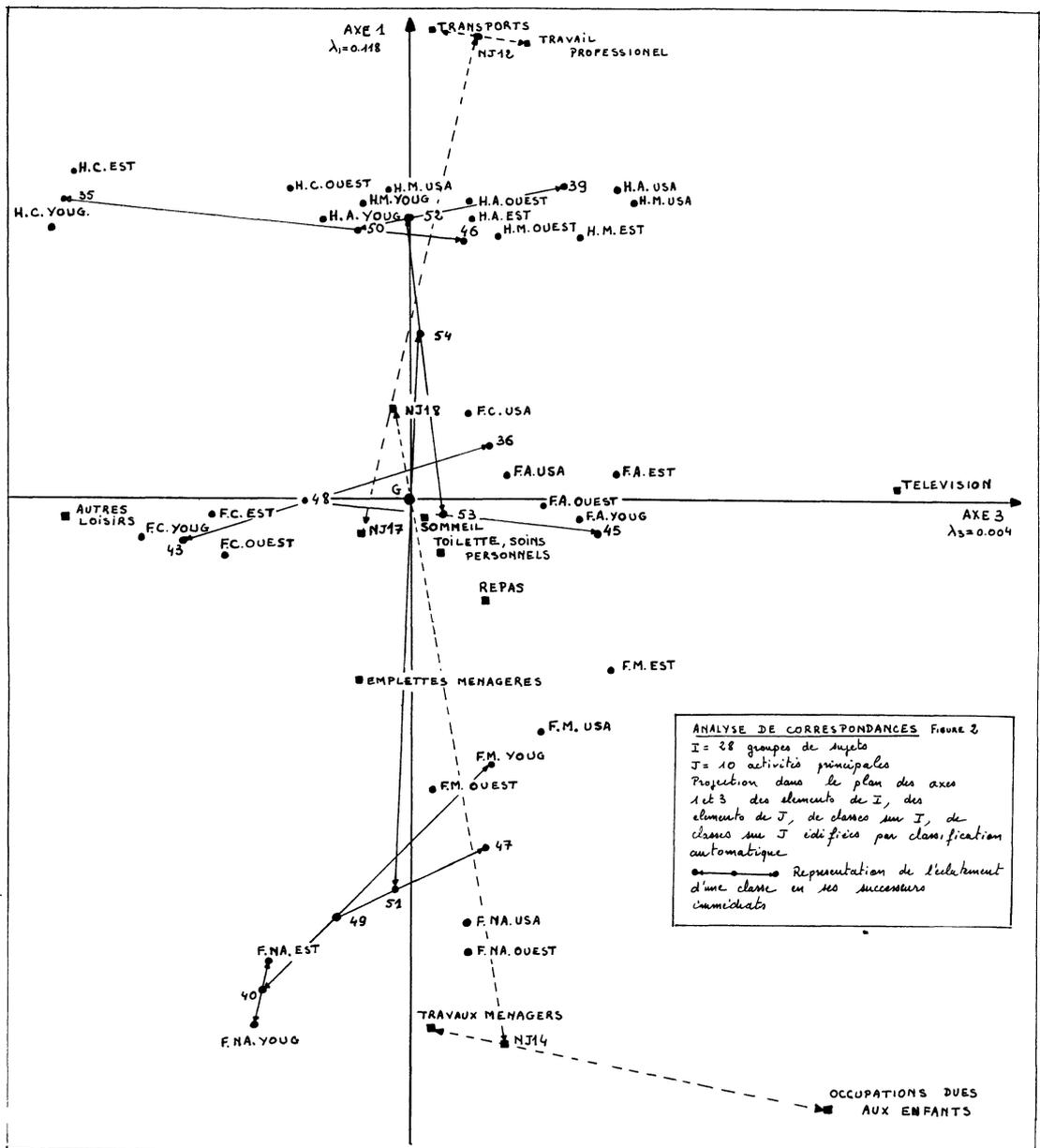
Représentation des classes k de B (J) dans l'espace factoriel de dimension 6

		JSPUJ						JSPUJ						JSPUJ								
		14F	CDM	CTR1	24F	CDM	CTR1	34F	CDM	CTR1	44F	CDM	CTR1	54F	CDM	CTR1	64F	CDM	CTR1			
11J	11J	-14	223	11	12	114	4	-5	44	2	-14	194	20	-11	132	33	-11	131	63	-9	47	143
12J	12J	44	444	421	-45	37	141	34	4	30	14	1	17	11	0	14	4	0	4	4	0	22
13J	13J	-24	240	2	17	123	11	0	1	0	-33	474	113	-10	54	33	-9	40	47	-7	24	123
14J	14J	-44	441	40	-13	34	207	43	4	37	14	0	4	14	0	14	10	0	14	2	0	3
15J	15J	-4	44	4	4	34	144	-14	474	344	44	74	42	10	3	11	1	0	0	10	3	9
16J	16J	-22	244	4	44	405	102	-34	311	124	-4	14	12	-4	4	4	-4	10	30	-1	1	14
17J	17J	-24	174	4	40	712	211	-19	44	41	-4	14	12	-4	4	14	-4	2	4	-1	1	13
18J	18J	102	441	74	20	44	31	-4	4	2	-1	0	1	-1	0	2	-1	0	2	0	0	-1

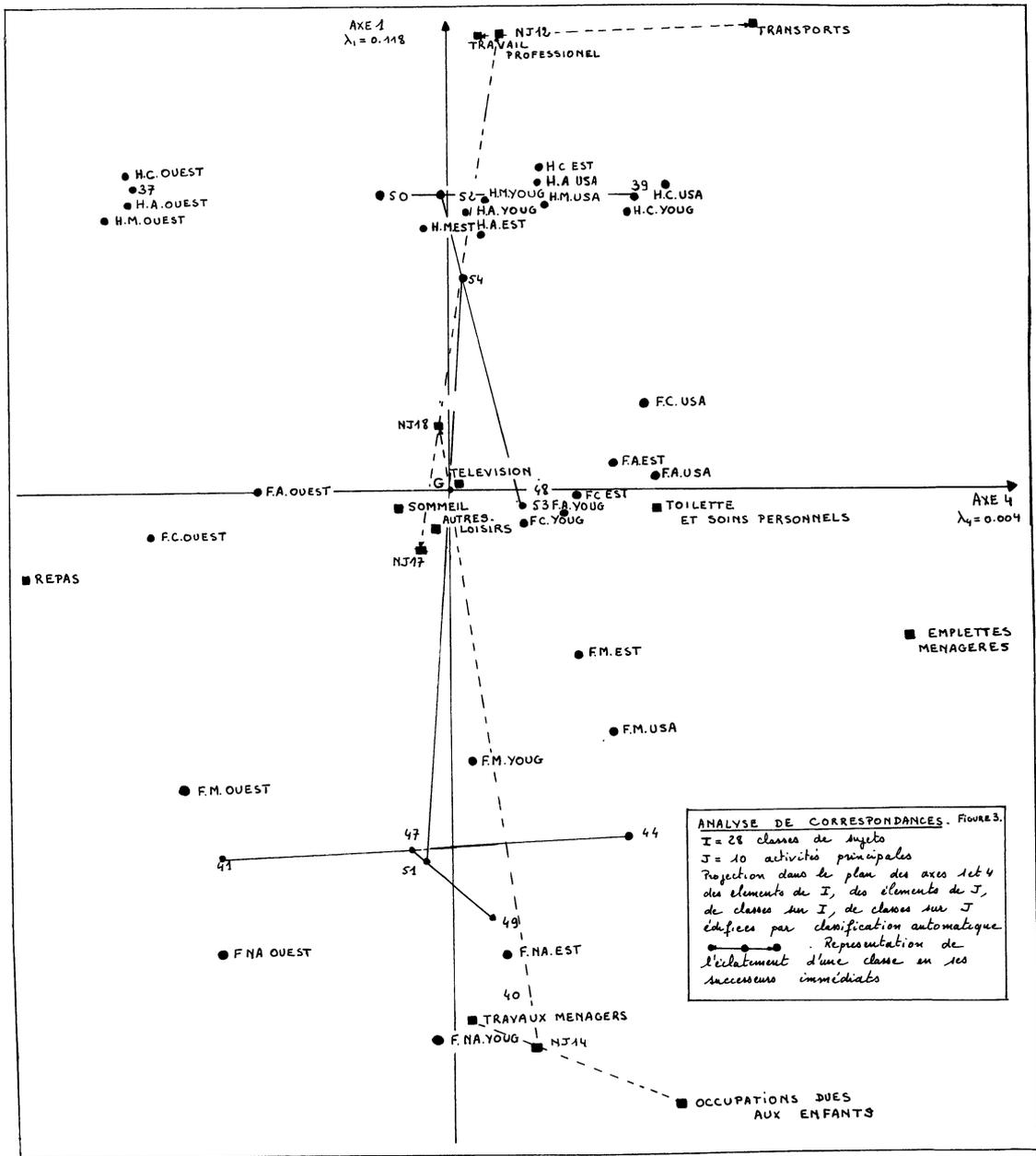
L'étude de la position de ces classes par rapport aux facteurs permet d'affiner l'interprétation des axes factoriels. On trouvera des compléments d'interprétation dans un prochain paragraphe.

Sur les graphiques suivants on a porté dans l'espace factoriel les éléments de I et de J, ainsi que leurs classes constituées en hiérarchie par la méthode exposée au § 5. On a fait partir de chaque nœud des hiérarchies des flèches qui vont dans la direction des classes, successeurs immédiats du nœud considéré.





ANALYSE DE CORRESPONDANCES FIGURE 2
 I = 28 groupes de sujets
 J = 10 activités principales
 Projection dans le plan des axes 1 et 3 des éléments de I, des éléments de J, de classes sur I, de classes sur J identifiés par classification automatique
 —•—•—•—•— Representation de l'échelonnement d'une classe en ses successeurs immédiats



6.3 Calcul des contributions sur une hiérarchie indicée.

6.3.1. Les contributions de $(i, j) \in I \times J$ à $\text{Lien}_2(I, J)$

On a
$$\text{Lien}_2(I, J) = \sum_{i \in I, j \in J} (f_{ij} - f_i f_j)^2 / f_i f_j = \sum_{i \in I, j \in J} C(i, j)$$

$C(i, j)$ est le terme général de cette double somme ; c'est la contribution de la paire (i, j) à $\text{Lien}_2(I, J)$. L'examen de ces nombres affectés du signe de $(f_{ij} - f_i f_j)$ permet de déterminer si le couple (i, j) intervient fortement dans Lien_2 .

Si $(f_{ij} - f_i f_j) < 0$ taux f_{ij} bas

Si $(f_{ij} - f_i f_j) > 0$ " f_{ij} élevé

Sur l'exemple, on a calculé pour chaque (i, j) de $I \times J$ les quantités suivantes :

$$n_{ij} = \text{signe}(f_{ij} - f_i f_j) \cdot C(i, j) / \text{Lien}_2(I, J)$$

	***1	***2	***3	***4	***5	***6	***7	***8	***9	**10
HAUS	6.0	7.5	17.6	1.7	0.1	0.0	0.0	0.1	6.0	0.3
FAUS	0.2	0.0	0.2	0.0	0.0	0.7	0.3	0.0	0.3	0.6
FAAU	4.4	6.0	17.6	19.3	3.6	0.3	0.1	0.0	3.8	2.0
HMUS	6.4	3.6	16.0	1.7	0.0	0.0	0.0	0.1	6.8	0.6
FMUS	16.6	3.9	7.6	4.0	2.6	0.3	0.0	0.0	2.0	0.2
HCUS	4.3	1.0	14.3	2.5	1.6	0.1	0.3	0.1	2.7	0.4
FCUS	0.3	0.1	2.6	0.7	1.0	1.4	0.4	0.0	1.1	0.0
HAWF	0.6	0.2	12.4	2.2	2.6	0.1	0.4	0.1	0.3	0.1
FAWF	0.0	0.3	0.3	0.0	0.0	0.0	0.5	0.1	0.2	2.2
FAAW	42.5	7.5	21.7	5.0	0.0	0.0	3.4	0.4	0.7	0.1
HMWF	0.0	0.1	12.1	1.7	3.1	0.1	1.0	0.1	0.5	0.2
FMWF	10.6	3.2	21.3	3.6	0.1	0.2	2.3	0.0	0.3	0.3
HCWF	9.7	0.4	15.7	3.6	2.1	0.3	0.4	0.1	0.0	0.5
FCWF	0.6	3.1	0.0	1.1	0.2	0.0	0.4	0.0	0.2	0.7
HAYO	0.4	3.5	4.2	1.0	0.5	0.0	0.1	0.1	0.9	0.1
FAYO	2.4	0.4	3.6	0.4	0.3	0.0	0.5	0.2	1.6	1.4
FMAY	40.6	7.0	70.2	1.6	1.1	0.1	0.1	0.1	1.6	0.2
HMYP	0.4	4.2	10.2	1.0	0.5	0.0	0.1	0.1	0.4	0.0
FMYP	9.2	1.4	23.5	2.1	0.1	0.1	0.0	0.0	1.2	0.6
HCYP	6.4	1.4	12.4	3.4	0.0	0.0	1.0	0.1	3.7	4.6
FCYP	0.1	0.0	0.6	0.3	0.0	0.0	0.2	0.0	3.1	1.1
HAFS	0.4	3.6	0.0	0.4	1.0	0.0	0.3	0.1	0.0	0.1
FAFS	3.0	0.4	1.4	0.2	0.0	0.0	0.6	0.1	1.3	4.1
FMAF	41.6	7.4	37.7	4.7	2.3	0.0	0.1	0.4	0.2	0.7
HMFS	0.6	2.7	7.7	0.4	1.6	0.0	0.2	0.1	0.5	0.4
FMFS	0.0	0.1	0.2	2.3	0.1	0.0	0.1	0.0	0.7	4.1
HCFS	7.3	4.6	16.4	3.5	0.4	0.0	0.4	0.0	1.8	3.6
FCFS	0.1	0.0	0.1	0.5	0.4	0.1	0.5	0.0	1.9	0.3

Les plus fortes contributions positives ou négatives se retrouvent pour quatre activités qui jouent un rôle prépondérant dans les budgets-temps (**1, **2, **3, **4). On ne doit pas pour autant négliger certains couples (i, j) intervenant dans d'autres activités que celles cités ci-dessus.

Par contre, la variable **8 (le temps consacré au sommeil) et la variable **6 (le temps consacré aux soins personnels) ne jouent pas un rôle prépondérant.

Ce premier tableau permet de répondre en partie aux questions : *Quelles sont les activités les plus pratiquées par les populations retenues ?* (contribution fortement positive).

Quelles sont les activités les moins pratiquées par les populations retenues ? (contribution fortement négative).

Résumé :

- Contributions fortement “négatives” de (i, j) à $\text{Lien}_2(I, J)$
 (FNAU, ***1), (FNAW, ***1), (FNAY, ***1), (FNAE, ***1) (HAWE, ***3), (HAUS, ***3), (HCUS, ***3), (HCWE, ***3), (HCES, ***3), (HMUS, ***3).
- Contributions fortement “positives” de (i, j) à $\text{Lien}_2(I, J)$
 (FMUS, ***1), (FMWE, ***1), (HMWE, ***1), (HMES, ***1)
 (FNAU, ***3), (FNAW, ***3), (FMWE, ***3), (FNAY, ***3)
 (FMYO, ***3), (FNAY, ***3), (FNAU, ***4), (FMUS, ***4).

6.3.2. Contributions signées des points de I aux classes construites sur J à partir d'un tableau $I \times J$ (et réciproquement).

Formules.

Dans la pratique, en classification on cherche à faire correspondre des classes d'éléments de I à des classes d'éléments de J. Dans le cas présent de l'exemple, peut-on associer à une classification des populations des activités discriminantes qui les caractériseraient ? En analyse factorielle, la représentation simultanée de I et J aide à répondre à la question en ce sens, qu'en gros, un groupe de points de I s'explique, par les éléments de J qui lui sont proches. En classification, sur un tableau de fréquences on peut parvenir à une explication intéressante en calculant les contributions absolues de chaque classe q d'une hiérarchie A (I) et la part qu'apportent les éléments j de J.

Supposons avoir construit une hiérarchie indicée totale sur l'ensemble I, soit A (I), par la méthode proposée au § 4 et Q une partition sur I, et q une classe de Q. on a :

$$\text{Lien}_2(Q, J) = \sum_{q \in Q} f_q \| f_j^q - f_j \|_{f_j}^2 = M^2(N_J(Q))$$

avec

$$\rho^2(q) = \| f_j^q - f_j \|_{f_j}^2$$

$\rho^2(q)$ est la distance de la classe q de Q au centre du Nuage $N_J(Q)$ ($N_J(Q)$ et $N_J(I)$ ont même centre).

Or :

$$\rho^2(q) = \sum_{j \in J} \left(\frac{f_{qj}}{f_q} - f_j \right)^2 / f_j$$

On pose

$$C_{qj} = \left(\frac{f_{qj}}{f_q} - f_j \right)^2 / f_j \quad \text{et} \quad S \text{ le signe de } (f_{qj} - f_j f_q)$$

C_{qj} est la contribution du point j de J à la distance de la classe q de Q au centre du nuage $N_J(I)$.

On étudie à la fois les quantités

$$\rho^2(q), (S \cdot C_{qj}) \text{ ou } (S \cdot C_{qj} / \rho^2(q)) \quad \forall j \in J \quad \forall q \in Q$$

$\rho^2(q)$ permet d'apprécier si la classe q est au centre (dans ce cas la classe q a un profil f_j^q proche du profil moyen f_j) ou très éloigné du centre.

Les autres quantités permettent de déterminer les points de j discriminants pour la classe q (ce sont les points j de J qui fournissent les plus fortes contributions c'est-à-dire que f_{qj} est très élevé ou très faible). Dans ce cas, si S est > 0 , ceci signifie que f_{qj} est élevé et si S est < 0 , ceci signifie que f_{qj} est faible.

Applications à l'exemple.

Classification sur I : Contributions des points j de J à la distance des classes q au centre de gravité du nuage N_j (I).

la colonne N contient les numéros des classes édifiées par l'algorithme de classification hiérarchiques.

la colonne RHO**2 contient la distance entre la classe de numéro N et le centre de gravité du nuage, soit $\rho^2(N)$

les colonnes suivantes contiennent les contributions relatives signées pour chacune des variables de l'ensemble sur lequel on édifie pas la classification.

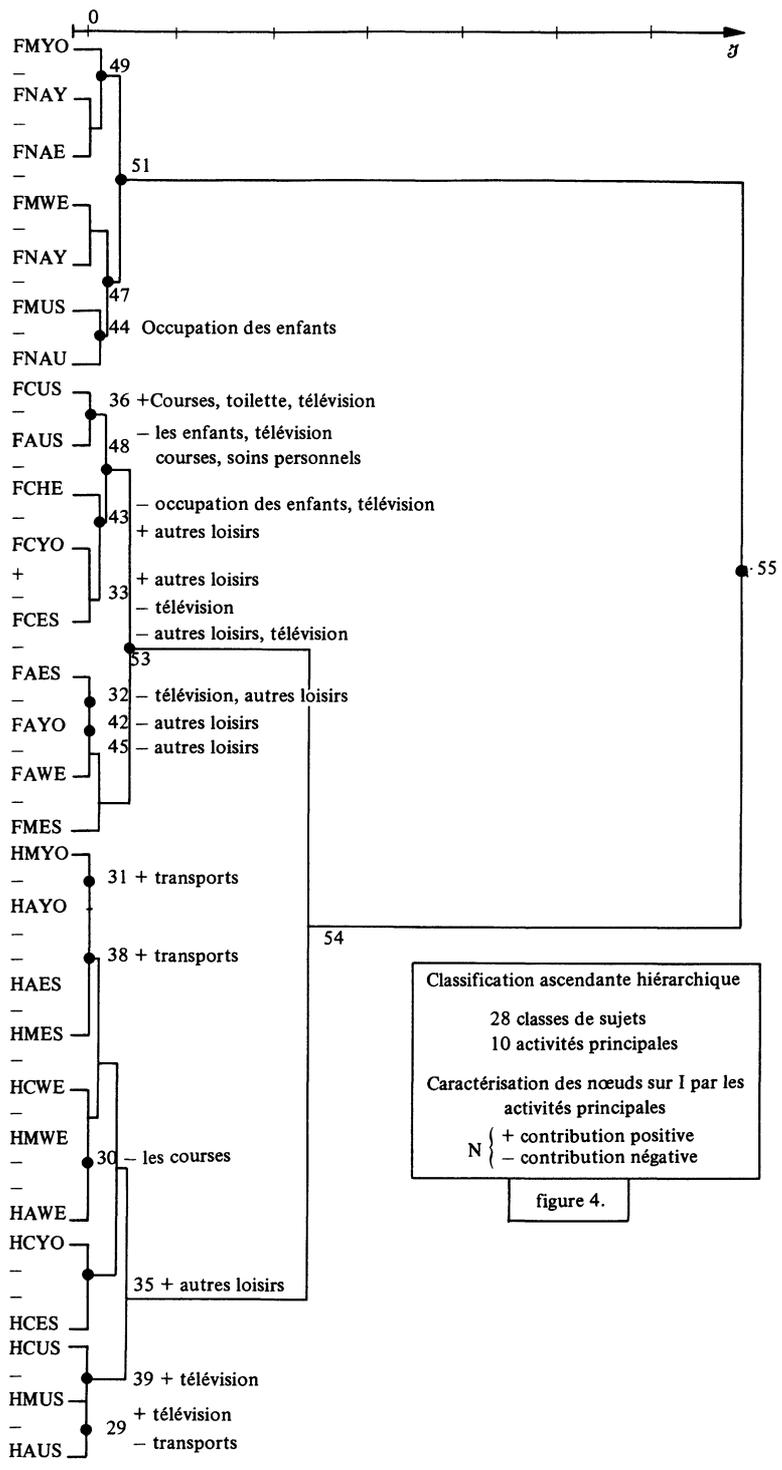
N	RHO**2	***1	***2	***3	***4	***5	***6	***7	***8	***9	***10
29	0.14	0.17	0.10	-0.49	-0.05	0.00	-0.00	-0.00	-0.00	0.18	-0.01
30	0.11	0.34	0.01	-0.43	-0.07	-0.10	-0.00	0.03	0.00	0.01	-0.01
31	0.10	0.37	0.15	-0.38	-0.04	-0.02	-0.00	-0.01	-0.00	-0.02	0.00
32	0.05	0.26	0.04	0.18	0.03	-0.01	-0.00	-0.04	-0.01	-0.11	-0.32
33	0.02	-0.01	0.00	0.08	-0.09	0.03	0.00	-0.08	0.00	-0.55	0.15
34	0.09	0.41	0.14	-0.36	-0.02	-0.06	-0.00	-0.01	-0.00	0.00	-0.01
35	0.14	0.19	0.09	-0.40	-0.10	-0.00	-0.00	-0.03	-0.00	-0.07	0.12
36	0.02	0.04	0.01	-0.23	-0.06	0.20	0.21	-0.07	-0.00	0.13	-0.05
37	0.12	0.32	0.01	-0.46	-0.08	-0.09	-0.01	0.03	0.00	0.01	-0.00
38	0.10	0.40	0.15	-0.38	-0.03	-0.04	-0.00	-0.01	-0.00	-0.00	-0.00
39	0.14	0.16	0.08	-0.53	-0.07	0.01	0.00	-0.00	-0.00	0.15	-0.00
40	0.44	-0.39	-0.07	0.48	0.03	0.02	-0.00	0.00	0.00	-0.01	0.00
41	0.28	-0.41	-0.09	0.35	0.08	-0.00	-0.00	0.04	0.00	0.01	0.00
42	0.03	0.27	0.01	0.17	0.01	-0.03	-0.00	-0.01	-0.00	-0.10	-0.39
43	0.01	-0.05	-0.09	0.03	-0.18	0.00	0.01	-0.00	0.03	-0.41	0.20
44	0.27	-0.43	-0.09	0.18	0.20	0.05	0.00	0.00	-0.00	0.04	0.01
45	0.04	0.14	0.00	0.31	0.04	-0.01	-0.00	-0.01	-0.00	-0.09	-0.40
46	0.10	0.38	0.07	-0.44	-0.05	-0.06	-0.00	0.00	-0.00	0.00	-0.00
47	0.27	-0.44	-0.09	0.28	0.14	0.01	0.00	0.01	0.00	0.02	0.01
48	0.01	-0.00	-0.05	-0.04	-0.34	0.15	0.17	-0.06	0.01	-0.12	0.07
49	0.34	-0.34	-0.06	0.55	0.03	0.01	-0.00	0.00	0.00	-0.01	0.00
50	0.10	0.35	0.08	-0.46	-0.07	-0.04	-0.00	-0.00	-0.00	-0.00	0.01
51	0.28	-0.41	-0.08	0.41	0.08	0.01	-0.00	0.01	0.00	0.00	0.00
52	0.10	0.31	0.08	-0.51	-0.07	-0.01	-0.00	-0.00	-0.00	0.00	0.00
53	0.01	0.13	-0.00	0.24	-0.01	0.01	0.03	-0.05	0.00	-0.25	-0.29
54	0.03	0.41	0.08	-0.41	-0.08	-0.01	0.00	-0.01	-0.00	-0.00	-0.00

L'interprétation d'un tel tableau permet de déterminer quelles variables j de J contribuent le plus à la formation des classes constituées en hiérarchie. Une première lecture du tableau des contributions relatives signées nous confirme l'importance des variables ***1 (temps passé au travail professionnel) et ***3 (travail ménager) comme activités discriminantes principales. (Il eut été surprenant, compte tenu du choix des variables de base, qu'il en fût autrement !). Une seconde lecture nous fait nous intéresser aux contributions d'importance moyenne (de l'ordre du quart des contributions les plus fortes) et qui apportent une explication à la constitution des classes, surtout aux niveaux inférieurs. Ce sont précisément ces variables qui font la différence à ces niveaux. On peut citer les exemples suivants :

classe 29	HMUS – HAUS	C+	Télévision, transport
classe 30	HMWE – HAWÉ	C–	Temps passé à faire les courses
classe 31	HMYO – HAYO	C+	Transports
classe 32	FAES – FAYO	C–	Télévision, autres loisirs
classe 33	FCYO – FCES	C–	Télévision
		C+	Autres loisirs
classe 35	HCYO – HCES	C+	Autres loisirs
classe 36	FCUS – FAUS	C+	Les courses, la toilette, la télévision
classe 38	HMYO – HAYO		
	HAES – HMES	C+	Transports
classe 39	HCUS – HMUS – HAUS	C+	Télévision
classe 42	FAES – FAYO – FAWÉ	C–	Autres loisirs
classe 43	FCES – FCYO – FCWE	C–	Occupations liées aux enfants, télévision
		C+	Autres loisirs
classe 44	FMUS – FNAU	C+	Occupations liées aux enfants
classe 45	FAES – FAYO – FAWÉ – FMES	C–	Autres loisirs
classe 48	FCUS – FAUS – FCWE	C–	Occupations liées aux enfants, télévision
		C+	Courses, les soins personnels
classe 53	Femmes célibataires Femmes actives FMES	C–	Autres loisirs, télévision

Sur ce résumé on n'a pas reporté les contributions des variables ***1 et ***3. Dans le but de mieux illustrer les résultats, on propose de reporter celles-ci sur la représentation de la hiérarchie en procédant de la façon suivante. A chaque nœud de la hiérarchie on associe leurs variables discriminantes positivement ou négativement en mettant + ou – ; quand les variables contribuent fortement on met V++ ou V-- (+ : taux fort, – : taux faible).

Cette représentation a l'avantage de faire porter sur un même graphique les éléments de I et de l'ensemble J. Ceci permet de pallier le désavantage en classification automatique d'une absence de représentation simultanée simple.

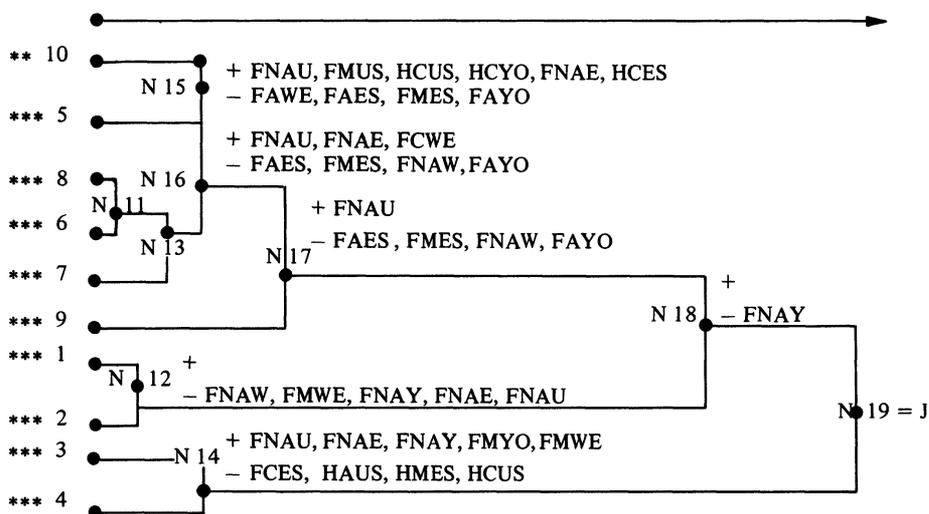


Classification sur J : Contributions du point i de I à la distance de la classe k de B (J) au centre de gravité N_i (K)

N	RMC**2	FAUS	FNAU	FMUS	HCUS	FCUS	HAWA	FAWE
11	0.00	-0.03	0.01	0.01	-0.03	0.00	-0.01	0.02
12	0.25	0.02	0.00	-0.14	0.02	-0.05	0.01	0.00
13	0.00	-0.01	-0.00	0.01	-0.01	0.00	-0.02	0.00
14	0.51	-0.04	-0.00	0.06	-0.04	0.03	-0.05	-0.01
15	0.03	-0.00	-0.00	0.12	-0.01	0.04	0.04	0.00
16	0.00	-0.01	-0.00	0.11	-0.02	0.03	0.01	0.00
17	0.00	0.00	0.00	0.16	0.00	0.05	0.04	0.01
18	0.01	0.04	0.00	-0.06	0.04	-0.03	0.05	0.01

N	RMC**2	FCYO	FAES	FNAE	FMES	FCES	HCF5	FCES
11	0.00	-0.04	-0.00	-0.02	-0.05	0.11	-0.02	-0.01
12	0.25	0.02	-0.00	0.03	0.01	-0.13	0.03	-0.00
13	0.00	-0.06	-0.01	-0.02	-0.05	0.06	-0.02	-0.01
14	0.51	-0.03	0.00	-0.02	0.00	0.06	-0.02	0.02
15	0.03	0.11	0.02	-0.01	-0.05	0.06	-0.04	-0.07
16	0.00	0.02	0.01	-0.03	-0.14	0.10	-0.05	-0.07
17	0.00	0.00	-0.00	-0.02	-0.14	0.06	-0.03	-0.04
18	0.01	0.03	-0.00	0.02	-0.00	-0.09	0.02	-0.02

FNAW	FMWE	FMWF	HCWE	FCWE	HAYO	FAYO	FNAY	HMYO	FMYO
0.11	0.01	-0.00	0.00	0.26	-0.04	-0.04	0.02	-0.04	-0.02
-0.13	0.02	-0.06	0.02	-0.01	0.03	0.01	-0.13	0.03	-0.02
0.19	0.03	0.03	0.01	0.18	-0.03	-0.07	0.01	-0.03	-0.01
0.04	-0.03	0.05	-0.04	-0.00	-0.02	0.01	0.15	-0.02	0.04
0.00	-0.04	0.00	-0.00	0.01	-0.00	-0.11	0.03	-0.00	-0.01
0.07	-0.01	0.02	0.00	0.08	-0.01	-0.15	0.04	-0.01	-0.02
0.04	-0.00	0.02	0.00	0.05	-0.02	-0.17	0.01	-0.02	0.03
-0.09	0.03	-0.05	0.04	0.00	0.02	-0.01	-0.15	0.02	-0.04



On procède pour cet ensemble de la même façon que précédemment pour l'ensemble I

Classe 11 (centrale)	Sommeil, toilette	C+ FNAW - FNAE - FCWE - FAWE C- FAES - FAYO - HCYO
Classe 12 (ρ^2 (q) élevé)	Travail, transport	C- FNAW - FMWE - FNAY - FNAE - FNAU C+
Classe 13 (centrale)	Sommeil, toilette, Repas	C+ FNAW - FCWE - FAWE C- FAYO - HCYO - FAES - HCES - FAYO
Classe 14 (ρ^2 (q) élevé)	Temps aux enfants, travail ménager	C+ FNAU - FNAE - FNAY - FMYO - FMWE C- HCES - HAUS - HMUS - HCUS
Classe 15	Courses, autres loisirs	C+ FNAU - FMUS - HCUS - HCYO - FNAE - HCES C- FAWE - FAES - FMES - FAYO
Classe 16	[15] U [13]	C+ FNAU - FNAE - FCWE C- FAES - FMES - FNAW - FAYO
Classe 17	[16] U {***9}	C+ FNAU C- FAES - FMES - FNAW - FAYO
Classe 18	[17] U [12]	C- FNAY

6.3.4. *Comparaison des classes produites en classification automatique aux axes factoriels.*

Formules.

Avec la méthode de classification décrite au § 5, on a pour une hiérarchie A (I) construite sur I les formules suivantes :

$$\mathcal{J}(n) = \frac{f_t \cdot f_s}{f_t + f_s} \| f_J^s - f_J^t \|_{f_J}^2 = \frac{f_t f_s}{f_t + f_s} \sum_{j \in J} \left(\frac{f_{sj}}{f_s} - \frac{f_{tj}}{f_t} \right)^2 / f_j$$

indice de stratification du noeud n

si t et s sont les successeurs immédiats de la classe n dans la hiérarchie binaire A (I).

$$\begin{aligned} \sum_{n \in \text{Nod}(A(I))} \mathcal{J}(n) &= \text{Lien}_2(I, J) = M^2(N_J(I)) = M^2(N_I(J)) \\ &= \sum_{\alpha \in A} \lambda_\alpha \end{aligned}$$

(cf. § 4 Analyse des correspondances)

Or :

$$\| f_J^s - f_J^t \|_{f_J}^2 = d^2(s, t) \text{ distance distributionnelle entre les centres de gravité de } s \text{ et } t \text{ de } A(I)$$

$$= \sum_{\alpha \in A} |F_\alpha(s) - F_\alpha(t)|^2$$

où

$$F_\alpha(s) = \sum_{i \in s} f_i F_\alpha(i) / \sum_{i \in s} f_i$$

et où s est centre de gravité de la classe s

en posant
$$p(n) = \frac{f_t \cdot f_s}{f_t + f_s}$$

on a :

$$\text{Lien}_2(I, J) = \sum_{n=2}^{n=2 \cdot \text{Card } J - 1} \cdot p(n) \left[\sum_{\alpha \in A} |F_\alpha(s) - F_\alpha(t)|^2 \right] = \sum_{\alpha \in A} \lambda_\alpha \quad (6.1)$$

en posant

$$I(\alpha, n) = p(n) \cdot |F_\alpha(t) - F_\alpha(s)|^2 \quad (6.2)$$

on a :

$$\lambda_\alpha = \sum_{n \in \text{Nod}(A(I))} I(\alpha, n) \quad \forall \alpha \in A \quad (6.3)$$

$$\mathcal{J}(n) = \sum_{\alpha \in A} I(\alpha, n) \quad \forall n \in \text{Nod}(A(I)) \quad (6.4)$$

$I(\alpha, n)$ représente l'inertie de la fourche issue du noeud n par rapport au facteur de rang α .

Dans la pratique, on étudiera les tableaux :

$$I(\alpha, n) / \lambda_\alpha \quad \text{pour } \alpha \in A^*, n \in \text{Nod}(A(I))$$

$\dim(A^*) \leq \dim(A)$

$$I(\alpha, n) / \mathcal{J}(n) \quad \alpha \in A^*, n \in \text{Nod}(A(I))$$

Applications à l'exemple (figures 1, 2, 3) :

1/ Comparaison des classes q de I et des facteurs sur I .

On étudie la qualité de la séparation d'une classe N par rapport à l'axe α en consultant les contributions relatives au diamètre de la classe N .

Par exemple, dans le cas présent et en se reportant à la description des hiérarchies on a les résultats suivants :

1 ^{er} Axe –	$I(55,1)$	séparation entre	N 51 et	N 54	(N 55 = N 51 \cup N 54)
	$I(54,1)$	”	”	N 53 ” N 52	(N 51 \cap N 54 = ϕ)
	$I(49,1)$	”	”	N 40 ”	FMYO
	$I(45,1)$	”	”	N 42 ”	FMES
	$I(44,1)$	”	”	FNAU”	FMUS
	$I(41,1)$	”	”	FNAW”	FMWE

Ces contributions consacrent à des niveaux d'importance inversement proportionnels à la hauteur du noeud dans la hiérarchie, la séparation entre population dite active et population dite inactive.

2 ^e Axe –	$I(51,2)$	séparation entre	N 49 et	N 47
	$I(52,2)$	”	”	N 50 ” N 39
	$I(53,2)$	”	”	N 48 ” N 45

Les contributions marquent la séparation entre les pays de l'Est et les pays de l'Ouest (à quelques nuances près – position intermédiaire des Européens de l'Ouest).

3 ^e Axe –	$I(30,3)$	séparation entre	HMWE et	HAWE
	$I(38,3)$	”	”	N 31 ” N 34
	$I(39,3)$	”	”	(HMUS, HAUS) et HCUS
	$I(37,3)$	”	”	(HMWE, HAWE) et HCWE
	$I(30,3)$	”	”	N 46 et N 35

Tableau des contributions absolues * 10 00

	F 1	F 2	F 3	F 4	F 5	F 6
N 29	0.00	0.00	0.00	0.00	0.00	0.00
N 30	0.00	0.00	0.01	0.00	0.00	0.00
N 31	0.00	0.00	0.01	0.00	0.00	0.00
N 32	0.03	0.01	0.00	0.01	0.01	0.01
N 33	0.00	0.01	0.02	0.01	0.02	0.01
N 34	0.00	0.00	0.04	0.01	0.01	0.00
N 35	0.04	0.00	0.02	0.01	0.01	0.03
N 36	0.07	0.07	0.01	0.00	0.01	0.00
N 37	0.01	0.01	0.18	0.01	0.00	0.00
N 38	0.01	0.01	0.22	0.00	0.00	0.00
N 39	0.01	0.03	0.31	0.04	0.02	0.08
N 40	0.13	0.24	0.01	0.01	0.05	0.05
N 41	0.49	0.05	0.01	0.00	0.00	0.00
N 42	0.00	0.30	0.01	0.36	0.05	0.01
N 43	0.00	0.24	0.01	0.53	0.04	0.05
N 44	0.76	0.13	0.01	0.00	0.03	0.00
N 45	0.85	0.00	0.01	0.04	0.00	0.00
N 46	0.01	0.21	0.00	0.99	0.12	0.05
N 47	0.06	0.53	0.02	0.86	0.05	0.16
N 48	0.43	0.22	0.82	0.32	0.03	0.00
N 49	1.40	0.41	0.24	0.00	0.02	0.00
N 50	0.02	0.01	1.95	0.44	0.01	0.00
N 51	0.32	2.13	0.37	0.05	0.41	0.07
N 52	0.01	1.70	0.84	0.42	0.40	0.08
N 53	0.15	2.33	1.39	0.01	0.13	0.00
N 54	19.43	1.76	0.07	0.14	0.19	0.23
N 55	94.82	0.51	0.00	0.01	0.05	0.03

Tableau des contributions relatives a $\lambda_\alpha, \alpha \in A : \frac{I(\alpha, N)}{\lambda_\alpha}$

	F 1	F 2	F 3	F 4	F 5	F 6
N 45	0.01	0.00	0.00	0.01	0.00	0.00
N 46	0.00	0.02	0.00	0.23	0.07	0.05
N 47	0.00	0.05	0.00	0.20	0.03	0.18
N 48	0.00	0.02	0.12	0.07	0.02	0.00
N 49	0.01	0.04	0.04	0.00	0.01	0.00
N 50	0.00	0.00	0.30	0.10	0.01	0.00
N 51	0.00	0.20	0.06	0.01	0.25	0.09
N 52	0.00	0.16	0.13	0.10	0.24	0.09
N 53	0.00	0.21	0.21	0.00	0.08	0.00
N 54	0.16	0.16	0.01	0.03	0.12	0.27
N 55	0.80	0.05	0.00	0.00	0.03	0.04

Tableau des contributions relatives

a l'indice de diamètre de la hiérarchie : $\frac{I(\alpha, N)}{\text{Indice}(N)}$

	F 1	F 2	F 3	F 4	F 5	F 6
N 30	0.01	0.01	0.73	0.13	0.03	0.01
N 31	0.17	0.24	0.41	0.01	0.00	0.17
N 32	0.46	0.16	0.00	0.13	0.08	0.09
N 33	0.06	0.15	0.21	0.13	0.21	0.07
N 34	0.03	0.04	0.57	0.07	0.07	0.01
N 35	0.31	0.01	0.14	0.05	0.07	0.21
N 36	0.44	0.41	0.05	0.00	0.04	0.01
N 37	0.06	0.03	0.81	0.03	0.00	0.02
N 38	0.03	0.03	0.74	0.01	0.00	0.00
N 39	0.02	0.05	0.63	0.07	0.05	0.16
N 40	0.25	0.47	0.03	0.01	0.10	0.09
N 41	0.82	0.08	0.02	0.00	0.00	0.00
N 42	0.00	0.40	0.02	0.48	0.06	0.01
N 43	0.01	0.27	0.01	0.59	0.05	0.06
N 44	0.83	0.14	0.01	0.00	0.03	0.00
N 45	0.91	0.00	0.01	0.04	0.00	0.00
N 46	0.01	0.15	0.00	0.69	0.08	0.03
N 47	0.04	0.31	0.01	0.50	0.03	0.09
N 48	0.23	0.12	0.44	0.17	0.02	0.00
N 49	0.69	0.19	0.11	0.00	0.01	0.00
N 50	0.01	0.00	0.80	0.18	0.00	0.00
N 51	0.10	0.63	0.11	0.01	0.12	0.02
N 52	0.00	0.49	0.24	0.12	0.12	0.02
N 53	0.04	0.57	0.34	0.00	0.03	0.00
N 54	0.89	0.08	0.00	0.01	0.01	0.01
N 55	1.00	0.01	0.00	0.00	0.00	0.00

Les contributions indiquent la séparation entre célibataires et non célibataires ; l'apparition d'actifs provient du fait qu'il existe un certain recouvrement des données.

4^e Axe – I (46,4) séparation entre (HMYO, HAYO, HMES, HAES)
 et (HMWE, HCWE, HAWÉ)
 I (47,4) ” ” (FMWE, FNAWE) et (FMUS, FNAUS)
 I (43,4) ” ” FCWE et (FCES, FCYO)

Les contributions indiquent une séparation des populations européennes de l'Ouest de toutes les autres populations.

Le tableau des contributions relatives à chaque facteur donne les résultats suivants :

1 ^{er} Axe	N 55, N 54 (5 fois moins)
2 ^e Axe	N 51, N 52, N 53, N 54
3 ^e Axe	N 50
4 ^e Axe	N 46, N 47

C'est la séparation entre les aînés et les benjamins de chacun de ces nœuds qui donne le sens général du facteur d'indice α considéré.

2/ Comparaison des classes k de J aux facteurs calculés sur J .

Tableau des contributions absolues * 1000

	F 1	F 2	F 3	F 4	F 5	F 6
N 11	0.01	0.05	0.02	0.31	0.01	0.09
N 12	0.02	0.00	0.01	0.44	0.06	0.47
N 13	0.24	0.10	0.07	1.11	0.00	0.02
N 14	0.18	0.54	0.63	0.08	1.10	0.11
N 15	0.86	0.01	0.62	1.32	0.34	0.03
N 16	0.09	0.86	1.70	0.82	0.05	0.01
N 17	0.12	2.67	2.73	0.00	0.02	0.10
N 18	44.65	3.99	0.49	0.11	0.04	0.01
N 19	72.04	2.55	0.27	0.03	0.03	0.01

Tableau des contributions relatives à $\lambda_\alpha, \alpha \in A : \frac{I(\alpha, N)}{\lambda_\alpha}$

	F 1	F 2	F 3	F 4	F 5	F 6
N 11	0.00	0.01	0.00	0.07	0.01	0.11
N 12	0.00	0.00	0.00	0.10	0.04	0.55
N 13	0.00	0.01	0.01	0.26	0.00	0.02
N 14	0.00	0.05	0.10	0.02	0.66	0.12
N 15	0.01	0.00	0.09	0.31	0.21	0.04
N 16	0.00	0.08	0.26	0.19	0.03	0.02
N 17	0.00	0.25	0.42	0.00	0.01	0.11
N 18	0.38	0.37	0.07	0.03	0.02	0.01
N 19	0.61	0.24	0.04	0.01	0.02	0.02

Tableau des contributions relatives à l'indice de diamètre

de la hiérarchie : $\frac{I(\alpha, N)}{\text{Indice}(N)}$

	F 1	F 2	F 3	F 4	F 5	F 6
N 11	0.01	0.09	0.03	0.51	0.02	0.15
N 12	0.02	0.00	0.01	0.41	0.06	0.44
N 13	0.15	0.06	0.04	0.69	0.00	0.01
N 14	0.07	0.20	0.24	0.03	0.42	0.04
N 15	0.26	0.00	0.19	0.40	0.11	0.01
N 16	0.03	0.24	0.47	0.22	0.01	0.00
N 17	0.02	0.46	0.48	0.00	0.00	0.02
N 18	0.90	0.08	0.01	0.00	0.00	0.00
N 19	0.96	0.03	0.00	0.00	0.00	0.00

Les résultats déduits de ces tableaux permettent de confirmer l'interprétation donnée au début de cette note.

- 1^{er} Axe Séparation entre trois groupes d'activités : N 17, N 12, N 14.
- 2^e Axe Séparation entre la télévision et les autres activités (autres loisirs, sommeil, repas, toilette).
- 3^e Axe Séparation entre télévision et autres loisirs
- 4^e Axe Séparation principale entre le sommeil, la toilette et le temps passé aux repas.

Conclusion : En reprenant ce qui a été dit en introduction, nous pensons que de tels calculs sont à même de contribuer tant à l'interprétation des facteurs de l'analyse des correspondances qu'à l'interprétation des classes, si on utilise un critère d'agrégation qui offre des points de comparaison avec l'analyse des correspondances. Comme toujours en analyse des données, l'utilisation de plusieurs méthodes qui offrent des points de vue différents est particulièrement recommandée.

7 – BIBLIOGRAPHIE

- [1] J.P. BENZECRI – L'Analyse des Données – Tome I ; La Taximonomie – Dunod 1973 ; L'Analyse des Données Tome II ; Correspondances Dunod 1973.
- [2] M. JAMBU – Introduction à l'Analyse des Données – les méthodes de classification automatique – *Consommation* n° 3 1973 Dunod.
- [3] M. JAMBU – Sur les indices de distances en vue de la construction d'une classification hiérarchique – *Consommation* n° 2 1974 Dunod.
- [4] M. JAMBU – Sur les critères d'agrégation utilisés en classification automatique – *Consommation* n° 4 1974.
- [5] M. SADAKA – Comparaison des classes hiérarchiques aux axes factoriels thèse de doctorat de 3^e cycle – 1974.
- [6] M. JAMBU – Programme de calcul des contributions mutuelles entre classes d'une classification et facteurs d'une analyse des correspondances. *Cahiers de l'Analyse des Données* n° 1 1976 – Dunod
- [7] M. JAMBU – Programme de calcul des aides à l'interprétation conjointes de deux classifications arborescentes construites sur une correspondances. Publications du laboratoire de statistique – Université P et M. Curie. Paris 1975.
- [8] M. JAMBU – Programme de construction de classification ascendante hiérarchique (CAH). In *l'Analyse des Données* de J.P. BENZECRI et collaborateurs Dunod – 2^e édition 1976. Tome I
- [9] N. TABET – Programme d'Analyse des Correspondances in *l'Analyse des Données* de J.P. BENZECRI et collaborateurs – Dunod 2^e édition – 1976. – Tome II.