

COMPARAISON
D'UN MODÈLE FACTORIEL
ET
D'UN MODÈLE HIÉRARCHIQUE
Application à l'analyse des budgets-temps

par

Michel JAMBU (*)

SOMMAIRE

1. Introduction.....	70
2. Présentation de l'exemple.....	70
3. Position du problème.....	72
4. L'analyse des correspondances. Rappel des notations et des résultats principaux.....	73
5. La classification automatique. Rappel des notations et des résultats principaux.....	78
6. Les aides à l'interprétation.....	84
Bibliographie.....	108

(*) Attaché de recherches au C.N.R.S.

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 œuvre en classification, aideront le taxinomiste tant dans l'interprétation des 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. PRÉSENTATION 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 étudiée, 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 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.

(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.

Composition de l'ensemble I

HAUS. hommes actifs des U.S.A.
FAUS. femmes actives des U.S.A.
FNAU. femmes non actives des U.S.A.
HMUS. hommes mariés des U.S.A.
FMUS. femmes mariées des U.S.A.
HCUS. hommes célibataires des U.S.A.
HAWE. hommes actifs des pays de l'Ouest
FAWE. femmes actives des pays de l'Ouest
FNAW. femmes non actives des pays de l'Ouest
HMWE. hommes mariés des pays de l'Ouest
FMWE. femmes mariées des pays de l'Ouest
HCWE. hommes célibataires des pays de l'Ouest
HAES. hommes actifs des pays de l'Est
FAES. femmes actives des pays de l'Est
FNAE. femmes non actives des pays de l'Est
HMES. hommes mariés des pays de l'Est
FMES. femmes mariées des pays de l'Est
HCES. hommes célibataires des pays de l'Est
HAYO. hommes actifs de Yougoslavie
FAYO. femmes actives de Yougoslavie
FNAY. femmes non actives de Yougoslavie
HMYO. hommes mariés de Yougoslavie
FMYO. femmes mariées de Yougoslavie
HCYO. hommes célibataires de Yougoslavie
FCUS. femmes célibataires des U.S.A.
FCWE. femmes célibataires des pays de l'Ouest
FCES. femmes célibataires des pays de l'Est
FCYO. femmes célibataires de Yougoslavie

Composition de l'ensemble J

***1. travail professionnel
***2. occupations dues ou liées au travail professionnel (transports...)
***3. travail ménager
***4. occupations liées aux enfants
***5. les courses
***6. la toilette, les soins personnels
***7. les repas
***8. le sommeil
***9. la télévision
**10. les autres loisirs

Tableau des données

	***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.	325.
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.
HAWÉ	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.	798.	58.	379.

3. POSITION DU PROBLÈME

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) ?

- à 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 RÉSULTATS PRINCIPAUX

4.1. Formules

Soient I et J deux ensembles finis et k_{ij} , $i \in I, j \in J$ le tableau des données correspondant. (Dans l'exemple présenté, 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.)

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

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

$$\text{avec } f_j \neq 0 \quad \text{avec } f_i \neq 0$$

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

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

Profil du point i sur J :

$$f_i^j = \left\{ \frac{f_{ij}}{f_j}, j \in J \right\}_i.$$

Profil du point j sur I :

$$f_j^i = \left\{ \frac{f_{ij}}{f_i}, i \in I \right\}_j.$$

- Distances distributionnelles entre i et i' et entre j et j' :

$$d^2(i, i') = \|f_i^j - f_{i'}^j\|_{f_j}^2 = \sum_{j \in J} \frac{((f_{ij}/f_j) - (f_{i'j}/f_j))^2}{f_j},$$

$$d^2(j, j') = \|f_j^i - f_{j'}^i\|_{f_i}^2 = \sum_{i \in I} \frac{((f_{ij}/f_i) - (f_{ij'}/f_i))^2}{f_i}.$$

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

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

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

- Facteurs et axes factoriels

soient

$$\begin{aligned} &\{F_\alpha(i), i \in I, \alpha \in A\}, && \{\varphi_\alpha^i, i \in I, \alpha \in A\}, \\ &\{G_\alpha(j), j \in J, \alpha \in A\}, && \{\varphi_\alpha^j, j \in J, \alpha \in A\}, \\ &\text{Card } A = \inf(\text{Card } J - 1, \text{Card } I - 1). \end{aligned}$$

On a les résultats suivants :

$$\begin{aligned} \sum_{j \in J} f_j \varphi_\alpha^j &= \sum_{i \in I} f_i \varphi_\alpha^i = \sum_{j \in J} f_j G_\alpha(j) = \sum_{i \in I} f_i F_\alpha(i) = 0, && \forall \alpha \in A, \\ \varphi_\alpha^i &= \frac{F_\alpha(i)}{\sqrt{\lambda_\alpha}}; && \varphi_\alpha^j = \frac{G_\alpha(j)}{\sqrt{\lambda_\alpha}}, \\ \sum_{j \in J} f_j \varphi_\alpha^j &= \sum_{i \in I} f_i \varphi_\alpha^i = 1; && \sum_{j \in J} f_j \varphi_\alpha^j \varphi_\beta^j = 0. \end{aligned}$$

Moment d'inertie du nuage $N_J(I)$ dans la direction de l'axe α :

$$\begin{aligned} \sum_{i \in I} f_i F_\alpha^2(i) &= \lambda_\alpha = \sum_{j \in J} f_j G_\alpha^2(j) \\ &= \sum_{\alpha \in A} \lambda_\alpha = \sum_{i \in I} f_i \sum_{j \in J} \frac{((f_{ij}/f_i) - f_j)^2}{f_j} = \text{Lien}_2(I, J) \quad (\text{cf. § 5}) \\ &= M^2(N_I(J)) = M^2(N_J(I)). \end{aligned}$$

- Formule de reconstitution des données

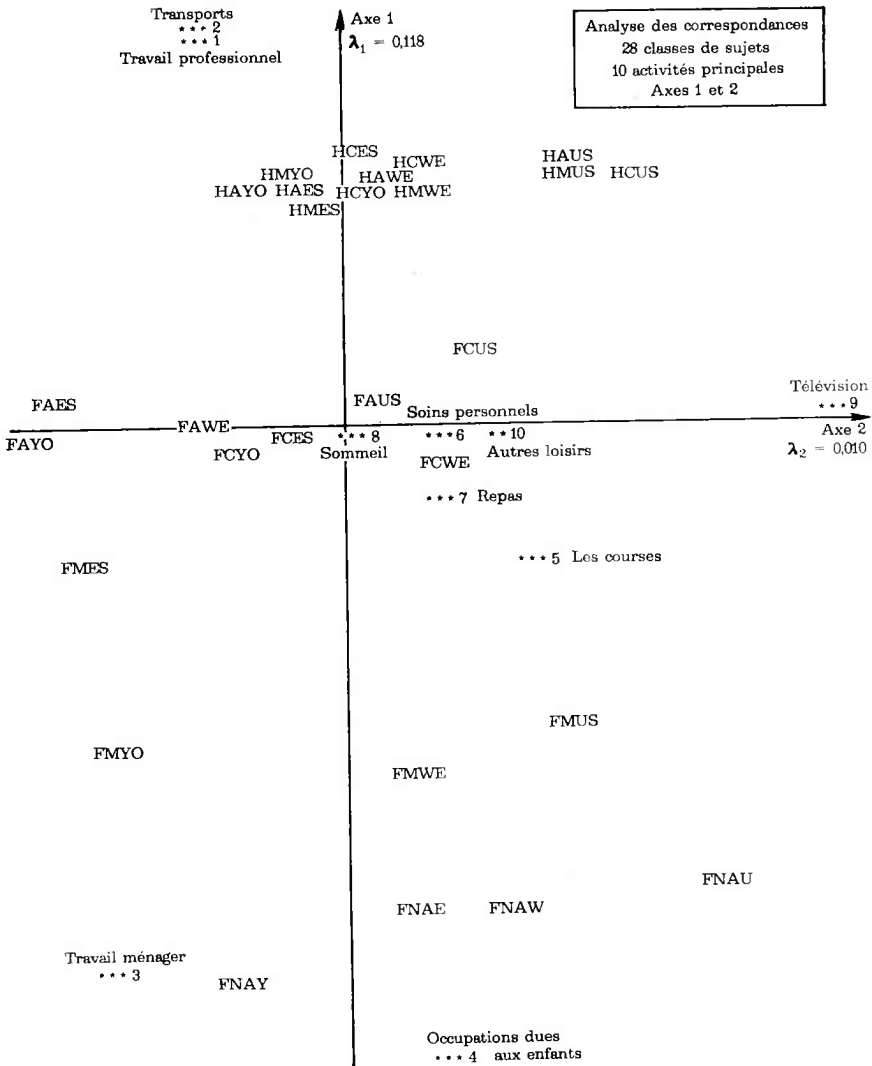
$$\begin{aligned} f_{ij} &= f_i f_j \left(1 + \sum_{\alpha \in A} \sqrt{\lambda_\alpha} \varphi_\alpha^i \varphi_\alpha^j\right), \\ f_{ij} &= f_i f_j \left(1 + \sum_{\alpha \in A} \frac{F_\alpha(i) G_\alpha(j)}{\sqrt{\lambda_\alpha}}\right). \end{aligned}$$

- Distances distributionnelles entre i et i' et entre j et j' :

$$\begin{aligned} d^2(i, i') &= \sum_{\alpha \in A} [F_\alpha(i) - F_\alpha(i')]^2, \\ d^2(j, j') &= \sum_{\alpha \in A} [G_\alpha(j) - G_\alpha(j')]^2, \end{aligned}$$

$$\left. \begin{aligned} F_{\alpha}(i) &= \frac{1}{\sqrt{\lambda_{\alpha}}} \sum_{j \in J} \frac{f_{ij}}{f_i} G_{\alpha}(j) \\ G_{\alpha}(j) &= \frac{1}{\sqrt{\lambda_{\alpha}}} \sum_{i \in I} \frac{f_{ij}}{f_j} F_{\alpha}(i) \end{aligned} \right\} \forall \alpha \in A, i \in I, j \in J.$$

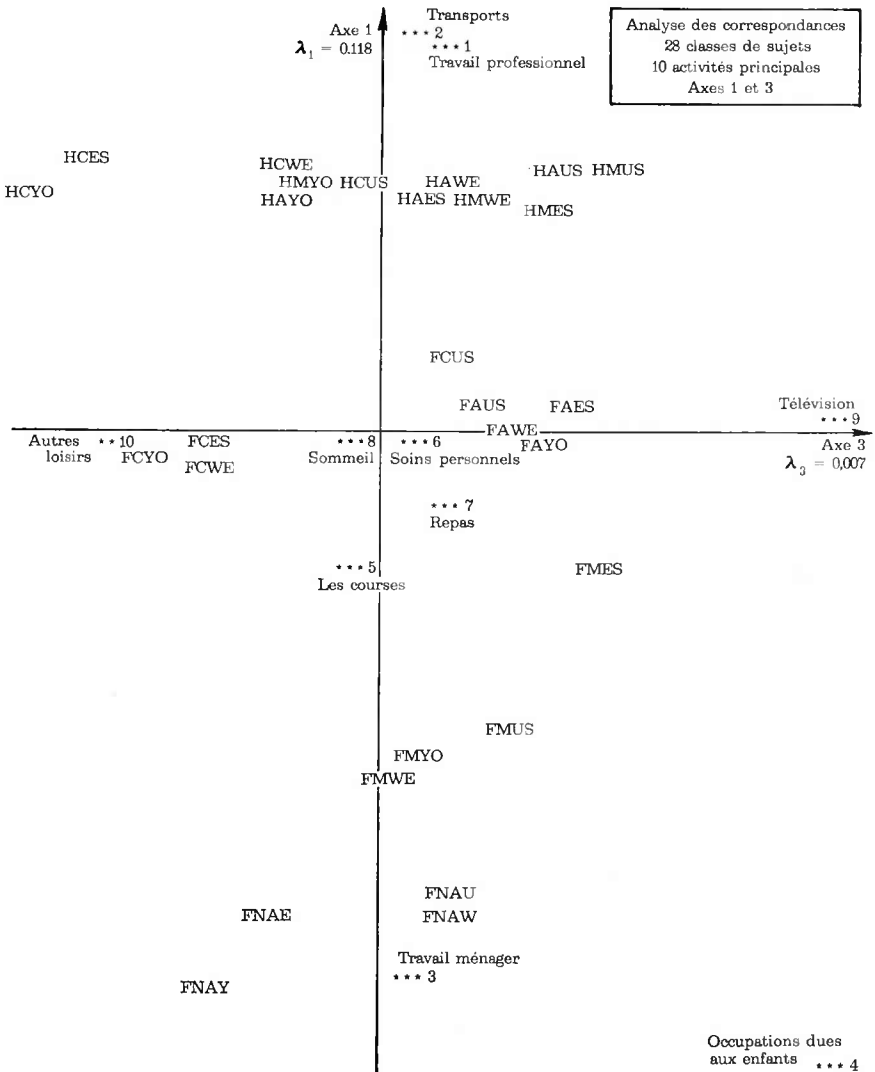
FIGURE 1
Représentation dans le plan factoriel (1, 2)



4.2. L'exemple

Les trois représentations graphiques suivantes illustrent les formules présentées au paragraphe précédent. On a représenté dans l'espace factoriel simultanément les éléments de I et de J . Les représentations graphiques sont effectuées dans les plans constitués des axes factoriels (1, 2), (1, 3) et (1, 4). Une interprétation sera précisée ultérieurement en fonction des calculs qu'on propose.

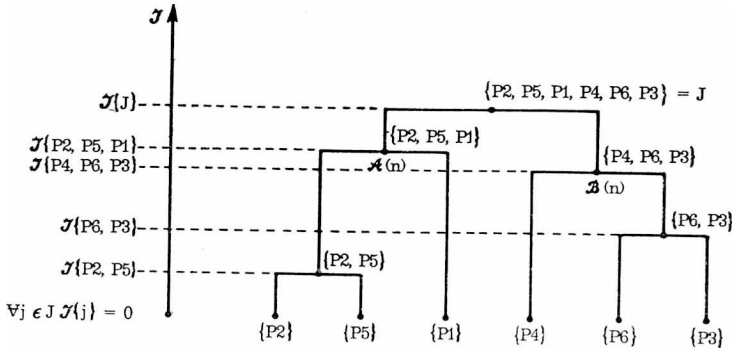
FIGURE 2
Représentation dans le plan factoriel (1, 3)



5. LA CLASSIFICATION AUTOMATIQUE

RAPPEL DES NOTATIONS ET DES RÉSULTATS 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) l'ensemble des minimaux de A ;

Nod (A) l'ensemble des éléments non minimaux de A ;

Pred (a, A) l'ensemble des prédécesseurs de l'élément a dans A ;

Suc (a, A) l'ensemble des successeurs de l'élément a dans A ;

Suci (a, A) l'ensemble des successeurs immédiats de a dans A .

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

$\mathcal{I}\{j\} = 0, \forall j \in J$.

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 $N_I(I)$ et $N_J(J)$ définis au paragraphe précédent :

On a

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

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 :

$$N_J(Q) = \left\{ \begin{array}{l} q \in Q, j\text{-ième coordonnée de } q \text{ dans } R_J = \frac{f_{qj}}{f_q} \\ \text{masse affectée au point } q = f_q \end{array} \right\},$$

$$N_I(K) = \left\{ \begin{array}{l} k \in K, i\text{-ième coordonnée de } k \text{ dans } R_I = \frac{f_{ik}}{f_k} \\ \text{masse affectée au point } k = f_k \end{array} \right\},$$

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

$$N_K(Q) = \left\{ \begin{array}{l} q \in Q, k\text{-ième coordonnée de } q \text{ dans } R_K = \frac{f_{qk}}{f_q} \\ \text{masse affectée au point } q = f_q \end{array} \right\},$$

$$N_Q(K) = \left\{ \begin{array}{l} k \in K, q\text{-ième coordonnée de } k \text{ dans } R_Q = \frac{f_{qk}}{f_k} \\ \text{masse affectée au point } k = f_k \end{array} \right\},$$

$$N_K(Q) = \{f_k^q, f_q\} \subset R_K; \quad N_Q(K) = \{f_Q^k, f_k\} \subset R_Q.$$

On définit

- $\text{Lien}_2(Q, K) = \|f_{QK} - f_Q f_K\|_{f_Q f_K}^2$

$$= \sum_{\substack{q \in Q \\ k \in K}} \frac{(f_{qk} - f_q f_k)^2}{f_q f_k}$$

$$= M^2(N_Q(K)) = M^2(N_K(Q)),$$

$$[f_k = \sum_{i \in K} f_i], \quad [f_q = \sum_{j \in Q} f_j], \quad [f_{qk} = \sum_{\substack{i \in k \\ j \in q}} f_{ij}];$$

- $\text{Lien}_2(Q, J) = \|f_{QJ} - f_Q f_J\|_{f_Q f_J}^2$

$$= \sum_{\substack{q \in Q \\ j \in J}} \frac{(f_{qj} - f_q f_j)^2}{f_q f_j}$$

$$= M^2(N_J(Q)) = M^2(N_Q(J)).$$

La quantité $\text{Lien}_2(Q, K)$ matérialise la dépendance globale entre deux partitions Q de I et K de 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.

Supposons l'étape de rang $(n-1)$ réalisée. Le sommet de la hiérarchie est la partition 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 :

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

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 $B(J)$ la hiérarchie de parties sur 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

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

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

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

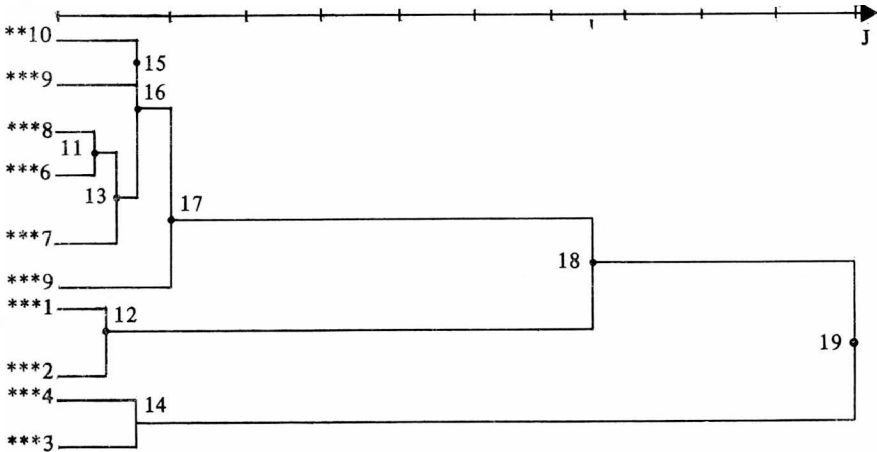
Les tableaux suivants donnent la description des hiérarchies constituées sur I et sur J .

A chaque nœud n sont associés son indice $\mathcal{I}(n)$, son taux d'inertie $\Gamma(n)$, ses 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 Card $J+1$ à 2.Card $J-1$ pour une hiérarchie binaire totale indicée construite sur un ensemble de Card J éléments.

5.3.1. Hiérarchie construite sur l'ensemble J des activités

NOEUD	TAUX	INDICE	AINE	BENJAMIN	POIDS	DESCRIPTION DE LA HIERARCHIE
11	0.007	0.001	8	6	2	***8 ***6
12	0.007	0.001	1	2	2	***1 ***2
13	0.014	0.002	11	7	3	***8 ***6 ***7
14	0.021	0.003	4	3	2	***4 ***3
15	0.021	0.003	10	5	2	**10 ***5
16	0.028	0.004	15	13	5	**10 ***5 ***8 ***6 ***7
17	0.042	0.006	16	9	6	**10 ***5 ***8 ***6 ***7 ***9
18	0.350	0.050	17	12	8	**10 ***5 ***8 ***6 ***7 ***9 ***1 ***2
19	0.520	0.075	18	14	10	**10 ***5 ***8 ***6 ***7 ***9 ***1 ***2 ***4 ***3

Agrégation par la maximisation du moment d'ordre 2 d'une partition



L'interprétation ⁽¹⁾. Le premier 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

(1) On laisse le soin au lecteur de confirmer ou d'infirmier 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.

activités. Pour les activités restantes, le deuxième 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 troisième niveau sépare la télévision des activités restantes. Ce qui signifie qu'elle constitue un point de différenciation pour les groupes de sujets retenus. Le quatrième niveau dépare, d'une part, certains types d'activités 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.

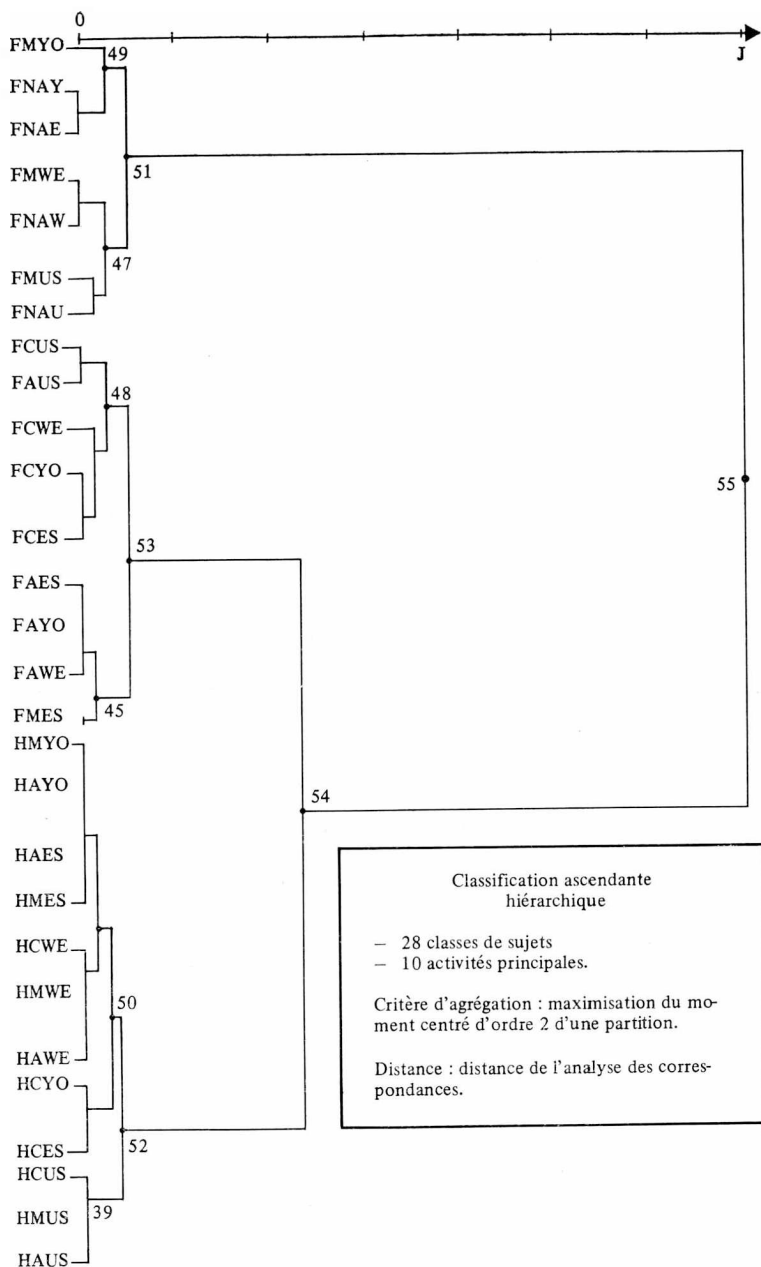
5.3.2. Hiérarchie construite sur l'ensemble I des groupes de sujets enquêtés

L'interprétation. On observe deux niveaux principaux :

- le premier 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 premier 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, les hommes, les célibataires et les femmes ayant une profession.

NOEUD	TAUX	INDICE	AINE	BENJAMIN	POIDS	DESCRIPTION DE LA HIERARCHIE
29	0.000	0.000	4	1	2	HMUS HAUS
30	"	"	11	8	2	HMWE HAWWE
31	"	"	18	15	2	HMYO HAYO
32	"	"	23	16	2	FAES FAYO
33	"	"	21	28	2	FCYO FCES
34	"	"	22	25	2	HAES HMES
35	"	"	20	27	2	HCYO HCES
36	"	"	7	2	2	FCUS FAUS
37	"	"	13	30	3	HCWE HMWE HAWWE HAES HMES
38	"	"	31	34	4	HMYO HAYO HAES HMES
39	"	"	6	29	3	HCUS HMUS HAUS
40	0.007	0.001	17	24	2	FNAY FNAE
41	"	"	12	10	2	FMWE FNAW
42	"	"	32	9	3	FAES FAYO FAWWE FMES
43	"	"	14	33	3	FCWE FCYO FCES
44	"	"	5	3	2	FMUS FNAU
45	"	"	42	26	4	FAES FAYO FAWWE FMES
46	"	"	38	37	7	FMYO HAYO HAES HMES
						HCWE HMWE HAWWE
47	0.014	0.002	41	44	4	FMWE FNAW FMUS FNAU
48	"	"	36	43	5	FCUS FAUS FCWE FCYO FCES
49	"	"	19	40	3	FMYO FNAY FNAE
50	"	"	46	35	9	HMYO HAYO HAES HMES HCWE
						HMWE HAWWE HCYO HCES
51	0.021	0.003	49	47	7	FMYO FNAY FNAE FMWE FNAW
						FMUS FNAU
52	"	"	50	39	12	HMYO HAYO HAES HMES HCWE HMWE
						HAWWE HCYO HCES HCUS HMUS HAUS
53	0.028	0.004	48	45	9	FCUS FAUS FCWE FCYO FCES FAES
						FAYO FAWWE FMES
54	0.154	0.022	53	52	21	FCUS FAUS FCWE FCYO FCES FAES
						FAYO FAWWE FMES HMYO HAYO HAES
						HMES HAWWE HCYO HCES HCUS HMUS
						HAUS HCWE HMWE
55	0.665	0.095	51	54	28	FMYO FNAY FNAE FMWE FNAW FMUS
						FNAU FCUS FAUS FCWE FCYO FCES
						FAES FMES HMYO HAYO HAES HMES
						HCWE HMWE HAWWE HCYO HCES
						HMUS HAUS FAYO FAWWE

● le deuxième 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 troisième 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 aux niveaux 1 et 2;

- le quatrième niveau sépare les femmes célibataires des femmes actives;

- le cinquième niveau sépare les hommes ou femmes des pays européens de l'Ouest des hommes ou femmes des États-Unis.

À 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.

6. LES AIDES A L'INTERPRÉTATION

On rappelle en premier lieu le calcul de contributions usité 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 $N_J(I)$:

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

On a :

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

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

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

- $f_i \cdot \rho^2(i)$ = 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 ;

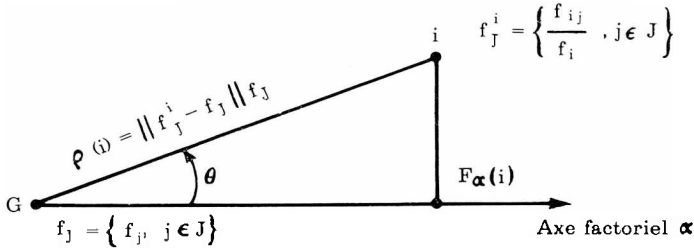
- $F_\alpha^2(i)/\rho^2(i)$ = contribution du point i à cause de la distance au sens du χ^2 entre i et le centre du nuage $N_J(I)$ – soit $\rho^2(i)$;

- $\text{Cos}^2 \theta = F_\alpha^2(i)/\rho^2(i)$ = contribution relative du facteur α au point i .

On a alors

$$\sum_{\alpha \in A} \frac{F_{\alpha}^2(i)}{\rho^2(i)} = 1.$$

On a les formules analogues pour les points j du nuage $N_I(J)$.



6.1.2. Applications à l'exemple

$$\lambda_1 = 0,118, \quad \lambda_2 = 0,010, \quad \lambda_3 = 0,007,$$

$$\lambda_4 = 0,004, \quad \lambda_5 = 0,002, \quad \lambda_6 = 0,001.$$

Tableau des facteurs, corrélations et contributions des points de I

II	$F_1(i)$			$F_2(i)$			$F_3(i)$			$F_4(i)$			$F_5(i)$			$F_6(i)$		
	1#F	COR	CTR	2#F	COR	CTR	3#F	COR	CTR	4#F	COR	CTR	5#F	COR	CTR	6#F	COR	CTR
HAUS	332	776	33	130	119	55	99	69	53	36	9	11	-32	8	23	50	18	101
FAUS	40	124	0	20	32	1	47	168	12	72	396	43	-46	172	47	-31	79	42
FNAU	-580	845	102	222	123	161	34	3	6	55	8	26	-85	18	152	-33	3	48
HMUS	333	765	33	124	106	50	114	90	70	28	5	7	-33	8	25	-59	24	141
FMUS	-375	825	43	137	109	61	55	18	17	65	25	35	-45	12	44	-41	10	71
HCUS	313	735	30	160	191	83	7	0	0	71	38	42	-63	30	86	-2	0	0
FCUS	103	356	3	81	218	21	25	20	3	75	188	47	-66	149	95	-24	21	26
HAWE	312	857	29	18	3	1	29	8	5	-121	130	123	-4	0	1	-14	2	9
FAWE	8	3	0	91	391	27	65	198	23	-78	294	53	-31	47	22	-34	58	51
FNAW	-602	948	109	84	18	23	31	2	5	-92	23	72	35	3	25	-26	2	27
HMWE	309	828	29	20	4	1	50	22	14	-130	148	143	0	0	0	-17	3	13
FMWE	-438	938	60	34	6	4	8	0	0	-96	46	82	23	3	12	-31	5	40
HCWE	334	877	34	35	10	4	-46	17	12	-108	93	99	4	0	0	-3	0	1
FCWE	-38	49	0	47	70	7	-89	259	43	-113	411	106	-58	110	73	-54	97	123
HAYO	305	932	28	-49	24	8	-41	18	10	11	1	1	45	21	44	8	1	3
FAYO	-6	1	0	-214	836	152	86	132	40	33	20	9	21	8	9	-6	1	2
FNAY	-695	952	146	-81	13	22	-93	17	47	0	0	0	-72	11	114	57	6	133
HMYO	316	954	30	36	13	5	24	6	3	14	2	2	44	18	40	19	4	15
FMYO	-405	865	50	-154	125	78	22	2	3	11	1	1	-13	1	4	28	4	32
HCYO	304	684	28	-4	0	0	-195	284	207	52	20	22	26	5	14	-16	2	11
FCYO	-32	49	0	67	212	15	-123	702	83	25	29	5	0	0	0	1	0	0
HAES	302	946	27	33	12	4	21	4	2	11	1	1	52	28	50	10	0	11
FAES	35	24	0	-190	737	120	90	163	44	55	60	25	4	0	0	-24	12	25
FNAE	-611	982	113	34	3	4	-65	11	24	19	1	3	-18	1	8	6	0	1
HMES	290	907	25	-19	4	1	71	55	28	-7	1	1	35	13	27	22	5	19
FMES	-167	418	8	-159	383	84	102	156	56	41	25	14	8	1	1	-21	7	20
HCES	354	802	38	6	0	0	-161	169	142	82	6	8	50	16	53	25	4	26
FCES	-16	19	0	-42	125	6	-93	601	48	48	158	19	-30	67	21	-16	19	11
			1000			1000			1000			1000			1000			1000

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

	G ₁ (j)			G ₂ (j)			G ₃ (j)			G ₄ (j)			G ₅ (j)			G ₆ (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
***1	486	953	372	- 98	39	167	38	6	40	- 1	0	0	- 3	0	1	- 15	1	58
***2	514	876	80	- 86	25	25	16	1	11	120	48	120	49	8	50	110	40	497
***3	-683	949	456	-154	49	253	19	1	7	96	0	1	- 17	1	23	20	20	53
***4	-804	809	76	54	4	4	243	74	124	86	9	24	280	98	648	- 72	7	85
***5	-171	343	11	108	136	49	13	2	1	196	446	407	- 66	51	119	- 23	7	31
***6	- 4	1	0	46	147	8	14	13	1	70	333	45	- 26	51	17	- 57	234	154
***7	- 89	182	3	60	81	16	34	26	9	-174	687	353	- 4	1	1	- 8	1	3
***8	- 16	214	1	7	38	2	- 8	53	4	- 24	457	49	- 9	78	21	- 8	32	17
***9	24	4	0	306	622	356	230	351	332	- 7	0	1	- 28	6	21	44	13	91
**10	- 13	6	0	95	279	121	-147	673	482		0	0	34	35	98	8	2	11
	1000			1000			1000			1000			1000			1000		

- 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 ;
- on a

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

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

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) :

— premier facteur :

FNAW-, FMWE-, FNAY-, FMYO-, FNAE-, FNAU-, ***3-, ***4-,
HAUS+, HMUS+, HCUS+, HAWE+, HMWE+, HCWE+, HAYO+,
HMYO+, HCYO+, HAES+, HMES+, HCES+, ***1+, ***2+;

— deuxième facteur :

HAUS+, FNAU+, HMUS+, FMUS+, HCUS+, ***5+, ***9+, **10+,
FAYO-, FMYO-, FAES-, FMES-, FAWE-, ***1-, ***2-, ***3-;

(¹) QLT (i) représente la qualité d'explication du point i (ou j) dans l'espace factoriel retenu.

— troisième facteur :

HAUS⁺, HMUS⁺, FAWÉ⁺, FAYO⁺, FAES⁺, FMES⁺, HMEs⁺, ***4⁺,
***9⁺,
FCWE⁻, FNAY⁻, HCYO⁻, FCYO⁻, FNAE⁻, HCES⁻, FCES⁻, **10⁻;

— quatrième facteur :

FCUS⁺, FAUS⁺, avec $F^2(i)/\rho^2(i)$ élevé ***5⁺, ***6⁺, FCWE⁻, FAWÉ⁻,
HMWE⁺, HAWÉ⁻, ***7⁻, ***8⁻, avec $F^2(i)/\rho^2(i)$ 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 États-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 États-Unis) du temps passé aux courses et aux soins personnels. Cette séparation est nette chez les femmes des deux groupes cités (cf. fig. 1, 2 et 3).

● *Remarque* : Dans un budget-temps, par construction du tableau des données, on a $f_i = \text{Cte}$. $\text{COR}_\alpha(j)$ représente le carré du coefficient de corrélation entre le facteur de rang α et la j -ième activité (cf. interprétation du quatrième 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 dont les centres de gravité sont représentés 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 $\{f_{sj}/f_s, j \in J\}$.

On calcule ses facteurs (ou ses coordonnées dans l'espace factoriel) de la façon suivante :

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

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

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

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{I} , on a

$$f_n = f_s + f_t; \quad f_{nj}^1 = f_{sj} + f_{tj}, \quad \forall j \in J.$$

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

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

ce qui entraîne

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

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

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

et

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

6.2.2. Applications à l'exemple

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 $f_q F_\alpha^2(q)$) 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 q de $A(I)$ dans l'espace factoriel de dimension 6

	ISUP	1#F	COR	CTR	2#F	COR	CTR	3#F	COR	CTR	4#F	COR	CTR	5#F	COR	CTR	6#F	COR	CTR
29	29	332	772	67	127	112	105	106	79	122	32	7	17	-33	8	48	54	21	240
30	30	311	841	58	19	3	2	40	14	17	-125	139	266	-2	0	0	-15	2	22
31	31	311	945	58	-42	18	12	-33	11	12	13	2	3	44	19	84	14	2	15
32	32	14	4	0	-202	801	271	88	149	83	44	37	32	12	3	6	-15	5	201
33	33	24	35	0	-55	179	20	-108	689	129	37	78	22	-14	14	10	-7	4	5
34	34	296	938	53	-26	8	5	46	23	23	1	0	0	44	21	82	19	4	29
35	35	329	754	65	0	0	0	-178	224	347	42	12	29	38	10	61	4	0	2
36	36	72	268	3	51	133	17	36	66	14	73	280	90	-56	170	138	-28	43	66
37	37	319	867	92	24	5	6	11	1	2	-120	124	363	0	0	0	-11	1	18
38	38	303	961	111	-34	13	16	6	0	1	7	1	2	44	20	166	16	3	43
39	39	326	785	96	138	140	186	68	34	76	45	15	51	-43	14	122	35	9	151
40	40	-653	977	258	-23	1	4	-79	15	69	9	0	1	-45	5	91	32	2	81
41	41	-518	949	165	58	12	23	19	1	4	-94	32	154	29	3	35	28	3	66
42	42	12	4	0	-165	795	271	80	186	104	3	0	0	-1	0	0	-21	14	60
43	43	-29	64	1	-21	35	5	-102	780	171	-12	12	4	-29	65	56	-22	41	67
44	44	-477	840	138	179	118	210	45	7	22	60	13	61	65	16	180	-37	5	118
45	45	-32	30	1	-164	748	354	86	203	159	12	4	5	0	0	0	-21	14	80
46	46	310	969	202	-9	1	2	8	1	3	-47	23	132	25	6	93	4	0	5
47	47	-498	934	303	118	53	185	32	4	22	-17	1	11	47	8	187	-4	0	4
48	48	11	23	0	7	10	1	-46	415	60	22	87	20	-40	307	176	-25	121	132
49	49	-571	973	295	-67	14	45	-45	6	34	10	0	3	-34	4	80	30	3	112
50	50	314	973	267	-6	1	2	-33	11	55	-27	8	58	28	8	149	4	0	6
51	51	-529	994	595	39	5	35	0	0	0	-5	0	2	12	0	20	10	0	30
52	52	317	988	363	29	8	33	-7	1	4	-8	1	9	10	1	26	12	1	70
53	53	-7	11	0	-68	746	141	12	22	7	17	48	23	-22	79	97	-23	90	211
54	54	178	994	200	-12	5	12	0	0	0	2	0	1	-3	0	7	-2	0	10

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).

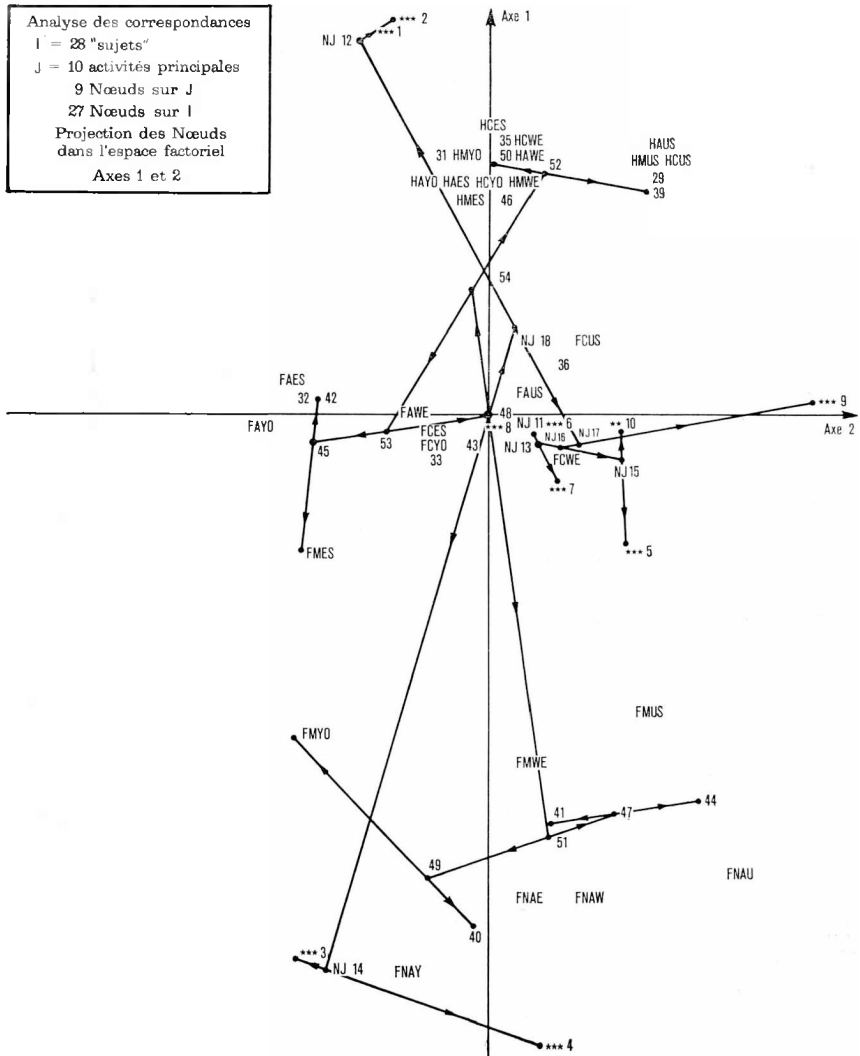
Représentation des classes k de $B(J)$ dans l'espace factoriel de dimension 6

JSUP	1#F	COR	CTR	2#F	COR	CTR	3#F	COR	HCTR	4#F	COR	CTR	5#F	COR	CTR	6#F	COR	CTR
NJ11	-15	223	1	12	116	4	-5	34	2	-14	199	20	-11	132	32	-11	131	63
NJ12	490	956	452	-96	37	191	34	5	39	18	1	17	11	0	15	4	0	4
NJ13	-24	750	2	17	123	11	0	1	0	-33	474	112	-10	54	33	-5	40	47
NJ14	-696	961	530	-131	35	207	43	4	37	14	0	6	14	0	15	10	0	15
NJ15	-51	96	4	98	334	169	-115	479	389	46	74	92	10	2	11	1	0	0
NJ16	-32	245	6	43	405	102	-36	311	129	-8	18	12	-4	5	8	-6	10	30
NJ17	-29	175	5	60	712	211	-19	83	41	-8	16	12	-5	8	16	-2	2	8
NJ18	103	961	79	20	35	31	-5	4	5	-1	0	1	-1	0	2	-1	0	2

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 paragraphe 5. On a fait partir de chaque nœud des hiérarchies des flèches qui vont dans la direction des classes, successeurs immédiats du nœud considéré.

FIGURE 4
Représentation des classes de $A(I)$ et $B(J)$ dans le plan (1, 2)
et pour les classes supérieures l'éclatement de ces classes



6.3. Calcul des contributions sur une hiérarchie indicée

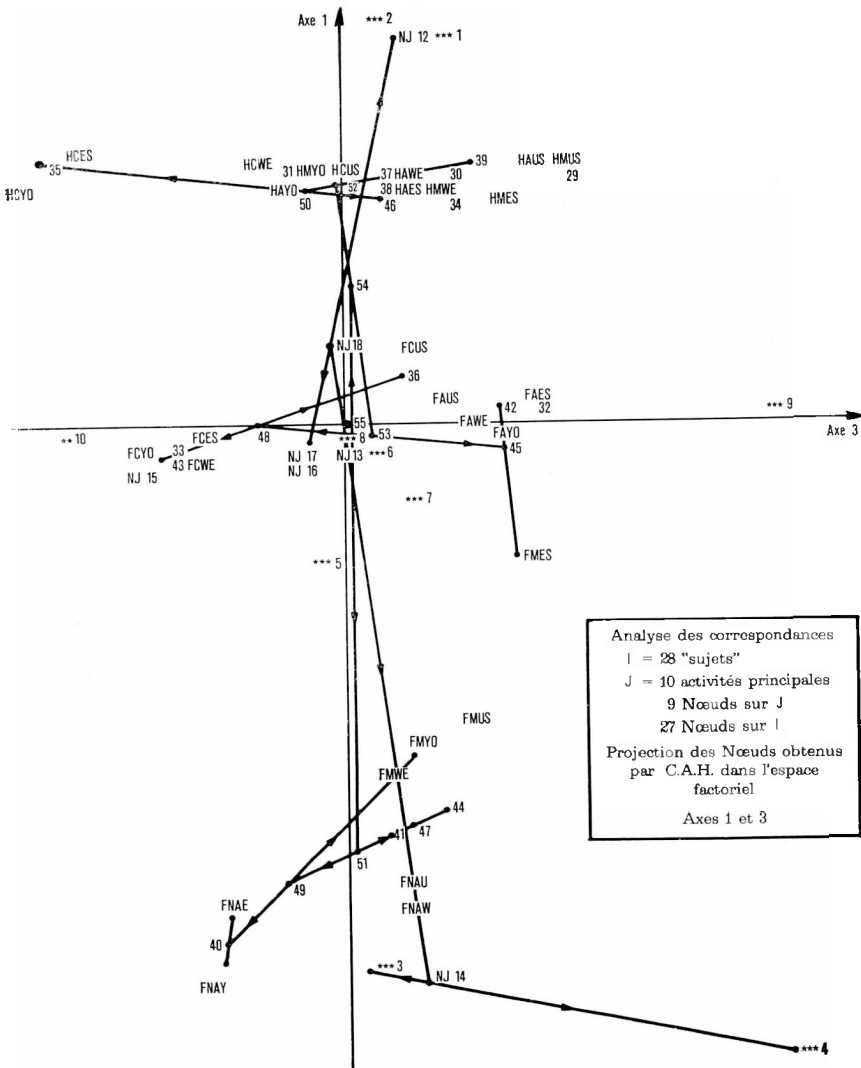
6.3.1. Les contributions de $(i, j) \in I \times J$ à $\text{Lien}_2(I, J)$

On a

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

FIGURE 5

Représentation des classes de $A(I)$ et $B(J)$ dans le plan (1, 3)



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

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

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 citées 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.

	**1	**2	**3	**4	**5	**6	**7	**8	**9	**10
HAUS	0.5	0.4	-1.8	-0.2	0.0	0.0	0.0	0.0	0.6	0.0
FAUS	0.0	0.0	0.0	-0.0	0.1	0.1	0.0	0.0	0.0	0.1
FNAU	-4.5	-0.9	1.8	1.8	0.4	0.0	0.0	0.0	0.4	0.2
HMUS	0.6	0.4	-1.7	-0.2	0.0	0.0	0.0	0.0	0.7	0.1
FMUS	1.7	-0.4	0.8	0.9	0.3	0.0	0.0	0.0	0.2	0.0
HCUS	0.4	0.1	-1.9	-0.3	0.2	0.0	0.0	0.0	0.3	0.0
FCUS	0.0	0.0	0.2	-0.1	0.1	0.1	0.0	0.0	0.1	0.0
HAWE	1.0	0.0	-1.2	-0.2	-0.3	0.0	0.1	0.0	0.0	0.0
FAWE	0.1	-0.0	0.0	-0.0	-0.1	0.0	0.1	0.0	0.0	0.2
FNAW	-4.2	-0.8	3.2	0.9	0.0	0.0	0.3	0.0	0.1	0.0
HMWE	1.0	0.0	-1.2	-0.2	-0.3	0.0	0.1	0.0	0.1	0.0
FMWE	2.0	-0.5	2.1	-0.4	0.0	-0.0	0.2	0.0	0.0	0.0
HCWE	0.9	0.0	-1.6	-0.3	-0.2	-0.0	0.0	0.0	0.0	0.0
FCWE	0.1	-0.3	0.0	-0.1	0.0	0.0	0.1	0.1	0.0	0.1
HAYO	0.9	0.4	0.9	-0.1	0.1	0.0	0.0	0.0	-0.1	0.0
FAYO	0.3	0.0	0.4	-0.0	0.0	0.0	0.0	0.0	-0.2	0.4
FNAY	-4.5	0.7	7.0	-0.1	0.1	0.0	0.0	0.0	-0.2	0.0
HMYO	0.9	0.4	1.0	-0.1	0.1	0.0	0.0	0.0	-0.0	0.0
FMYO	0.8	-0.1	3.4	-0.2	0.0	0.0	0.0	0.0	-0.1	0.1
HCYO	0.6	0.2	-1.2	-0.3	0.0	0.0	0.1	0.0	-0.4	0.5
FCYO	0.0	0.0	0.1	-0.0	0.0	0.0	0.0	0.0	-0.3	0.1
HAES	0.9	0.4	0.9	-0.0	-0.1	0.0	0.0	0.0	-0.0	0.0
FAES	0.4	0.0	0.1	-0.0	0.0	0.0	0.1	0.0	-0.1	0.4
FNAE	-4.2	0.7	3.8	-0.5	0.2	0.0	0.0	0.0	0.0	0.1
HMES	1.0	0.3	0.8	-0.0	-0.2	0.0	0.0	0.0	0.1	0.0
FMES	0.0	-0.0	0.9	-0.2	0.0	0.0	0.0	0.0	-0.1	0.4
HCES	0.7	0.5	-1.6	-0.3	0.0	0.0	0.1	0.0	-0.2	0.4
FCES	0.0	0.0	0.0	-0.0	0.0	0.0	0.0	0.0	-0.2	0.0

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 paragraphe 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

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

On pose

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

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 $\rho^2(q)$, (SC_{qj}) ou $(SC_{qj}/\rho^2(q))$, $\forall j \in J, \forall q \in Q$.

$\rho^2(q)$ 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 $N_j(I)$.

Contributions absolues signées

N	RHO**2	***1	***2	***3	***4	***5	***6	***7	***8	***9	**10
29	0.14	0.02	0.01	-0.07	-0.01	0.00	-0.00	-0.00	-0.00	0.03	-0.00
30	0.11	0.04	0.00	-0.05	-0.01	-0.01	-0.00	0.00	0.00	0.00	-0.00
31	0.10	0.04	0.02	-0.04	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.00
32	0.05	0.01	0.00	0.01	0.00	-0.00	-0.00	-0.00	-0.00	-0.01	-0.02
33	0.02	-0.00	0.00	0.00	-0.00	0.00	0.00	-0.00	0.00	-0.01	0.00
34	0.09	0.04	0.01	-0.03	-0.00	-0.01	-0.00	-0.00	-0.00	0.00	-0.00
35	0.14	0.03	0.01	-0.06	-0.01	-0.00	-0.00	-0.00	-0.00	-0.01	0.02
36	0.02	0.00	0.00	-0.00	-0.00	0.00	0.00	-0.00	-0.00	0.00	-0.00
37	0.12	0.04	0.00	-0.05	-0.01	-0.01	-0.00	0.00	0.00	0.00	-0.00
38	0.10	0.04	0.01	-0.04	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
39	0.14	0.02	0.01	0.07	-0.01	0.00	0.00	0.00	-0.00	-0.02	-0.00
40	0.44	-0.17	-0.03	0.21	0.11	0.01	-0.00	0.00	0.00	-0.00	0.00
41	0.28	-0.12	-0.02	0.10	0.02	0.00	-0.00	0.01	0.00	0.00	0.00
42	0.03	0.01	0.00	0.01	0.00	0.00	-0.00	-0.00	-0.00	-0.00	-0.01
43	0.01	-0.00	-0.00	0.00	-0.00	-0.00	0.00	-0.00	0.00	-0.01	0.00
44	0.27	-0.12	-0.02	0.05	0.05	0.01	0.00	0.00	0.00	0.01	0.00
45	0.04	0.00	0.00	0.01	0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.01
46	0.10	0.04	0.01	-0.04	-0.01	-0.01	-0.00	0.00	-0.00	0.00	-0.00
47	0.27	-0.12	-0.02	0.07	0.04	0.00	0.00	0.00	0.00	0.01	0.00
48	0.01	-0.00	-0.00	-0.00	-0.00	0.00	0.00	-0.00	0.00	-0.00	0.00
49	0.34	-0.11	-0.02	0.18	0.01	0.00	-0.00	0.00	0.00	-0.00	0.00
50	0.10	0.04	0.01	-0.05	-0.01	-0.00	-0.00	-0.00	-0.00	-0.00	0.00
51	0.28	-0.12	-0.02	0.11	0.02	0.00	-0.00	0.00	0.00	0.00	0.00
52	0.10	0.03	0.01	-0.05	-0.01	-0.00	-0.00	-0.00	-0.00	0.00	0.00
53	0.01	0.00	-0.00	-0.00	0.00	0.00	0.00	-0.00	0.00	-0.00	-0.00
54	0.03	0.01	0.00	-0.01	-0.00	-0.00	0.00	-0.00	-0.00	-0.00	-0.00

La colonne N contient les numéros des classes.

La colonne RHO**2 contient $\rho^2(N)$.

Les autres colonnes contiennent les contributions absolues signées.

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,

Contributions relatives signées

N	RHO*2	***1	***2	***3	***4	***5	***6	***7	***8	***9	**10
29	0.14	0.17	0.10	-0.48	-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	0.09	-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	0.36	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

surtout aux niveaux inférieurs. Ce sont précisément ces variables qui font la différence à ces niveaux. On peut citer les exemples suivants :

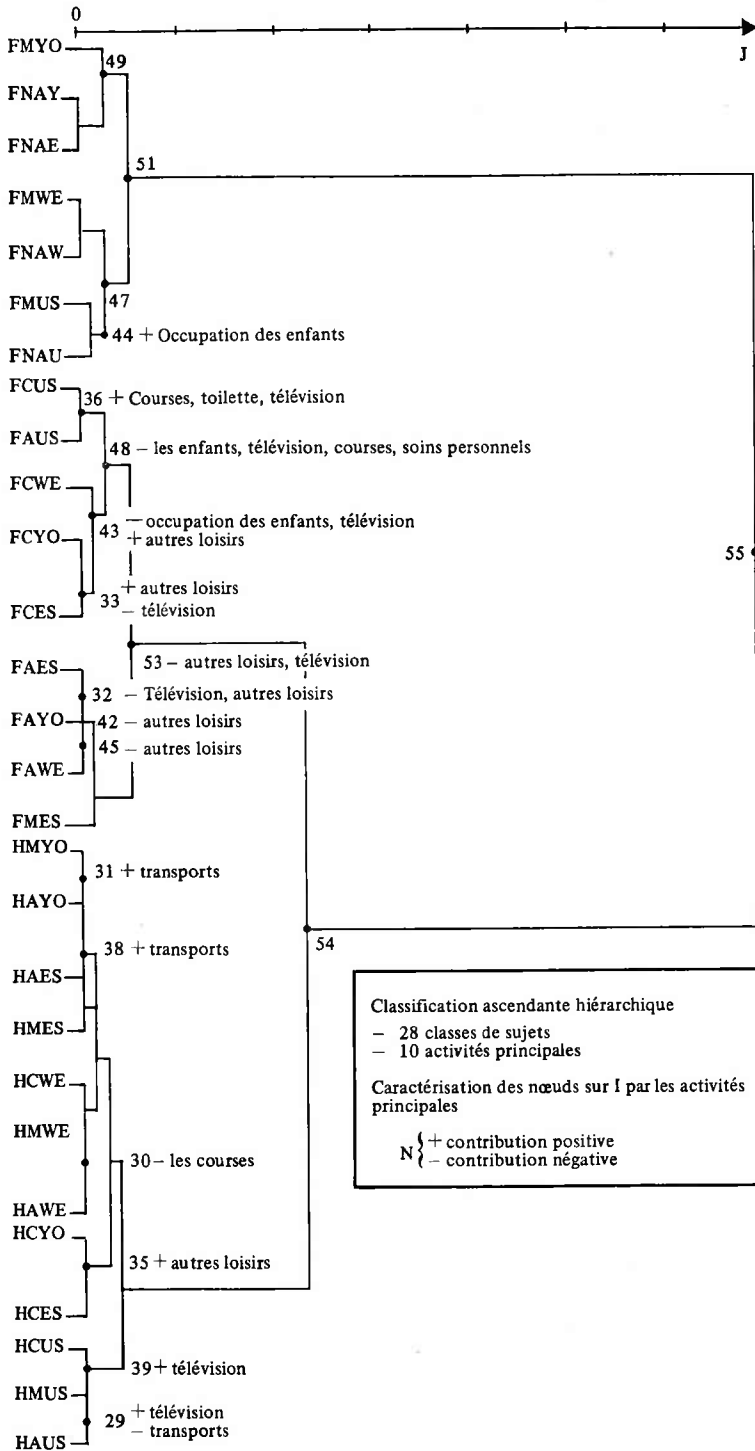
classe 29	HMUS — HAUS	C ⁺ télévision, transports
classe 30	HMWE — HAWE	C ⁻ temps passé à faire les courses
classe 31	HMYO — HAYO	C ⁺ transports
classe 32	FAES — FAYO	C ⁻ télévision, autres loisirs
classe 33	FCYO — FCES	C ⁻ télévision
		C ⁺ autres loisirs
classe 35	HCYO — HCES	C ⁺ autres loisirs
classe 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 — FAWE	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 — FAWE — 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

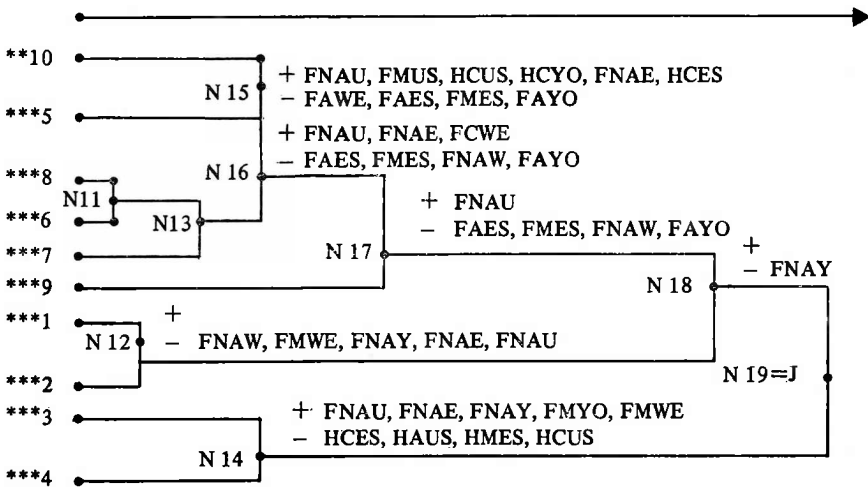
Sur ce résumé, on n'a pas reporté les contributions des variables ***1 et 3***. Dans le but de mieux illustrer les résultats, on propose de reporter ceux-ci sur la représentation de la hiérarchie en procédant de la façon suivante. A chaque nœud de la hiérarchie, on associe leurs variables discriminantes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

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	RHO**2	HAUS	FAUS	FNAU	HMUS	FMUS	HCUS	FCUS	HAWE	FAWE
11	0.00	-0.03	0.01	0.01	-0.03	0.00	-0.01	0.02	0.01	0.04
12	0.25	0.02	0.00	-0.14	0.02	-0.05	0.01	0.00	0.02	0.00
13	0.00	-0.01	-0.00	0.01	-0.01	0.00	-0.02	0.00	0.03	0.04
14	0.51	-0.04	-0.00	0.06	-0.04	0.03	-0.05	-0.01	-0.03	-0.00
15	0.03	-0.00	-0.00	0.12	-0.01	0.04	0.04	0.00	-0.03	-0.08
16	0.00	-0.01	-0.00	0.11	-0.02	0.03	0.01	0.00	-0.00	-0.01
17	0.00	0.00	0.00	0.16	0.00	0.05	0.03	0.01	-0.00	-0.02
18	0.01	0.04	0.00	-0.06	0.04	-0.03	0.05	0.01	0.03	-0.00
N	RHO**2	HAYO	FCYO	HAES	FAES	FNAE	HMES	FMES	HCES	FCES
11	0.00	-0.04	-0.00	-0.02	-0.05	0.11	-0.02	-0.01	-0.01	0.02
12	0.25	0.02	-0.00	0.03	0.01	-0.13	0.03	-0.00	0.03	-0.00
13	0.00	-0.06	-0.01	-0.02	-0.05	0.06	-0.02	-0.01	-0.04	-0.00
14	0.51	-0.03	0.00	-0.02	0.00	0.09	-0.02	0.02	-0.04	0.00
15	0.03	0.11	0.02	-0.01	-0.09	0.06	-0.04	-0.07	0.05	0.02
16	0.00	0.02	0.01	-0.03	-0.13	0.10	-0.05	-0.07	0.01	0.01
17	0.00	0.00	-0.00	-0.02	-0.14	0.06	-0.03	-0.08	0.00	0.00
18	0.01	0.03	-0.00	0.02	-0.00	-0.09	0.02	-0.02	0.04	-0.00
	FNAW	HMWE	FMWE	HCWE	FCWE	HAYO	FAYO	FNAY	HMYO	FMYO
	0.11	0.01	-0.00	0.00	0.26	-0.04	-0.08	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	0.08
	0.09	-0.03	0.05	-0.04	-0.00	-0.02	0.01	0.15	-0.02	-0.01
	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
	0.08	-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	-0.08



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.



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

classe 11	sommeil, toilette	C ⁺	FNAW — FNAE — FCWE — FAWE
(centrale)		C ⁻	FAES — FAYO — HCYO
classe 12	travail, transports	C ⁻	FNAW — FMWE — FNAY — FNAE
[ρ ² (q) élevé]			FNAU
		C ⁺	
classe 13	sommeil, toilette,	C ⁺	FNAW — FCWE — FAWE
(centrale)	repas	C ⁻	FAYO — HCYO — FAES — HCES
			FAYO
classe 14	temps aux enfants,	C ⁺	FNAU — FNAE — FNAY — FMYO
[ρ ² (q) élevé]	travail ménager		FMWE
		C ⁻	HCES — HAUS — HMUS — HCUS
classe 15	courses, autres	C ⁺	FNAU — FMUS — HCUS — HCYO
loisirs			FNAE — HCES
		C ⁻	FAWE — FAES — FMES — FAYO
classes 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

6.3.3. Comparaisons des classes q de Q , partition de I , et des classes s de K , partition de I

Pour améliorer l'interprétation de la constitution des classes, on peut effectuer les calculs suivants :

● Calcul des Liens entre partitions successives créées par classification sur I et sur J .

On a

$$\begin{aligned} \text{Lien}_2(Q, K) &= \|f_{QK} - f_Q f_K\|_{f_Q f_K}^2 \\ &= \sum_{\substack{q \in Q \\ k \in K}} \frac{(f_{qk} - f_q f_k)^2}{f_q f_k} \\ &= \sum_{q \in Q} f_q \|f_K^q - f_K\|_{f_K}^2 = \sum_{q \in Q} f_q \rho_K^2(q) \\ &= \sum_{k \in K} f_k \|f_q^k - f_q\|_{f_q}^2 = \sum_{k \in K} f_k \rho_Q^2(k), \end{aligned}$$

$\rho_K^2(q)$ est la distance de la classe q de Q au centre de gravité du nuage $N_K(Q)$.

Les quantités intéressantes sont :

a) le tableau des Liens entre partitions de I et de J constituées séparément sur les hiérarchies $A(I)$ et $B(J)$;

b) les profils

$$\begin{aligned} f_q^k &= \left\{ \frac{f_{qk}}{f_k}, k \in K \right\}, \\ f_k^q &= \left\{ \frac{f_{qk}}{f_q}, q \in Q \right\}; \end{aligned}$$

c) les rapports $\{f_{qk}/f_q f_k; q \in Q, k \in K\}$.

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

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

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

indice
de stratification
du noeud n

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

On a de plus

$$\sum_{n \in \text{Nod}(A(I))} \mathcal{J}(n) = \text{Lien}_2(I, J) = M^2(N_J(I)) = M^2(N_I(J)) \\ = \sum_{\alpha \in A} \lambda_\alpha \quad (\text{cf. } \S 4 \text{ Analyse des correspondances}).$$

Or

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

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

a) Tableau des dépendances globales entre partitions de $I(Q)$ et de $J(K)$ déduites de $A(I)$ et de $B(J)$

$$L(Q, K) = \text{Lien}_2(Q, K)$$

		K=1	K=2	K=3	K=4	K=5	K=6	K=7	K=8	K=9
		NJ 11	NJ 12	NJ 13	NJ 14	NJ 15	NJ 16	NJ 17	NJ 18	L(Q,J)
Q=1	NJ 29	0.143	0.142	0.140	0.137	0.134	0.130	0.125	0.075	0.143
	NJ 30	0.143	0.142	0.140	0.137	0.134	0.130	0.125	0.075	0.143
	NJ 31	0.143	0.142	0.140	0.137	0.134	0.130	0.125	0.075	0.143
	NJ 32	0.143	0.141	0.140	0.137	0.134	0.130	0.125	0.075	0.143
	NJ 33	0.142	0.141	0.140	0.137	0.134	0.130	0.125	0.075	0.143
	NJ 34	0.142	0.141	0.140	0.137	0.134	0.130	0.125	0.075	0.143
	NJ 35	0.142	0.141	0.140	0.137	0.134	0.130	0.125	0.075	0.143
	NJ 36	0.142	0.141	0.139	0.137	0.134	0.130	0.124	0.075	0.143
	NJ 37	0.142	0.141	0.139	0.137	0.133	0.130	0.124	0.075	0.143
	NJ 38	0.142	0.141	0.139	0.136	0.133	0.130	0.124	0.075	0.142
	NJ 39	0.141	0.140	0.139	0.136	0.133	0.130	0.124	0.075	0.142
	NJ 40	0.141	0.140	0.138	0.136	0.133	0.129	0.124	0.075	0.141
	NJ 41	0.140	0.139	0.137	0.135	0.132	0.129	0.123	0.074	0.141
	NJ 42	0.139	0.138	0.137	0.135	0.131	0.128	0.123	0.074	0.140
NJ 43	0.138	0.138	0.136	0.134	0.131	0.128	0.123	0.074	0.139	
NJ 44	0.137	0.137	0.135	0.133	0.130	0.127	0.122	0.074	0.138	
NJ 45	0.137	0.136	0.135	0.132	0.129	0.126	0.121	0.073	0.137	
NJ 46	0.135	0.135	0.134	0.132	0.129	0.126	0.121	0.073	0.136	
NJ 47	0.134	0.133	0.133	0.131	0.128	0.125	0.120	0.073	0.134	
NJ 48	0.132	0.132	0.131	0.129	0.127	0.124	0.120	0.073	0.132	
NJ 49	0.130	0.129	0.129	0.127	0.125	0.122	0.118	0.072	0.130	
NJ 50	0.127	0.127	0.127	0.125	0.123	0.121	0.118	0.072	0.128	
NJ 51	0.124	0.124	0.123	0.123	0.121	0.119	0.117	0.071	0.124	
NJ 52	0.121	0.120	0.120	0.119	0.118	0.117	0.116	0.071	0.121	
NJ 53	0.117	0.116	0.116	0.116	0.114	0.114	0.113	0.069	0.117	
Q=27	NJ 54	0.095	0.095	0.094	0.094	0.094	0.093	0.093	0.053	0.095
Q=28	L(I,K)	0.143	0.142	0.140	0.137	0.134	0.130	0.125	0.075	0.143

où

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

et où s est centre de gravité de la classe s en posant

$$p(n) = \frac{f_t f_s}{f_t + f_s},$$

on a

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

En posant

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

b 1) Tableau des profils $f_q^k = \{f_{qk}/f_k, k \in K\}, q \in Q$

	NJ 11	NJ 12	NJ 13	NJ 14	NJ 15	NJ 16	NJ 17	NJ 18	J
NI 29	0.07	0.10	0.07	0.02	0.07	0.07	0.07	0.08	0.07
NI 30	0.07	0.10	0.07	0.02	0.06	0.07	0.07	0.08	0.07
NI 31	0.07	0.11	0.07	0.03	0.07	0.07	0.07	0.08	0.07
NI 32	0.07	0.09	0.07	0.09	0.05	0.06	0.08	0.07	0.07
NI 33	0.07	0.07	0.07	0.08	0.08	0.07	0.07	0.07	0.07
NI 34	0.07	0.11	0.07	0.03	0.06	0.07	0.07	0.08	0.07
NI 35	0.07	0.10	0.07	0.02	0.09	0.07	0.07	0.08	0.07
NI 36	0.07	0.08	0.07	0.06	0.07	0.07	0.07	0.07	0.07
NI 37	0.11	0.15	0.11	0.03	0.09	0.11	0.11	0.12	0.11
NI 38	0.14	0.21	0.14	0.06	0.13	0.14	0.13	0.15	0.14
NI 39	0.10	0.15	0.10	0.02	0.11	0.11	0.11	0.12	0.11
NI 40	0.07	0.00	0.07	0.16	0.08	0.08	0.08	0.06	0.07
NI 41	0.07	0.01	0.08	0.14	0.08	0.08	0.08	0.06	0.07
NI 42	0.11	0.13	0.10	0.13	0.08	0.10	0.09	0.10	0.11
NI 43	0.11	0.10	0.11	0.11	0.12	0.11	0.11	0.11	0.11
NI 44	0.07	0.01	0.07	0.13	0.09	0.08	0.08	0.06	0.07
NI 45	0.14	0.16	0.14	0.19	0.11	0.13	0.13	0.14	0.14
NI 46	0.25	0.36	0.25	0.10	0.23	0.24	0.24	0.27	0.25
NI 47	0.15	0.03	0.15	0.27	0.16	0.16	0.16	0.13	0.14
NI 48	0.18	0.17	0.18	0.16	0.19	0.18	0.18	0.18	0.18
NI 49	0.11	0.02	0.11	0.24	0.12	0.11	0.11	0.09	0.11
NI 50	0.32	0.46	0.32	0.12	0.31	0.32	0.31	0.35	0.32
NI 51	0.26	0.05	0.26	0.51	0.28	0.27	0.27	0.21	0.25
NI 52	0.42	0.61	0.42	0.14	0.42	0.42	0.42	0.47	0.43
NI 53	0.32	0.34	0.32	0.35	0.30	0.31	0.31	0.32	0.32
NI 54	0.74	0.95	0.74	0.49	0.72	0.73	0.73	0.79	0.75
I	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

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 du nœud n par rapport à l'indice α .

Dans la pratique, on étudiera les tableaux :

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

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

A^* = espace factoriel retenu.

b 2) Tableau des profils $f_k^q = \{f_{ak}/f_a, q \in Q, k \in K\}$

	NJ 11	NJ 12	NJ 13	NJ 14	NJ 15	NJ 16	NJ 17	NJ 18	J
NI 29	0.36	0.31	0.40	<u>0.03</u>	0.18	0.58	0.66	0.97	1.00
NI 30	0.37	0.31	0.44	<u>0.04</u>	0.16	0.59	0.64	0.96	1.00
NI 31	0.35	0.33	0.40	<u>0.05</u>	0.19	0.58	0.62	0.95	1.00
NI 32	0.35	0.28	0.39	<u>0.17</u>	0.14	0.53	0.99	0.83	1.00
NI 33	0.37	0.22	0.41	0.14	0.21	0.62	0.65	0.86	1.00
NI 34	0.36	0.33	0.40	<u>0.06</u>	0.16	0.56	0.61	0.94	1.00
NI 35	0.36	0.32	0.39	<u>0.03</u>	0.24	0.63	0.69	0.97	1.00
NI 36	0.38	0.24	0.42	<u>0.10</u>	0.19	0.61	0.66	0.90	1.00
NI 37	0.37	0.31	0.43	<u>0.04</u>	0.17	0.60	0.65	0.96	1.00
NI 38	0.36	0.33	0.40	<u>0.06</u>	0.18	0.57	0.61	0.94	1.00
NI 39	0.36	0.31	0.40	<u>0.03</u>	0.19	0.60	0.67	0.97	1.00
NI 40	0.38	<u>0.01</u>	0.44	<u>0.30</u>	0.23	0.66	0.69	0.70	1.00
NI 41	0.38	<u>0.04</u>	0.45	0.26	0.20	0.65	0.70	0.74	1.00
NI 42	0.36	0.27	0.41	0.16	0.14	0.54	0.57	0.84	1.00
NI 43	0.38	0.20	0.43	0.13	0.21	0.64	0.66	0.87	1.00
NI 44	0.37	<u>0.05</u>	0.42	0.23	0.24	0.66	0.72	0.77	1.00
NI 45	0.36	0.25	0.41	0.17	0.14	0.55	0.58	0.83	1.00
NI 46	0.36	0.32	0.41	<u>0.05</u>	0.17	0.59	0.63	0.95	1.00
NI 47	0.37	<u>0.04</u>	0.44	<u>0.24</u>	0.22	0.65	0.71	0.76	1.00
NI 48	0.38	0.22	0.42	0.12	0.20	0.63	0.65	0.88	1.00
NI 49	0.37	<u>0.05</u>	0.43	0.29	0.21	0.63	0.66	0.71	1.00
NI 50	0.36	0.32	0.41	<u>0.05</u>	0.19	0.60	0.63	0.95	1.00
NI 51	0.37	<u>0.05</u>	0.43	<u>0.26</u>	0.21	0.65	0.69	0.74	1.00
NI 52	0.36	<u>0.32</u>	0.41	<u>0.04</u>	0.19	0.60	0.64	0.96	1.00
NI 53	0.37	0.23	0.41	0.14	0.18	0.59	0.62	0.86	1.00
NI 54	0.36	0.28	0.41	<u>0.08</u>	0.18	0.59	0.63	0.92	1.00
I	0.37	0.22	0.42	0.13	0.19	0.61	0.65	0.87	1.00

c) Tableau des rapports $\{f_{qk}/f_q f_k, k \in K, q \in Q\}$

	NJ 11	NJ 12	NJ 13	NJ 14	NJ 15	NJ 16	NJ 17	NJ 18	J
NI 29	0.97	1.41	0.97	0.23	0.94	0.96	1.01	1.11	1.00
NI 30	1.01	1.41	1.05	0.34	0.83	0.98	0.99	1.10	1.00
NI 31	0.97	1.48	0.96	0.42	0.98	0.96	0.95	1.09	1.00
NI 32	0.95	1.26	0.94	1.29	0.72	0.87	0.85	0.96	1.00
NI 33	1.01	0.97	0.99	1.06	1.13	1.03	1.00	0.99	1.00
NI 34	0.97	1.47	0.96	0.48	0.86	0.93	0.94	1.08	1.00
NI 35	0.97	1.42	0.94	0.26	1.25	1.04	1.00	1.11	1.00
NI 36	1.02	1.07	1.00	0.80	1.01	1.00	1.02	1.03	1.00
NI 37	1.01	1.40	1.04	0.30	0.88	0.99	1.00	1.10	1.00
NI 38	0.97	1.48	0.96	0.45	0.92	0.95	0.94	1.08	1.00
NI 39	0.97	1.37	0.97	0.21	1.01	0.98	1.03	1.12	1.00
NI 40	1.04	0.05	1.05	2.31	1.18	1.09	1.07	0.81	1.00
NI 41	1.03	0.20	1.08	1.98	1.05	1.07	1.08	0.85	1.00
NI 42	0.98	1.20	0.98	1.22	0.73	0.90	0.89	0.97	1.00
NI 43	1.03	0.92	1.03	1.01	1.11	1.05	1.03	1.00	1.00
NI 44	1.01	0.20	1.02	1.79	1.24	1.09	1.12	0.88	1.00
NI 45	0.98	1.14	0.97	1.31	0.74	0.90	0.89	0.95	1.00
NI 46	0.99	1.45	0.99	0.39	0.90	0.97	0.97	1.09	1.00
NI 47	1.02	0.20	1.05	1.89	1.14	1.08	1.10	0.87	1.00
NI 48	1.03	0.98	1.02	0.92	1.07	1.03	1.02	1.01	1.00
NI 49	1.02	0.23	1.02	2.22	1.09	1.04	1.02	0.82	1.00
NI 50	0.98	1.44	0.98	0.36	0.98	0.98	0.98	1.09	1.00
NI 51	1.02	0.21	1.04	2.03	1.12	1.06	1.07	0.85	1.00
NI 52	0.98	1.42	0.98	0.32	0.99	0.98	0.99	1.10	1.00
NI 53	1.01	1.05	1.00	1.10	0.92	0.97	0.96	0.99	1.00
NI 54	0.99	1.26	0.99	0.65	0.96	0.98	0.98	1.05	1.00
I	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Applications à l'exemple

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 :

premier axe :

$I(55,1)$ séparation entre $N 51$ et $N 54$ ($N 55 = N 51 \cup N 54$)

$I(54,1)$ séparation entre $N 53$ et $N 52$ ($N 51 \cap N 54 = \emptyset$)

$I(49,1)$ séparation entre $N 40$ et FMYO

$I(45,1)$ séparation entre $N 42$ et FMES

$I(44,1)$ séparation entre FNAU et FMUS

$I(41,1)$ séparation entre FNAW et FMWE

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

Tableau des contributions absolues *1000

	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	1.48	0.41	0.24	0.00	0.02	0.00
N 50	<u>0.02</u>	0.01	<u>1.95</u>	0.44	0.01	0.00
N 51	0.32	<u>2.13</u>	<u>0.37</u>	0.05	0.41	0.07
N 52	0.01	<u>1.70</u>	0.84	0.42	0.40	0.08
N 53	0.15	<u>2.33</u>	<u>1.39</u>	0.01	0.13	0.00
N 54	<u>19.43</u>	<u>1.76</u>	<u>0.07</u>	0.14	0.19	0.23
N 55	<u>94.82</u>	0.51	0.00	0.01	0.05	0.03

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

	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	0.00	0.00	0.00	0.00
N 31	0.00	0.00	0.00	0.00	0.00	0.00
N 32	0.00	0.00	0.00	0.00	0.00	0.01
N 33	0.00	0.00	0.00	0.00	0.01	0.01
N 34	0.00	0.00	0.01	0.00	0.00	0.00
N 35	0.00	0.00	0.00	0.00	0.01	0.04
N 36	0.00	0.01	0.00	0.00	0.00	0.00
N 37	0.00	0.00	0.03	0.00	0.00	0.00
N 38	0.00	0.00	0.03	0.00	0.00	0.00
N 39	0.00	0.00	0.05	0.01	0.01	0.09
N 40	0.00	0.02	0.00	0.00	0.03	0.05
N 41	0.00	0.00	0.00	0.00	0.00	0.00
N 42	0.00	0.03	0.00	0.08	0.03	0.01
N 43	0.00	0.02	0.00	0.12	0.03	0.06
N 44	0.01	0.01	0.00	0.00	0.02	0.00
N 45	0.01	0.00	0.00	0.01	0.00	0.00
N 46	0.00	0.02	0.00	<u>0.23</u>	0.07	0.05
N 47	0.00	0.05	0.00	<u>0.20</u>	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	<u>0.30</u>	0.10	0.01	0.00
N 51	0.00	<u>0.20</u>	<u>0.06</u>	0.01	<u>0.25</u>	0.09
N 52	0.00	<u>0.16</u>	<u>0.13</u>	0.10	<u>0.24</u>	0.09
N 53	0.00	<u>0.21</u>	<u>0.21</u>	0.00	0.08	0.00
N 54	<u>0.16</u>	<u>0.16</u>	0.01	0.03	0.12	<u>0.27</u>
N 55	<u>0.80</u>	0.05	0.00	0.00	0.03	0.04

Tableau des contributions
relatives à l'indice de diamètre de la hiérarchie : $I(\alpha, N)/\text{Indice}(N)$

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

deuxième axe :

$I(51,2)$ séparation entre N 49 et N 47

$I(52,2)$ séparation entre N 50 et N 39

$I(53,2)$ séparation entre N 48 et 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).

troisième axe :

$I(30,3)$ séparation entre HMWE et HAWE

$I(38,3)$ séparation entre N 31 et N 34

$I(39,3)$ séparation entre (HMUS, HAUS) et HCUS

$I(37,3)$ séparation entre (HMWE, HAWE) et HCWE

$I(30,3)$ séparation entre N 46 et N 35

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.

quatrième axe :

$I(46,4)$ séparation entre (HMYO, HAYO, HMES, HAES)
et (HMWE, HCWE, HAWE)

$I(47,4)$ séparation entre (FMWE, FNAWE) et (FMUS, FNAUS)

$I(43,4)$ séparation entre 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 :

premier axe : $N 55, N 54$ (5 fois moins)

deuxième axe : $N 51, N 52, N 53, N 54$

troisième axe : $N 50$

quatrième 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	<u>44.65</u>	3.99	0.49	0.11	0.04	0.01
N 19	<u>72.04</u>	2.55	0.27	0.03	0.03	0.01

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

	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	<u>0.55</u>
N 13	0.00	0.01	0.01	<u>0.26</u>	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	<u>0.31</u>	0.21	0.04
N 16	0.00	0.08	<u>0.26</u>	0.19	0.03	0.02
N 17	0.00	<u>0.25</u>	<u>0.42</u>	0.00	0.01	0.11
N 18	<u>0.38</u>	<u>0.37</u>	0.07	0.03	0.02	0.01
N 19	<u>0.61</u>	0.24	0.04	0.01	0.02	0.02

Tableau des contributions relatives
à l'indice de diamètre de la hiérarchie : $I(\alpha, N)/\text{Indice}(N)$

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

premier axe : séparation entre trois groupes d'activités : N 17, N 12, N 14;

deuxième axe : séparation entre la télévision et les autres activités (autres loisirs, sommeil, repas, toilette);

troisième axe : séparation entre télévision et autres loisirs;

quatrième axe : séparation principale entre le sommeil, la toilette et le temps passé aux repas.

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] BENZÉCRI (J. P.), *L'Analyse des Données*, tome I, *La Taxinomie*, Dunod, Paris, 1973, *L'Analyse des Données*, tome II, *Correspondances*, Dunod, Paris, 1973.
- [2] JAMBU (M.), *Introduction à l'Analyse des Données. Les méthodes de classification automatique*, *Consommation*, n° 3, 1973, Dunod, Paris.
- [3] JAMBU (M.), *Sur les indices de distances en vue de la construction d'une classification hiérarchique*, *Consommation*, n° 2, 1974, Dunod, Paris.
- [4] JAMBU (M.), *Sur les critères d'agrégation utilisés en classification automatique*, *Consommation*, n° 4, 1974.
- [5] SADAKA (M.), *Comparaison des classes hiérarchiques aux axes factoriels*, Thèse de doctorat de 3^e cycle, 1974.
- [6] JAMBU (M.), *Programme de calcul des contributions mutuelles entre classes d'une hiérarchie et facteurs d'une correspondance*. Cahiers de l'Analyse des Données, n° 1, 1976, Dunod, Paris.
- [7] JAMBU (M.), *Programme de construction de classification ascendante hiérarchique (CAH)*. In *L'Analyse des Données* de J. P. Benzécri et coll., Dunod, Paris, 2^e édition, 1976.
- [8] TABET (N.), *Programme d'Analyse des Correspondances* in *L'Analyse des Données* de J. P. Benzécri et coll., Dunod, Paris, 2^e édition, 1976.

RÉSUMÉS - ABSTRACTS

L'ENQUÊTE DE 1970 SUR LES CONSOMMATIONS MÉDICALES, par A. et A. MIZRAHI. *Consommation*, 2-1976, avril-juin 1976, pages 3 à 67.

Les enquêtes de consommation médicale effectuées auprès des ménages sont un instrument essentiel de l'analyse aussi bien dans les domaines économique et social qu'épidémiologique. On présente ici une synthèse des travaux publiés par l'équipe de micro-économie médicale du C.R.E.D.O.C. sur l'enquête médicale de 1970 (janvier 1970 à janvier 1971). Les consommations étaient observées pendant 12 semaines pour chaque ménage. On observe 4,3 % d'abandons et 7,6 % de refus, et les 7 393 ménages de l'échantillon sont bien représentatifs de la population française des ménages ordinaires.

Divers modes de protection couvrent une part plus ou moins importante des dépenses médicales. La Sécurité Sociale protège 95,5 % de la population dont 48,7 % bénéficient de prestations mutualistes; le taux de mutualistes croît avec le revenu, la catégorie socio-professionnelle, le niveau d'instruction. Il reste 1,60 % de personnes ne bénéficiant d'aucune couverture : ce sont souvent des personnes défavorisées.

Parmi les facteurs qui conditionnent la consommation médicale (en niveau et en structure), l'âge et le sexe jouent un rôle prépondérant aussi bien en ce qui concerne l'hospitalisation que les soins ambulatoires. Les facteurs économiques et sociaux ont également une influence notable; en niveau, certaines couches de la population sous-consomment gravement de tous les types de soins; lorsque les contraintes financières deviennent moins sévères, la consommation médicale augmente, et on observe une saturation à partir des couches moyennes : pour les groupes plus favorisés, aux inégalités, quant au niveau de la

THE 1970 SURVEY OF THE CONSUMPTION OF MEDICAL CARE, by A. and A. MIZRAHI. *Consommation*, 2-1976, April-June 1976, pages 3 to 67.

Households surveys of the consumption of medical care are an essential tool of the economic, social or epidemiological analysis. The paper is a synthesis of work already published by the C.R.E.D.O.C. research unit on medical micro-economics on the results of the 1970 survey (made from January 1970 to January 1971). Each household was under survey during 12 weeks. There were 4,3 % drop out and 7,6 % refused to cooperate, and the 7 393 households of the sample are representative of the french population.

A more or less important share of medical expenses are taken in charge through several systems of insurance. 95,5 % of the population is under the coverage of Social Security, 45,7 % of which also have the benefits of mutualist organisms; the rate of mutualists, increases with income, occupational group and level of education. 1,60 % of the population has no coverage : they are often poor people.

Among the factors that influence the consumption of medical care (both its level and its structure) age and sex play a predominant part for hospitalization as well as out-patient care. Economic and social factors also have a notable influence; certain groups have a seriously low level of consumption of any type of care; when financial constraints become less severe, consumption increases, there is a saturation from middle class upwards: for the affluent social groups, inequality in the structure of medical care succeeds to inequality as to its level; for instance, when a professional sees a physician,

consommation médicale globale, succèdent des inégalités quant aux structures; ainsi, par exemple, quand un cadre supérieur voit un médecin, une fois sur deux, il s'agit d'un spécialiste, un ouvrier une fois sur quatre, un salarié agricole une fois sur six seulement.

Près de 3 séances de médecins sur 4 sont effectuées par des généralistes; 32 % des enquêtés ont vu au moins une fois un généraliste et 12 % au moins une fois un spécialiste. Les séances d'électro-radiologie aux malades ambulatoires sont effectuées, pour la plupart, par des radiologues (46,5 %), mais également par des généralistes, des spécialistes, des dentistes. Les analyses de laboratoire sont essentiellement des examens hématologiques (41,9 %) et des examens chimiques des urines (37,9 %); les motifs les plus fréquents des analyses sont les maladies endocrines et métaboliques et la pathologie cardio-vasculaire. Une acquisition pharmaceutique (achat un jour donné d'un produit bien spécifié) est constituée en moyenne de 1,37 unités de vente; il s'agit essentiellement de thérapeutiques cardio-vasculaires et digestives et d'antalgiques. Les soins de généralistes, auxquels sont associés les soins infirmiers et les produits pharmaceutiques prescrits, sont relativement bien diffusés et accessibles à presque toute la population; par contre, les soins qui contribuent le plus à la croissance des consommations médicales, et dont l'aspect technique est plus accentué, atteignent davantage les couches les plus favorisées.

Une étude approfondie des prix effectivement pratiqués permet, d'une part, par comparaison aux tarifs, d'estimer des taux de dépassement en 1970 (ils sont environ de 6 % pour les généralistes et de 12 % pour les spécialistes), d'autre part, de construire une estimation globale de la consommation et de sa structure en valeur.

Une description détaillée de l'offre de soins médicaux montre que si 36 % des communes ne disposent pas sur place des services d'un généraliste, elles en disposent toujours à moins de 20 km; à moins de 40 km, on trouve toujours un chirurgien, un radiologue, un anesthésiste. La zone la plus significative correspond à un cercle de 20 km de rayon; c'est au niveau de l'existence des praticiens que se trouvent les disparités les plus importantes, les différences dans les densités s'en déduisent.

Enfin, l'enquête permet différentes mesures de l'état de santé de la population, selon le niveau de l'approche (nosologique, fonctionnelle, ou des répercussions

once out of twice, it is a specialist; the proportion is once out of four in the case of a blue collar, once out of six in the case of a farm-worker.

Nearly 3 physician visits out of four are with a G.P.; 32 % of the answerers had seen a G.P. at least once, and 12 % at least once a specialist. Electro-radiology care is mostly given at the radiologist's (46,5 %), but also at the G.P.'s, specialists' or dentists'. Laboratory tests are mostly blood tests (41,9 %) and chemical tests of urine (37,9 %); the most frequent motives of laboratory tests are endocrine and metabolic diseases and cardio-vascular pathology. The acquisition of drugs (the purchase on a given day of a specified product) consists of an average 1,37 packages; they are mostly cardio-vascular, digestive and analgic therapeutics. The care by G.P.s to which are associated care by nurses and prescribed drugs, are comparatively well distributed and accessible to all the population; but affluent groups benefit more of the care that contributes the most to the increase of the consumption of medical care, and that has the most advanced technical aspect.

A study of the prices actually payed, as compared to the tariff rates shows the rate of over-pricing in 1970 (about 6 % for G.P.s and 12 % for specialists), an estimate of the consumption of medical care and its structure at real price is deduced.

A detailed description of the offer of medical care indicates that whereas 36 % of villages do not have a G.P. on the spot, there is always one within a distance of 20 km; within a distance of 40 km, there is always a surgeon, a radiologist and an anesthetist. The most significant area corresponds to a 20 km radius circle; the most significant differences are at the level of the existence of physicians, differences in densities can be deduced.

Last, the results can be used to measure in different ways the state of health of the population according to the approach (nosologic, functional or socio-economic

socio-économiques). Ainsi, 0,68 % de l'échantillon ne sort jamais (dont 0,19 % de grabataires) et on peut estimer à 11 le nombre de jours d'interruption d'activité par personne et par an. La relation avec l'âge et le sexe des taux de personnes atteintes par les différentes maladies dépend du type de pathologie considéré. Les enquêtes auprès d'échantillons aléatoires importants de ménages apportent des informations sur la morbidité de toute la population, elles permettent de cadrer les données obtenues à d'autres sources : statistiques administratives ou enquêtes auprès des producteurs.

Étant donné la très grande richesse des données recueillies, l'enquête médicale de 1970 a permis de développer une gamme de recherches extrêmement variées; encore, l'ensemble du champ n'a-t-il pas été entièrement exploré, et plusieurs projets ouvrent la voie aux recherches à venir.

effects). For instance, 0,68 % of the sample never goes out of doors (0,19 % of which are bed-ridden), the number of days a year lost for work per capita is 11. The relation between age, sex and the rates of people having different illnesses depends on the kind of disease. Households surveys of large random samples yield information on the morbidity of the whole population, the results give a frame to data gathered through other sources : administrative statistics or surveys of producers.

Because of the richness of the collected data, the 1970 medical care survey is at the root of a large range of various research projects; the whole field has not yet been explored but several schemes open the road to further research.

COMPARAISON D'UN MODÈLE FACTORIEL ET D'UN MODÈLE HIÉRARCHIQUE. Application à l'analyse des budgets-temps, par M. JAMBU. *Consommation*, 2-1976, avril-juin 1976, pages 69 à 108.

A FACTORIAL MODEL COMPARED TO A HIERARCHICAL MODEL. Application to the time-budget analysis, by M. JAMBU. *Consommation*, 2-1976, April-June 1976, pages 69 to 108.

L'analyse des correspondances a le mérite de faire figurer sur un même graphique deux ensembles *I* et *J* en relation. Ainsi, interprète-t-on chaque structure par l'autre. Cependant, les taxinomistes doivent souvent fixer en un système rigide de classes la vision spatiale que leur offre l'analyse des correspondances.

Through correspondance analysis, 2 sets *I* and *J* that are related can figure on the same graph. Each structure can therefore be interpreted by the other. But taxinomists often have to fix into a rigid system of classes the spatial vision given through correspondance analysis.

L'auteur propose ici un ensemble de calculs qui permettent de comparer deux types d'analyses qui, utilisées conjointement, enrichissent considérablement l'interprétation soit des facteurs soit des classes. L'auteur illustre ces calculs d'un exemple concret portant sur une étude internationale des budgets-temps, de façon à permettre au lecteur de reprendre ces calculs et leur interprétation statistique.

The author explains the calculations to compare the two types of analysis which, when jointly applied, considerably enrich the interpretation either of the factors or of the classes. An application is given on the results of a time-budget survey, to allow the reader to go over the calculations and their statistical interpretation.

*Parce que
l'avènement de l'ordinateur
exige la rénovation de
l'enseignement de la statistique :*

statistique et informatique appliquées

par

L. Lebart et J.-P. Fénélon

3^e édition 1975, 448 pages, 155 x 240 : 58 F

et son complément

exercices commentés de statistique et informatique appliquées

par l'équipe

Ronald Céhessat

1^{re} édition 1976, 432 pages, 155 x 240 : 58 F

dunod

Prenez "FAITS ET CAUSES" avec

les cahiers de

L'ANALYSE Des DONNÉES

LES FAITS : c'est la multitude des observations, des relevés, des résultats d'expériences ou de sondages qu'accumule le spécialiste d'une science de l'homme ou de la nature : biologiste, écologiste, psychologue, historien, archéologue, économiste, spécialiste d'études de marché,...

LES CAUSES : c'est la signification profonde des phénomènes observés telle que l'analyse statistique multidimensionnelle permet de la dégager en toute **neutralité** statistique.

LES CAHIERS DE L'ANALYSE DES DONNÉES, revue trimestrielle sous la direction du Professeur J.-P. Benzécri, publiée avec le concours du CNRS - Nombreux exemples d'application, exposé des méthodes, programmes de calcul, chroniques,...

Revue trimestrielle 4 × 96 pages - Format 15 × 24

RENSEIGNEMENTS ET ABONNEMENTS BP 119 - 93104 Montreuil Cedex

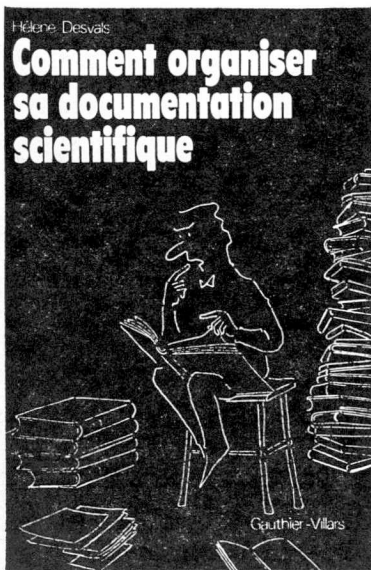
Tarif France 150 F - Étranger 180 F

publié avec le concours du CNRS

dunod

NOUVEAUTÉ

gauthier-villars



Hélène Desvals

Ingénieur documentaliste au C.N.R.S.

comment organiser sa documentation scientifique

Enfin un guide pratique étudié pour tous les scientifiques: les moyens et les outils de la documentation scientifique moderne.

Comment s'en servir rationnellement: les références bibliographiques. Les revues «secondaires». Les articles de périodiques. Le fichier personnel. Les bibliothèques. Les citations dans les mémoires scientifiques. Les communications personnelles. Les tirés à part. Les sources d'information: les revues signalétiques et analytiques. Les thèses. Les ouvrages. Les congrès. Les brevets. Les compilations. La diffusion personnalisée des informations: les services de renseignements «S.V.P.». La documentation commerciale. Editeurs et Libraires. Reproduction et Reprographie. Microfilms. Microcopies. Traductions. Stockage des informations. Classement des documents. Les mots clés. Elaboration d'un fichier personnel: fiches simples, fiches à encoches marginales, fiches à perforations centrales et à sélection visuelle, trucs pour petits budgets.

240 pages, 15,5 x 24, broché, 54 illust.

A lire aussi:

R. GUILLOUX: Réseaux et Systèmes de Documentation

J. REBOUL: Du bon usage des bibliographies.



OFFRE SPECIALE RESERVEE AUX ABONNES ET LECTEURS

BULLETIN DE COMMANDE

à renvoyer aux Editions DUNOD/GAUTHIER-VILLARS,
17, rue Rémy Dumoncel, 75680 PARIS Cedex 14

- Je désire bénéficier de votre offre spéciale réservée aux lecteurs de la revue et recevoir votre ouvrage:

Hélène Desvals: Comment organiser sa documentation scientifique
au prix de 39 F franco

- Je désire recevoir le catalogue de vos publications dans les disciplines suivantes:

.....

NOM et PRENOM:

PROFESSION:

ADRESSE:

CODE POSTAL: VILLE: PAYS:

C.R.E.

REVUES SCIENTIFIQUES ET TECHNIQUES SCIENTIFIC AND TECHNICAL JOURNALS

TARIF 1976 et années antérieures.

Subscription rates for 1976 and former years.

Les abonnements sont payables d'avance et partent de janvier.

All subscriptions are payable in advance and start in January.

GAUTHIER-VILLARS

	France	Étranger FF.
Annales d'Histochemie (4 N ^{os}) - (4 issues), Vol. 21.....	200	230
Annales de l'École Normale Supérieure (4 N ^{os}) - (4 issues), Vol. 9...	290	350
Annales de Physique Biologique et Médicale (4 N ^{os}) - (4 issues), Vol. 10.	130	170
Annales de l'Institut Henri-Poincaré (2 séries);		
Série A (Physique théorique) 2 tomes (8 N ^{os}) - (8 issues), Vol. 24-25.	270	320
Série B (Calcul Prob. et Statist.) 1 tome (4 N ^{os}) - (4 issues), Vol. 12..	140	170
Bulletin des Sciences Mathématiques (4 N ^{os}) - (4 issues), Vol. 100...	230	290
Comptes Rendus de l'Académie des Sciences :		
Hebdomadaire - 2 tomes par an. Weekly - 2 vol/year. Vol. 282-283		
— 1 ^{er} fascicule - Séries AB : Sciences mathématiques et physiques...	780	1 300
— 2 ^e fascicule - Série C : Sciences chimiques.....	680	1 130
— 3 ^e fascicule - Série D : Sciences naturelles.....	750	1 300
Les trois fascicules : AB - C - D.....	1 600	2 800
Journal de Mathématiques pures et appliquées (4 N ^{os}) - (4 issues), Vol. 55.	250	290
Journal de Mécanique (5 N ^{os}) - (5 issues), Vol. 15.....	260	320
Mathématiques et Sciences Humaines (4 N ^{os}) - (4 issues), Vol. 14...	98	114
Œcologia Plantarum (4 N ^{os}) - (4 issues), Vol. 11.....	145	185
Physiologie Végétale (4 N ^{os}) - (4 issues), Vol. 14.....	185	230
Revue d'Écologie et de Biologie du Sol (4 N ^{os}) - (4 issues), Vol. 13.	180	230
Revue de Chimie Minérale (6 N ^{os}) - (6 issues), Vol. 13.....	270	330

DUNOD

Automatisme (8 N ^{os}) - (8 issues).....	135	160
Bulletin d'Informations scientifiques et techniques du C.E.A. (11 N ^{os}) - (11 issues).....	105	128
Consommation (4 N ^{os}) - (4 issues).....	90	115
Construction (11 N ^{os}) - (11 issues).....	120	160
French Railway Techniques (4 N ^{os}) - (4 issues).....	80	105
Journal International de Psychologie (4 N ^{os}) - (4 issues).....	80	95
Matériaux et Constructions, essais et recherches (6 N ^{os}) - (6 issues)...	150	180
Radioprotection (4 N ^{os}) - (4 issues).....	80	98
Revue Française du Marketing (6 N ^{os}) - (6 issues).....	125	145
Revue Française d'Automatique, Informatique, Recherche Opérationnelle (15 N ^{os}) - (15 issues).....	275	320
Revue Générale des Chemins de Fer (11 N ^{os}) - (11 issues).....	120	145
L'Informatique	260	285
— 11 N ^{os} (uniquement par MM. les libraires)		
— 11 issues (subscription through booksellers only)		

NOUVEAUTÉ :

Les Cahiers de l'Analyse des données, (4 numéros) - (4 issues) :
150 F (France) — 180 F (Étranger).

Le renouvellement automatique en fin d'abonnement n'est pas prévu.
Standing orders are not available.

CDR DUNOD-GAUTHIER-VILLARS - B.P. n° 119 93104 Montreuil

CONSOMMATION (ANNALES DU C. R. E. D. O. C.)

1973

- N° 1. — Rôle des valeurs et politique sociale. — A qui profite l'impôt ? Mythes et réalités. — Les entreprises financières en mutation face au commerce de l'épargne. — Les leçons d'une enquête sur les petits commerçants âgés. — Cheminements aléatoires et modèles systématiques d'intervention. Bourse des valeurs de Paris. — Les dépenses de soins médicaux au Canada de 1957 à 1969.
- N° 2. — Consommation des ménages et consommation publique « divisible ». — Inflation et processus de décision. — Vers une description du mode de vie au moyen d'indicateurs.
- N° 3. — Un indicateur de morbidité. — Rémunère-t-on les études ? — Les immigrés : réflexions sur leur insertion sociale et leur intégration juridique. — Introduction à l'analyse des données; les méthodes de classification automatique.
- N° 4. — Un premier bilan de la redistribution des revenus en France : les impôts et cotisations sociales à la charge des ménages en 1965.

1974

- N° 1. — Recherche et politique sociale. — Les facteurs démographiques et la croissance des consommations médicales. — La justice civile, sa place dans la société française.
- N° 2. — La consommation pharmaceutique en 1970. — Une définition des dépenses d'éducation des familles. — L'utilisation des études à long terme dans la planification française. — Sur les indices de distances en vue de la construction d'une classification hiérarchique.
- N° 3. — L'essentiel ou le résidu : le cas de la planification urbaine. — Diffusion des consommations médicales de ville dans la population en 1970. — Les grèves dans l'économie française.
- N° 4. — Aspects géographiques du système des soins médicaux. Analyse des données départementales. — Vieillesse et classe sociale. L'exemple des paysans bénéficiaires de l'I.V.D. et celui des petits commerçants. — Sur les critères d'agrégation utilisés en classification automatique.

1975

- N° 1. — Structure et inégalité des patrimoines. — L'appréciation monétaire d'un surplus dans la consommation alimentaire de différentes catégories sociales. — Quelques critères de comparaison des hiérarchies indicées produites en classification automatique.
- N° 2. — Vers une évaluation de la consommation réelle des ménages. — La justice distributive de l'école. — L'orientation du dépouillement de certaines enquêtes par l'analyse des correspondances multiples.
- N° 3. — Pour une définition de l'environnement. — Note sur la rentabilité du patrimoine des ménages salariés et inactifs.
- N° 4. — Comparaison des prix et des tarifs des services médicaux 1960-1975. — Refus et approbation systématiques dans les enquêtes par sondage. — Notation et classement : deux méthodes de recueil de données.

1976

- N° 1. — Analyse de contenu et contenu de l'analyse. — Application d'un problème de classification avec contrainte de contiguïté.

SOMMAIRE DES PROCHAINS NUMÉROS

La nouvelle théorie du consommateur et la formation des choix. — La représentation des artisans dans la statistique officielle de la France depuis un siècle. — Les bénéficiaires des dépenses de l'Éducation Nationale en 1965 et 1970.

sommaire

Éditorial..... 2

ÉTUDES

ANDRÉE ET ARIÉ MIZRAHI

L'enquête de 1970 sur les consommations médicales. 3

MICHEL JAMBU

Comparaison d'un modèle factoriel et d'un modèle
hiérarchique..... 69

RÉSUMÉS-ABSTRACTS..... 109

**CENTRE DE RECHERCHES
ET DE DOCUMENTATION
SUR LA CONSOMMATION**

45, bd Vincent-Auriol, PARIS-13^e
Tél. 707-97-59

1976 n° 2
Avril-Juin