

# REVUE DE STATISTIQUE APPLIQUÉE

JEAN-JACQUES DENIMAL

## **Application de l'analyse interne multiple à l'étude d'un tableau de contingence à 3 entrées**

*Revue de statistique appliquée*, tome 42, n° 4 (1994), p. 25-37

[http://www.numdam.org/item?id=RSA\\_1994\\_\\_42\\_4\\_25\\_0](http://www.numdam.org/item?id=RSA_1994__42_4_25_0)

© Société française de statistique, 1994, tous droits réservés.

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

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

# APPLICATION DE L'ANALYSE INTERNE MULTIPLE À L'ÉTUDE D'UN TABLEAU DE CONTINGENCE À 3 ENTRÉES

Jean-Jacques DENIMAL

Laboratoire de Statistique et Probabilités  
Université des sciences et Technologies de Lille  
59655 - Villeneuve d'Ascq cedex, France

## RÉSUMÉ

Dans le cadre de l'analyse des correspondances, et appliquant des méthodes développées dans le cas de données structurées, cet article propose une nouvelle approche, permettant une analyse plus complète d'un tableau de contingence à 3 entrées.

**AMS Classification :** 62XX07

**Mots-clés :** *Analyse des correspondances – analyse interne multiple – tableau de contingence à 3 entrées.*

## SUMMARY

Using correspondence analysis and some methods developed in the case of structured data, this paper introduces a new approach to study a three way contingency table.

**Keywords :** *Correspondence analysis – multiple internal analysis – three way contingency table.*

## 1. Introduction

### 1.1. Analyse interne multiple (Rappels)

Dans le cadre de l'analyse des correspondances, l'analyse interne multiple (introduite par CAZES-MOREAU en 1991) s'obtient à partir d'un tableau de contingence multiple  $k_{J_1 \times J_2 \times \dots \times J_p}$ , défini à partir d'un produit d'ensembles finis  $J_1 \times J_2 \times \dots \times J_p$ , chaque ensemble  $J_k$  ( $1 \leq k \leq p$ ) étant supposé muni d'une partition  $P_{J_k}$ . D'autre part, cet ensemble  $J_k$  sera considéré comme l'ensemble des modalités d'une variable qualitative définie sur un ensemble  $I$  d'individus. Au tableau de contingence multiple précédent, sera associé le tableau disjonctif complet  $k_{IC}$  croisant  $I$  et la réunion  $C = \bigcup_{k=1}^p J_k$ .

Par définition, l'analyse interne multiple de  $k_{J_1 \times J_2 \times \dots \times J_p}$ , relativement aux partitions  $\{P_{J_k}/1 \leq k \leq p\}$  est l'analyse factorielle intra-classe du tableau disjonctif complet  $k_{IC}$  relativement à la partition  $P_C$  de  $C$  définie par la réunion des partitions  $P_{J_k}$  ( $P_C = \bigcup_{k=1}^p P_{J_k}$ ).

### 1.2. Conséquences

Lorsque  $p = 2$ , on retrouve l'analyse interne classique (introduite par CAZES - CHESSEL - DOLEDEC en 1988), mettant en jeu un tableau de contingence classique  $k_{J_1 J_2}$  croisant 2 ensembles  $J_1$  et  $J_2$ , munis respectivement des partitions  $P_{J_1}$  et  $P_{J_2}$ . Ces 2 partitions définissent sur  $k_{J_1 J_2}$  une structure en blocs, et l'on montre que l'analyse interne de  $k_{J_1 J_2}$  a pour but de ne retenir que la structure interne des blocs, et peut s'interpréter comme l'analyse des correspondances de  $k_{J_1 J_2}$ , l'influence des 2 partitions  $P_{J_1}$  et  $P_{J_2}$  étant supprimée. On trouvera, dans l'article cité ci-dessus, une application de la méthode en hydrobiologie.

### 1.3. Applications

Nous allons appliquer l'analyse interne multiple, dans le cadre d'un tableau ternaire particulier (tableau de contingence à 3 entrées)  $k_{J_1 \times J_2 \times J_3}$ ;  $J_1, J_2, J_3$  représentant les ensembles de modalités de trois couples de variables qualitatives  $(v_1, v_2), (v_2, v_3), (v_3, v_1)$  définies sur un ensemble  $I$  d'individus.

$P_1, P_2, P_3$  représentant les partitions de  $I$ , définies par  $v_1, v_2, v_3$ , et notant, d'une part  $c_r, c_s, c_l$  ( $(r, s, l)$  désignant une permutation de  $\{1, 2, 3\}$ ) 3 classes appartenant respectivement à  $P_r, P_s, P_l$ , et d'autre part  $n_{rsl} = \text{card}(c_r \cap c_s \cap c_l)$  (de même,  $n_{rs} = \text{card}(c_r \cap c_s) \dots$  etc), nous montrerons que l'analyse interne multiple de ce tableau ternaire  $k_{J_1 \times J_2 \times J_3}$ , relativement à des partitions bien choisies de  $J_1, J_2, J_3$ , va permettre de comparer les rapports  $\frac{n_{rsl}}{n_{rs}}$  et  $\frac{n_{rl}}{n_r}$ , autrement dit les lois conditionnelles associées, et fournira une méthode d'analyse de ce tableau ternaire, dont les résultats viendront compléter ceux issus de l'analyse des correspondances du tableau disjonctif complet  $k_{IC}$  croisant  $I$  et la réunion  $C = J_1 \cup J_2 \cup J_3$ .

Dans le cas particulier où les 3 variables qualitatives  $v_1, v_2, v_3$  précédentes n'ont chacune que 2 modalités, cette analyse interne multiple (équivalente à une analyse en composantes principales normée, § 3.3.) sera illustrée par une application à un jeu de données (§ 4).

## 2. Le tableau analysé : définition et propriétés

### 2.1. Définition

On considère, donc, sur un ensemble  $I$  d'individus, trois variables qualitatives  $v_1, v_2, v_3$  définissant 3 partitions de  $I$ , notées  $P_1, P_2, P_3$ . Trois nouvelles partitions

de  $I$ , notées  $P_{12}, P_{23}, P_{31}$  sont obtenues par intersection :

$$P_{12} = P_1 \cap P_2 \quad P_{23} = P_2 \cap P_3 \quad P_{31} = P_3 \cap P_1$$

La partition  $P_1$  de  $I$  (de même, pour les partitions  $P_2$  et  $P_3$ ) sera également considérée comme partition de  $P_{12}$  ou  $P_{31}$ , et notée de la même façon.

On se propose, dans cet article, d'étudier les 2 analyses internes multiples (dont nous montrerons l'équivalence) du tableau ternaire croisant  $P_{12}, P_{23}$  et  $P_{31}$ , relatives pour la première analyse aux partitions  $P_1, P_2, P_3$  (et pour la seconde, aux partitions  $P_2, P_3, P_1$ ).  $C$  désignant la réunion des 3 partitions  $P_{12}, P_{23}, P_{31}$ , ces 2 analyses internes se définissent comme les 2 analyses intra-classes du tableau disjonctif complet  $k_{IC}$  croisant  $I$  et  $C$ ,  $C$  étant muni, pour l'une de la partition  $P_C = P_1 \cup P_2 \cup P_3$ , pour l'autre de la partition  $P_C^* = P_2 \cup P_3 \cup P_1$ .

On a alors le résultat suivant :

### 2.2. Théorème

*Les 2 analyses intra-classes du tableau disjonctif complet  $k_{IC}$ , croisant  $I$  et  $C$ ,  $C$  étant muni de l'une des 2 partitions  $P_C$  et  $P_C^*$ , sont équivalentes, en ce sens qu'elles génèrent les mêmes facteurs sur  $I$ , associés aux mêmes valeurs propres.*

**Démonstration :** (voir DENIMAL, thèse Paris-Dauphine (1993)).

### 2.3. Notations et conséquences

De façon générale, on pose :  $(r, s, l)$  une permutation de  $\{1, 2, 3\}$ ,  $C = P_{rs} \cup P_{sl} \cup P_{lr}$  et  $P_C = P_r \cup P_s \cup P_l$  la partition de  $C$  associée.

$c_{rs}$  étant une classe de  $P_{rs}$ , et  $c_r$  la classe de  $P_C$  contenant  $c_{rs}$ , l'analyse factorielle intra-classe du tableau disjonctif complet  $k_{IC}$  relative à la partition  $P_C$  de  $C$ , s'effectue en soumettant à l'analyse des correspondances le tableau  $K_{IC}$  suivant (construit à partir de  $k_{IC}$ ) :  $\forall i \in I, \forall c_{rs} \in C$

$$K(i, c_{rs}) = k(i, c_{rs}) - \frac{k(i, c_r)k(c_{rs})}{k(c_r)} + \frac{\text{card}(c_{rs})}{\text{card}(I)}$$

où l'on a posé :

$$k(i, c_r) = \sum_{\substack{c_{rs} \subset c_r \\ c_{rs} \in P_{rs}}} k(i, c_{rs}) ; k(c_{rs}) = \sum_{i \in I} k(i, c_{rs}) ; k(c_r) = \sum_{i \in I} k(i, c_r)$$

En conséquence,  $k(i, c_r)$  vaut 1 si  $i \in c_r$  et 0 sinon, tandis que  $k(c_{rs}) = \text{card}(c_{rs})$  et  $k(c_r) = \text{card}(c_r)$ .

Considérant le nuage  $\mathcal{N}(C)$  des « colonnes » du tableau  $K_{IC}$ , on vérifie, tout d'abord, que le centre de gravité de  $\mathcal{N}(C)$  vaut  $f_I = (f_i)_{i \in I}$  où  $f_i = \frac{1}{\text{card}(I)}$ , et que

le vecteur joignant le profil de  $c_{rs}$  à  $f_I$  est noté et vaut :

$$\left( \frac{k(i, c_{rs})}{k(c_{rs})} \right)_{i \in I} - \left( \frac{k(i, c_r)}{k(c_r)} \right)_{i \in I} = f_I^{c_{rs}} - f_I^{c_r}$$

On se propose, dans la propriété donnée ci-dessous, de calculer différents produits scalaires, entre vecteurs de  $R^{\text{card}(I)}$  du type  $f_I^{c_{rs}} - f_I$ , vis-à-vis de la métrique du CHI2, de centre  $f_I$ , (qui est la métrique usuelle, au facteur  $\text{card}(I)$  près).

#### 2.4. Propriétés

Avec les notations précédentes, on a les résultats suivants qui se vérifient facilement :

$$a) \quad PS1 = \|f_I^{c_{rs}} - f_I^{c_r}\|^2 = \text{card}(I) \cdot \frac{\text{card}(c_r) - \text{card}(c_{rs})}{\text{card}(c_{rs}) \cdot \text{card}(c_r)}$$

$$b) \quad PS2 = \langle f_I^{c_{rs}} - f_I^{c_r}, f_I^{c_{lr}} - f_I^{c_{rl}} \rangle = \|f_I^{c_{lr}} - f_I^{c_{rl}}\|^2 \left[ \frac{\text{card}(c_{rsl})}{\text{card}(c_{rs})} - \frac{\text{card}(c_{rl})}{\text{card}(c_r)} \right]$$

(où l'on a posé :  $c_{rsl} = c_{rs} \cap c_l$ ).

$$c) \quad PS3 = \langle f_I^{c_{rs}} - f_I^{c_r}, f_I^{c_{lr}^*} - f_I^{c_{rl}} \rangle = -\frac{\text{card}(I)}{\text{card}(c_l)} \left[ \frac{\text{card}(c_{rsl})}{\text{card}(c_{rs})} - \frac{\text{card}(c_{rl})}{\text{card}(c_r)} \right]$$

(où  $c_{lr}^*$  est une classe de  $P_{lr}$  d'intersection vide avec  $c_r$ ).

$$d) \quad PS4 = \langle f_I^{c_{rs}} - f_I^{c_r}, f_I^{c_{rs}^*} - f_I^{c_r^*} \rangle = 0$$

(où  $c_{rs}^*$  et  $c_r^*$  sont des classes de  $P_{rs}$  et de  $P_r$  respectivement d'intersection vide avec  $c_{rs}$  et  $c_r$ ).

$$e) \quad PS5 = \langle f_I^{c_{rs}} - f_I^{c_r}, f_I^{\hat{c}_{rs}} - f_I^{c_r} \rangle = \frac{-\text{card}(I)}{\text{card}(c_r)}$$

(où  $c_{rs}$  et  $\hat{c}_{rs}$  sont 2 classes différentes de  $P_{rs}$  incluses dans  $c_r$ , donc d'intersection vide).

#### 2.5. Remarques

En utilisant les résultats a) et b) de la propriété § 2.4., il apparaît que le cosinus de l'angle  $\theta = (f_I^{c_{rs}} - f_I^{c_r}, f_I^{c_{lr}} - f_I^{c_{rl}})$  vaut :

$$\cos(\theta) = \frac{\frac{\text{card}(c_{rsl})}{\text{card}(c_{rs})} - \frac{\text{card}(c_{rl})}{\text{card}(c_r)}}{\|f_I^{c_{rs}} - f_I^{c_r}\| / \|f_I^{c_{lr}} - f_I^{c_{rl}}\|}$$

Autrement dit, si les normes des 2 vecteurs sont peu différentes, l'angle  $\theta$  sera d'autant plus petit que la différence  $\frac{\text{card}(c_{rsi})}{\text{card}(c_{rs})} - \frac{\text{card}(c_{rl})}{\text{card}(c_r)}$  sera plus grande.

D'autre part, ces 2 rapports seront égaux si et seulement si les 2 vecteurs  $f_I^{c_{rs}} - f_I^{c_r}$  et  $f_I^{c_{lr}} - f_I^{c_l}$  sont orthogonaux.

**2.6. Conséquences**

Du résultat a) de la propriété § 2.4., se déduit l'inertie  $In$  du tableau  $K_{IC}$  :

$$In = \frac{1}{3}[\text{card}(P_{12}) + \text{card}(P_{23}) + \text{card}(P_{31}) - \text{card}(P_1) - \text{card}(P_2) - \text{card}(P_3)]$$

Cette inertie  $In$  n'est autre que l'inertie intra-classes de  $k_{IC}$ ,  $C$  étant muni de la partition  $P_C$  définie au § 2.3. L'inertie intra-classes  $In$  peut donc être comparée à l'inertie totale du tableau disjonctif complet  $k_{IC}$  qui vaut :

$$\frac{1}{3}[\text{card}(P_{12}) + \text{card}(P_{23}) + \text{card}(P_{31}) - 3]$$

**3. Étude d'un cas particulier où chacune des trois partitions  $P_1, P_2, P_3$  de  $I$  n'est composée que de 2 classes**

**3.1. Notations**

Dans ce cas particulier, les partitions de  $I, P_{12}, P_{23}, P_{31}$  possèdent chacune 4 classes dont les notations sont précisées ci-dessous :

				P <sub>2</sub>							
				c2(1)	c2(2)						
P <sub>1</sub>	c1(1)	c12(11)	c12(12)								
	c1(2)	c12(21)	c12(22)								
				P <sub>12</sub>							

				P <sub>3</sub>							
				c3(1)	c3(2)						
P <sub>2</sub>	c2(1)	c23(11)	c23(12)								
	c2(2)	c23(21)	c23(22)								
				P <sub>23</sub>							

				P <sub>1</sub>							
				c1(1)	c1(2)						
P <sub>3</sub>	c3(1)	c31(11)	c31(12)								
	c3(2)	c31(21)	c31(22)								
				P <sub>31</sub>							

Le nuage  $\mathcal{N}(C)$  des «colonnes» du tableau  $K_{IC}$  est muni de la partition  $P_C = P_1 \cup P_2 \cup P_3$ ; ce qui est résumé dans le tableau ci-dessous :

$P_{12}$ partitionnée par $P_1$	$\underbrace{c12(11) \quad c12(12)}_{c1(1)}$	$\underbrace{c12(21) \quad c12(22)}_{c1(2)}$
$P_{23}$ partitionnée par $P_2$	$\underbrace{c23(11) \quad c23(12)}_{c2(1)}$	$\underbrace{c23(21) \quad c23(22)}_{c2(2)}$
$P_{31}$ partitionnée par $P_3$	$\underbrace{c31(11) \quad c31(12)}_{c3(1)}$	$\underbrace{c31(21) \quad c31(22)}_{c3(2)}$

**3.2. Etude du nuage  $\mathcal{N}(C)$  issu du tableau  $K_{IC}$  (§ 2.3 et 2.4)**

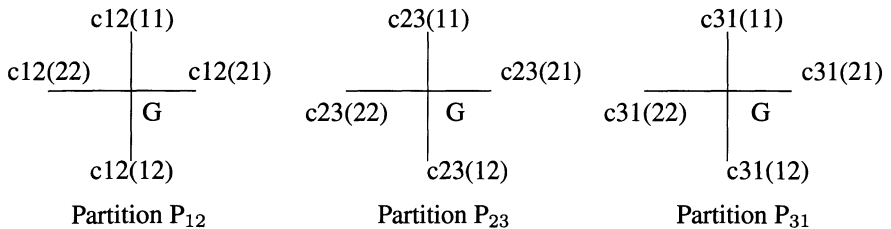
Le centre de gravité global de  $\mathcal{N}(C)$ , avec  $C = P_{12} \cup P_{23} \cup P_{31}$ , est aussi, par construction, centre de gravité de chacun des sous-nuages associés à chacune des partitions  $P_{12}, P_{23}, P_{31}$ . (ce qui est une propriété classique de l'analyse intra-classe).

On obtient, ainsi, en notant de la même manière, une classe et son centre de gravité, et en considérant, par exemple, la partition  $P_{12}$  (qui comporte les 2 classes  $\{c12(11), c12(12)\}$  et  $\{c12(21), c12(22)\}$ ) :

a)  $G$  désignant le centre de gravité global de  $\mathcal{N}(C)$ ,  $c12(11), c12(12)$  et  $G$  (et de même  $c12(21), c12(22)$  et  $G$ ) sont alignés.

b) Les vecteurs  $[c12(11), c12(12)]$  et  $[c12(21), c12(22)]$  sont orthogonaux (ce qui est une conséquence de la propriété d) § 2.4.).

Ainsi, le sous nuage de  $\mathcal{N}(C)$  composé des éléments de  $P_{12}$  se présente sous la forme d'un quadripole, et le nuage complet  $\mathcal{N}(C)$  sous la forme de 3 quadripoles.



Quant à la position de ces quadripoles les uns par rapport aux autres, il suffit de reprendre la remarque § 2.5. Ainsi, l'angle  $\theta$  formé par les vecteur  $[G, c12(11)]$  et  $[G, c23(11)]$  sera d'autant plus petit que le rapport  $\frac{\text{card}[c12(11) \cup c23(11)]}{\text{card}(c23(11))}$  est grand par rapport à  $\frac{\text{card}(c12(11))}{\text{card}(c2(1))}$  (si les 2 vecteurs ont des normes peu différentes).

### 3.3. Théorème

Considérons le tableau  $K_{IC}$  croisant  $I$  et  $C = P_{12} \cup P_{23} \cup P_{31}$  (§ 2.3) et celui noté  $K[c12(11), c12(21); c23(11), c23(21); c31(11), c31(21)]$  dont les colonnes associées aux classes ainsi précisées sont extraites du premier tableau.

On montre, alors, que l'analyse des correspondances de  $K_{IC}$  (Analyse 1) et l'analyse en composantes principales normée du second tableau (Analyse 2) sont équivalentes.

De manière précise, les analyses 1 et 2 ont respectivement pour triple  $(I, m_I, d_{II})$  et  $(I, m_I, d_{II}^*)$ , où  $m_i$ , poids d'un élément  $i$  de  $I$ , vaut  $\frac{1}{\text{card}(I)}$ , et où  $d_{II}$  et  $d_{II}^*$ , désignant respectivement les tableaux des distances au carré entre élément de  $I$  pour les analyses 1 et 2, vérifient  $d_{II} = \frac{1}{3}d_{II}^*$ .

**Démonstration :** (voir DENIMAL, thèse Paris-Dauphine (1993)).

### 3.4. Remarques

a) L'intérêt de ce théorème est de montrer que l'on peut obtenir les résultats de l'analyse des correspondances de  $K_{IC}$ , en soumettant à l'ACP normée un tableau de dimensions plus petites (le nombre de colonnes ayant été réduit de moitié).

b) Ce théorème permet, au niveau de l'interprétation des résultats de l'AFC de  $K_{IC}$ , de procéder comme en ACP normée.

c) Enfin, ce résultat est à rapprocher de celui analogue, concernant tout tableau disjonctif complet, pour lequel chaque question ne comporte que 2 modalités.

## 4. Application à un jeu de données

### 4.1. Présentation des données

Il s'agit d'un tableau disjonctif complet croisant un ensemble  $I$  de 20000 individus et les 3 variables suivantes : sexe, acide urique, consommation d'alcool, chacune d'elles ayant 2 modalités. Ces 3 variables généreront par croisement 2 à 2, 3 nouvelles variables qui auront, par conséquent, chacune 4 modalités.

Nous donnons, ci-dessous, les définitions et les notations de ces 6 variables, ainsi que le tableau de contingence croisant l'ensemble des modalités des 3 variables obtenues par croisement, avec lui-même :

Variables	Qualitatives	Modalités des variables	
Définitions	Notations	Définitions	Notations
Sexe	Sexe	Homme	H
		Femme	F
Acide Urique	AU	AU < 280	AU-
		AU > 280	AU+
Consommation d'alcool	AL	AL=0	AL-
		AL > 0	AL+



		Variable AL				Variable AU	
		AL-	AL+			AU-	AU+
Variable	H	(H,AL-)	(H,AL+)	Variable	H	(H,AU-)	(H,AU+)
Sexe	F	(F,AL-)	(F,AL+)	Sexe	F	(F,AU-)	(F,AU+)

Variable AU			
		AU-	AU+
Variable	AL-	(AL-,AU-)	(AL-,AU+)
AL	AL+	(AL+,AU-)	(AL+,AU+)

	(H, AU-)	(H, AU+)	(F, AU-)	(F, AU+)	(AL-, H)	(AL-, F)	(AL+, H)	(AL+, F)	(AU-, AL-)	(AU-, AL+)	(AU+, AL-)	(AU+, AL+)
(H,AU-)	2413	0	0	0	1211	0	1202	0	1211	1202	0	0
(H,AU+)	0	7477	0	0	2437	0	4940	0	0	0	2537	4940
(F,AU-)	0	0	7333	0	0	5389	0	1944	5389	1944	0	0
(F,AU+)	0	0	0	2777	0	1669	0	1108	0	0	1669	1108
(AL-,H)	1211	2537	0	0	3748	0	0	0	1211	0	2537	0
(AL-,F)	0	0	5389	1669	0	7058	0	0	5389	0	1669	0
(AL+,H)	1202	4940	0	0	0	0	6142	0	0	1202	0	4940
(AL+,F)	0	0	1944	1108	0	0	0	3052	0	1944	0	1108
(AU-,AL+)	1211	0	5389	0	1211	5389	0	0	6600	0	0	0
(AU-,AL-)	1202	0	1944	0	0	0	1202	1944	0	3146	0	0
(AU+,AL-)	0	2537	0	1669	2537	1669	0	0	0	0	4206	0
(AU+,AL+)	0	4940	0	1108	0	0	4940	1108	0	0	0	6048

## 4.2. Les analyses réalisées

4.2.1. La première est l'analyse des correspondances du tableau disjonctif complet  $k_{IC}$  croisant l'ensemble  $I$  des 20000 individus et l'ensemble  $C$  des 12 modalités des 3 variables (sexe, acide urique), (acide urique, alcool) et (alcool, sexe).

4.2.2. Les deux autres analyses sont les analyses intra-classes de  $k_{IC}$   $C$  étant muni de l'une des 2 partitions notées  $P_1$  et  $P_2$  :

Partitions de	$J = (\text{sexe}, AU) \cup (AU, AL) \cup (AL, \text{sexe})$
Partition $P_1$	$P_1 = (\text{sexe}) \cup (AU) \cup (AL)$
Partition $P_2$	$P_2 = (AU) \cup (AL) \cup (\text{sexe})$

(Considérant, par exemple, la partition  $P_1$  de  $J$ , la variable qualitative sexe est alors considérée comme une partition en 2 classes des 4 modalités du couple (sexe, AU))

D'autre part, on sait que, dans une analyse intra-classe, le vecteur joignant le centre de gravité  $G$  de  $\mathcal{N}(C)$  à l'un des éléments de ce nuage (calculé à partir du tableau analysé, noté  $K_{IC}$  (voir § 2.3.)) représente (pour le tableau initial, ici  $k_{IC}$ ) le vecteur joignant cet élément au centre de gravité de la classe dans laquelle il se trouve. Ainsi, en appliquant les résultats du § 3.2., le nuage  $\mathcal{N}(C)$ , obtenu pour chacune de ces 2 analyses intra-classes (équivalentes d'après le théorème § 2.2) se présente sous la forme des 3 quadripôles suivants :

	Analyse intra classe de $K_{IJ}$ , $J$ étant muni de la partition $P_1$	Analyse intra-classe de $K_{IJ}$ , $J$ étant muni de la partition $P_2$
Quadripôle 1	[H,(H,AU-)]; [H,(H,AU+)] [F,(F,AU-)]; [F,(F,AU+)]	[AU-,(AU-,F)]; [AU-,(AU-,H)] [AU+,(AU+,F)]; [AU+,(AU+,F)]
Quadripôle 2	[AU-,(AU-,AL-)]; [AU-,(AU-,AL+)] [AU+,(AU+,AL-)]; [AU+,(AU+,AL+)]	[AL-,(AL-,AU-)]; [AL-,(AL-,AU+)] [AL+,(AL+,AU-)]; [AL+,(AL+,AU+)]
Quadripôle 3	[AL-,(AL-,H)]; [AL-,(AL-,F)] [AL+,(AL+,H)]; [AL+,(AL+,F)]	[H,(H,AL-)]; [H,(H,AL+)] [F,(F,AL-)]; [F,(F,AL+)]

(Où, par exemple, la notation [H,(H,AU+)] représente le vecteur joignant le centre de gravité de la classe  $H = \{(H,AU-), (H,AU+)\}$  à la modalité (H,AU+))

### 4.3. Résultats des analyses

#### 4.3.1. Analyse du tableau disjonctif complet $k_{IC}$ (graphique 1)

Le plan (1,2) obtenu représente 49% de l'inertie du tableau étudié, dont 28% pour le premier axe. Celui-ci s'explique à partir des variables sexe et acide urique et oppose les modalités  $F$  et  $AU-$  aux modalités  $H$  et  $AU+$ . Le second axe, quant à lui, met en opposition les 2 modalités de la variable alcool.

#### 4.3.2. Analyse intra-classe de $k_{IC}$ , $C$ étant muni de la partition $P_1$ (graphique 2)

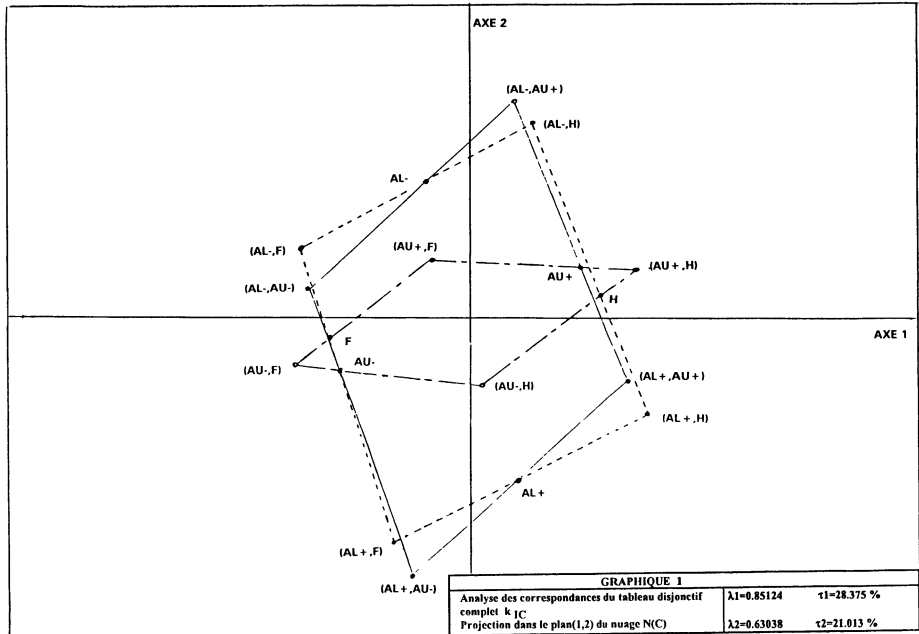
Compte tenu du théorème § 2.2., il apparaît que les résultats de la seconde analyse intra classe de  $k_{IC}$  (relative à la partition  $P_2$ ) peuvent s'obtenir en mettant le tableau  $K_{IC}$  issu de cette dernière, en colonnes supplémentaires au tableau  $K_{IC}$  issu

de la première analyse intra-classe de  $k_{IC}$  (relative à  $P_1$ ). On trouvera, au graphique 3, la représentation de  $\mathcal{N}(C)$  issue de la seconde analyse intra-classe de  $k_{IC}$  (relative à  $P_2$ ).

Nous nous limiterons, ci-dessous, à l'interprétation des résultats de la première analyse intra-classe (relative à  $P_1$ , graphique 2).

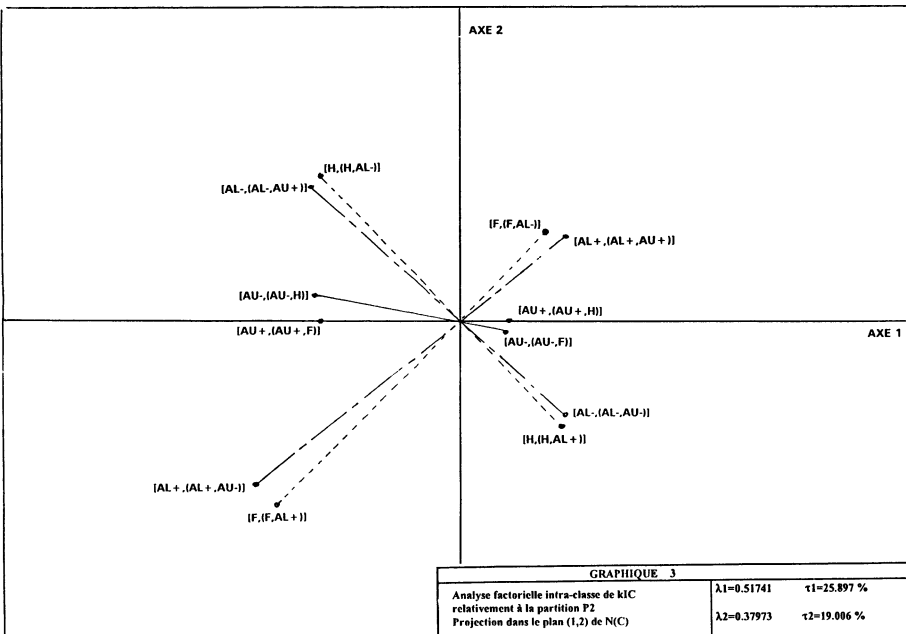
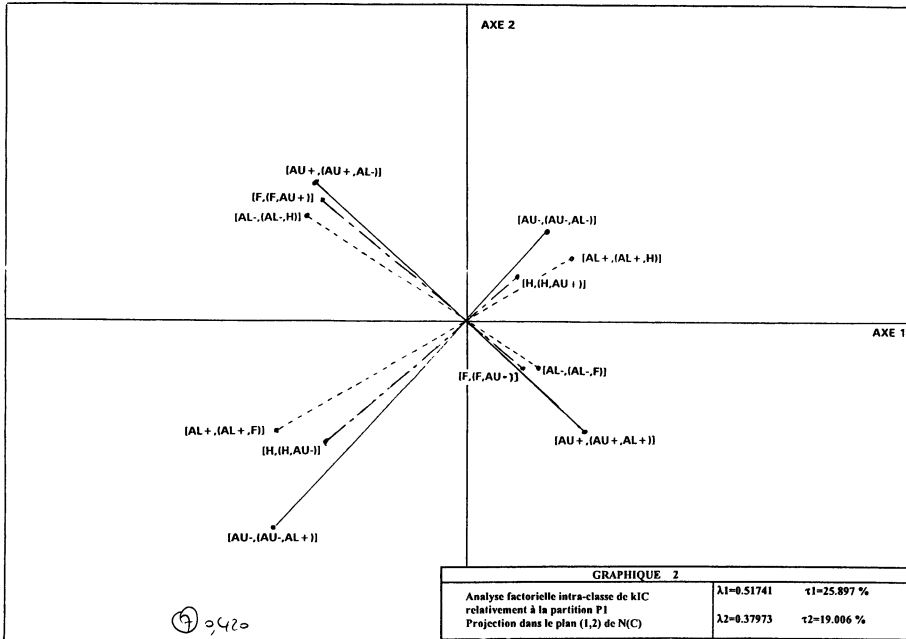
On peut faire, tout d'abord, les 2 remarques suivantes :

a) En appliquant les résultats du § 2.6, il apparaît que l'inertie du tableau disjonctif complet  $k_{IC}$  vaut 3, et celle du tableau  $K_{IC}$  associé vaut 2 (autrement dit, 66% de l'inertie de  $k_{IC}$ ).



b) Bien que le plan (1,2) soit le meilleur plan représentant le nuage  $\mathcal{N}(C)$  des 3 quadrupoles, les angles droits de ces derniers risquent d'être déformés après projection.

Enfin, en appliquant la propriété § 2.4, l'étude du plan (1,2) (graphique 2), (représentant 44% de l'inertie du tableau analysé) génère les résultats suivants :



	Couples vecteurs faisant un angle petit	Normes des vecteurs au carré	Propriétés utilisées §3.4 Chap.4	Significations
F1<0; F2<0	[AI+, (AL+, F)] [AU-, (AU-, AL+)]	1.088 1.086	b)	$\frac{\text{Card}(AL+, F, AU-)}{\text{Card}(AL+, F)} \geq \frac{\text{Card}(AU-, AL+)}{\text{Card}(AL+)}$  (1944/3052) = 0.63 > (3146/9194) = 0.34
	[H, (H, AU-)] [AU-, (AU-, AL+)]	1.575 1.086	b)	$\frac{\text{Card}(AL+, H, AU-)}{\text{Card}(AL+, AU-)} \geq \frac{\text{Card}(AU-, H)}{\text{Card}(AU-)}$  (1202/2413) = 0.49 < (6142/9890) = 0.62
	[H, (H, AU-)] [AL+, (AL+, F)]	1.575 1.088	c)	$\frac{\text{Card}(AL+, H, AU-)}{\text{Card}(H, AU-)} \geq \frac{\text{Card}(AL-, H)}{\text{Card}(H)}$  (1202/2413) = 0.49 < (6142 :9890) = 0.62
F1>0; F2<0	[F, (F, AU-)] [AU+, (AU+, AL+)]	0.188 0.336	c)	$\frac{\text{Card}(AL+, F, AU+)}{\text{Card}(AL+, AU+)} \geq \frac{\text{Card}(AU+, F)}{\text{Card}(AU+)}$  (1108/6048) = 0.18 < (2777/10254) = 0.27
	[AL-, (AL-, F)] [F, (F, AU-)]	0.245 0.188	b)	$\frac{\text{Card}(AL+, F, AU-)}{\text{Card}(F, AU-)} \geq \frac{\text{Card}(AL-, F)}{\text{Card}(F)}$  (5389/7333) = 0.73 > (7058/10110) = 0.69

Ainsi, concernant le premier couple décrit ci-dessus, le petit angle qu'il forme, signifie que la probabilité d'avoir un taux d'acide urique faible ( $AU-$ ) sachant que l'on est de sexe féminin et consommatrice d'alcool ( $F, AL+$ ) est sensiblement plus élevée que la probabilité d'avoir un taux d'acide urique faible ( $AU-$ ) sachant que l'on est consommateur d'alcool ( $AL+$ ). Ainsi, le fait de se placer dans la population des femmes, augmente cette dernière probabilité.

## 5. Conclusion

L'analyse interne multiple, qui est appliquée dans cet article, est réalisée en soumettant à l'analyse des correspondances, un certain tableau  $K_{IC}$  (§ 2.3) croisant un ensemble  $I$  d'individus, et un ensemble  $C$ , construit à partir des modalités de

$p = 3$  variables qualitatives (définies sur  $I$ ). On trouvera, dans la thèse (Denimal, Paris-Dauphine 1993), d'une part une définition d'une classe de tableaux, dans laquelle se retrouvent le tableau  $K_{IC}$  précédent, ainsi que le tableau disjonctif complet classique, et d'autre part une généralisation de l'ensemble des propriétés et théorèmes classiques de l'analyse des correspondances multiples.

### Références Bibliographiques

- BENZECRI, J.P. (1983) : Analyse de l'inertie intra-classe par l'analyse d'un tableau de contingence. CAD VIII n°3, p. 351-358.
- CAZES P. (1986) : Correspondances entre 2 ensembles et partitions de ces 2 ensembles. CAD XI n°3, p. 335-340.
- CAZES P. - CHESSEL D. - DOLEDEC S. (1988) : L'analyse des correspondances interne d'un tableau partitionné. Son usage en hydrobiologie, RSA XXXVI n°1, p. 39-54.
- CAZES P. - MOREAU, J. (1991) : Analysis of contingency table in which the rows and columns have a graph structure dans Symbolic Numeric Data Analysis and Learning, édité par E. DÍDAY et Y. LECHEVALLIER, Nova Science Publishers Inc, p. 271-280.
- DENIMAL J.J. (1991) : Essai de modélisation de la consommation d'alcool dans une population. Journal d'Alcoologie n°2, (HCEIA), p. 73-100.
- DENIMAL J.J. (1993) Analyse des interactions entre  $k$  partitions définies sur l'ensemble  $I$  d'un tableau de contingence  $k_{IJ}$ . Thèse Paris-Dauphine.
- ESCOFIER B. - DROUET, D. (1983) : Analyse de la différence entre plusieurs tableaux de fréquences. CAD VIII n°4, p. 491-499.
- MOREAU J. (1990) : Tableaux, blocs et graphes. Analyses de tableaux structurés. Centre Vaudois de Recherches Pédagogiques n°90901 Lausanne.
- MOREAU J. (1992) : Analyse des données structurées par des graphes. Cas de l'analyse des correspondances. Thèse EPFL Lausanne.