



HAL
open science

Croisement de classifications floues

I. C. Lerman

► **To cite this version:**

I. C. Lerman. Croisement de classifications floues. Annales de l'ISUP, 1979, XXIV (1-2), pp.13-46.
hal-04081223

HAL Id: hal-04081223

<https://hal.science/hal-04081223>

Submitted on 25 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CROISEMENT DE CLASSIFICATIONS FLOUES

PAR I.C. LERMAN

Laboratoire de Statistiques - I.R.I.S.A. - Université de RENNES I.

I - INTRODUCTION

Le problème s'est posé dans l'analyse des résultats de notre méthode de classification hiérarchique appliquée à un ensemble d'attributs de description résultant d'une étude (enquête ou expérience) à caractère psycho-sociologique. Dans les questionnaires supportant ce type d'étude, l'ensemble des questions peut être scindé en deux parties. La première, qui a un caractère physique, permet une description purement sociologique du sujet ; par exemple son âge, sa catégorie socio-professionnelle, ses activités, son lieu de résidence, etc... La deuxième partie qui a un caractère psychologique, décrit la réaction du sujet à des questions de nature comportementale où l'estimation subjective de l'individu peut intervenir ; par exemple, relativement à une étude pour l'élaboration d'indicateurs socio-économiques liés à la qualité de l'eau d'un lac, on a des questions de perception tels que : "Appréciez vous la présence de roseaux ou vous gênent-ils ?" ou bien "Comment jugez-vous la couleur de l'eau ? (claire, transparente, trouble, sombre)". Un traitement global par la classification de l'ensemble de tous les attributs montre que la majeure partie des variables purement sociologiques peut se séparer dans une même classe pour se structurer et la partie restante de ce type de variables se mêle à des variables à caractère psychologique perturbant une organisation plus claire et plus cohérente de l'ensemble des seules variables de comportement. De sorte, qu'à notre avis, il faut de façon fondamentale un traitement séparé de chacune des deux parties du questionnaire ; les liens entre les variables sociologiques étant beaucoup plus contrastés que ceux, entre les variables psychologiques.

Cependant, le spécialiste psycho-sociologue pose inévitablement la question : comment se répartissent les différentes catégories d'individus séparées par une même variable sociologique, entre les différents types de comportement dégagés par le traitement de l'ensemble des variables d'attitude ?

Ainsi, considérons l'exemple de l'étude à laquelle nous faisons allusion ci-dessus. On y a obtenu, à la limite, deux principaux types de comportement : "partisans d'une nature aménagée" et "partisans d'une nature sauvage" ; à travers ces deux aspects, diverses sous-classes ont mis en évidence différentes fonctions de l'eau du lac : "sportive", "écologique", "ludico-sociale" et "eau à regarder". Il est dans ces conditions intéressant de savoir comment se répartissent à travers les différents types de perception du lac, les différentes classes d'âge ou bien, les différentes catégories socio-professionnelles.

En considérant la classification des attributs associés aux caractéristiques socio-individuelles, le spécialiste peut poser une question plus élaborée que la précédente quant à la situation des types extraits d'une telle classification par rapport à ceux de comportement. Désignons par

$$\pi(S) = \{S_i / 1 \leq i \leq k\}$$

et

$$\pi(C) = \{C_j / 1 \leq j \leq h\} \quad , (1)$$

les classifications respectives de l'ensemble S des attributs socio-individuels et de l'ensemble C des attributs de comportement qui définissent respectivement les deux typologies. $\pi(S)$ (resp. $\pi(C)$) peut correspondre à la classification produite à un niveau donné de l'arbre des classifications sur S (resp. C).

Une réponse "faible" à la question posée consisterait à établir la table de dimension $k \times h$

$$\{P(S_i, C_j) / 1 \leq i \leq k, 1 \leq j \leq h\} \quad , (2)$$

où $P(S_i, C_j)$ est l'indice de proximité entre les classes S_i et C_j ; celui là même qui a prévalu dans la construction ascendante de chacun des arbres de classification sur C et sur S . Rappelons brièvement que dans notre méthode cet indice est basé sur la vraisemblance du lien défini par la plus grande valeur observée d'une statistique de proximité entre deux éléments appartenant respectivement aux deux classes à comparer.

Une première raison de la faiblesse de cette réponse est due à ce que l'indice P utilisé ne tient pas le plus étroitement compte de la nature des deux structures à comparer qui sont deux classifications polythétiques de deux ensembles disjoints d'attributs descriptifs. Ce type d'indice est en effet utilisé dans notre méthode pour la formation des classes quelle que soit la nature des données ; qu'il s'agisse de la classification des variables descriptives ou bien des objets. D'autre part, un bon indice pour l'émergence des classes n'est pas nécessairement pertinent pour comparer les classes respectives de deux classifications déjà formées ; d'ailleurs les nombres de la table (2) doivent être tous trop faibles pour tirer une conclusion significative.

S'il s'agit de répondre à la première question du psychosociologue où il s'agit d'analyser comment se répartissent les différentes catégories d'individus séparées par une même variable sociologique, entre les différents types de comportement ; ce que nous proposons est un équivalent pour la classification de l'introduction d'éléments supplémentaires en analyse factorielle des correspondances. Mais cet équivalent est plus précis car d'une part, on a une plus grande maîtrise de l'interprétation d'une classe d'attributs et d'autre part, on a immédiatement sous la forme d'une table de nombres les valeurs des associations des différentes modalités de la variable extérieure définissant une partition, avec les différentes classes polythétiques d'attributs. Notre approche est plus générale puisqu'elle permet de situer l'une par rapport à l'autre deux classifications polythétiques sur respectivement, deux ensembles disjoints d'attributs de description.

II - CLASSIFICATION DICHOTOMIQUE "NETTE" ASSOCIEE A UN ATTRIBUT ET CLASSIFICATION DICHOTOMIQUE "FLOUE" ASSOCIEE A UNE CLASSE D'ATTRIBUTS.

Désignons une fois pour toutes par E l'ensemble défini par l'échantillon (supposé représentatif) de la population étudiée. La donnée d'un attribut de description a détermine sur E une partition $\pi(a)$ en deux classes E_a et E_a^- où E_a (resp. E_a^-) est formée des objets de E où l'attribut a est présent (resp. absent).

Associations à la partie E_a (resp. E_a^-) de E , sa fonction indicatrice ϕ_a (resp. ϕ_a^-) :

$$(\forall x \in E) : \phi_a(x) = 1 \text{ (resp. } \phi_a(x) = 0) \iff x \in E_a \text{ (resp. } x \in E_a^-)$$

$$\text{et } \phi_a^-(x) = 1 \text{ (resp. } \phi_a^-(x) = 0) \iff x \in E_a^- \text{ (resp. } x \in E_a) ;$$

on a

$$(\forall x \in E) : \phi_a(x) + \phi_a^-(x) = 1 \quad (1)$$

La donnée de l'attribut de description a définit ce que nous appellerons une classification dichotomique "nette" $\pi(a)$ caractérisée par la fonction indicatrice ϕ_a ou ϕ_a^- .

Considérons à présent un ensemble fini A , de cardinal coté $c(A)$, d'attributs de description de E . Dans les applications A est une classe d'attributs "proches" extraite par exemple d'un arbre des classifications sur un ensemble plus vaste d'attributs, définissant un profil sociologique ou d'attitude. Un même individu x de E peut être plus ou moins affecté par un tel profil ; il l'est d'autant plus qu'il possède un plus grand nombre d'attributs de A . Il est par conséquent naturel d'associer à chaque x de E , la proportion d'attributs de A que possède x ; soit

$$\phi_A(x) = \frac{1}{c(A)} \sum_{a \in A} \phi_a(x) \quad (2)$$

De la même manière à chaque x de E , on associe

$$\phi_{\bar{A}}(x) = \frac{1}{c(A)} \sum_{a \in A} \phi_a^-(x) \quad (3)$$

qui représente la proportion d'attributs de \bar{A} possédés par x ; où \bar{A} est obtenu à partir de A en associant à chaque attribut de A , sa négation.

On a une relation équivalente à la relation (1) ci-dessus :

$$(\forall x \in E) : \phi_A(x) + \phi_{\bar{A}}(x) = 1 \quad (4)$$

La fonction d'appartenance relative ϕ_A au profil défini par la classe A d'attributs, définit ce que nous appellerons une classification dichotomique "floue" où à chaque élément x de E , on associe le couple de nombres positifs de somme 1 $(\phi_A(x), \phi_A^-(x))$.

Signalons ici que bien que nous associons à chaque objet de E un degré d'appartenance (nombre compris entre 0 et 1) à un type ; nous n'avons nullement besoin de nous référer au formalisme introduit par L.A. ZADEH (cf. [7]), auquel d'ailleurs nous ne concédons ici aucune valeur opérationnelle.

III - k-CLASSIFICATION "NETTE" et k-CLASSIFICATION "FLOUE"

Une k-classification "nette" sur E est une partition en k classes au plus sur E . Dans le contexte de ce travail une telle partition se trouve associée à un caractère descriptif présentant k modalités ; où l'ensemble des modalités mutuellement exclusives est sans structure. Désignons par $\{1, 2, \dots, k\}$ l'ensemble des codes des modalités d'un tel caractère de description c . c peut être défini comme une application de E dans $\{1, 2, \dots, k\}$, associant à chaque x de E , le code $c(x)$ de la modalité du caractère présente chez x .

$E_j = c^{-1}(j)$ est la partie de E formée des objets possédant la modalité codée j , $1 \leq j \leq k$. L'ensemble des parties $\{E_1, E_2, \dots, E_k\}$ forme une partition en k -classes au plus de E . En désignant par ϕ_j la fonction indicatrice de E_j :

$$(\forall x \in E) : \phi_j(x) = 1 \text{ (resp. } 0) \iff x \in E_j \text{ (resp. } x \notin E_j), 1 \leq j \leq k.$$

On a dans ces conditions une relation équivalente à celle (1) du paragraphe précédent

$$(\forall x \in E) : \sum_{1 \leq j \leq k} \phi_j(x) = 1 \quad (1)$$

Le contexte dans lequel la notion de classification dichotomique "floue" peut s'étendre est le suivant : Soit un ensemble C de caractères

descriptifs où l'ensemble des modalités mutuellement exclusives d'un même caractère est considéré sans structure. On suppose que le nombre k de modalités d'un même caractère c de C est le même quel que soit ce dernier. On considère une classification par proximité avec contrainte de l'ensemble B des attributs de description associé à l'ensemble des modalités des différents caractères de C , en attachant à chaque modalité un attribut. La contrainte consiste à imposer à chaque classe d'attributs de contenir exactement une modalité de chacun des caractères. Il s'agit donc d'une classification de B en exactement k classes $\{B_1, B_2, \dots, B_k\}$ où $B_j, 1 \leq j \leq k$, de cardinal $p = \text{card}(C)$, contient exactement une modalité de chaque c de C .

Comme ci-dessus (cf. formule (2) § II), nous associons à chaque classe B_j la fonction $\phi_{B_j}(x)$ définie sur E comme suit

$$\phi_{B_j}(x) = \frac{1}{p} \sum_{b \in B_j} \phi_b(x) \quad ; \quad (2)$$

associant à chaque objet x de E , la proportion d'attributs de B_j présents chez x . $\phi_{B_j}(x)$ définit le degré d'appartenance de l'objet x au type décrit par la classe B_j d'attributs descriptifs.

De façon analogue à la relation (4) ci-dessus, on a la relation

$$(\forall x \in E) : \sum_{1 \leq j \leq k} \phi_{B_j}(x) = 1 \quad . \quad (3)$$

Ayant précisé les qualificatifs "nette" et "floue" à propos d'une classification, ils ne seront plus dans la suite de ce texte encadrés de guillemets.

IV - CROISEMENT DE CLASSIFICATIONS NETTES. INTRODUCTION AUX AUTRES CROISEMENTS

Soit un couple de partitions sur E , associé en l'occurrence à un couple de caractères descriptifs présentant respectivement k et h modalités.

Désignons par

$$P = \{E_1, E_2, \dots, E_k\} \quad (1)$$

et

$$Q = \{F_1, F_2, \dots, F_h\}$$

ces deux partitions.

La partition croisée $P \wedge Q$ est définie par

$$P \wedge Q = \{E_r \cap F_s / 1 \leq r \leq k, 1 \leq s \leq h\} \quad (2)$$

Soit ϕ_r , $1 \leq r \leq k$, la fonction indicatrice de la classe E_r et soit ψ_s , $1 \leq s \leq h$, la fonction indicatrice de la classe F_s . La table de contingence associée à (P, Q) est la table des cardinaux des classes de $P \wedge Q$; soit

$$\{n_{rs} / 1 \leq r \leq k, 1 \leq s \leq h\} \quad (3)$$

où

$$n_{rs} = \text{card}(E_r \cap F_s) = \sum_{x \in E} \phi_r(x) \psi_s(x)$$

$$1 \leq r \leq k, 1 \leq s \leq h.$$

Désignons par $I = \{1, 2, \dots, i, \dots, n\}$ l'ensemble codant E . De la sorte nous écrirons le contenu de la case (r, s) du tableau de contingence précédent sous la forme

$$n_{rs} = \sum_{1 \leq i \leq n} \phi_r(i) \psi_s(i) \quad (4)$$

$$1 \leq r \leq k, 1 \leq s \leq h.$$

On attache au tableau de contingence deux marges ; l'une colonne et l'autre, ligne. La r -ème ligne de la marge colonne contient

$$n_{r.} = \sum_{1 \leq i \leq n} \phi_r(i) \quad (5)$$

$$1 \leq r \leq k,$$

et la s -ème colonne de la marge ligne contient

$$n_{.s} = \sum_{1 \leq i \leq n} \psi_s(i) \quad (6)$$

Le problème posé est celui de la mesure orientée de l'intensité des liens entre les classes de la partition P et celles de la partition Q . Commençons par considérer le cas $k = h = 2$ où le contenu de l'une des cases, par exemple n_{11} , détermine la totalité de la table de contingence 2×2 .

Conformément à notre démarche générale (cf [5] chap. 2), nous associons à n_{11} la variable aléatoire N_{11} obtenue en fixant $P = \{E_1, E_2\}$ (resp. $Q = \{F_1, F_2\}$) et en associant à Q (resp. P) une partition aléatoire dans l'ensemble, muni d'une probabilité uniformément répartie, de toutes les partitions en classes étiquetées de type $(n_{.1}, n_{.2})$ (resp. (n_1, n_2)). On remarquera que la donnée d'une telle partition aléatoire Q' (resp. P') est équivalente à la donnée d'une partie aléatoire F'_1 (resp. E'_1) dans l'ensemble, muni d'une probabilité uniformément répartie, de toutes les parties de E de même cardinal $n_{.1}$ (resp. n_1). En désignant par ψ'_1 (resp. ϕ'_1) la fonction indicatrice de F'_1 (resp. E'_1); la variable aléatoire N_{11} peut se mettre sous l'une des deux formes équivalentes

$$\sum_{i \in I} \phi_1(i) \psi'_1(i) \quad (7)$$

ou bien

$$\sum_{i \in I} \phi'_1(i) \psi_1(i) \quad (7')$$

La distribution de la v.a. (7) est la même que celle de

$$\sum_{1 \leq i \leq n} \phi_1(i) \psi_1(\sigma(i)) \quad (8)$$

où σ est un élément aléatoire pris uniformément au hasard dans l'ensemble G_n de toutes les permutations sur $(1, 2, \dots, n)$ qui est de cardinal $n!$. Lorsque σ parcourt G_n , chaque expression (7) se retrouve dans (8) $n_{.1}! \times n_{.2}!$ fois exactement.

De la même façon la v.a. (7') est la même que celle

$$\sum_{1 \leq i \leq n} \phi_1(\sigma(i)) \psi_1(i) \quad (8')$$

La distribution commune des variables aléatoires (8) et (8') est exactement hypergéométrique de moyenne $(n_1 \cdot n_{.1} / n)$ et de variance $n_1 \cdot n_2 \cdot n_{.1} \cdot n_{.2} / n^2 (n-1)$.

La loi de la v.a. centrée réduite

$$(N_{11} - (n_1 \cdot n_{.1} / n)) / (n_1 \cdot n_2 \cdot n_{.1} \cdot n_{.2} / n^2 (n-1))^{1/2} \quad (9)$$

tend, dans des conditions très générales et bien connues, vers la loi normale centrée réduite $\mathcal{N}(0,1)$. Le carré de la valeur observée de (9) peut se mettre sous la forme

$$\sum_{1 \leq s \leq 2} \sum_{1 \leq r \leq 2} q_{rs}^2 \quad (10)$$

où

$$q_{rs} = (n_{rs} - (n_r \cdot n_{.s} / n)) / \sqrt{n_r \cdot n_{.s} / n} \quad (11)$$

(10) est, dans l'hypothèse d'absence de lien et pour n "assez" grand, une réalisation d'une v.a. du χ^2 à un seul degré de liberté.

Le nombre q_{rs} ($1 \leq r \leq 2$, $1 \leq s \leq 2$) définira la mesure orientée de l'intensité du lien entre la modalité r du caractère dichotomique associé à la partition P et la modalité s du caractère dichotomique associé à la partition Q .

Considérons à présent la situation la plus générale d'un tableau de contingence $k \times h$ associé au couple de partitions (P, Q) (cf. (1), (2), et (3)). L'hypothèse d'absence de lien fixe P (resp. Q) et associée à Q (resp. P) un élément Q' (resp. P') aléatoire dans l'ensemble, muni d'une probabilité uniformément ré-

partie. Sous une telle hypothèse, la probabilité d'un tableau donné tel que (3) se met sous la forme

$$\left(\prod_{1 \leq r \leq k} n_{r.}! \prod_{1 \leq s \leq h} n_{.s}! \right) / (n! \prod_{\substack{1 \leq r \leq k \\ 1 \leq s \leq h}} n_{rs}!) \quad (12)$$

On démontre (cf. [4]) que tout tableau de contingence $k \times h$ peut se réduire à une suite de $(k-1)(h-1)$ tables 2×2 de telle sorte que la probabilité (12) de l'ensemble du tableau puisse s'exprimer comme un produit des probabilités des tables 2×2 . A chacune de ces tables élémentaires peut être associée la statistique du χ^2 , notée χ^2 pour faire référence à la table t et qui peut s'exprimer comme le carré t d'une statistique telle que (9). Dans ces conditions

$$\sum_{1 \leq t \leq (k-1)(h-1)} \chi^2$$

est asymptotiquement une v.a. du χ^2 à $(k-1)(h-1)$ degrés de liberté. D'autre part, il s'agit asymptotiquement de la statistique du χ^2 attachée à l'ensemble du tableau $k \times h$, dont la valeur observée se présente sous la forme bien connue

$$\sum_{1 \leq s \leq h, 1 \leq r \leq k} q_{rs}^2 \quad (13)$$

avec l'expression (11) de $q_{rs} = (\text{observé-espéré}) / \sqrt{\text{espéré}}$.

Comme dans le cas 2×2 , le nombre q_{rs} ($1 \leq r \leq k, 1 \leq s \leq h$) définira la mesure orientée de l'intensité du lien entre la r -ème modalité du caractère à k valeurs associé à la partition P et la s -ème modalité du caractère à h valeurs associé à la partition Q .

D'ailleurs, c'est précisément un indice de la même forme que q_{rs} que nous adoptons, avant la référence à un indice de vraisemblance, entre attributs de description, dans notre méthode de classification hiérarchique basée sur la vraisemblance des liens.

Devant une table de contingence représentant le croisement de deux caractères descriptifs ; la pratique du spécialiste en Sciences Humaines a longtemps consisté à calculer la valeur observée du χ^2 ($\chi^2_{ob.}$) et à se limiter au rejet ou non de l'hypothèse d'indépendance ou d'absence de lien entre les deux caractères globalement pris.

Or l'importance d'une valeur de $\chi^2_{ob.}$ qui n'entraîne pas nécessairement le rejet de l'hypothèse d'indépendance peut n'être due qu'à quelques rares cases du tableau de contingence où la valeur de q_{rs} est très prononcée de façon positive ou négative ; ce qui exprime des liens locaux qui ne sont nullement à négliger.

Par conséquent, on associera au tableau de contingence (3), le tableau des nombres

$$\{q_{rs} / 1 \leq r \leq k, 1 \leq s \leq h\} \quad (14)$$

pour découvrir et interpréter les liens éventuels entre les modalités de l'un des caractères et celles de l'autre.

L'idée de ce travail consiste à découvrir et à justifier des tables d'indices équivalents aux q_{rs} pour les deux situations suivantes

- . croisement d'une classification nette et d'une classification floue.
- . croisement de deux classifications floues.

Ces nouveaux indices que nous noterons x_{rs} seront formellement analogues aux q_{rs} ; nous les justifierons d'autre part d'un point de vue statistique par la propriété correspondante : la somme de leurs carrés est, dans le cadre d'une hypothèse d'absence de lien à caractère permutational, la réalisation d'une v.a. du χ^2 .

Nous commencerons par considérer le cas du croisement d'une classification dichotomique nette associée à un seul attribut de description et

d'une classification dichotomique floue associée à une classe d'attributs (cf. § II). Nous envisagerons ensuite le croisement de deux classifications dichotomiques floues respectivement associées à deux classes disjointes d'attributs de description.

L'étude de ces deux seuls cas suffit à donner une réponse complète au problème posé dans l'introduction. Soit en effet $\{A_1, A_2, \dots, A_k\}$ une classification par proximité d'un ensemble A d'attributs descriptifs tel que l'opposé \bar{a} d'un même attribut a de A , n'appartienne pas à A , pour tout a de A . Considérons alors le problème de situer les classes d'une classification nette en h classes de E associée à un caractère de description b à h modalités notées b_1, b_2, \dots, b_h . La table de croisement est de dimension $k \times h$ et le contenu x_{rs} de la case (r, s) ($1 \leq r \leq k, 1 \leq s \leq h$) correspond au contenu de la case $(1, 1)$ de la table 2×2 de croisement

| | | |
|-------------|-------|-------------|
| | b_s | \bar{b}_s |
| A_r | | |
| \bar{A}_r | | |

figure 1.

où \bar{A}_r est l'ensemble des attributs opposés à ceux de A_r , et où \bar{b}_s est l'attribut opposé à celui b_s défini par la s -ème modalité du caractère b .

Considérons maintenant un couple de classifications $\{A_1, A_2, \dots, A_k\}$ et $\{B_1, B_2, \dots, B_h\}$ sur deux ensembles disjoints A et B d'attributs de description tels que l'opposé \bar{a} (resp. \bar{b}) d'un même a (resp. b) de A (resp. B), pour tout a de A (resp. b de B). Le croisement de ce couple de classifications polythétiques conduit à une table de dimension $k \times h$ et le contenu x_{rs} de la case (r, s) ($1 \leq r \leq k, 1 \leq s \leq h$) correspond au contenu de la case $(1, 1)$ de croisement des deux classifications dichotomiques floues suivantes (voir schéma) où \bar{A}_r (resp. \bar{B}_s) est formé des attributs opposés à ceux de A_r (resp. B_s) (cf. § II).

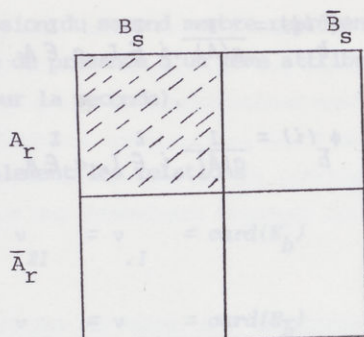


figure 2.

Néanmoins, pour cette dernière situation de croisement du couple de classifications $(\{A_1, A_2, \dots, A_k\}, \{B_1, B_2, \dots, B_h\})$; nous considérerons au niveau global de la table $k \times h$ de croisement, une hypothèse d'absence de lien à caractère permutatif, pour l'étude des statistiques associées aux nombres v_{rs} qui sont les correspondants des nombres n_{rs} du croisement de classifications nettes.

V - CROISEMENT D'UNE CLASSIFICATION DICHOTOMIQUE NETTE ET D'UNE CLASSIFICATION DICHOTOMIQUE FLOUE.

La situation est celle schématisée par la figure 1 ci-dessus où nous désignerons par A la classe d'attributs définissant la classification dichotomique floue et par b l'attribut définissant la classification dichotomique nette.

En associant à b les fonctions indicatrices ϕ_b et $\phi_{\overline{b}}$ et à A les fonctions d'appartenance relative ϕ_A et $\phi_{\overline{A}}$ (cf. § II); les nombres v_{rs} correspondants aux nombres n_{rs} du croisement de classifications nettes, se mettent sous la forme

$$v_{11} = \sum_{i \in I} \phi_A(i) \phi_b(i) = \frac{1}{c(A)} \sum_{i \in I} \sum_{a \in A} \phi_a(i) \phi_b(i)$$

$$v_{12} = \sum_{i \in I} \phi_A(i) \phi_{\overline{b}}(i) = \frac{1}{c(A)} \sum_{i \in I} \sum_{a \in A} \phi_a(i) \phi_{\overline{b}}(i)$$

$$\begin{aligned}
 v_{21} &= \sum_{i \in I} \frac{\phi(i)}{A} \quad \phi(i) = \frac{1}{c(A)} \sum_{i \in I, a \in A} \frac{\phi(i)}{a} \phi(i) \\
 v_{22} &= \sum_{i \in I} \frac{\phi(i)}{A} \quad \phi(i) = \frac{1}{c(A)} \sum_{i \in I, a \in A} \frac{\phi(i)}{a} \phi(i)
 \end{aligned}
 \tag{1}$$

où $c(A)$ est, rappelons-le, le cardinal de A .

On peut également écrire

$$\begin{aligned}
 v_{11} &= \frac{1}{c(A)} \sum_{a \in A} \text{card}(E_a \cap E_b) \\
 v_{12} &= \frac{1}{c(A)} \sum_{a \in A} \text{card}(E_a \cap E_{\bar{b}}) \\
 v_{21} &= \frac{1}{c(A)} \sum_{a \in A} \text{card}(E_a^- \cap E_b) \\
 v_{22} &= \frac{1}{c(A)} \sum_{a \in A} \text{card}(E_a^- \cap E_{\bar{b}})
 \end{aligned}
 \tag{2}$$

où, rappelons le, $\{E_a, E_a^-\}$ (resp. $\{E_b, E_{\bar{b}}\}$) est la partition en deux classes définie par l'attribut a (resp. b).

On a les relations

$$\begin{aligned}
 v_{11} + v_{12} &= v_1 = \frac{1}{c(A)} \sum_{a \in A} \text{card}(E_a) \\
 v_{21} + v_{22} &= v_2 = \frac{1}{c(A)} \sum_{a \in A} \text{card}(E_a^-)
 \end{aligned}
 \tag{3}$$

L'expression du second membre représente la moyenne sur A de la fréquence absolue de présence d'un même attribut (a pour la première relation et \bar{a} pour la seconde).

On a également les relations

$$v_{11} + v_{21} = v_{.1} = \text{card}(E_b) \quad (4)$$

$$v_{12} + v_{22} = v_{.2} = \text{card}(E_{\bar{b}})$$

Enfin

$$v_{1.} + v_{2.} = v_{.1} + v_{.2} = v_{..} = n = \text{card}(E) \quad (5)$$

Ainsi, la connaissance du contenu de l'une des cases du tableau 2×2 des v_{rs} , détermine celui des autres.

Dans ces conditions, les nombres x_{rs} correspondants aux nombres q_{rs} ($1 \leq r \leq 2$, $1 \leq s \leq 2$) (cf. formule (11) § IV) s'expriment sous la forme suivante

$$x_{rs} = (v_{rs} - (v_{r.} v_{.s}/n)) / \sqrt{v_{r.} v_{.s}/n} \quad (6)$$

$$1 \leq r \leq 2, 1 \leq s \leq 2.$$

x_{rs} ($1 \leq r \leq 2$, $1 \leq s \leq 2$) définira la mesure orientée de l'intensité du lien entre le profil d'attitude défini par la classe A_r et l'attribut b_s , où

$$A_1 = A, \quad A_2 = \bar{A}, \quad b_1 = b \text{ et } b_2 = \bar{b}.$$

Des expressions équivalentes à celles (1), (2), (3), (4), (5) et (6) auraient pu être considérées dans le cadre plus général du croisement d'une

k-classification nette et d'une h-classification floue (cf. § III). Toutefois, nous ne les considérerons pas ici compte tenu du caractère contraignant de la définition d'une h-classification floue qui ne peut être associée qu'à des données de forme particulière (caractères de description ayant exactement h modalités) par un algorithme sous contrainte déterminant h classes dont chacune emprunte exactement une modalité de chaque caractère.

Nous allons à présent justifier d'un point de vue statistique l'introduction des expressions x_{rs} , $1 \leq r \leq 2$, $1 \leq s \leq 2$, (cf. (6)). De manière analogue à la comparaison, dans l'hypothèse d'absence de lien à caractère permutatif, d'un couple d'attributs (cf. formules (8) et (8') § IV) ; à la statistique v_{11} nous associons les deux v.a. duales

$$X(\sigma) = \sum_{i \in I} \phi_A(i) \phi_B(\sigma(i)) \quad (7)$$

$$Y(\sigma) = \sum_{i \in I} \phi_A(\sigma(i)) \phi_B(i)$$

où σ est un élément aléatoire dans l'ensemble G_n , muni d'une probabilité uniformément répartie, de toutes les permutations sur I (il y en a $n!$).

$\phi_A(i)$ est la fonction sur I définie par la formule (2) du paragraphe II ci-dessus.

Les deux v.a. duales $X(\sigma)$ et $Y(\sigma)$ sont de même loi, asymptotiquement normale (pour $n \rightarrow \infty$) dans le cadre des conditions suffisantes définies par le théorème de Wald, Wolfowitz et Wothe (cf. par exemple chap. 3 de [6]) qu'on pourra directement vérifier ici. La convergence vers la loi normale est en réalité, le plus souvent, très rapide.

Les expressions de la moyenne et de la variance communes sont

$$\begin{aligned} \mathcal{E}(X(\sigma)) &= \mathcal{E}(Y(\sigma)) = n \mu(\phi_A) \mu(\phi_B) \\ \text{var}(X(\sigma)) &= \text{var}(Y(\sigma)) = \frac{n^2}{(n-1)} \text{var}(\phi_A) \text{var}(\phi_B) \end{aligned} \quad (8)$$

où $\mu(\phi_A)$ (resp. $\mu(\phi_b)$) est la moyenne de la fonction ϕ_A (resp. ϕ_b) sur I et où $\text{var}(\phi_A)$ (resp. $\text{var}(\phi_b)$) est la variance de la fonction ϕ_A (resp. ϕ_b) sur I .

On a

$$\begin{aligned}\mu(\phi_A) &= \frac{1}{n} \sum_{i \in I} \phi_A(i) = \frac{1}{c(A)} \sum_{a \in A} p_a \\ \mu(\phi_b) &= \frac{1}{n} \sum_{i \in I} \phi_b(i) = p_b\end{aligned}\quad (9)$$

où p_a (resp. p_b) est la proportion d'éléments de l'ensemble E des objets (codé par I) où l'attribut a (resp. b) est présent.

D'autre part, on a

$$\text{var}(\phi_A) = \frac{1}{n} \sum_{i \in I} [\phi_A(i)]^2 - \left[\frac{1}{c(A)} \sum_{a \in A} p_a \right]^2$$

où $[\phi_A(i)]^2$ peut se mettre sous la forme

$$\frac{1}{c(A)^2} \left\{ \sum_{a \in A} \phi_a(i) + \sum_{(a,a') \in A^{[2]}} \phi_{a \wedge a'}(i) \right\}$$

où $A^{[2]}$ désigne l'ensemble des couples d'éléments de A à composantes distinctes ($A^{[2]} = A \times A - \mathcal{S}(A \times A)$ où \mathcal{S} est la diagonale) et où $\phi_{a \wedge a'}$ est la fonction indicatrice de $E_a \cap E_{a'}$. De sorte que

$$\text{var}(\phi_A) = \frac{1}{c(A)^2} \left\{ \sum_{(a,a') \in A \times A} (p_{a \wedge a'} - p_a p_{a'}) \right\} \quad (10)$$

où $p_{a \wedge a'} = \text{card}(E_a \cap E_{a'})/n$ est la proportion d'éléments de E où les deux attributs a et a' sont présents.

D'autre part,

$$\text{var}(\phi_b) = p_b (1 - p_b) \quad (11)$$

Par conséquent

$$\begin{aligned} \mathcal{G}(X(\sigma)) &= \mathcal{G}(Y(\sigma)) = \frac{n}{c(A)} \left(\sum_{a \in A} p_a \right) p_b \\ &= \frac{1}{nc(A)} \left(\sum_{a \in A} n(a) \right) n(b) \\ \text{var}(X(\sigma)) &= \text{var}(Y(\sigma)) = \frac{1}{(n-1)} \left(\frac{n}{c(A)} \right)^2 \left[\sum_{A \times A} (p_{a \wedge a'} - p_a p_{a'}) \right] p_b (1 - p_b) \\ &= \frac{1}{(nc(A))^2} \left[\sum_{A \times A} (n(a \wedge a') - \frac{n(a)n(a')}{n}) n(b) n(b) \right] \end{aligned} \quad (12)$$

où $n(a)$, $n(a')$, $n(b)$, $n(b')$ et $n(a \wedge a')$ sont les cardinaux respectifs de E_a , $E_{a'}$, E_b , $E_{b'}$, et $E_a \cap E_{a'}$.

La statistique v_{11} centrée réduite s'écrit dans ces conditions

$$\frac{\sqrt{(n-1)} \sum_{a \in A} (p_{a \wedge b} - p_a p_b)}{\sqrt{\left[\sum_{A \times A} (p_{a \wedge a'} - p_a p_{a'}) \right] p_b (1 - p_b)}} \quad (13)$$

dont le carré peut être considéré, dans l'hypothèse d'absence de lien décrite ci-dessus, comme la réalisation d'une v.a. du χ^2 à 1 degré de liberté.

Nous allons à présent vérifier que

$$\sum_{\substack{1 \leq r \leq 2 \\ 1 \leq s \leq 2}} x_{rs}^2 \quad (14)$$

(cf. (6)) est à un coefficient multiplicatif près, le carré de la statistique (13). Ce coefficient ne dépendra que des marges de la table de croisement 2×2 ; c'est-à-dire, de caractéristiques liées à la distribution sur E de b d'une part et de A d'autre part.

On peut écrire

$$x_{11} = x_{Ab} = \sqrt{\frac{n}{c(A)}} \left[\sum_{a \in A} (p_{a \wedge b} - p_a p_b) \right] / \sqrt{\sum_{a \in A} p_a p_b}$$

$$x_{12} = x_{A\bar{b}} = \sqrt{\frac{n}{c(A)}} \left[\sum_{a \in A} (p_{a \wedge \bar{b}} - p_a p_{\bar{b}}) \right] / \sqrt{\sum_{a \in A} p_a p_{\bar{b}}}$$

$$x_{21} = x_{\bar{A}b} = \sqrt{\frac{n}{c(A)}} \left[\sum_{a \in A} (p_{\bar{a} \wedge b} - p_{\bar{a}} p_b) \right] / \sqrt{\sum_{a \in A} p_{\bar{a}} p_b}$$

$$x_{22} = x_{\bar{A}\bar{b}} = \sqrt{\frac{n}{c(A)}} \left[\sum_{a \in A} (p_{\bar{a} \wedge \bar{b}} - p_{\bar{a}} p_{\bar{b}}) \right] / \sqrt{\sum_{a \in A} p_{\bar{a}} p_{\bar{b}}}$$

(15)

et vérifier que le carré du numérateur de x_{rs} est le même pour tout $r=1,2$ et tout $s=1,2$; on tiendra compte pour cela des relations suivantes

$$p_{\bar{a}} = (1 - p_a), \quad p_{\bar{b}} = (1 - p_b), \quad p_{\bar{a} \wedge b} = p_b - p_{a \wedge b}$$

et

$$p_{\bar{a} \wedge \bar{b}} = 1 - p_a - p_b + p_{a \wedge b}.$$

Par conséquent, la somme des carrés (14) se met sous la forme

$$\frac{n}{c(A)} \left[\sum_{a \in A} (p_{a \wedge b} - p_a p_b) \right]^2 \times \left\{ \left(\sum_A p_a p_b \right)^{-1} + \left(\sum_A p_a p_{\bar{b}} \right)^{-1} + \left(\sum_A p_{\bar{a}} p_b \right)^{-1} + \left(\sum_A p_{\bar{a}} p_{\bar{b}} \right)^{-1} \right\} \quad (16)$$

qui, après simplification, s'écrit

$$n \left[\sum_{a \in A} (p_{a \wedge b} - p_a p_b) \right]^2 / \left(\sum_{a \in A} p_a \right) \left(\sum_{a \in A} p_a^- \right) p_b p_b^- \quad (17)$$

En assimilant $(n-1)$ à n ; l'expression (17) est, au coefficient multiplicatif

$$\left[\sum_{(a, a') \in A \times A} (p_{a \wedge a'} - p_a p_{a'}) \right] / \left[\sum_{(a, a') \in A \times A} p_a p_{a'} \right]$$

le carré de l'expression (13). D'où le résultat qu'on peut énoncer sous la forme d'un théorème.

Théorème. La somme des carrés des statistiques x_{rs} , $1 \leq r \leq 2$, $1 \leq s \leq 2$, (cf. (6)) est dans l'hypothèse d'absence de lien à caractère permutational (cf. (7)), la réalisation d'une v.a. du χ^2 à 1 degré de liberté. x_{rs} représente la contribution orientée à la mesure de la dépendance par une statistique type χ^2 entre la classification dichotomique floue et celle, nette.

VI - CROISEMENT ENTRE DEUX CLASSIFICATIONS DICHOTOMIQUES FLOUES.

La situation que nous allons considérer maintenant, schématisée par la figure 2 du paragraphe IV précédent, est analogue mais plus générale que celle ci-dessus (§ V). Toutefois, nous avons voulu distinguer les deux situations car d'une part elles répondent à des objectifs d'analyse différents de par leur nature ; d'autre part, en procédant graduellement, on se rend compte en quoi les calculs peuvent différer en passant du croisement entre deux classifications dichotomiques nettes à celui où l'une est floue et l'autre nette et à celui enfin, où les deux sont floues.

Soient A et B les deux classes disjointes d'attributs définissant respectivement les deux classifications dichotomiques floues à croiser. Associons à A (resp. B) les fonctions d'appartenance relative ϕ_A et ϕ_A^- (resp. ϕ_B et ϕ_B^-) (cf. § II). Les nombres v_{rs} , $1 \leq r \leq 2$, $1 \leq s \leq 2$, deviennent ici

$$v_{11} = v_{AB} = \sum_{i \in I} \phi_A(i) \phi_B(i) = \frac{1}{c(A)c(B)} \sum_{i \in I, (a,b) \in A \times B} \phi_a(i) \phi_b(i)$$

$$= \frac{1}{c(A)c(B)} \sum_{i \in I_1} \sum_{(a,b) \in A \times B} \phi_{a \wedge b}^{(i)} = \frac{1}{c(A)c(B)} \sum_{(a,b) \in A \times B} \text{card}(E_a \cap E_b)$$

$$v_{12} = \frac{v}{\overline{AB}} = \frac{1}{c(A)c(B)} \sum_{i \in I_1} \sum_{(a,b) \in A \times B} \phi_{a \wedge \bar{b}}^{(i)} = \frac{1}{c(A)c(B)} \sum_{(a,b) \in A \times B} \text{card}(E_a \cap E_{\bar{b}})$$

$$v_{21} = \frac{v}{\overline{AB}} = \frac{1}{c(A)c(B)} \sum_{i \in I_1} \sum_{(a,b) \in A \times B} \phi_{\bar{a} \wedge b}^{(i)} = \frac{1}{c(A)c(B)} \sum_{(a,b) \in A \times B} \text{card}(E_{\bar{a}} \cap E_b)$$

$$v_{22} = \frac{v}{\overline{AB}} = \frac{1}{c(A)c(B)} \sum_{i \in I_1} \sum_{(a,b) \in A \times B} \phi_{\bar{a} \wedge \bar{b}}^{(i)} = \frac{1}{c(A)c(B)} \sum_{(a,b) \in A \times B} \text{card}(E_{\bar{a}} \cap E_{\bar{b}}) \quad (1)$$

D'autre part, on obtient

$$v_{1.} = v_{A.} = v_{11} + v_{12} = \frac{1}{c(A)} \sum_{a \in A} n(a) = \frac{n}{c(A)} \sum_{a \in A} P_a$$

$$v_{2.} = v_{A.} = v_{21} + v_{22} = \frac{1}{c(A)} \sum_{a \in A} n(\bar{a}) = \frac{n}{c(A)} \sum_{a \in A} P_{\bar{a}}$$

$$v_{.1} = v_{.B} = v_{11} + v_{21} = \frac{1}{c(B)} \sum_{b \in B} n(b) = \frac{n}{c(B)} \sum_{b \in B} P_b$$

$$v_{.2} = v_{.B} = v_{12} + v_{22} = \frac{1}{c(B)} \sum_{b \in B} n(\bar{b}) = \frac{n}{c(B)} \sum_{b \in B} P_{\bar{b}}$$

$$n = v_{1.} + v_{2.} = v_{.1} + v_{.2}.$$

Dans ces conditions, le nombre x_{rs} ($1 \leq r \leq 2$, $1 \leq s \leq 2$) qui définit la mesure orientée de l'intensité du lien entre le profil d'attitude défini par la classe A_r ($A_1 = A$ et $A_2 = \bar{A}$) et celui défini par la classe B_s ($B_1 = B$ et $B_2 = \bar{B}$), s'exprime par la formule analogue à celle (6) du paragraphe précédent

$$x_{rs} = (v_{rs} - (v_{r.} v_{.s}/n)) / \sqrt{v_{r.} v_{.s}/n} \quad (3)$$

$$1 \leq r \leq 2, 1 \leq s \leq 2.$$

Ici encore, compte tenu des relations (2), la connaissance de l'une des cases du tableau 2×2 des v_{rs} , détermine celle des autres. Pour justifier d'un point de vue statistique l'introduction des x_{rs} , $1 \leq r \leq 2$, $1 \leq s \leq 2$; associons à la statistique observée v_{11} , les deux v.a. duales

$$\begin{aligned} X(\sigma) &= \sum_{i \in I} \phi_A(i) \phi_B(\sigma(i)) \\ Y(\sigma) &= \sum_{i \in I} \phi_A(\sigma(i)) \phi_B(i) \end{aligned} \quad (4)$$

où σ est un élément aléatoire dans l'ensemble G_n , muni d'une probabilité uniformément répartie, de toutes les permutations sur I (il y en a $n!$).

Les deux v.a. duales $X(\sigma)$ et $Y(\sigma)$ sont de même loi, asymptotiquement normale (pour $n \rightarrow \infty$), dans le cadre des conditions suffisantes définies par le théorème de Wald, Wolfowitz et Noether (cf. chap. 3 de [6]) qu'on pourra directement vérifier ici. Cette convergence est, le plus souvent, très rapide.

Avec des notations que l'on comprend et qu'on a déjà utilisé au paragraphe précédent; les expressions de la moyenne et de la variance commune sont

$$\begin{aligned} \mathcal{U}(X(\sigma)) &= \mathcal{U}(Y(\sigma)) = n \mu(\phi_A) \mu(\phi_B) \\ \text{var}(X(\sigma)) &= \text{var}(Y(\sigma)) = \frac{n^2}{(n-1)} \text{var}(\phi_A) \text{var}(\phi_B) \end{aligned} \quad (5)$$

Compte tenu de calculs déjà effectués (cf. formules (9) et (10 § V), on a

$$\begin{aligned} \mu(\phi_A) &= \frac{1}{c(A)} \sum_{a \in A} P_a \\ \mu(\phi_B) &= \frac{1}{c(B)} \sum_{b \in B} P_b \end{aligned} \quad (6)$$

$$\text{var}(\phi_A) = \frac{1}{c(A)^2} \left\{ \sum_{(a,a') \in A \times A} (p_{a \wedge a'} - p_a p_{a'}) \right\} \quad (7)$$

où, rappelons le, $p_{a \wedge a'} = \text{card}(E_a \cap E_{a'})/n$.

De même,

$$\text{var}(\phi_B) = \frac{1}{c(B)^2} \left\{ \sum_{(b,b') \in B \times B} (p_{b \wedge b'} - p_b p_{b'}) \right\} \quad (7')$$

Par conséquent

$$\psi(X(\sigma)) = \psi(Y(\sigma)) = \frac{n}{c(A)c(B)} \sum_{(a,b) \in A \times B} p_a p_b \quad (8)$$

$$\begin{aligned} \text{var}(X(\sigma)) = \text{var}(Y(\sigma)) = \frac{1}{(n-1)} \left(\frac{n}{c(A)c(B)} \right)^2 & \left\{ \sum_{A \times A} (p_{a \wedge a'} - p_a p_{a'}) \right\} \\ & \times \left\{ \sum_{B \times B} (p_{b \wedge b'} - p_b p_{b'}) \right\} \quad (9) \end{aligned}$$

La statistique v_{11} centrée réduite s'écrit alors ici

$$\frac{\sqrt{(n-1)} \sum_{A \times B} (p_{a \wedge b} - p_a p_b)}{\sqrt{\left[\sum_{A \times A} (p_{a \wedge a'} - p_a p_{a'}) \right] \left[\sum_{B \times B} (p_{b \wedge b'} - p_b p_{b'}) \right]}} \quad (10)$$

dont le carré peut être considéré, dans l'hypothèse d'absence de lien qui a prévalu à l'introduction des v.a. (4), comme la réalisation d'une v.a. du χ^2 à 1 degré de liberté.

Comme au paragraphe précédent, nous allons maintenant vérifier que

$$\sum_{\substack{1 \leq r \leq 2 \\ 1 \leq s \leq 2}} z_{rs}^2$$

(cf. (3), (2) et (1)) est à un coefficient multiplicatif près le carré de la statistique (10). Ce coefficient ne dépendra que des marges de la table de croisement 2×2 ; c'est-à-dire, de caractéristiques liées à la distribution sur E de la classe A d'attributs d'une part et de celle, B d'autre part.

Reprenons le détail des expressions de x_{rs} , $1 \leq r \leq 2, 1 \leq s \leq 2$.

$$x_{11} = x_{AB} = \sqrt{\frac{n}{c(A)c(B)}} \left[\sum_{AxB} (p_{a\lambda b} - p_a p_b) \right] / \sqrt{\sum_{AxB} p_a p_b}$$

$$x_{12} = x_{AB} = \sqrt{\frac{n}{c(A)c(B)}} \left[\sum_{AxB} (p_a b - p_a p_b) \right] / \sqrt{\sum_{AxB} p_a p_b}$$

$$x_{21} = x_{\bar{A}\bar{B}} = \sqrt{\frac{n}{c(A)c(B)}} \left[\sum_{AxB} (p_{\bar{a} b} - p_a p_b) \right] / \sqrt{\sum_{AxB} p_a p_b}$$

$$x_{22} = x_{\bar{A}\bar{B}} = \sqrt{\frac{n}{c(A)c(B)}} \left[\sum_{AxB} (p_{\bar{a} \bar{b}} - p_a p_b) \right] / \sqrt{\sum_{AxB} p_a p_b}$$

(12)

Encore une fois, en tenant compte des relations

$$p_{\bar{a}} = (1-p_a), \quad p_{\bar{b}} = (1-p_b), \quad p_{\bar{a}\lambda b} = p_b - p_{a\lambda b} \quad \text{et} \quad p_{\bar{a}\lambda \bar{b}} = 1-p_a - p_b + p_{a\lambda b};$$

on vérifie que le carré du numérateur de x_{rs} est le même pour tout $r = 1, 2$ et tout $s = 1, 2$. Donc, la somme des carrés (11) s'exprime par la formule

$$\frac{n}{c(A)c(B)} \left(\sum_{AxB} (p_{a\lambda b} - p_a p_b) \right)^2 \left\{ \left[\sum_{AxB} p_a p_b \right]^{-1} \left[\sum_{AxB} p_a p_b \right]^{-1} + \left[\sum_{AxB} p_a p_b \right]^{-1} + \left[\sum_{AxB} p_a p_b \right]^{-1} \right\} \quad (13)$$

qui, après simplification, s'écrit

$$n \left[\sum_{A B} (p_{a\lambda b} - p_a p_b) \right]^2 / \left(\sum_A p_a \right) \left(\sum_A p_{\bar{a}} \right) \left(\sum_B p_b \right) \left(\sum_B p_{\bar{b}} \right)$$

En confondant $(n-1)$ et n ; cette expression est, au coefficient multiplicatif

$$\frac{\left[\sum_{A \times A} (p_{a \wedge a'} - p_a p_{a'}) \right]}{\sum_{A \times A} p_a p_{a'}} \times \frac{\left[\sum_{B \times B} (p_{b \wedge b'} - p_b p_{b'}) \right]}{\sum_{B \times B} p_b p_{b'}}$$

le carré de l'expression (10). D'où le théorème

Théorème. La somme des carrés des statistiques x_{rs} , $1 \leq r \leq 2$, $1 \leq s \leq 2$, (cf. (3)) est dans l'hypothèse d'absence de lien à caractère permutational (cf. (4)), la réalisation d'une v.a. du χ^2 à 1 degré de liberté. x_{rs} représente la contribution orientée à la mesure de la dépendance par une statistique type χ^2 entre deux classifications dichotomiques floues.

VII - CROISEMENT ENTRE DEUX TYPOLOGIES

Soient A et B deux ensembles disjoints d'attributs de description tels que l'opposé \bar{a} (resp. \bar{b}) d'un même attribut a de A (resp. b de B) n'appartient pas à A (resp. B) pour tout a de A (resp. b de B). Pour fixer les idées, on peut reprendre l'exemple de l'introduction où, relativement à une enquête à caractère psycho-sociologique, A est formé des attributs de comportement et B des caractéristiques socio-individuelles. Désignons par

$$\{A_1, A_2, \dots, A_r, \dots, A_k\} \quad (\text{resp. } \{B_1, B_2, \dots, B_s, \dots, B_h\})$$

une classification par proximité de A (resp. B) où une même classe A_r ($1 \leq r \leq k$) (resp. B_s ($1 \leq s \leq h$)) définit un profil de comportement (resp. sociologique).

La situation peut être schématisée par la table de croisement

| | B_1 | \dots | B_s | \dots | B_h |
|---------|-------|---------|-------|---------|-------|
| A_1 | | | | | |
| \cdot | | | | | |
| \cdot | | | | | |
| A_r | | | | | |
| \cdot | | | | | |
| \cdot | | | | | |
| A_k | | | | | |

où, rappelons le, à l'intersection de la r -ème ligne et de la s -ème colonne ($1 \leq r \leq k$, $1 \leq s \leq h$) on place le nombre x_{rs} qui est la contribution orientée à la statistique du χ^2 à 1 degré de liberté, attachée à la table de croisement des deux classifications dichotomiques floues respectivement associées à A_r et B_s (cf. figure 2 § IV ci-dessus).

Considérons la suite, dans un ordre fixé, des couples

$$\{(X_{rs}(\sigma_{rs}), Y(\sigma_{rs})) / 1 \leq r \leq k, 1 \leq s \leq h.\} \quad (1)$$

de v.a. de même loi, respectivement associées aux statistiques observées

$$v_{rs} = \sum_{i \in I} \phi_{A_r}(i) \phi_{B_s}(i) \quad (2)$$

$1 \leq r \leq k$, $1 \leq s \leq h$. (cf. formule (4) § VI).

Dans l'optique développée ci-dessus, les différents σ_{rs} ($1 \leq r \leq k$, $1 \leq s \leq h$) sont des éléments aléatoires indépendants de l'ensemble G_n , muni d'une probabilité uniformément répartie, de toutes les permutations sur $(1, 2, \dots, n)$.

Comme nous l'avons annoncé à la fin du paragraphe IV, nous allons à présent considérer une hypothèse d'absence de lien à caractère plus restrictif, définie globalement au niveau de toute la table de croisement. Ainsi aux v_{rs} ($1 \leq r \leq k, 1 \leq s \leq h$), on associe la suite de couples

$$\{ (X_{rs}(\sigma), Y_{rs}(\sigma)) \quad / \quad 1 \leq r \leq k, 1 \leq s \leq h \} \quad (3)$$

de v.a.; de même loi (cf. formule (4) § VI) où σ , cette fois-ci le même pour toutes les cases du tableau, est un élément aléatoire dans l'ensemble G_n , muni d'une probabilité uniformément répartie, de toutes les permutations sur $(1, 2, \dots, n)$.

Puisque la loi de $Y_{rs}(\sigma)$ est la même que celle de $X_{rs}(\sigma)$; contentons nous de considérer la suite de v.a.

$$\{ X_{rs}(\sigma) / 1 \leq r \leq k, 1 \leq s \leq h \} \quad (4)$$

Considérons une combinaison linéaire quelconque de cette suite de v.a.; soit

$$\sum_{\substack{1 \leq r \leq k \\ 1 \leq s \leq h}} \beta_{rs} X_{rs}(\sigma) \quad (5)$$

que nous écrivons sous perte de généralité sous la forme

$$\sum_{\substack{1 \leq r \leq k \\ 1 \leq s \leq h}} \alpha_r \alpha_s X_{rs}(\sigma) = \sum_{i \in I} \left\{ \sum_{1 \leq r \leq k} \alpha_r \phi_{A_r}(i) \right\} \left\{ \sum_{1 \leq s \leq h} \alpha_s \phi_{B_s}[\sigma(i)] \right\} \quad (6)$$

qui est une v.a. asymptotiquement normale (pour $n \rightarrow \infty$) dans le cadre des conditions suffisantes qui ont un caractère assez général de Wald, Wolfowitz et Noether (cf. chap. 3 de [6]).

Il en résulte que la distribution jointe de la suite (4) de v.a. est asymptotiquement normale dans le sens suivant : s'il existe une sous suite de p v.a. dont la forme limite de la matrice V des covariances est définie positive; alors la forme limite de la distribution jointe de la fréquence est la

fonction densité

$$f(K) = K \exp \left(-\frac{1}{2} t_z' V z \right) \quad (7)$$

où z est un point de R^p et où K est une constante. (cf. par exemple lemme 2.3. chap. 5 de [4]).

Nous allons dans ces conditions montrer que, de façon très générale, on peut s'attendre à ce que la matrice W des covariances de la totalité des $k \times h$ v.a. de la suite (4), soit définie positive.

1 - Calcul de la matrice W des covariances

Vu qu'on s'intéresse au comportement limite de la matrice W , on n'hésitera pas à confondre au terme des calculs $(n-1)$ et n . On a déjà calculé au paragraphe précédent la variance d'une même v.a. $X_{rs}(\sigma)$ (cf. formule (9) § VI) que nous allons réécrire :

$$\text{var} [X_{rs}(\sigma)] = \left(\frac{\sqrt{n}}{c(A_r)c(B_s)} \right)^2 \left\{ \sum_{A_r \times A_r} (p_{a\lambda a'} - p_a p_{a'}) \right\} \left\{ \sum_{B_s \times B_s} (p_{b\lambda b'} - p_b p_{b'}) \right\} \quad (5)$$

avec des notations explicitées au paragraphe précédent.

Calculons à présent, de façon générale

$$\mathcal{C}[X_{rs}(\sigma) X_{r's'}(\sigma)] \quad (6)$$

pour $1 \leq r, r' \leq k$ et $1 \leq s, s' \leq h$.

Le produit sous le signe [] du second membre de (6) s'écrit

$$\left\{ \sum_{i \in I} \frac{1}{c(A_r)c(B_s)} \sum_{(a,b) \in A_r \times B_s} \phi_a(i) \phi_b[\sigma(i)] \right\}$$

$$\times \left\{ \sum_{i' \in I} \frac{1}{c(A_{r'})c(B_{s'})} \sum_{(a',b') \in A_{r'} \times B_{s'} \phi_{a'}(i') \phi_{b'}[\sigma(i')] \right\} \quad (7)$$

$$= \frac{1}{c(A_{p'})c(A_p)c(B_s)c(B_{s'})} \sum_{(a,a',b,b') \in A_p \times A_{p'} \times B_s \times B_{s'}} \left\{ \sum_{i \in I} \phi_{a \wedge a'}(i) \phi_{b \wedge b'}[\sigma(i)] \right. \\ \left. + \sum_{(i,i') \in I^{[2]}} \phi_a(i) \phi_{a'}(i') \phi_b[\sigma(i)] \phi_{b'}[\sigma(i')] \right\} \quad (8)$$

où $I^{[2]}$ est l'ensemble des couples à composantes distinctes d'éléments de I .

L'opérateur espérance \mathcal{U} désignant

$$\frac{1}{n!} \sum_{\sigma \in G_n},$$

où G_n est l'ensemble des $n!$ permutations sur I ; on a, en utilisant un raisonnement par symétrie, la valeur suivante de l'espérance mathématique du contenu de l'accolade de (8) :

$$\frac{1}{n} n(a \wedge a')n(b \wedge b') + \frac{1}{n(n-1)} [n(a)n(a') - n(a \wedge a')] [n(b)n(b') - n(b \wedge b')] \quad (9)$$

où $n(a), n(a'), n(b), n(b'), n(a \wedge a')$ et $n(b \wedge b')$ représentent respectivement $\text{card}(E_a), \text{card}(E_{a'}), \text{card}(E_b), \text{card}(E_{b'}), \text{card}(E_a \cap E_{a'})$ et $\text{card}(E_b \cap E_{b'})$.

D'autre part, compte tenu de la formule (8) du paragraphe VI précédent

$$\mathcal{U}[X_{rs}(\sigma)] = \frac{1}{nc(A_{p'})c(B_s)} \sum_{(a,b) \in A_p \times B_s} n_a n_b \quad (10)$$

Par conséquent, on a l'expression suivante de la covariance

$$Y_{r'p's's'} = \text{cov}[X_{rs}(\sigma), X_{r'p's'}(\sigma)] = \frac{1}{c(A_{p'})c(A_{p'})c(B_s)c(B_{s'})} \sum_{A_p \times A_{p'} \times B_s \times B_{s'}} \left\{ \frac{1}{n-1} \right. \\ \left. [n(a \wedge a')n(b \wedge b') - n(a \wedge a') \frac{n(b)n(b')}{n} - n(b \wedge b') \frac{n(a)n(a')}{n} + \frac{n(a)n(a')n(b)n(b')}{n^2}] \right\} \quad (11)$$

D'où

$$\gamma_{pp'ss'} = \frac{1}{\sqrt{n}} \frac{1}{c(A_p)c(A_{p'})} \sum_{A_p \times A_{p'}} [n(a \wedge a') - \frac{n(a)n(a')}{n}]$$

$$\times \frac{1}{\sqrt{n}} \frac{1}{c(B_s)c(B_{s'})} \sum_{B_s \times B_{s'}} [n(b \wedge b') - \frac{n(b)n(b')}{n}] \quad (12)$$

qui peut aussi se mettre sous la forme

$$\frac{\sqrt{n}}{c(A_p)c(A_{p'})} \sum_{A_p \times A_{p'}} [p(a \wedge a') - p(a)p(a')]$$

$$\times \frac{\sqrt{n}}{c(B_s)c(B_{s'})} \sum_{B_s \times B_{s'}} [p(b \wedge b') - p(b)p(b')] \quad (13)$$

on vérifie bien que $\gamma_{pp'ss'}$ est la variance de $X_{p's}$ (σ) conformément à la formule (9) du paragraphe précédent.

2 - Caractère défini positif de W.

Considérons à présent, associée à une typologie $\{C_t / 1 \leq t \leq l\}$ qui est une classification polythétique d'un ensemble C d'attributs de description, la famille de vecteurs de R^l

$$\{(\phi_{C_1}(i), \dots, \phi_{C_t}(i), \dots, \phi_{C_l}(i)) / i \in I\} \quad (14)$$

où, rappelons le, $\phi_{C_t}(i)$ est la proportion d'attributs de la classe C_t possédés par l'individu i . Introduisons la condition (c) que cette famille de vecteurs engendre par combinaison linéaire l'espace R^l . (c) signifie qu'il existe un sous ensemble $\{i_t / 1 \leq t \leq l\}$ de l points de I , tel que l'ensemble des l vecteurs

$$\{(\phi_{C_1}(i_t), \dots, \phi_{C_l}(i_t)) / 1 \leq t \leq l\} \quad (15)$$

constitue un système libre.

Une telle condition est en réalité très peu restrictive si on tient compte de la définition intuitive d'une classification polythétique (cf. chap. 3 de [5]). Il suffit, mais c'est loin d'être nécessaire, que les individus i_1, i_2, \dots, i_l correspondent à des types "purs" mutuellement distincts ; c'est-à-dire que les attributs présents chez l'individu i_t appartiennent exclusivement à la classe C_t , $1 \leq t \leq l$.

Nous avons dans ces conditions la propriété suivante.

Lemme 1 :

Si la classification $\{A_r / 1 \leq r \leq k\}$ satisfait la condition (c), la matrice $k \times k$ de terme général

$$\alpha_{rr'} = \frac{1}{c(A_r)c(A_{r'})} \sum_{A_r \times A_{r'}} \left[\frac{n(a \wedge a') - \frac{n(a)n(a')}{n}}{n} \right], \quad (16)$$

$1 \leq r, r' \leq k$, est définie positive.

En tenant compte de la relation

$$\sum_{i \in I} [\phi_a(i) - p(a)] [\phi_{a'}(i) - p(a')] = \frac{n(a \wedge a') - \frac{n(a)n(a')}{n}}{n}, \quad (17)$$

où $p(a) = \frac{1}{n} \sum_{i \in I} \phi_a(i)$; on peut écrire

$$\alpha_{rr'} = \sum_{i \in I} [\phi_{A_r}(i) - p(A_r)] [\phi_{A_{r'}}(i) - p(A_{r'})], \quad (18)$$

où $p(A_r) = \frac{1}{c(A_r)} \sum_{a \in A_r} p(a)$ et $p(A_{r'}) = \frac{1}{c(A_{r'})} \sum_{a' \in A_{r'}} p(a')$

Dans ces conditions, la forme quadratique

$$\sum_{\substack{1 \leq r \leq k \\ 1 \leq r' \leq k}} \alpha_{rr'} x_r x_{r'} \quad (19)$$

peut se mettre sous la forme

$$\sum_{i \in I} \sum_{\substack{1 \leq r \leq k \\ 1 \leq r' \leq k}} [\phi_{A_r}(i) - p(A_r)] [\phi_{A_{r'}}(i) - p(A_{r'})] x_r x_{r'}$$

$$= \sum_{i \in I} \left(\sum_{1 \leq r \leq k} [\phi_{A_r}(i) - p(A_r)] x_r \right)^2 \quad (20)$$

par conséquent, la forme quadratique (19) est positive. Si de plus la condition (c) est remplie relativement à la typologie $\{A_1, \dots, A_k\}$; il existe un sous ensemble $\{i_t / 1 \leq t \leq k\}$ de k points sur I , tel que l'ensemble des k vecteurs de R^k

$$\{(\phi_{A_1}(i_t), \phi_{A_2}(i_t), \dots, \phi_{A_k}(i_t)) / 1 \leq t \leq k\}$$

forme un système libre; par conséquent, tel que l'ensemble des k vecteurs

$$\{(\phi_{A_1}(i_t) - p(A_1), \phi_{A_2}(i_t) - p(A_2), \dots, \phi_{A_k}(i_t) - p(A_k)) / 1 \leq t \leq k\}$$

forme également un système libre.

Dans cette situation, la nullité de l'expression (20) implique celle de chacune des composantes x_r du vecteur $(x_1, x_2, \dots, x_r, \dots, x_k)$. La forme quadratique (19) est donc définie.

Supposons à présent que chacune des deux matrices carrées $k \times k$ et $h \times h$ de termes généraux respectifs $\alpha_{rr'}$, (cf. formule (16)) et $\beta_{ss'}$,

$$\beta_{ss'} = \frac{1}{c(B_s)c(B_{s'})} B_s \times B_{s'} \left(n(b \wedge b') - \frac{n(in(b \vee))}{n} \right) ; \quad (22)$$

soient définies positives. Nous allons voir immédiatement qu'il en est de même de la matrice $(k \times h) \times (k \times h)$ de terme général $\gamma_{rr'ss'}$, (cf. formule (12)). Au coefficient $1/n$ près, on a

$$\gamma_{rr'ss'} = \alpha_{rr'} \beta_{ss'} \quad (23)$$

Soit $(z_{rs} / 1 \leq r \leq k, 1 \leq s \leq h)$ un vecteur quelconque de $R^{k \times h}$ et considérons la forme quadratique

$$\sum_{(r, r', s, s')} \gamma_{rr'ss'} z_{rs} z_{r's'} \quad (24)$$

Sans perte de généralité, on peut et d'une infinité de façons d'ailleurs, poser

$$z_{rs} = x_r y_s \text{ pour tout } r, 1 \leq r \leq k \text{ et tout } s, 1 \leq s \leq h.$$

L'expression (24) précédente devient

$$\left\{ \sum_{(r, r')} \alpha_{rr'} x_r x_{r'} \right\} \left\{ \sum_{(s, s')} \beta_{ss'} y_s y_{s'} \right\} \quad (25)$$

Chacune des formes quadratiques entre accolades étant positive ; il en est de même de celle (24). D'autre part, la nullité de l'expression (24) entraîne nécessairement celle de l'un des facteurs entre accolades de (25) qui implique alors

$$z_{rs} = 0 \text{ pour tout } r, 1 \leq r \leq k \text{ et tout } s, 1 \leq s \leq h.$$

Donc la forme quadratique (24) est définie positive. D'où le théorème.

Théorème. Si chacune des typologies $\{A_r / 1 \leq r \leq k\}$ et $\{B_s / 1 \leq s \leq h\}$ vérifie la condition (c) ci-dessus définie, alors la matrice des covariances W est définie positive.

Corollaire. La loi jointe des $k \times h$ v.a. $X_{rs}(o)$ (cf. (4) ci-dessus) est asymptotiquement normale de dimension $k \times h$.

La programmation et première analyse des résultats du croisement sur des données réelles en pédagogie mathématique ont été réalisées par Mme HARDOUIN (cf. [3]). Le programme a également été appliqué par T. CHANTREL (cf. [2]) dans le cadre de l'analyse de volumineuses données en psycho-sociologie. Dans tous les cas, les résultats concrets ont beaucoup intéressé le spécialiste des données.

BIBLIOGRAPHIE

- [1] J.P. BENZECRI - "L'analyse des données" Tome II (T II.A. n°2 § 2.4), Dunod, Paris (1973).
- [2] T. CHANTREL. - "Nouvelle approche dans la classification et représentation d'un vaste ensemble d'échelles et profils d'attitude. Application à des données en économie rurale et en psycho-sociologie". Thèse à paraître en 1979. Université de Rennes I, I.R.I.S.A.
- [3] M. HARDOUIN. - "Analyse informatique et statistique du croisement entre classifications floues sur un exemple de pédagogie mathématique". Rapport I.R.I.S.A. à paraître.
- [4] H.O. LANCASTER. - "The chi-squared distribution", John Wiley (1969).
- [5] I.C. LERMAN. - "Reconnaissance et classification des structures finies en analyse des données", volume I, théorie et méthodes, rapport n°70 I.R.I.S.A., Rennes (1977).
- [6] M.L. PURI et P.K. SEN. - "Nonparametric methods in multivariate analysis" John Wiley (1971).
- [7] L.A. ZADEH. - "Fuzzy sets", Information and control 8 (1965).

Reçu en Décembre 1978

(1)

(2)