

SUR LE CODAGE BARYCENTRIQUE LINÉAIRE PAR
MORCEAUX D'UN QUESTIONNAIRE

RUFIN BIDOUNGA

ABSTRACT. Our study is focused on the research of a coding that would keep as best as can be the real configuration of a questionnaire containing ordered categorical variables which are graded according to approbation. The piecewise barycentric codings with picewise acting pivots and an active zero describes both an improvement and and a generalisation of the personal equation coding keeping the order of given rating. As we have given to the middle pivot a distinct action on the different belonging functions for the four manners of response, we have established a more refined coding which distinguishes the ways to separate the the initial data $x(i,j)$ on the intervals $[-a,0]$ and $[0,b]$. Hence, the re-creating formula of the initial data are more precise. The acting zero leads to an infinite numbers of coding characterized by homotopical and homologous ways.

Résumé.

Notre étude se focalise sur la recherche d'un codage conservant au mieux les configurations réelles d'un questionnaire clos à variables qualitatives ordonnées sur une échelle d'approbation. Les codages barycentriques linéaires par morceaux à pivots agissant par morceaux avec zéro actif traduisent une amélioration et une généralisation du codage suivant l'équation personnelle conservateur de l'ordre des modalités. Ayant conféré au pivot central *zéro* une action distincte sur les différentes fonctions d'appartenance aux quatre modalités, nous avons établi un codage plus fin distinguant les manières d'éclater les notes recadrées $x(i,j)$ sur les intervalles $[-a,0]$ et $[0,b]$. Ainsi les formules de reconstitution des données initiales se sont révélées plus précises. Le zéro actif fait apparaître une infinité de codages qui sont déterminés par des chemins homotopes et homologues.

2000 *Mathematics Subject Classification.* Primary 94A50;62-07; Secondary 62H17; 62H25.

Key words and phrases. **English** : Data analysis ; Questionnaire ; Pivoting method ; Contingency table ; Correspondence analysis; Piecewise linearity. **Français** : Analyse des données ; Questionnaire ; usage des pivots ; Tableau contingence ; Analyse des correspondances; Codage barycentrique ; Linéarité par morceaux.

1. PROBLEMATIQUE

Face à un questionnaire dont les questions sont quelque peu embarrassantes, les sujets adoptent souvent une attitude préservant leur pensée véritable.

Dans ce travail, nous essayons de mettre en évidence une stratégie du comportement des réponses au dépens du sens vis-à-vis d'une variable ordonnée d'un questionnaire. Pour cela, nous mettrons à profit les mérites des codages flous tout en mettant en évidence les atouts et les insuffisances du codage disjonctif complet qui, généralement, pour l'étude des sondages d'opinions, ne prend en compte ni les relations de proximité ou de similitude entre les modalités de réponses, ni les particularités individuelles des réponses des sujets (Cazes [10]-[12]).

Notre problématique est la recherche d'un codage conservant au mieux les configurations réelles d'un questionnaire à variables qualitatives ordonnées. Ainsi, proposons-nous un codage suivant l'équation personnelle, qui est un codage barycentrique, linéaire par morceaux et organisé sur plusieurs modalités (Benzecri [1]-[5]). Nous notons que ce codage atténue l'effet stratégique en éliminant les particularités individuelles dans l'utilisation des variables. Entre autre, notre préoccupation est de respecter d'une part, la conservation de la distance (au sens de la métrique de chi-deux) entre individus dans l'espace des modalités, et d'autre part, la conservation de la structure ordonnée des modalités d'une variable ordonnée (Mezzoughi [17]).

C'est ainsi que le codage qui nous est proposé, a pour objectif le respect de la configuration réelle du questionnaire et l'acquisition d'une information représentative du sens des réponses des personnes interrogées.

2. INTRODUCTION

Dans cette partie, nous reprenons le codage suivant l'équation personnelle conservateur de l'ordre des modalités. Nous l'étudions dans un cas de questionnaire à variables qualitatives ordonnées, continues ou discrètes qui éclate chaque réponse d'un individu i (ou bien une note attribuée à un individu i) à la variable j , de code $x(i, j)$ sur quatre modalités : $j_{<}$, $j_{<<}$, $j_{>}$, $j_{>>}$. Nous voulons, à travers cette étude, définir les valeurs $k(i, j_{<})$; $k(i, j_{<<})$; $k(i, j_{>})$; $k(i, j_{>>})$ obtenues en fonction de l'ensemble des réponses de chaque sujet. Pour illustrer une telle situation, il suffit d'imaginer un exemple concernant des individus qui expriment leurs opinions sur le SIDA suivant une échelle d'approbation à quatre modalités. On propose la phrase :

le SIDA est une maladie mortelle

et l'on désire une réponse parmi les choix suivants :

"tout à fait d'accord" > "d'accord" > "pas d'accord" > "nullement d'accord".

3. CONCEPTION ANALYTIQUE DU NOUVEAU CODAGE ET METHODOLOGIE

3.1. **Cas du pivot central inactif.** La configuration décrite par le codage disjonctif complet est un tétraèdre régulier. Les points obtenues par codage flou (voir Gallego [16]) sont localisés sur la ligne polygonale $i_1i_2i_3i_4$. En "déployant" le tétraèdre dans un plan, nous obtenons la figure ci-après (FIG. 1).

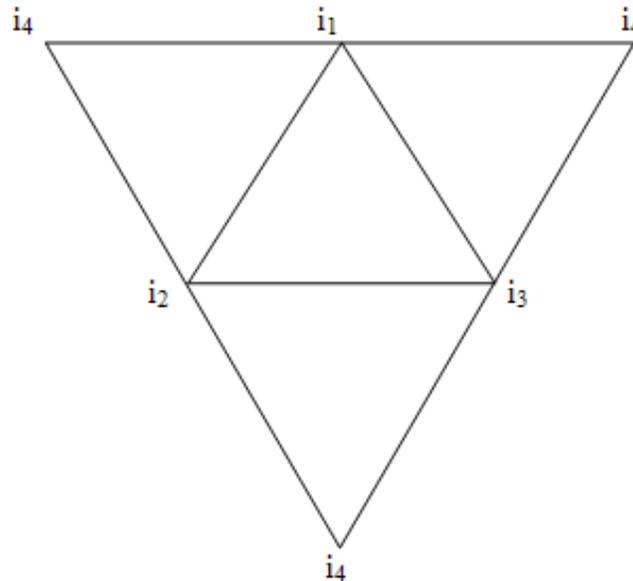


FIGURE 1. Les quatres faces du tétraèdre vues dans un même plan.

Nous pouvons constater que les points i_1 , i_2 , i_3 et i_4 sont tous équidistants les uns les autres; ce qui ne respecte ni l'ordre des modalités, ni leurs distances selon l'échelle d'approbation établie. Pour remédier à cela, nous nous intéressons aux parties intérieures des différentes faces du tétraèdre (c'est-à-dire les triangles $i_1i_2i_3$, $i_1i_2i_3i_4$, $i_1i_2i_3i_4, i_1i_2i_3i_4$).

Nous voulons passer de i_1 à i_4 en respectant à la fois la progression des distances et l'ordre des modalités. En cela, nous ferons un premier travail sur la face $i_1i_2i_3$.

L'observation du tétraèdre régulier ne se justifiant que dans un cas de modalités équipondérées, nous admettrons qu'il en sera ainsi pour déterminer la ligne polygonale des points supposés "valable" (FIG. 2)

Pour passer de i_1 à i_2 , nous évitons formellement l'arête $[i_1, i_2]$ pour ne pas être en permanence à égale distance des points i_3 et i_4 . D'une manière générale, nous éviterons systématiquement toutes les arêtes du tétraèdre. Le point i_5 est placé exprès sur la médiatrice du segment $[i_1, i_2]$, il est donc à mi-distance entre i_1 et i_2 . Si nous notons O le centre de gravité du triangle $i_1i_2i_3$, et A le milieu de $[i_1, i_2]$, nous pouvons voir que : $d(i_1, i_5) = d(i_2, i_5) < d(i_3, i_5)$. Cette inégalité se justifiant par le fait que O étant équidistant des sommets i_1, i_2 et i_3 , on observe que $d(O, i_3) < d(i_5, i_3)$.

De même, si nous regardons le quatrième sommet i_4 qui se projette orthogonalement au point O , nous pouvons remarquer que : $d(i_4, i_5) > d(i_5, i_3)$ car la distance entre i_4 et O est minimale dans le triangle $i_1i_2i_3$. Le tétraèdre étant régulier, cette distance minimale est égale à $d(i_3, A)$.

Partant de ces observation, nous pouvons noter que tout point i_k du segment $[i_1, i_5]$, nous avons $d(i_1, i_k) \leq d(i_1, i_5)$. En refaisant le même raisonnement à partir du point i_4 , nous pouvons placer le point i_8 sur la médiatrice du segment $[i_3, i_4]$. Nous pouvons aussi définir les points O' et B sur le triangle $i_1i_2i_3$. Par symétrie, nous pourrions noter que pour tout point i_k du segment $[i_4, i_8]$, nous avons $d(i_4, i_8) < d(i_4, i_8)$.

A l'issue de ce travail, il nous suffit de relier les points i_5 et i_8 par un segment n'appartenant ni au plan $i_1i_2i_3$ ni au plan $i_2i_3i_4$. Nous voyons finalement que pour tout point i_k de l'ensemble $[i_1, i_5] \cup [i_5, i_8] \cup [i_8, i_4]$, on a : $d(i_1, i_k) \leq d(i_1, i_4)$ et $d(i_4, i_k) \leq d(i_4, i_1)$. Ainsi, la ligne polygonale $i_1i_5i_8i_4$ est l'ensemble des points supposés "valables" pour le respect de la progression des distances et celui de l'ordre des modalités (cf FIG. 22). En conséquence, il nous faut définir les coordonnées des points i_5 et i_8 telles que les distances (au sens de la métrique de chi-deux) entre les points i_3 et i_5 puis entre les points i_2 et i_8 soient les plus faibles possibles en admettant comme perpendiculaire les vecteurs i_1i_2 et i_3i_5 d'une part, les vecteurs i_2i_8 et i_3i_4 d'autre part.

Pour que les distances réelles entre individus soient conservées, dans le nuage des points issus du codage barycentrique linéaire par morceaux à pivots agissant par morceaux, sans être influencées par les masses des modalités, nous supposerons que les masses $k_{j<}, k_{j<<}, k_{j>}, k_{j>>}$ sont quelconques et fixées à l'avance aux modalités $j_{<}, j_{<<}, j_{>}, j_{>>}$.

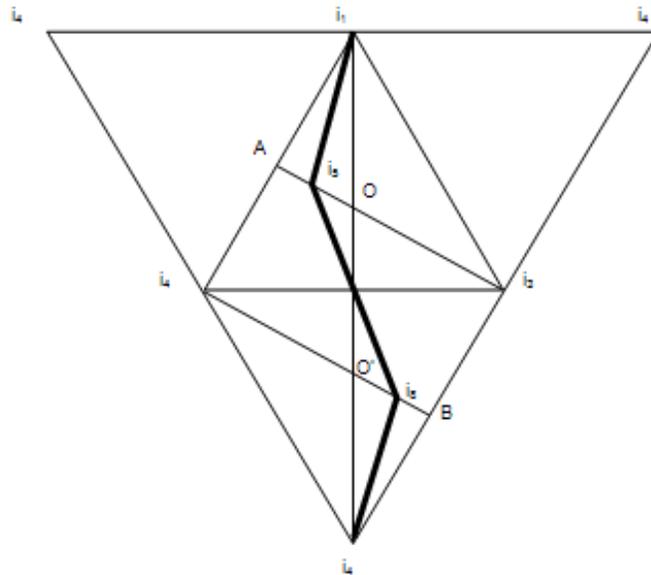


FIGURE 2. Les sommets du tétraèdre décrivent les modalités pures $j_{<<}, j_{<}, j_{>}, j_{>>}$ sous codage disjonctif complet. Les points issus du codage barycentrique linéaire par morceaux à pivots agissant par morceaux sont sur les segments en tracé épais. La représentation de la perpendicularité doit être vue au sens de la métrique du chi-deux.

Soient $i_1(1, 0, 0, 0)$; $i_2(0, 1, 0, 0)$; $i_3(0, 0, 1, 0)$; $i_4(0, 0, 0, 1)$; $i_5(1 - 2a, a, a, 0)$; $i_8(0, b, b, 1 - 2b)$

Nous reconnaitrons que, s'agissant d'un codage flou, les coordonnées sont toutes positives et astreintes à une somme égale à 1. Les vecteurs précités ont pour coordonnées :

$i_1i_2(-1, 1, 0, 0)$; $i_2i_8(0, b - 1, b, 1 - 2b)$; $i_3i_4(0, 0, -1, 1)$; $i_3i_5(1 - 2a, a, a - 1, 0)$

D'après les conditions évoquées, on obtient, (au sens de la métrique du khi-deux) :

$$\begin{aligned} & \langle i_3i_5, i_1i_2 \rangle = 0 \\ \Rightarrow & -a/k_{j<} + (1 - 2a)/k_{j<<} = 0 \\ \Rightarrow & a(1/k_{j<} + 2/k_{j<<}) = 1/k_{j<} \\ \Rightarrow & a = k_{j<} / (k_{j<<} + 2/k_{j<}) \end{aligned}$$

et

$$\begin{aligned} & \langle i_2i_8, i_3i_4 \rangle = 0 \\ \Rightarrow & -b/k_{j>} + (1 - 2b)/k_{j>>} = 0 \end{aligned}$$

$$\begin{aligned} \Rightarrow b(1/k_{j>} + 2/k_{j>>}) &= 1/k_{j>} \\ \Rightarrow b &= k_{j<}/(k_{j>>} + 2k_{j>}) \end{aligned}$$

Les valeurs de a et b sont définies en fonction des masses $k_{j<}$, $k_{j<<}$, $k_{j>}$, $k_{j>>}$ qui sont elles-mêmes calculées à partir du tableau recadré issu du codage linéaire par morceaux suivant l'équation personnelle avec simple ou double recadrage. Dans le cas du double recadrage, il suffit de remplacer les modalités $j_{<}$, $j_{<<}$, $j_{>}$, $j_{>>}$ respectivement par les modalités $\varphi_{<}$, $\varphi_{<<}$, $\varphi_{>}$, $\varphi_{>>}$ qui sont relatives au tableau-recadré en lignes et en colonnes.

Par ailleurs, les coordonnées des points i_5 et i_8 peuvent être retrouvées en considérant l'ordre des modalités des réponses. En cela, il nous faut rappeler que s'agissant d'un codage linéaire par morceaux, la contribution absolue d'une question q est donnée par la formule :

$$ctr(q) = 1/CardQ \left(\sum \left\{ \sum \{x(i, j)/k_j, i \in l\} j \in Jq \right\} - 1 \right)$$

Pour chaque individu i vu à l'intérieur du tétraèdre sous codage suivant l'équation personnelle en quatre modalités, on a au moins trois modalités, dont les valeurs sont non nulle. Ainsi, dans la somme qui donne l'inertie, trouve-t-on quatre termes dont deux consécutifs sont impérativement non nuls.

$$x^2(i, j)/k_j + (1 - 2x(i, j))^2/k_{j+1} + x^2(i, j)/k_{j^*} = \xi(x(i, j))$$

(la notation j^* étant attribuée à une modalité n'étant pas forcément consécutive aux deux autres).

Ayant choisi les individus qui nous intéressent sur la médiatrice du segment reliant les modalités consécutives, il nous est facile de noter que la fonction $(k(i, j))$ est minimale pour

$$\begin{aligned} x(i, j)/k_j + (1 - 2x(i, j))/k_{j+1} &= 0 \\ \Rightarrow x(i, j) &= k_j/(k_{j+1} + 2k_j) \end{aligned}$$

En remplaçant j par $j + 1$ par les modalités $j_{<<}$ et $j_{<}$ par les modalités $j_{>}$, $j_{>>}$, on retrouve simultanément les valeurs a et b .

On peut donc choisir cinq pivots recadrés -1 , $-a$, 0 , b , 1 (voir fig.3). Les codes recadrés $x(i, j)$ varient du pivot -1 au pivot 1 (pivots extrêmes) en passant par les pivots $-a$, 0 , et b (pivots intermédiaires) tels que :

- -1 : pivot d'extrême gauche agissant si $x(i, j) \in [-1, -a] \cup [b, 1]$
- 0 : pivot central n'agissant pas si $x(i, j) \in [-1, -a] \cup [b, 1]$
- 1 : pivot d'extrême droite agissant si $x(i, j) \in [-1, -a] \cup [b, 1]$

-a : pivot de gauche agissant si $x(i, j) \in [-a, b]$
 0 : pivot central n'agissant pas si $x(i, j) \in [-a, b]$
 b : pivot de droite agissant si $x(i, j) \in [-a, b]$
 Si $x(i, j)$ est une note recadrée entre -1 et 1, elle peut alors être recodée sur les quatres modalités comme suit :
 Si $-1 \leq x(i, j) \leq -a$, on a :

$$\begin{cases} k(i, j_{<<}) &= -x(i, j) \\ k(i, j_{<}) &= [1 + x(i, j)](1 - 2a)/(1 - a) \\ k(i, j_{>}) &= a[1 + x(i, j)]/(1 - a) \\ k(i, j_{>>}) &= 0 \end{cases}$$

Si $-a \leq x(i, j) \leq 0$, on a :

$$\begin{cases} k(i, j_{<<}) &= a - [a + x(i, j)]a/(a + b) \\ k(i, j_{<}) &= 1 - 2a - [a + x(i, j)](1 - 2a - b)/(a + b) \\ k(i, j_{>}) &= a + (1 - 2b - a)[a + x(i, j)]/(a + b) \\ k(i, j_{>>}) &= b[a + x(i, j)]/(a + b) \end{cases}$$

Si $0 \leq x(i, j) \leq b$, on a :

$$\begin{cases} k(i, j_{<<}) &= a - [a + x(i, j)]a/(a + b) \\ k(i, j_{<}) &= 1 - 2a - [a + x(i, j)](1 - 2a - b)/(a + b) \\ k(i, j_{>}) &= a + (1 - 2b - a)[a + x(i, j)]/(a + b) \\ k(i, j_{>>}) &= b[a + x(i, j)]/(a + b) \end{cases}$$

Si $B \leq x(i, j) \leq 1$, on a : Pour retrouver le cas de double recadrage, il suffit de remplacer les modalités $j_{<<}$, $j_{<}$, $j_{>}$, $j_{>>}$ respectivement par les modalités $\varphi_{<<}$, $\varphi_{<}$, $\varphi_{>}$, $\varphi_{>>}$.

Nous pouvons noter également que ce codage laisse le pivot central inactif; ce qui explique le fait que $x(i, j)$ soit éclatée de la même manière sur les intervalles $[-a, 0]$ et $[0, b]$.

Représentations graphiques du codage

Pour construire les différentes fonctions d'appartenances aux quatres modalités, nous devons tenir compte du caractère arbitraire des masses. Chacune des configurations observées est donc fonction des masses choisies.

Soit f_1 , la fonction d'appartenance à la première modalité $j_{<<}$, f_2 , la fonction d'appartenance à la deuxième modalité $j_{<}$, f_3 et f_4 respectivement celles d'appartenance aux modalités $j_{>}$ et $j_{>>}$.

Si nous supposons que les valeurs $x(i, j)$ recadrées sur $[-1, 1]$ indiquent les abscisses sur l'axe horizontal et si nous admettons que les valeurs

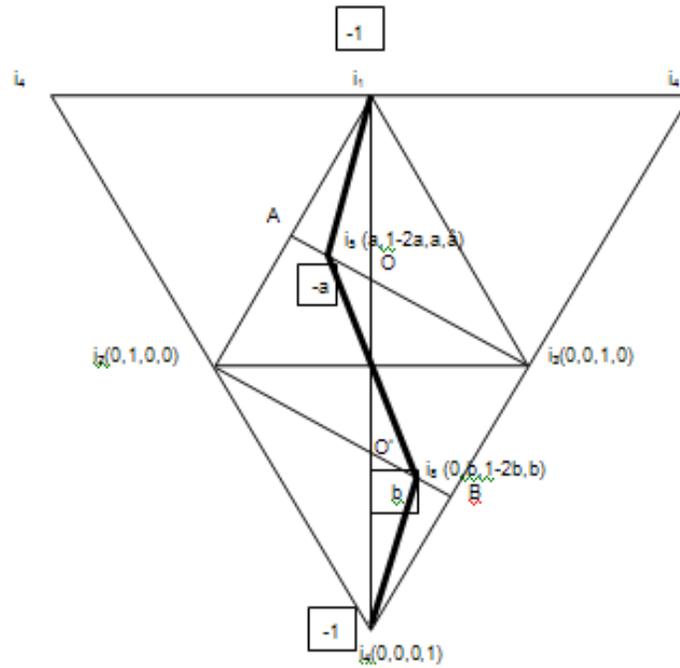


FIGURE 3. Observation en tracé épais du nuage résultant du codage barycentrique linéaire par morceaux à pivots agissant par morceaux dans un cas modèle de codage en quatre modalités équipondérées. Les valeurs pivots sont celles qui sont encadrées.

définies $[0,1]$ après recodage indiquent les ordonnées, nous pouvons représenter ce codage selon les configurations décrites par les figures 4 et 5 qui suivent :

3.2. Exemple explicatif du codage. Si nous nous intéressons au cas particulier où toutes les modalités sont équipondérées, nous aurons : $a = b = 1/3$.

Les coordonnées des principaux points retenus seront :
 $i_1(1, 0, 0, 0)$; $i_2(0, 1, 0, 0)$; $i_3(0, 0, 1, 0)$; $i_4(0, 0, 0, 1)$; $i_5(1/3, 1/3, 1/3, 0)$;
 $i_6(0, 1/3, 1/3, 1/3)$

Si nous définissons les points i_{10} et i'_5 et i'_8 comme milieux respectifs des segments $[i_5, i_6]$, $[i_1, i_5]$, $[i_8, i_4]$ nous pourrons leur accorder les coordonnées suivantes : $i_{10}(1/6, 1/3, 1/3, 1/6)$; $i'_5(2/3, 1/6, 1/6, 0)$; $i'_8(0, 1/6, 1/6, 2/3)$. (voir FIG. 6)

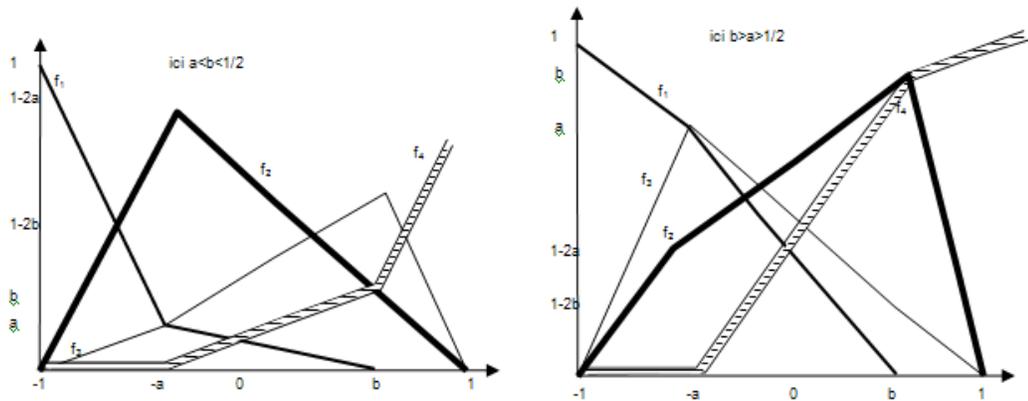


FIGURE 4. Représentations des fonctions d'apparence aux quatre modalités selon le cogage barycentrique linéaire par morceaux à pivots agissant par morceaux. Chacune des modalités intermédiaires aux pivots $-a$ et b doit être interprétée comme étant partagée sur les trois pivots $-a$, 0 et b . (Remarquons que l'échelle des abscisses est différente de celle des ordonnées).

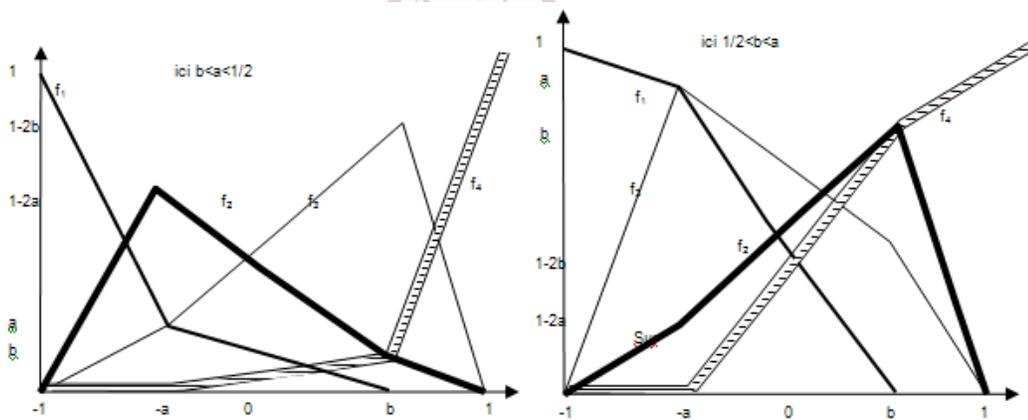


FIGURE 5. Représentations des fonctions d'apparence aux quatre modalités selon le cogage barycentrique linéaire par morceaux à pivots agissant par morceaux. Chacune des modalités intermédiaires aux pivots $-a$ et b doit être interprétée comme étant partagée sur les trois pivots $-a$, 0 et b . (Remarquons que l'échelle des abscisses est différentes de celle des ordonnées).

Si nous supposons que les codes $x(i, j)$ sont assimilés à des abscisses sur l'axe horizontal, nous pouvons représenter ce codage sur la figure 7 où les pivots agissent par morceaux. Les codes recadrés $x(i, j)$ sont alors éclatés de manière suivante :

Si $-1 \leq x(i, j) \leq -1/2$, on a :

$$\begin{cases} k(i, j_{<<}) &= -x(i, j) \\ k(i, j_{<}) &= 1/2[1 + x(i, j)] \\ k(i, j_{>}) &= 1/2[1 + x(i, j)] \\ k(i, j_{>>}) &= 0 \end{cases}$$

Si $-1/2 \leq x(i, j) \leq 0$, on a :

$$\begin{cases} k(i, j_{<<}) &= 1/3 - 1/2[1/3 + x(i, j)] \\ k(i, j_{<}) &= 1/3 \\ k(i, j_{>}) &= 1/3 \\ k(i, j_{>>}) &= 1/2[1/3 + x(i, j)] \end{cases}$$

Si $0 \leq x(i, j) \leq 1/2$, on a :

$$\begin{cases} k(i, j_{<<}) &= 1/3 - 1/2[1/3 + x(i, j)] \\ k(i, j_{<}) &= 1/3 \\ k(i, j_{>}) &= 1/3 \\ k(i, j_{>>}) &= 1/2[1/3 + x(i, j)] \end{cases}$$

Si $1/2 \leq x(i, j) \leq 1$, on a :

$$\begin{cases} k(i, j_{<<}) &= 0 \\ k(i, j_{<}) &= 1/3 - 1/2[x(i, j) - 1/3] \\ k(i, j_{>}) &= 1/3 - 1/2[x(i, j) - 1/3] \\ k(i, j_{>>}) &= x(i, j) \end{cases}$$

Le mérite de ce codage éclaté en quatre modalités est de nous montrer qu'en dehors des individus codés en disjonctif complet (situés sur les sommets du tétraèdre) qui ont un jugement exclusif et formel, les individus repérés sur le nuage sont influencés au moins par trois modalités de réponse.

L'individu i'_5 pense que le SIDA est une maladie mortelle, il est convaincu de cette affirmation dont il refoule la négation. En fait, l'individu se refuse de croire que la maladie ne soit pas mortelle bien qu'il y ait un doute (minime certes) dans sa réponse.

L'individu i_5 a une attitude similaire à celle de i'_5 tout en étant équitablement partagé entre les trois modalités "tout à fait d'accord", "d'accord" et "pas d'accord". C'est un individu qui pense que le SIDA est une maladie mortelle mais qui demeure sceptique quant à la confirmation de cette affirmation.

L'individu i_{10} a une attitude mitigée entre l'accord et le désaccord.

Il évite copieusement les positions extrêmes. C'est une personne qui semble indécise quant à l'attitude à adopter.

L'individu i_8 ne pense pas que le SIDA est une maladie mortelle, mais il manque d'assurance sur sa conception. Ainsi, s'inscrit en lui un "accord" voilé mais suffisamment perceptible.

L'individu i_8 ne pense pas que le SIDA n'est pas une maladie mortelle, mais le caractère mortel le trouble quand même. Ce qui explique une présence de doute qu'il n'arrive pas à exclure entièrement.

Les représentations graphiques des fonctions d'appartenance aux quatre modalités décrites par ce cas particulier nous montre que les fonctions f_2 et f_4 (relatives aux modalités intermédiaires) sont identiques (fig 7). Ce qui nous conduit à admettre que le codage en trois modalités suffirait largement pour expliquer cette situation.

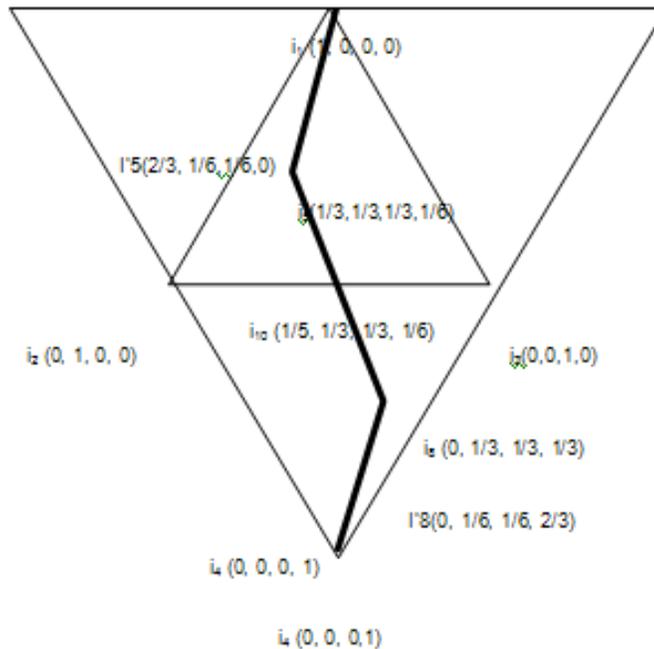


FIGURE 6. Observation en tracé épais du nuage résultant du codage barycentrique linéaire par morceaux à pivots agissant par morceaux avec O actif dans un cas modèle de codage en quatre modalités

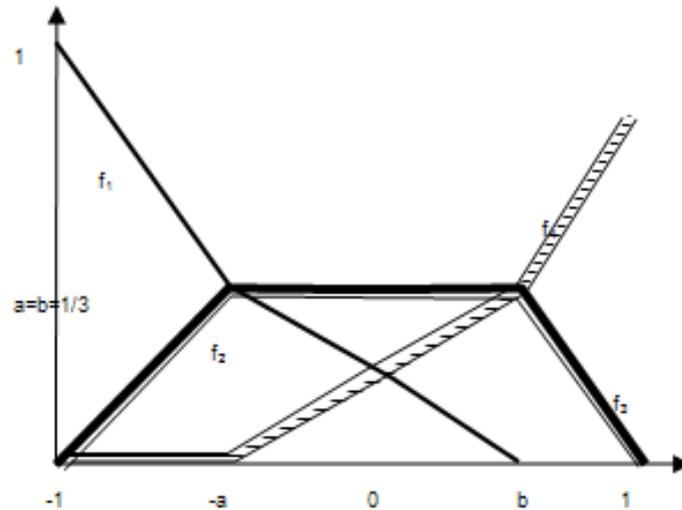


FIGURE 7. Représentations des fonctions d'apparence aux quatre modalités

3.2.1. **Extention au pivot central actif.** En dépit du caractère fort intéressant du codage établi dans le chapitre précédent, nous constatons que l'inactivité du pivot central zéro, nous empêche d'entrevoir certaines nuances entre les différentes fonctions d'appartenance. Ainsi, après avoir admis (comme dans le cas précédent) la ligne polygonale $i_1 i_5 i_8 i_4$ comme étant celui des points supposés *valable*, nous ajoutons le point i_{10} situé à l'intérieur du tétraèdre à égale distance des points i_2 et i_3 tout en étant à distance minimale de i_4 (voir fig 8). Par conséquent : $i_{10}(e, 1 - 3e, e, e)$ et $i_4 i_{10}(e, 1 - 3e, e, e - 1)$.

$$\begin{aligned} \langle i_4 i_{10}, i_2 i_3 \rangle &= 0 \\ \Rightarrow -(1 - 3e)/k_{j<} + e/k_{j>} &= 0 \\ \Rightarrow e(3/k_{j<} + 1/k_{j>}) &= 1/k_{j<} \\ \Rightarrow e &= k_{j>} / (1/k_{j<} + 3/k_{j>}) \end{aligned}$$

Cette valeur de e peut encore être facilement retrouvée à l'aide de l'expression de la contribution absolue d'une question dans le cas de codage linéaire par morceaux. Par ailleurs, si nous voulons réécrire cette valeur dans le cas de double recadrage, il nous suffit de remplacer les modalités $j_{<}$ et $j_{>}$ par les modalités $j_{>}$ et $j_{<}$.

De ce fait, nous pouvons définir les pivots intermédiaires a, e, b . Après recadrage, nous retrouvons cinq pivots centrés qui, cette fois-ci,

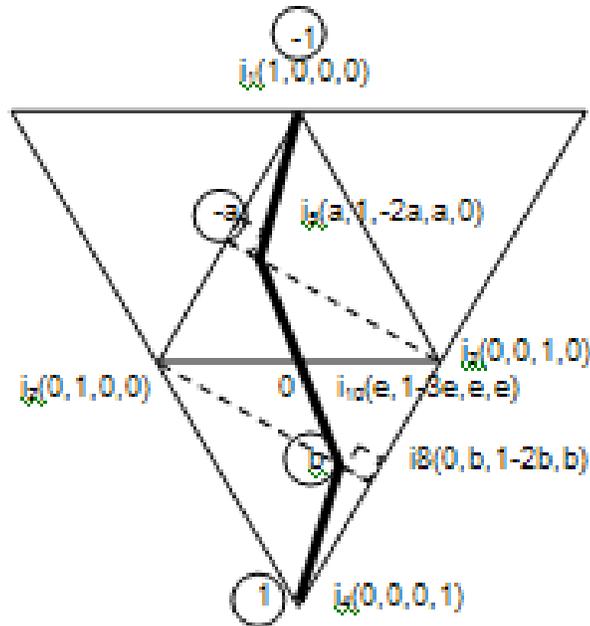


FIGURE 8. Observation en tracé épais du nuage résultant du codage barycentrique linéaire par morceaux à pivot agissant par morceaux avec O actif dans un cas modèle de codage en quatre modalités équipondérées. Les valeurs pivots sont celles qui sont encadrées

agissant par morceaux.

Les codes recadrés $x(i, j)$ varient du pivot -1 (pivots extrêmes) en passant par les pivots $-a$, 0 et b (pivots intermédiaires) tels que :

- -1 : pivot d'extrême gauche agissant si $x(i, j) \in [-1, -a] \cup [b, 1]$
- 0 : pivot central agissant si $x(i, j) \in [-1, -a] \cup [b, 1]$
- 1 : pivot d'extrême droite agissant si $x(i, j) \in [-1, -a] \cup [b, 1]$
- $-a$: pivot gauche agissant si $x(i, j) \in [-a, b]$
- 0 : pivot central agissant si $x(i, j) \in [-a, b]$
- b : pivot central agissant si $x(i, j) \in [-a, b]$

Etant donnée une note $x(i, j)$ comprise entre -1 et 1 , elle peut être recodée sur quatre modalités de la manière suivante :

Si $-1 \leq x(i, j) \leq -a$, on a :

$$\begin{cases} k(i, j_{<<}) &= x(i, j) \\ k(i, j_{<}) &= [1 + x(i, j)](1 - 2a)/(1 - a) \\ k(i, j_{>}) &= [1 + x(i, j)]/(1 - a) \\ k(i, j_{>>}) &= 0 \end{cases}$$

Si $-a \leq x(i, j) \leq 0$, on a :

$$\begin{cases} k(i, j_{<<}) &= a + [a + x(i, j)](e - a)/a \\ k(i, j_{<}) &= 1 - 2a + [a + x(i, j)](2a - 3e)/a \\ k(i, j_{>}) &= a + [a + x(i, j)](e - a)/a \\ k(i, j_{>>}) &= e[a + x(i, j)]/a \end{cases}$$

Si $0 \leq x(i, j) \leq b$, on a :

$$\begin{cases} k(i, j_{<<}) &= e - x(i, j)e/b \\ k(i, j_{<}) &= 1 - 3e - x(i, j)(1 - 3e - b)/b \\ k(i, j_{>}) &= e + x(i, j)(1 - 2b - e)/b \\ k(i, j_{>>}) &= e + x(i, j)(b - e)/b \end{cases}$$

Si $b \leq x(i, j) \leq 1$, on a :

$$\begin{cases} k(i, j_{<<}) &= 0 \\ k(i, j_{<}) &= b - [x(i, j) - b]b/(1 - b) \\ k(i, j_{>}) &= 1 - 2b - (1 - 2b)[x(i, j) - b]/(1 - b) \\ k(i, j_{>>}) &= x(i, j) \end{cases}$$

Pour retrouver le cas de double recadrage, il suffit de remplacer les modalités $j_{<}, j_{<<}, j_{>}, j_{>>}$ respectivement par les modalités $\varphi_{<}, \varphi_{<<}, \varphi_{>}, \varphi_{>>}$

Nous pouvons remarquer que cette manière d'éclater les codes recadrés $x(i, j)$ nous évite d'obtenir la même formule si :

$-a \leq x(i, j) \leq 0$ ou bien $0 \leq x(i, j) \leq b$ ceci nous révèle qu'il s'agit d'un codage qui est plus fin.

3.3. Représentations graphiques du codage. Si nous définissons f_1, f_2, f_3, f_4 comme étant respectivement les fonctions d'appartenance aux modalités $j_{<}, j_{<<}, j_{>}, j_{>>}$, nous pouvons alors les représenter graphiquement selon les masses $k_{j_{<}}, k_{j_{<<}}, k_{j_{>}}, k_{j_{>>}}$ choisies (voir figures 9 et 10).

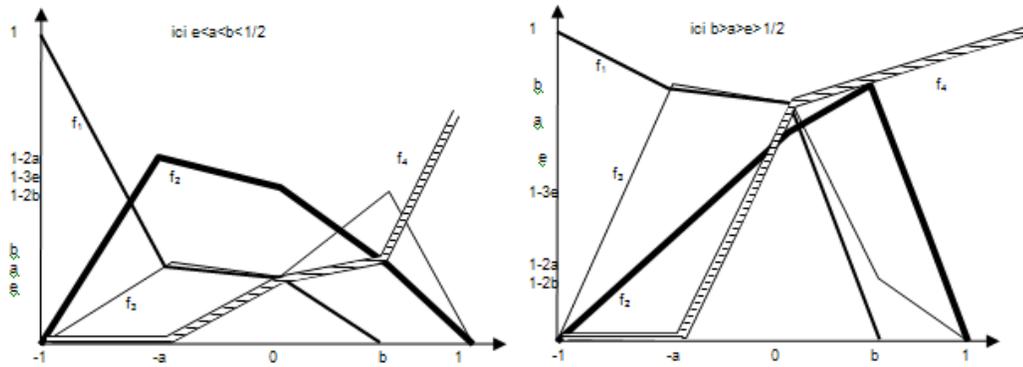


FIGURE 9. Représentations des fonctions d'appartenances aux quatre modalités selon le codage barycentrique linéaire par morceaux à pivots agissant par morceaux avec O actif. Chacune des modalités intermédiaires aux pivots $-a$ et b doit être interprétée comme étant partagée sur les trois pivots $-a$, 0 , et b . (Remarquons que l'échelle des abscisse est différentes de celle des ordonnées)

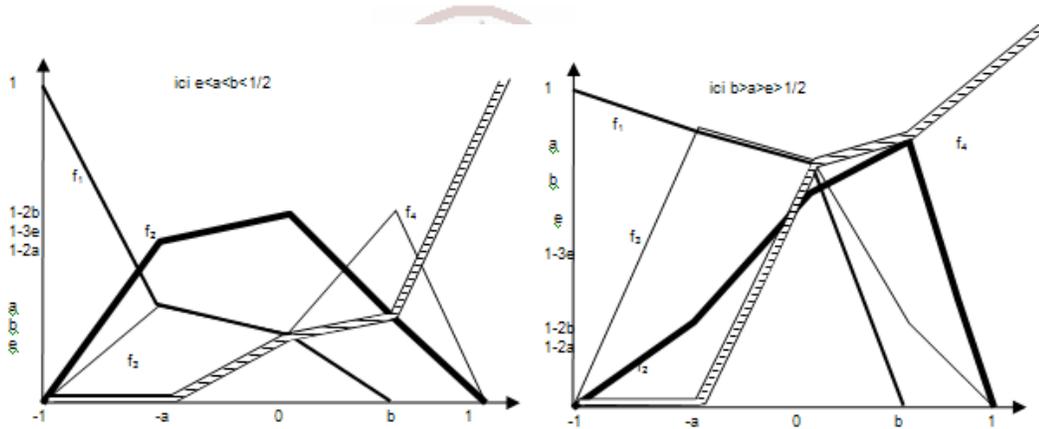


FIGURE 10. Représentations des fonctions d'appartenances aux quatre modalités selon le codage barycentrique linéaire par morceaux à pivots agissant par morceaux avec O actif. Chacune des modalités intermédiaires aux pivots $-a$ et b doit être interprétée comme étant partagée sur les trois pivots $-a$, 0 , et b . (Remarquons que l'échelle des abscisse est différentes de celle des ordonnées)

3.4. **Exemple explicatif du codage.** Si nous regardons le cas particulier où toutes les masses sont identiques, nous aurons : $a = b = 1/3$ et $e = 1/4$.

Toute note recadrée $x(i, j)$ peut être recodée comme suit :

Si $-1 \leq x(i, j) \leq -1/3$, on a :

$$\begin{cases} k(i, j_{<<}) &= -x(i, j) \\ k(i, j_{<}) &= 1/2[1 + x(i, j)] \\ k(i, j_{>}) &= 1/2[1 + x(i, j)] \\ k(i, j_{>>}) &= 0 \end{cases}$$

Si $-1/3 \leq x(i, j) \leq 0$, on a :

$$\begin{cases} k(i, j_{<<}) &= 1/4[1 - x(i, j)] \\ k(i, j_{<}) &= 1/4[1 - x(i, j)] \\ k(i, j_{>}) &= 1/4[1 - x(i, j)] \\ k(i, j_{>>}) &= 1/4[1 + 3x(i, j)] \end{cases}$$

Si $0 \leq x(i, j) \leq 1/3$, on a :

$$\begin{cases} k(i, j_{<<}) &= 1/4[1 - 3x(i, j)] \\ k(i, j_{<}) &= 1/4[1 + x(i, j)] \\ k(i, j_{>}) &= 1/4[1 + x(i, j)] \\ k(i, j_{>>}) &= 1/4[1 + x(i, j)] \end{cases}$$

Si $1/3 \leq x(i, j) \leq 1$, on a :

$$\begin{cases} k(i, j_{<<}) &= 0 \\ k(i, j_{<}) &= 1/3 - 1/2x(i, j) \\ k(i, j_{>}) &= 1/3 - 1/2x(i, j) \\ k(i, j_{>>}) &= x(i, j) \end{cases}$$

Nous pouvons, à travers ce codage, remarquer que les individus i_5 et i_8 (fig. 11) ont des attitudes diamétralement opposées ; ce qui se justifie par leurs codages respectifs. Entre autre, la richesse d'information apportée par ce codage en quatre modalités s'explique par l'observation permanente de deux ou trois modalités qui se complètent pour dégager la tendance prédominante qui souvent, est nuancée par une singularité. Cette observation facilite énormément l'interprétation.

L'individu i_{10} est très sceptique et il donne un avis uniformément partagé par les quatre modalités de réponse. En réalité, il ne se prononce pas sur la question ; c'est plutôt une personne sur qui subsiste un doute ; peut être même un espoir du contraire.

L'individu i_{10} est convaincu que le SIDA n'est pas une maladie mortelle malgré sa conviction, le doute est toujours présent.

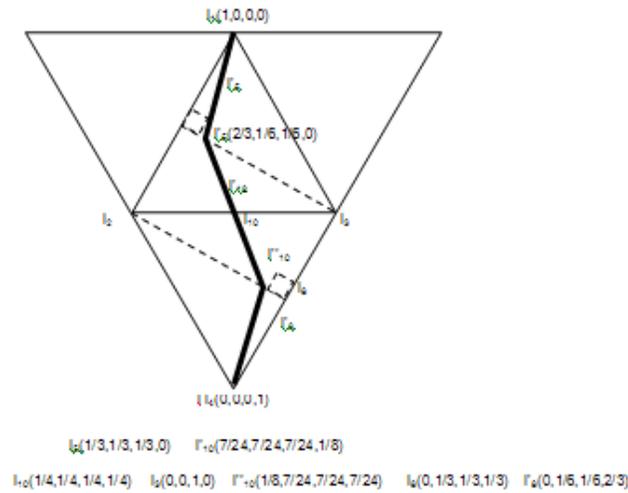


FIGURE 11. Observation en tracé épais du nuage résultant du codage barycentrique linéaire par morceaux à pivot agissant par morceaux avec actif O actif dans le cas particulier où les masses sont identiques

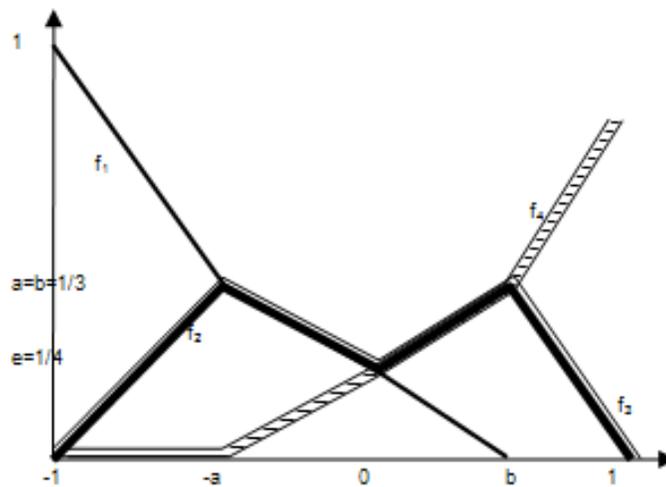


FIGURE 12. Représentations des fonctions d'appartenances aux quatre modalités selon le codage barycentrique linéaire par morceaux à pivots agissant par morceaux avec actif O actif dans le cas particulier où les masses sont identiques

Les représentations graphiques des fonctions d'appartenance aux quatre modalités décrites par ce cas particulier nous montrent distinctement l'action du pivot central "zero" (FIG. 10).

4. UN CONCEPT DE GÉNÉRALISATION DE CES CODAGES.

Le point i_5 décrit le segment $]A, O]$. Le chapitre 3.1.1 nous permet de définir les points $i_5(t)$ qui découlent des différentes positions du point i_5 dont les coordonnées pourraient être aisément obtenues selon la relation :

$$x_{i_5(t)} = (1 - t)x_5 + tx_0, \text{ avec } t \in [0, 1]$$

De même, nous pouvons repérer, selon la symétrie centrale de centre G (centre de gravité du tétraèdre) dans le plan AOG les individus $i_8(t)$ qui décrivent le segment $]B, O']$. On obtient finalement une infinité de lignes polygonales $C_t = i_1 i_5(t) i_8(t) i_3$ de points *valables*. Chacune de ces lignes polygonales C_t nous permet d'établir un codage barycentrique linéaire par morceaux obtenus. Ceci fera l'objet d'une prochaine publication.



REFERENCES

- [1] BENZECRI, J.P.(1982). Construction d'une classification ascendante hiérarchique par la recherche en chaînes de voisins réciproques *CAD vol VII, n°2*, pp. 209-218.
- [2] BENZECRI, J.P.(1972). *Sur l'analyse des tableaux binaires associés à une correspondance multiple*, (Note multigraphiée), *LS*.
- [3] BENZECRI, J.P. (1973). *Analyse des Données : l'Analyse des Correspondances (Tome 2, Dunod)*
- [4] BENZECRI, J.P. et BENZEKRI, F.(1989). Codage linéaire par morceaux et équation personnelle. (*CAD vol. XIV, n°3*, pp. 331-336.
- [5] BENZECRI, J.P. et BENZECRI, F.(1989). Le codage linéaire par morceaux : réalisation et applications. *CAD vol. XIV n°2*, pp 203-210.
- [6] BENZECRI, J.P. et , CAZES, P.(1978). Problème sur la classification. *CAD vol. III n°1*.
- [7] BERTIER, P. et CAZES P.(1975). Analyse des données multidimensionnelles. P.U.F., Paris.
- [8] CAILLIEZ F. et PAGES J.P.(1976). *Introduction à l'analyse des données*. Smash, Paris. 1976)
- [9] CARTAN H.(1957). *Quelques questions de topologie*. Séminaire Henry Cartan, 9.
- [10] CAZES P.(1977). Etude dans un questionnaire d'un modèle de non réponse. *CAD vol.I, n°2* pp161-172.
- [11] CAZES P.(1980). Analyse de certain tableau rectangulaires décomposés en blocs. *CAD vol. V n°2*.
- [12] CAZES P.(1980). Problème sur l'analyse des questionnaires. *CAD vol. II, n°1*, pp 73-77.
- [13] CAZES P.(1990). Codage d'une variable en vue de l'analyse des correspondances. *RSA 1990, XXXVIII(3)*.
- [14] ESCOFIER-CORDIER B.(1969). Analyse des correspondances. *Cahier de buro, n°13*.
- [15] ESCOFIER B. et PAGES, J.(1988). *Analyse factorielle simple et multiple*. Dunod, Paris.
- [16] GALLEGO F.(1982). Codage flou en analyse des correspondances. *CAD vol VII, n°2* pp 413-430.
- [17] M.MEZZOUGH (1990). *Le codage suivant l'équation personnelle conservateur de l'ordre des modalités*. Thèse de Doctorat, Université Paris VI.

UNIVERSITÉ : MARIEN NGOUABI BRAZZAVILLE, DÉPARTEMENT DE SCIENCES EXACTES
E-mail address: rufbid@yahoo.fr