



# Etude empirique sur la notion de région homogène

Bernard Fustier

## ► To cite this version:

Bernard Fustier. Etude empirique sur la notion de région homogène. [Rapport de recherche] Institut de mathématiques économiques (IME). 1978, 30 p., figures, bibliographie. hal-01527166

**HAL Id: hal-01527166**

**<https://hal.science/hal-01527166>**

Submitted on 24 May 2017

**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.

RP 210 (2)  
**I.M.E.**

EQUIPE DE RECHERCHE ASSOCIEE AU C.N.R.S.

DOCUMENT DE TRAVAIL

N° 25

Etude empirique sur la notion de région homogène.

Bernard FUSTIER

Avril 1978



**INSTITUT DE MATHEMATIQUES ECONOMIQUES**



**UNIVERSITE DE DIJON**

**FACULTE DE SCIENCE ECONOMIQUE ET DE GESTION**

**4, BOULEVARD GABRIEL — 21000 DIJON**

18 DEC. 1978

N° 25

Etude empirique sur la notion de région homogène.

Bernard FUSTIER

Avril 1978



Le but de cette Collection est de diffuser rapidement une première version de travaux afin de provoquer des discussions scientifiques. Les lecteurs désirant entrer en rapport avec un auteur sont priés d'écrire à l'adresse suivante :

INSTITUT DE MATHÉMATIQUES ÉCONOMIQUES

4 Boulevard Gabriel - 21000 - DIJON - France.

## TRAVAUX DEJA PUBLIES

- N°1 Michel PREVOT : Théorème du point fixe. Une étude topologique générale  
(juin 1974)
- N°2 Daniel LEBLANC : L'introduction des consommations intermédiaires dans  
le modèle de LEFEBER (juin 1974)
- N°3 Colette BOUNON : Spatial Equilibrium of the Sector in Quasi-Perfect  
Competition (september 1974)
- N°4 Claude PONSARD : L'imprécision et son traitement en analyse économique  
(septembre 1974)
- N°5 Claude PONSARD : Economie urbaine et espaces métriques (septembre 1974)
- N°6 Michel PREVOT : Convexité (mars 1975)
- N°7 Claude PONSARD : Contribution à une théorie des espaces économiques  
imprécis (avril 1975)
- N°8 Aimé VOGT : Analyse factorielle en composantes principales d'un caractère  
de dimension-n (juin 1975)
- N°9 Jacques THISSE et Jacky PERREUR : Relation between the Point of Maximum  
Profit and the Point of Minimum Total Transportation  
Cost : A Restatement (juillet 1975)
- N°10 Bernard FUSTIER : L'attraction des points de vente dans des espaces  
précis et imprécis (juillet 1975)
- N°11 Régis DELOCHE : Théorie des sous-ensembles flous et classification en  
analyse économique spatiale (juillet 1975)
- N°12 Gérard LASSIBILLE et Catherine PARRON : Analyse multicritère dans un  
contexte imprécis (juillet 1975)
- N°13 Claude PONSARD : On the Axiomatization of Fuzzy Subsets Theory (july 1975)
- N°14 Michel PREVOT : Probability Calculation and Fuzzy Subsets Theory  
(august 1975)

- N° 15 Claude PONSARD : Hiérarchie des places centrales et graphes - flous  
(avril 1976)
- N° 16 Jean-Pierre AURAY et Gérard DURU : Introduction à la théorie des espaces multiflous (avril 1976)
- N° 17 Roland LANTNER, Bernard PETITJEAN et Marie-Claude PICHERY : Jeu de simulation du circuit économique (Août 1976)
- N° 18 Claude PONSARD : Esquisse de simulation d'une économie régionale : l'apport de la théorie des systèmes flous (septembre 1976)
- N° 19 Marie-Claude PICHERY : Les systèmes complets de fonctions de demande (avril 1977)
- N° 20 Gérard LASSIBILLE et Alain MINGAT : L'estimation de modèles à variables dépendante dichotomique - La sélection universitaire et la réussite en première année d'économie (avril 1977)
- N° 21 Claude PONSARD : La région en analyse spatiale (mai 1977)
- N° 22 Dan RALESCU : Abstract Models for Systems Identification (juin 1977)
- N° 23 J. MARCHAL et F. POULON : Multiplicateur, graphes et chaînes de Markov (décembre 1977).
- N° 24 Pietro BALESTRA : Determinant and Inverse of a sum of matrices with applications in economics and statistics (avril 1978)

## ETUDE EMPIRIQUE SUR LA NOTION DE REGION HOMOGENE

### Contribution à une typologie des régions de programme

#### INTRODUCTION

Consacré à une "étude empirique sur la notion de région homogène", le présent document s'érige en marge de toute conceptualisation théorique de la région pour ne considérer que ses aspects purement opératoires. Dans cette optique n'aurait-il pas été plutôt préférable de traiter de la région nodale, de la région polarisée, de la région fonctionnelle ?

Les différents qualificatifs utilisés pour caractériser ce dernier type de région sont légitimes ; ils aident à révéler un contenu riche et font comprendre que les économistes se détournent progressivement de la région homogène (J.L. GUIGOU [10] ; M. MOUGEOT [14]).

Pourtant, région homogène et région nodale ne sont pas des notions tellement éloignées. Dans certains cas, elles peuvent naître d'un même instrument d'analyse (J.C. CHEVAILLER [2]) et, d'un certain point de vue, elles demeurent complémentaires "si l'on veut bien considérer que la première établit des relations d'équivalence sur l'espace considéré et que la seconde s'attache à construire des relations d'ordre sur des classes d'équivalence" (C. PONSARD [15] p. 12).

Cette dernière remarque nous amène à une critique fondamentale de la région homogène dont la mise en évidence, conduit généralement à une partition de l'espace national qui servira ensuite de cadre de référence à l'étude de la région nodale. Or, la réalité économique, en général et les différents sous-ensembles issus d'une classification régionale, en particulier, ne se laissent pas séparer aussi facilement en classes d'équivalence. Mais il faut souligner que cette critique malgré toute l'importance qu'il convient de lui accorder, ne possède pas un caractère inéluctable dans la mesure où un aspect particulier (mais digne d'intérêt) de la région homogène n'est pas, par nature, réductible à des classes d'équivalence. En effet, la notion de région

homogène recouvre deux autres notions qu'il convient de distinguer en fonction du nombre de critères retenus pour réaliser un découpage de l'espace de référence.

Il y a, tout d'abord, les régions qui résultent d'une segmentation de l'espace selon un critère unique et qui donnent lieu à une partition dans la mesure où le critère de classification (ou critère spécifique) peut-être considéré, à juste titre, comme une relation d'équivalence. Ce type de région homogène que P. MORAN [ 13 ] p. 33 , range sous le vocable de "région spécifique" relève des méthodes de taxinomie (segmentations, classifications). En revanche, les régions dont l'obtention font intervenir un nombre plus important de critères et qui résultent le plus souvent d'un regroupement de régions spécifiques, font appel à des algorithmes de réduction des données (les analyses factorielles principalement) qui ne proposent pas directement à l'analyste une partition de l'espace de référence (même si on se croit obligé de définir des classes d'équivalence au vu des résultats obtenus). Ce second type de région que P. MORAN (ibid. p. 33) suggère d'appeler "région synthétique " possède, par définition, un contenu beaucoup plus riche que la région spécifique.

Lorsqu'elle est définie à partir des circonscriptions administratives existantes, la région synthétique permet, en outre, la mise en évidence d'entités économiques plus vastes que les précédentes et fournit aux divers responsables de l'aménagement du territoire un instrument de travail qui leur permettront d'établir "le diagnostic de situations complexes et de définir les actions et moyens spécifiques à mettre en oeuvre" (A. DORE et G. PELISSONNIER [3] p. 9).

L'introduction d'un nombre important de critères présente, en effet l'avantage de caractériser la structure de chaque région synthétique et peut, par conséquent, aider l'analyste à formuler des hypothèses et à rechercher des explications sur les mutations socio-économiques des régions.

Pour ces raisons, la présente étude retient la région synthétique comme type privilégié de région homogène.

Après avoir exposé aussi précisément que possible les termes du problème à résoudre, nous montrerons comment il s'avère possible de définir un système concret de régions homogènes en retenant des méthodes appropriées et en essayant de jouer sur la complémentarité des techniques.

## 0 - POSITION DU PROBLEME

### 0 - 1 - Données

- Soit I un ensemble d'espaces élémentaires :

$$I = \{1 \dots i \dots m\}$$

il représente ici les 21 régions françaises de programme (Provence-Côte-D'Azur et Corse ne formant qu'une seule région)

- On considère un ensemble J dont les éléments se réfèrent à des branches économiques selon la nomenclature retenue par l'INSEE

$$J = \{1 \dots j \dots n\}$$

il comprend 22 branches industrielles et la branche agricole (cf. annexe 1)

- On forme la matrice de terme général  $x_{ij}$  associée à l'ensemble produit  $I \times J$  :

$x_{ij}$  = valeur ajoutée brute aux prix du marché pour la branche j dans la région i en 1970.

### Remarque :

Les chiffres sont extraits des comptes de branches élaborés par l'INSEE (pour les branches industrielles c.f. les "Collections de l'INSEE" R - 21. p.p. 59 - 69 ; pour la branche agricole c.f. les "Collections de l'INSEE" R - 15. p. 93). Le tableau des données est reproduit dans un travail antérieur qui a servi de base à la présente étude (B. FUSTIER et B. ROUGET [6] ).

### 0 - 2 - Notations et définitions

D'après ce qui précède :

$\sum_j x_{ij} = x_{i.}$  = valeur ajoutée par l'ensemble des branches au niveau de la région i.



Donc :

$f_{ij} = \frac{x_{ij}}{x_i}$  = contribution relative de  $j$  à la valeur ajoutée par l'ensemble des branches au niveau de la région  $i$ .

On vérifie que :

$$0 \leq f_{ij} \leq 1, \quad \forall j \in J \text{ et } \sum_j f_{ij} = 1$$

La suite  $(f_{ij})_{j \in J}$  est appelée "profil économique" de la région  $i$ .

Remarque :

Pour éviter une surcharge inutile de notations, on omettra la référence à l'indice  $i$  : le profil  $(f_j)_{j \in J}$  caractérise une région donnée.

### 0 - 3 - Objectifs

Dans un premier temps, on se propose d'effectuer un découpage de l'espace national en regroupant les régions de programme en fonction des ressemblances entretenues par les profils économiques. L'analyse factorielle des correspondances, étant donné le caractère multidimensionnel des profils, semble particulièrement bien adaptée à l'établissement d'une telle typologie qui ne doit pas coïncider nécessairement avec une partition de l'espace de référence.

Si l'on veut, pour des considérations pratiques liées aux contraintes du planificateur, éliminer les inconvénients de l'appartenance simultanée d'une même région de programme à des regroupements différents, il semble plus judicieux d'utiliser un critère extérieur à la première analyse ; l'établissement d'une partition fera l'objet de la seconde étape de l'étude en tenant compte du niveau de spécialisation régional. La définition d'un indicateur de spécialisation sera alors proposée.

## I - TYPOLOGIE PROVISoire SELON LE DEGRE DE RESSEMBLANCE DE LA STRUCTURE PRODUCTIVE DES REGIONS

### I - 1 - Application de l'analyse factorielle des correspondances

#### I - 1 - 1 - But

L'analyse factorielle des correspondances possède essentiellement un caractère descriptif en tant qu'algorithme de réduction des données. Elle permet également de déceler les causes de certaines proximités (ou ressemblances\*) et peut, dans certains cas, être utilisée " comme moyen d'investigation des relations causales" (J.L. GUIGOU [9] p. 47).

Dans la présente étude, on s'intéresse uniquement à l'aspect descriptif de la méthode . L'effort portera plus précisément sur la description des proximités des éléments de I eu égard à la position des éléments de J. Le problème consiste donc à maximiser les ressemblances régionales en fonction de leur structure productive traduite par les profils économiques.

Cette façon de procéder est du reste compatible avec la réintroduction possible, après coup, d'autres critères, économiques ou non qui viendront illustrer "les composantes de la typologie"(c.f. annexe 1). Ces critères illustratifs ne participent pas à la détermination des axes factoriels : ils sont projetés sur les axes ou sur les plans factoriels en fonction de leurs corrélations avec le profil des branches \*\*

#### I - 1 - 2 - Méthodologie

Convenons que la délimitation des régions homogènes sur les différents plans factoriels s'avère particulièrement difficile dans la mesure où l'appartenance d'un point à un regroupement est, dans bien des cas , décidée d'une façon plus ou moins subjective. D'autre part, nous savons

---

\* Dans ce type d'analyse factorielle, nous savons que la proximité entre 2 points d'un même ensemble est mesurée par la distance du khi-deux qui rend compte des proximités entre forme des profils (indépendamment des niveaux ou tailles)

\*\* D'une façon plus précise, la projection d'un critère illustratif revient à faire une régression multiple de ce critère sur les axes factoriels préalablement déterminés.

que deux points assez éloignés l'un de l'autre dans l'hyperplan initial, peuvent occuper des positions voisines dans le plan factoriel et faire partie, par conséquent et abusivement, du même regroupement \* .

Pour tenter de surmonter ces difficultés, les utilisateurs de l'analyse factorielle ont souvent recours à d'autres méthodes d'analyse des données. En particulier, l'analyse arborescente (A. GROSMANGIN [8]) semble très bien répondre aux préoccupations précédentes \*\*. Mais lorsqu'il s'agit de mettre en évidence une typologie tout en essayant d'interpréter économiquement les sous-ensembles qui la composent, le problème se complique et ne paraît pas devoir relever de l'application d'une quelconque technique complémentaire à la précédente.

Pour cette raison la recherche de la typologie s'est accompagnée, comme nous allons le voir maintenant, de plusieurs analyses factorielles en éliminant à chaque fois les points (régions ou branches) jugés "aberrants"\*\*\*. Pour chaque analyse, les sous-ensembles de régions retenus font l'objet d'un examen particulier sur chacun des axes factoriels extraits puis sur chacun des plans. Cette procédure quelque peu fastidieuse permet, toutefois, de tester la stabilité de la typologie mise à jour.

## I - 2 - Résultats

- La première analyse que nous présentons à titre d'exemple (fig 1), prend en compte la totalité des éléments de I et de J. Elle montre le caractère "aberrant" du profil économique de la Lorraine et, dans une moindre mesure, du profil du Nord ( la somme des contributions absolues de ces 2 régions à la construction du premier axe factoriel est considérable, presque 83 %).

---

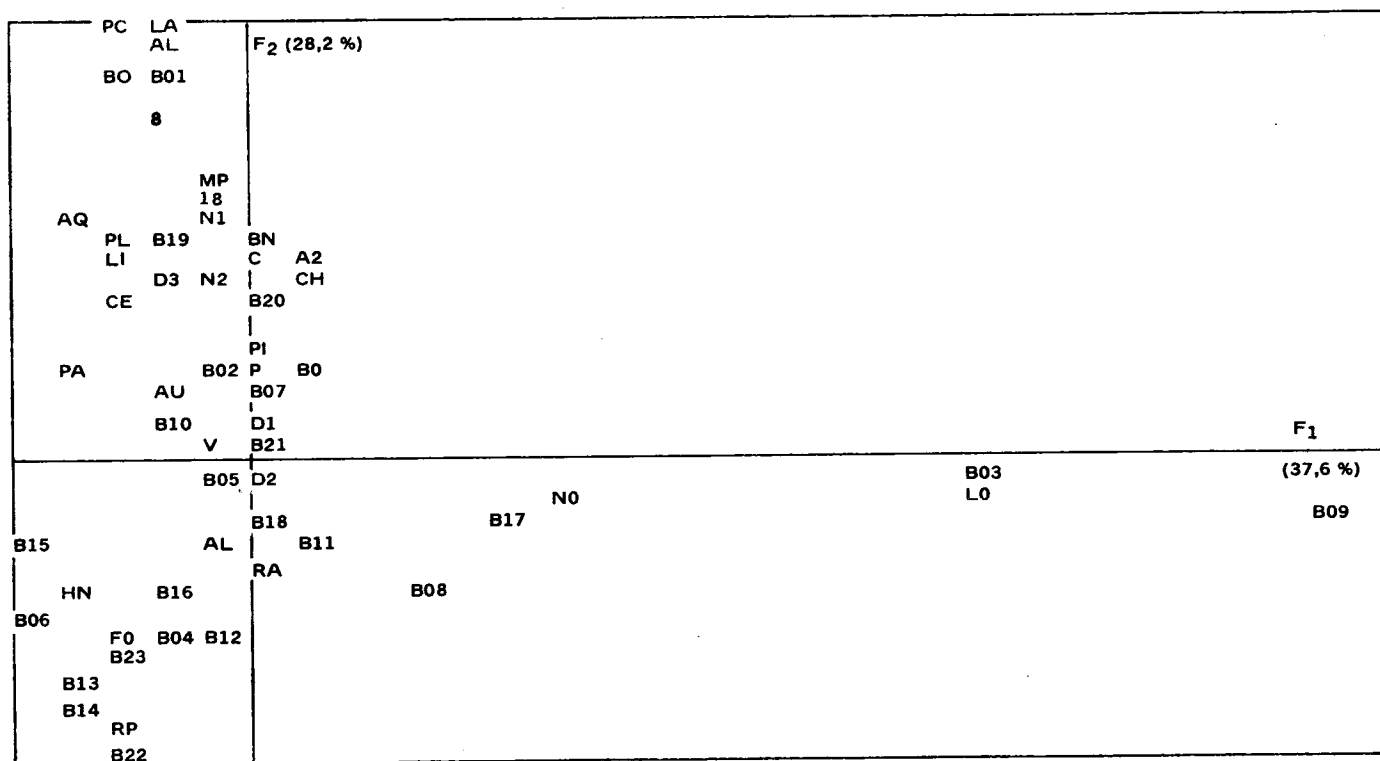
\* En moyenne, cet inconvénient est assez rare, car on constate une bonne stabilité des images obtenues par projection (J.P. BALLADUR [1] p. 74)

\*\* Cette technique (appelée quelquefois, mais à tort, "analyse hiérarchique" considère le nuage de points dans le même système de coordonnées que celui utilisé par l'analyse factorielle des correspondances, mais au lieu de le projeter, elle cherche à caractériser directement les proximités.

\*\*\* Le carré de l'abscisse d'un point sur un axe, pondéré par sa masse, mesure le moment d'inertie du point par rapport au plan perpendiculaire à l'axe et passant par le centre de gravité. Au centre l'inertie est nulle. Plus l'inertie est grande, plus le point se situe à l'extérieur du plan factoriel loin des autres points : son profil est très différent des autres d'où le terme "point aberrant".

FIGURE 1

PLAN FACTORIEL DETERMINE PAR LES DEUX PREMIERS AXES DE L'ANALYSE N°1



- Pour la raison évoquée ci-dessus, le Nord et la Lorraine sont éliminés de la deuxième analyse. Elles sont simplement projetées à titre de "régions illustratives" \* selon la même technique que les "critères illustratifs".

- La troisième analyse élimine la branche agricole dont la discrimination est apparue importante dans l'analyse précédente. Cette branche est projetée ici en "critère illustratif" ainsi qu'un plus grand nombre de régions. La présente analyse précise davantage la typologie ressortant des précédentes.

- La recherche des caractéristiques de la typologie obtenue fait l'objet d'une quatrième analyse effectuée, non pas pour tester les regroupements déjà réalisés, le nombre des données éliminées devenant trop important, mais pour préciser la structure économique de chaque région homogène (c.f. annexe 2).

La figure 2 reproduite ci-après représente le découpage de l'espace national effectué sur la base des analyses précédentes. Il ne correspond pas à une partition sur ce dernier. Ainsi le Centre appartient à deux régions homogènes : "Ceinture de l'Ouest" et "France Moyenne". Il en va de même pour la Basse Normandie. D'autre part, on constate que des régions possédant une forte individualité ne semblent pas devoir s'associer avec d'autres, c'est le cas pour la région parisienne, la Franche-Comté et la région Rhône-Alpes.

## II - PARTITION DE L'ESPACE NATIONAL FONDÉE SUR LE NIVEAU DE SPECIALISATION REGIONALE

### II - 1 - Mesures de la spécialisation régionale

#### II - 1 - 1 - Degré de "désordre" d'un profil économique

L'entropie \*\* semble particulièrement bien adaptée à la mesure du degré de désordre pouvant affecter le profil économique d'une région et on peut se demander si, à la limite, la définition d'un indicateur

---

\* Elles sont projetées sur les axes en fonction de leurs corrélations avec les profils des autres régions retenues dans l'analyse principale.

\*\* Pour les différents sens donnés au concept d'entropie, voir par exemple : H. GUITTON [1] ; B. FUSTIER [7] p. 426 et s.

TYPLOGIE

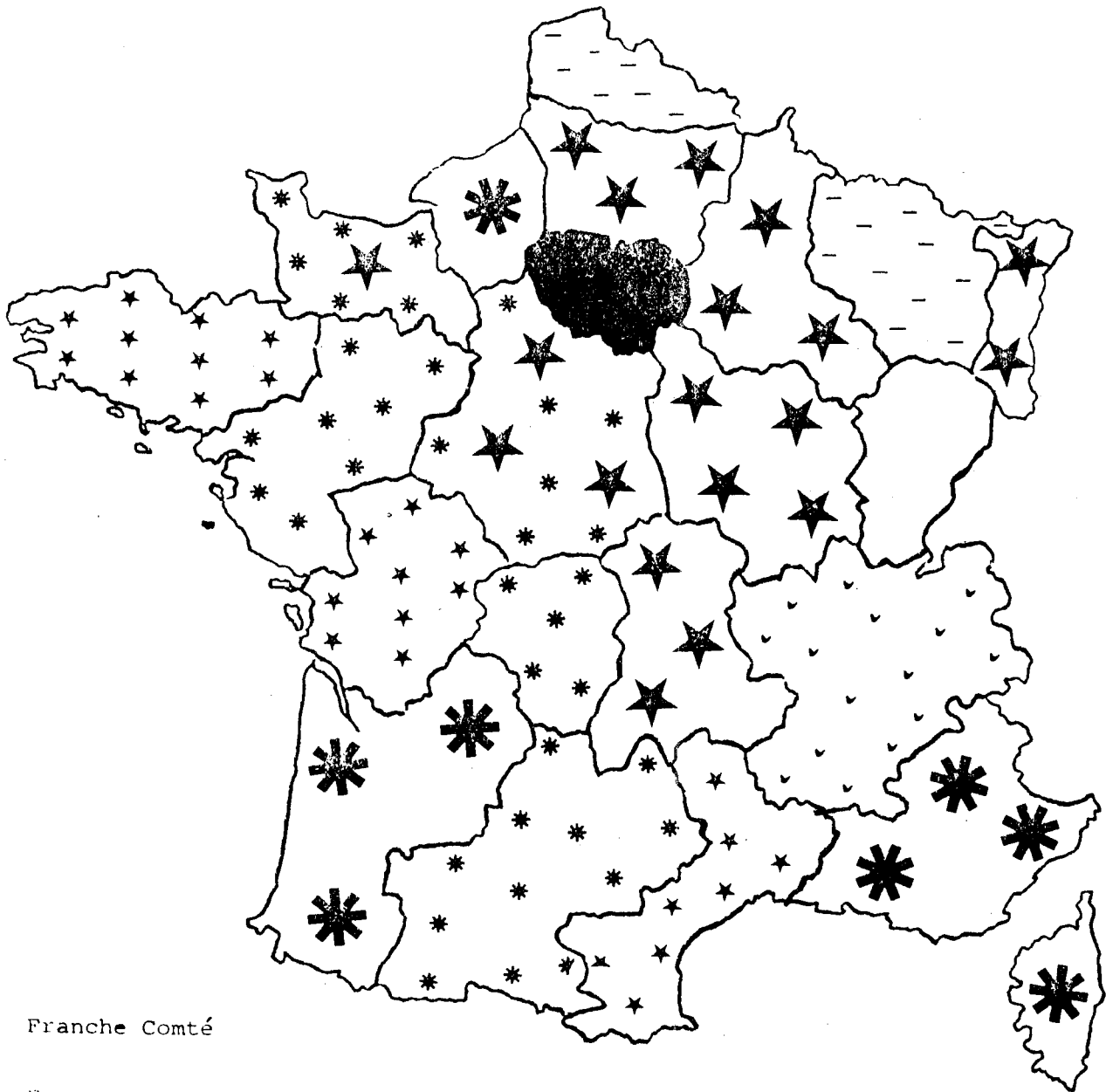


FIGURE 2  
TYPLOGIE PROVISoire

de spécialisation peut en résulter ?

Par définition l'entropie  $H$  d'une région donnée a pour expression (quelle que soit la base du logarithme) :

$$(1) \quad H = - \sum_{j=1}^n f_j \log f_j$$

On démontre (B. FUSTIER [7] p.p. 748-49) que  $H$  prend ses valeurs dans l'intervalle  $[0, \log n]$ . Pour les besoins de l'exposé, les résultats suivants sont rappelés :

cas 1 :  $H = 0 \Leftrightarrow f_{j_1} = 1$  et  $f_j = 0, \forall j \neq j_1$ . Le profil économique de la région considérée est constitué par l'unité et des 0 ailleurs : la région est complètement spécialisée.

cas 2 :  $H = \log n \Leftrightarrow f_j = \frac{1}{n}, \forall j \in J$ . Toutes les branches interviennent également dans le profil économique de la région considérée : la région est complètement diversifiée.

Mais en dehors de ces cas limites, la structure mathématique de l'entropie ne lui confère pas une forme idéale pour rendre compte du niveau de spécialisation régionale. On constate, en effet, que des régions associées à des profils économiques assez dissemblables, et ayant par conséquent des niveaux de spécialisation différents, possèdent en fin de compte la même entropie \*.

Le problème est de la même nature que celui posé par la structure mathématique de toute moyenne arithmétique. D'ailleurs, l'entropie est, au signe près, la moyenne arithmétique du caractère  $\log f$  pondérée par les fréquences  $f_j$  :

$$(2) \quad -H = \sum_j f_j \log f_j$$

Cette remarque nous invite à définir l'écart-type  $\sigma_1$  du caractère  $\log f$  :

$$(3) \quad \sigma_1 = \sqrt{\sum_j f_j (\log f_j + H)^2}$$

---

\* Cette affirmation a été testée sur des exemples hypothétiques par les étudiants du D.E.S.S. "Analyse Economique de Branche" (1977-78)

Ou en développant :

$$(4) \quad \sigma_1 = \sqrt{\sum_j f_j \log^2 f_j - H^2}$$

$\sigma_1$  mesure donc la dispersion du logarithme des éléments  $f_j$  d'un profil économique donné autour de leur moyenne arithmétique  $-H$ . Par conséquent, des régions de même entropie, mais associées à des profils économiques différents, pourront être jugées plus ou moins spécialisées selon les valeurs de  $\sigma_1$ .

Mais dans les cas extrêmes (cas 1 et 2), donc sans appel, l'écart-type doit normalement s'annuler. En effet, on démontre (annexe 3) que :

$$(5) \text{ cas 1, cas 2} \Rightarrow \sigma_1 = 0$$

C'est dans cette perspective que la définition d'un indicateur de spécialisation doit être tentée. Mais auparavant, il convient de donner une interprétation statistique plus précise de  $\sigma_1$ .

## II - 1 - 2 - Interprétation de l'écart-type défini sur le logarithme des données.

Etant donné que  $\sum_j f_j = 1$ , la moyenne géométrique, notée  $f_g$ , du caractère  $f$  pondéré par les fréquences  $f_j$  possède une formule relativement simple :

$$(6) \quad f_g = \prod_j f_j^{f_j}$$

Son logarithme a pour expression :

$$(7) \quad \log f_g = \sum_j f_j \log f_j$$

En rapprochant (2) de (7) on obtient :

$$(8) \quad \log f_g = -H$$

Par conséquent,  $\sigma_1$  peut encore s'exprimer de la façon suivante :

$$(9) \quad \sigma_1 = \sqrt{\sum_j f_j (\log f_j - \log f_g)^2}$$

Cette nouvelle expression de  $\sigma_1$  fait intervenir explicitement la dispersion du logarithme des données autour du logarithme de leur moyenne



géométrique, et nous amène à la définition de l'écart-géométrique, noté  $\sigma_2$  tel que :

$$(10) \quad \sigma_2 = \exp \sigma_1$$

L'écart-géométrique est donc le nombre  $\sigma_2$  dont le logarithme est égal à  $\sigma_1$ . Il mesure la dispersion du caractère f autour de sa moyenne géométrique  $f_g$  (pondérée par les  $f_j$ ).

En rapprochant (5) et (10) on vérifie immédiatement que l'écart-géométrique est égal à l'unité lorsque la région est complètement spécialisée ou complètement diversifiée :

$$(11) \quad \text{cas 1, cas 2} \Rightarrow \sigma_2 = 1$$

Remarquons bien que la moyenne géométrique possède ici une signification économique intéressante.

En effet, le nombre F bien connu des utilisateurs du concept d'entropie s'interprète ici comme le nombre de branches de taille égale conduisant à l'entropie observée H. Or son inverse est justement égal à la moyenne géométrique calculée à partir des éléments constituant le profil économique de la région en question. Montrons le.

Par définition :

$$(12) \quad F = \exp H$$

Or d'après (8)

$$(13) \quad f_g = \exp - H$$

Donc :

$$(14) \quad f_g = \frac{1}{F}$$

Par conséquent, si la valeur de H est obtenue par un nombre théorique F de branches de taille égale,  $f_g$  représente alors la contribution relative d'une de ces branches à la valeur ajoutée totale au niveau de la région.

En d'autres termes, la valeur observée H est calculée d'après le profil économique  $(f_j)_{j \in J}$  de la région. Elle peut-être obtenue également à partir d'un profil théorique constitué par des fréquences dont les valeurs sont égales à 0 ou  $f_g$ .

On vérifie que  $\sigma_2$  prend la valeur 1 lorsque le profil économique observé correspond au profil théorique ainsi défini. Les cas 1 et 2 de la relation (11) ne sont, en fait, que des cas particuliers \*. Par conséquent,  $\sigma_2$  possèdent des valeurs identiques que la région soit complètement spécialisée, ou complètement diversifiée, ou encore dans d'autres situations intermédiaires pourvu que le profil économique observé coïncide avec le profil théorique.

L'écart-géométrique ne peut donc pas constituer, à lui seul, un indicateur de spécialisation.

### II - 1 - 3 - Définition d'un indicateur de spécialisation.

Puisque  $\sigma_2$  mesure la dispersion des données autour de leur moyenne géométrique, une solution possible consisterait à définir un indicateur de spécialisation par analogie avec le coefficient de variation utilisé habituellement en statistique, c'est-à-dire par le rapport :

$$(15) \quad \alpha = \frac{\sigma_2}{f_g}$$

ou d'après (14) :

$$(16) \quad \alpha = F \sigma_2$$

Mais cet indicateur présente une difficulté certaine d'interprétation dans la mesure où un coefficient de variation fait intervenir la dispersion d'un caractère exprimée par rapport à sa moyenne arithmétique. Ainsi, le coefficient  $\gamma_1$  défini ci-dessous est un coefficient de variation au sens statistique du terme :

$$(17) \quad \gamma_1 = \frac{\sigma_1}{-H}$$

En revenant à un caractère exprimé dans un système de coordonnées arithmétiques, on définit le nombre  $\gamma_2$  tel que :

---

\* En effet, lorsque la région est complètement spécialisée, son profil économique est constitué par l'unité et des 0 ailleurs ; on sait que  $H = 0$  donc  $F = 1$  et  $f_g = 1$ . Le profil théorique obtenu est identique au profil observé. De même lorsque la région est complètement diversifiée, son profil comporte des fréquences dont les valeurs sont toutes égales à  $1/n$  ; comme  $H = \log n$ , alors  $F = n$  et  $f_g = 1/n$ .

$$(18) \quad \boxed{\gamma_2 = \exp \gamma_1}$$

$\gamma_2$  peut être considéré comme un indicateur du niveau de spécialisation régionale. Il prend ses valeurs dans l'intervalle  $[0, 1]$  et varie en fonction directe du niveau de diversification atteint par les régions. On démontre (annexe 3) les propriétés suivantes :

(19) cas 1  $\Rightarrow \gamma_2 = 0$  (région complètement spécialisée)

(20) cas 2  $\Rightarrow \gamma_2 = 1$  (région complètement diversifiée)

## II - 2 - Résultats

### II - 2 - 1 - Diversification des régions françaises.

L'annexe 4 présente le préordre établi sur les régions de programme en fonction de leur niveau de spécialisation donné par les valeurs de  $\gamma_2$  dont la moyenne arithmétique assez élevée (0,645) préjuge, en fait, d'un bon degré de diversification des régions françaises.

Les régions les plus diversifiées forment un ensemble comprenant la Haute-Normandie, le Nord, la région parisienne, la région Rhône-Alpes et l'Alsace.

Parmi les régions accusant une tendance certaine à la spécialisation on compte la Bretagne, le Languedoc, le Poitou-Charente et la Basse-Normandie.

Les autres régions forment un bloc intermédiaire.

### II - 2 - 2 - Complémentarité de l'analyse factorielle des correspondances et de l'indice de spécialisation.

L'indice de spécialisation retenu fournit un renseignement global sur le profil économique d'une région sans préciser, toutefois, les branches à l'origine de la valeur observée de  $\gamma_2$ . On sait, par exemple, que la Bretagne est la région française la plus spécialisée, mais l'indice ne donne pas la ou les quelques branche(s) responsable(s) de cet état de fait. L'analyse factorielle des correspondances permet de pallier cette carence d'information.

En effet, sur un plan factoriel, la proximité croisée entre un point-région et un point-branche donne une idée assez précise sur la

composition structurelle des profils associés à ces points. Une région sera d'autant plus proche d'une branche donnée que celle-ci interviendra fortement dans le profil de la région, et réciproquement. Aussi peut-on affirmer que la Bretagne est fortement spécialisée dans le domaine agricole, que la Basse-Normandie est caractérisée principalement par l'industrie sidérurgique et que le Centre, dans une moindre mesure, l'est par les industries agricoles et alimentaires.

Mais il convient de souligner que des régions situées à une des extrémités du plan factoriel et positionnées à proximité d'une branche économique donnée, ne sont pas des régions nécessairement spécialisées dans leur ensemble. C'est le cas pour le Nord, par exemple, qui vient se projeter vers l'industrie textile. En effet, le Nord est la première région française, en termes de valeur ajoutée, pour l'industrie textile ce qui tend à expliquer sa position "aberrante" sur le plan factoriel, mais son profil économique n'est pas constitué uniquement par cette activité. Les autres branches sont bien représentées dans leur ensemble d'où l'aspect diversifié de son économie régionale révélé par la valeur relativement importante de  $\lambda_2$ . L'analyse factorielle apporte une information supplémentaire : le Nord accuse une spécialisation relative pour l'industrie textile au niveau national.

## II - 2 - 3 - Obtention d'une partition

La Basse-Normandie et le Centre appartiennent simultanément à deux regroupements différents ("France moyenne" et "Ceinture de l'Ouest") ; cet état de fait peut entraîner certaines conséquences jugées inopportunes par le planificateur lorsqu'il s'agit de mettre en oeuvre des moyens spécifiques et des actions adaptées à chaque région homogène.

Pour tenter de simplifier la tâche du praticien, il semble que l'exclusion de ces deux régions de programme de la "France moyenne" au profit de la "Ceinture de l'ouest" renforce l'homogénéité des regroupements tout en réalisant une partition sur l'espace national.

En effet, les résultats précédents montrent que la Basse-Normandie et, à un moindre degré, le Centre sont spécialisés dans des activités affectant un taux de croissance en régression par rapport aux autres branches économiques (M. FONDEVIOLE [5] p. 54). Ajoutés à la "Ceinture de l'Ouest", la Basse-Normandie et le Centre ressemblent davantage aux autres

régions de ce regroupement caractérisées également par des taux d'activité assez faibles (industries du cuir, papiers et cartons...). La "France Moyenne" regroupe maintenant 5 régions de programme (Alsace, Auvergne, Picardie, Bourgogne et Champagne) caractérisées par des branches économiques en expansion comme l'industries chimique et la transformation des métaux (M. FONDEVIOLE. *ibid* ).

Pour terminer, signalons que la délimitation d'un bloc de régions économiquement défavorisées ne constitue pas un non sens économique dans la mesure où le handicap commun peut amener les pouvoirs publics à proposer des solutions d'ensemble qui peuvent se concrétiser par une politique de développement régional (Y. DURRIEU [4] p. 73).

#### CONCLUSION :

Afin d'éviter le piège de la subjectivité qui accompagne trop souvent l'interprétation des résultats issus d'une analyse factorielle (C. LABROUSSE [12] p. 108 ; M. MOUGEOT [14]), le découpage de l'espace proposé ici fait intervenir une seconde technique dont la complémentarité avec la première a été soulignée dans les développements précédents.

Ajoutons que les résultats donnés par l'indicateur de spécialisation tendent à prouver que la notion d'homogénéité retenue par l'analyse factorielle n'implique pas nécessairement des regroupements de régions spécialisées dans une même activité. La "France du Nord et de la Lorraine" regroupe une région assez diversifiée (Lorraine) et une autre qui l'est beaucoup plus (Nord). La région parisienne qui forme à elle seule une région homogène est très diversifiée. De même pour la région Rhône-Alpes et, dans une moindre mesure, pour la Franche-Comté.

Homogénéité ne signifie pas spécialisation.

Cette confusion résulte généralement d'une assimilation trop rapide des notions de régions spécifique et synthétique à l'intérieur du concept

plus vaste de région homogène.

Par nature, une région spécifique est une région spécialisée lorsque le critère unique de sériation coïncide avec une branche d'activité. En revanche, la région synthétique résulte d'une analyse multicritère, d'une combinaison de facteurs multiples en un critère synthétique sur lequel est fondé la typologie. De ce point de vue l'analyse factorielle des correspondances constitue un outil particulièrement bien adapté à la mise en évidence des régions homogènes (synthétiques).

Mais l'analyse statistique multidimensionnelle ne doit pas rester l'instrument exclusif des problèmes de régionalisation. Les indicateurs socio-économiques et la comptabilité nationale peuvent être utilisées à cette fin (A. DORE et G. PELISSONNIER [3] p. 9 et s.). Toutefois, l'application de la théorie des sous-ensembles flous apparaît, ici, constituer l'apport le plus intéressant puisqu'elle donne des algorithmes appropriés au traitement de l'imprécision des frontières régionales et des contenus qu'elles enveloppent (C. PONSARD [15] ). Les résultats qu'elle propose sont, en général, beaucoup plus riches que ceux donnés par l'analyse factorielle (B. FUSTIER [7] p. 704), et les travaux de Ph. TRAN QUI [16] sur la classification régionale floue, semblent également corroborer cette impression.

## ANNEXE 1

### Codes

#### Régions

RP (Région Parisienne) - CH (Champagne) - PI (Picardie)  
HN (Hte Normandie) - CE (Centre) - BN (Basse Normandie)  
BO (Bourgogne) - NO (Bord) - LO (Lorraine) - AL (Alsace)  
FC (Fanche-Comté) - PL (Pays de la Loire) - BR (Bretagne)  
PC (Poitou-Charente) - AQ (Aquitaine) - MP (Midi-Pyrénées)  
LI (Limousin) - RA (Rnônes-alpes) - AU (Auvergne) - LA (Languedoc)

#### Branches

01 (Agriculture) 02 (Industries agricoles et alimentaires)  
03 (Combustibles minéraux solides) 04 (Gaz) 05 (Electricité  
eau et divers) 06 (Pétrole, gaz naturel et carburants) 07 (Ma-  
tériaux de construction) 08 (Verre) 09 (Minerais de fer et indus-  
trie sédérurgique) 10 (Minerais et métaux non ferreux)  
11 (Première transformation des métaux) 12 (Industries mécaniques)  
13 (Industries électriques) 14 (Automobile) 15 (Construction  
navale, aéronautique, armement) 16 (Industrie chimique)  
17 (Industrie textile) 18 (Habillement) 19 (Industrie du cuir)  
20 (Industrie du bois) 21 (Papiers et cartons) 22 (Presse et  
édition) 23 (Industries diverses).

#### Critères illustratifs

##### 1) Analyses I, II et III

S : Superficie de la région (Annuaire statistique INSEE 1972)

P : Population légale (Annuaire statistique INSEE 1972)

A<sub>1</sub> : Répartition de la population active dans le secteur primaire

A<sub>2</sub> : Répartition de la population active dans le secteur secondaire

A<sub>3</sub> : Répartition de la population active dans le secteur tertiaire

C : Taux de chômage total

B : Nombre de débits de boissons en 1969 (collections de l'INSEE)

V : Immatriculation des voitures particulières et commerciales  
neuves (Annuaire statistique INSEE)

- Niveau de diplômes en pourcentage de la population  
de 15 ans ou plus ni scolaire, ni apprentie :  
(collection INSEE R 16-17)

N<sub>1</sub> : BEPC, BE ou BEPS et CAP, brevets ou aucun diplômes déclarés

N<sub>2</sub> : Diplôme de niveau supérieur au bac complet

- Nombre de condamnés par régions suivant la nature  
des délits : (compte général du Ministère de la  
Justice 1970)

D<sub>1</sub> : vols

D<sub>2</sub> : Chèques sans provision

D<sub>3</sub> : Fraudes commerciales

D<sub>4</sub> : Conduite en état d'ivresse

D<sub>5</sub> : Total de tous les délits

R : Répartition régionale des crédits distribués par le système  
bancaire en 1970 (Bulletin trimestriel de la Banque de France  
n° 3.3.1972)

## 2) Analyse IV

Variables citées précédemment avec en outre :

L 40 : Vacances des habitants de la région (collection de l'INSEE 71)

L 41 : Journée de vacances que les gens prennent dans la région  
(niveau touristique)

L 42 : Dépenses d'aide sociale par habitant Bulletin de statistique

L 43 : Lits d'hôpitaux et d'hospices publics Ministère de la santé  
1970 publique 1970.

L 44 : Impôts sur le revenu des personnes physiques (INSEE 71)

L 45 : Trafic total de téléphone (INSEE 71)

L 46 : Redevance de T.V. (INSEE 71)

### Remarque :

Toutes les variables illustratives sont en pourcentage.



## ANNEXE 2

### Caractéristique de la typologie

Regroupe- ments	Régions	Caractéristiques après éli- mination des branches et des régions	Caracté- ristiques moyennes	Critères illus- tratifs
France moyenne	AL	Encore moyenne	02, 07	P, B D <sub>1</sub> , D <sub>2</sub> V, L 43 L 44, N <sub>2</sub>
	AU	Industrie chimique (Miche- lin)		
	PI	Transformation des métaux		
	BO	Industrie chimique (Kodak)		
	CH	Industrie textile		
	CE	Industries agricoles et alimentaires		
	BN	Minerai de fer et indus- trie sidérurgique		
France agricole	PC	Régions illustratives	01	S A <sub>1</sub>
	BR			
	LA			
Ceinture de l'Ouest	BN	Minerai de fer et indus- trie sidérurgique	19,20	A <sub>2</sub> , A <sub>3</sub> N <sub>1</sub> , N <sub>2</sub> B, C D <sub>3</sub> , D <sub>4</sub>
	PL	Electricité, eau, divers		
	CE	Industries agricoles et alimentaires		
	LI	Papiers et cartons		
	MP	Industries du cuir		
France Portuaire	HN	Industrie électrique	06,15	
	PA	Pétrole et industries agricoles et alimen- taires		
	AQ	Construction navale, aéro- nautique et armement		

France Rhodanienne	RA		11,14,17,09	
France du Nord et de la Lorraine	I.O NO	Régions illustratives	03,09,17	
Région Parisienne	RP	Région illustrative	12,13,14,22,23	R
Franche Comté	FC	Industries diverses	14	

### ANNEXE 3

#### I - Valeurs de $\sigma_1$ pour le cas 1 et 2

$$\sigma_1 = \sqrt{\sum_j f_j \log^2 f_j - H^2}$$

##### cas 1

$$\forall j \neq j', f_j = 0, \text{ donc } \sum_{j \neq j'} f_j \log^2 f_j \rightarrow 0$$

$$f_{j'} = 1, \text{ donc } f_{j'} \log^2 f_{j'} = 0$$

Comme  $H = 0$ , alors  $\sigma_1 = 0$

##### cas 2

$$\forall j \in J, f_j = \frac{1}{n}, \text{ donc } \sum_j f_j \log^2 f_j = \log^2 n$$

Comme  $H = \log n$ , alors  $\sigma_1 = 0$

#### II - Valeurs de $\gamma_2$ pour les cas 1 et 2

$$\gamma_2 = \exp \frac{\sigma_1}{-H}$$

##### cas 1

D'après ce qui précède  $\frac{\sigma_1}{-H}$  apparaît comme une forme indéterminée qu'il convient de lever.

a) Considérons un profil économique à 2 composante  $n = 2$  :

On se donne un nombre  $\varepsilon$  tel que :  $0 \leq \varepsilon \leq 1$

$$\frac{\sigma_1}{-H} = -\sqrt{\frac{f_1 \log^2 f_1 + f_2 \log^2 f_2}{H^2}} - 1$$

avec  $f_1 = 1 - \varepsilon$  et  $f_2 = \varepsilon$

Comme  $H^2 = [(1 - \varepsilon) \log (1 - \varepsilon) + \varepsilon \log \varepsilon]^2$ , il convient de déterminer la limite de l'expression  $E$  quand  $\varepsilon \longrightarrow 0$

$$E = \frac{(1 - \varepsilon) \log (1 - \varepsilon) + \varepsilon \log \varepsilon}{[(1 - \varepsilon) \log (1 - \varepsilon) + \varepsilon \log \varepsilon]^2} \sim \frac{\varepsilon^2 + \varepsilon \log \varepsilon}{\varepsilon^2 (1 + \log \varepsilon)^2}$$

$$E \sim \frac{1}{\varepsilon}$$

$$\text{Donc } \lim_{\varepsilon \rightarrow 0} E = + \infty$$

$$\text{Puisque } \frac{\sigma_1}{-H} = - \sqrt{E - 1}, \quad \lim_{\varepsilon \rightarrow 0} \frac{\sigma_1}{-H} = - \infty$$

$$\text{Par conséquent : } \gamma_2 = 0$$

b) Par récurrence on généralise le résultat précédent pour  $n > 2$

### cas 2

$$\text{Trivial. } \gamma_2 = 1$$

ANNEXE 4

Classement des régions de programme  
selon leur niveau de spécialisation

N.B. : Les valeurs de  $\chi^2$  sont indiquées entre parenthèses

Bretagne (0,45)  
Languedoc (0,47)  
Poitou-Charente (0,52)  
Basse-Normandie (0,58)

Lorraine, Midi-Pyrénées, Aquitaine (0,62)

Centre (0,63)

Pays de la Loire, Auvergne, Champagne (0,64)

Franche-Comté (0,65)  
Limousin (0,66)  
Picardie (0,67)  
Bourgogne (0,68)  
Provence (0,69)  
Haute-Normandie (0,74)  
Nord (0,75)

Paris, Rhône-Alpes, Alsace (0,76)

### Bibliographie

- [1] BALLADUR (J.P.) : "Analyse factorielle des correspondances"  
Annales de l'INSEE n° 14 - 1970.
- [2] CHEVAILLER (J.C.) : "Classification en analyse économique spatiale"  
Paris - Cujas - 1974
- [3] DORE (A) et PELISSONNIER (G) : " Etablissement d'une typologie  
des régions françaises".  
L'économie du Centre-Est- n° 78 - 1977 - p.P.7-16
- [4] DURRIEU (Y) : "Régionaliser la France - Capitalisme ou socialisme"  
Paris - Mercure de France - 1969
- [5] FONDEVIOLE (M) : "Vues statistiques rapides sur l'industrie  
française en 1971".  
(Economie et Statistique - n° 35 - juin 1972  
p.p. 54 - 60.
- [6] FUSTIER (B) et ROUGET (B) : " Analyse économique de branche et  
recherche d'une typologie des régions françaises"  
Faculté de Science Economique et de Gestion  
Dijon - mai 1976 (document de travail du sémi-  
naire d'Analyse Economique de Branche)ronéoté.
- [7] FUSTIER (B) : "Interactions humaines et espace économique"  
Thèse de doctorat d'Etat - Faculté de Science  
Economique et de Gestion - Dijon - 1977.
- [8] GROSMANGIN (A) : " Construction d'une nomenclature économique  
d'activités"  
Thèse de doctorat 3ème cycle - Paris - ISUP  
1970
- [9] GUIGOU (J.L.) : "Théorie économique et transformation de l'espace  
agricole" - Tome 1.  
T.E.M. Paris - Gauthier - Villars - 1972
- [10] GUIGOU (J.L.) : "A propos de l'article de Michel MOUGEOT ([14])"  
A paraître dans la revue d'Economie Politique.
- [11] GUITTON (H) : "Entropie et gaspillage"  
Paris - Cujas 1975.
- [12] LABROUSSE (C) : "Introduction à l'économétrie"  
Paris-Dunod- 1972.
- [13] MORAN (P) : "L'analyse spatiale en science économique"  
Paris - Cujas - 1966

- [14] MOUGEOT (M) : "Analyse factorielle et régionalisation ou l'ir-réductibilité des particularismes régionaux"  
Faculté de Droit et des Sciences Economiques  
Besançon - Document de travail n° 11 -  
Janvier 1977 - (A paraître dans la Revue  
d'Economie Politique).
- [15] PONSARD (C) : " La région en analyse spatiale"  
IME - document de travail n° 21 - mai 1977
- [16] TRAN QUI (Ph) : "Les régions économiques floues. Application au  
cas de la France"(à paraître Collection  
de l'I.M.E. Sirey).