

Quelques réflexions sur l'exploitation statistique de données d'observatoires socio-économiques

Marie Piron
Statisticienne

Un observatoire est supposé être un outil pour observer le changement. Quelle que soit, ensuite, la définition accordée à la notion d'observatoire, et quelle qu'en soit la fonction, les problèmes méthodologiques s'articulent autour de la mise en place des données et sur l'exploitation de ces systèmes d'information.

Nous considérons, ici, acquise la mise en place des données et nous présentons une réflexion générale sur la structure et l'exploitation statistique des données d'observatoires socio-économiques qui recouvrent un bon nombre de cas de figures méthodologiques. Seront présentés en annexes des aspects plus techniques d'une démarche de l'exploitation statistique.

■ Observatoires socio-économiques

Nous nous plaçons, ici, à un niveau relativement fin pour observer le changement. L'intérêt est d'identifier des groupes d'individus et de comprendre leurs comportements compte tenu d'une réalité socio-économique complexe.

Dans ce contexte, les observatoires socio-économiques permettent l'évaluation d'un changement, par exemple de l'impact de mesures politiques prises à un niveau global, en s'intéressant aux effets de ces mesures sur une ou plusieurs populations pré-identifiées ; est-ce que cette politique a déjà des effets ; vont-ils dans le sens souhaité ? Tout ceci permet éventuellement de réorienter la politique en question. On peut aussi vouloir évaluer une population pour mettre en place des mesures. De ce fait la population d'étude sera souvent pré-identifiée suivant des critères privilégiés qui définiront des partitions sociales ou/et géographique de la population en question.

Dans les PED, l'objectif des observatoires socio-économiques réside actuellement dans la nécessité de saisir les effets des plans d'ajustement structurel. Ils visent à une meilleure connaissance des populations concernées par les mesures prises, de leur évolution, des modifications de comportement. Ainsi, espère-t-on en déduire des stratégies d'adaptation, de reconversion face à une situation de crise mais également définir des indicateurs d'évolution ou de changements économiques ou sociaux pour élaborer alors « un système de suivi d'indicateurs pertinents sur l'évolution économique et sociale de la population et des groupes qui la composent » (CHARMES, 1992).

Dans ce cadre, les observatoires ont des fonctions diverses : définir des typologies de population sur plusieurs périodes, définir des indicateurs qui répondent mieux à des propositions plus directes plus efficaces, déduire des mécanismes, saisir qualitativement le changement.

Structure de l'observatoire

Comment à partir de ces attentes, définir et concevoir un observatoire ?

On considère, dans cette optique, un observatoire socio-économique comme un dispositif de collecte de données qui permet le suivi d'une population. Il doit répondre à des impératifs de permanence, de représentativité et de pertinence nécessaire à l'observation du changement.

– Permanence (intégration de la dimension temporelle). L'idée même qui ressort d'un observatoire est la continuité dans l'observation suivant une certaine périodicité. Dans l'absolu, cette continuité est

permanente. Mais il faut avoir les moyens d'en assurer le suivi (c'est un travail à part entière qui nécessite des structures adaptées). Le plus souvent, il s'agit davantage d'une notion de durabilité (où la notion de permanence est limitée dans le temps). On peut parler alors d'enquêtes à passages répétés.

Il s'agit par conséquent de bien fixer la durée du suivi et la périodicité des passages.

– Représentativité (élaboration des échantillons). Si l'observatoire n'est pas exhaustif, l'échantillon doit respecter les caractéristiques de la population étudiée. Or le suivi d'une population sous-entend sa permanence et, dans la réalité, il faut tenir compte des mouvements de population ; d'où la nécessité d'un protocole d'échantillonnage qui prévoit souvent de renouveler régulièrement l'échantillon (compromis entre un échantillon indépendant et échantillon de suivi ou panel).

De par les objectifs de l'étude mais aussi pour assurer une meilleure représentativité, on s'appuiera sur une ou plusieurs partitions de la population. La mouvance des individus à l'intérieur des classes d'une partition apparaît comme un des aspects fondamentaux de l'analyse du changement.

Par ailleurs, il faut s'assurer d'un effectif suffisamment important (mais dont on sait qu'il est inutile qu'il soit trop élevé) de l'échantillon. Pour cela, il faut tenir compte, à la fois, de l'effectif de chaque groupe de population identifié afin de pouvoir effectuer des analyses séparées d'une part, et de celui du panel afin de rendre significatif le suivi des populations, d'autre part. Pour donner un ordre de grandeur, indiquons que l'effectif d'un groupe d'individus susceptible d'être étudié séparément doit être supérieur à trente.

– Pertinence (conception des questionnaires et structure d'information pour l'observation). Cette notion repose à la fois sur les choix d'unités d'observation et d'analyse et sur une structure de questionnaires relativement légère. La notion de pertinence nécessite, d'une période à l'autre, une bonne identification : des questions répétitives qui vont être les éléments de comparaison entre les différents passages pour mettre en évidence une évolution ; des jeux de questions différentes qui vont être les éléments pour évaluer qualitativement le changement.

Dans ce dernier cas, il faut alors intégrer de nouveaux descripteurs, au fur et à mesure de l'évolution du phénomène suivi. Par conséquent les questionnaires conçus pour des observatoires socio-économiques sont amenés à évoluer « doucement » et doivent être adaptés aux vues des résultats des passages précédents.

Concrètement, il s'agit de la mise en place méthodique d'un dispositif d'enquêtes à passages répétés. Celles-ci posent un certain nombre de problèmes spécifiques :

- quel est le protocole d'échantillonnage adapté pour à la fois capter des évolutions individuelles et représenter une population mouvante ?
- comment faire évoluer un questionnaire simple de façon à percevoir des changements qualitatifs dans une population dont les structures évoluent ?

Les exigences de permanence, de représentativité et de pertinence des observatoires sont bien envisagées dans le cadre des protocoles statistiques des enquêtes socio-économiques (BLAIZEAU, DUBOIS, 1989). De plus, il est fréquent, au sein même de l'enquête, d'avoir plusieurs niveaux d'observation emboîtés ce qui accroît la complexité de la structure de la base d'information. Par exemple, on peut vouloir disposer des caractéristiques relatives à des quartiers, à certains logements de ces quartiers, à tous les ménages composant ces logements et aux personnes appartenant à ces ménages, chacune de ces entités constituant des types d'unités d'observation emboîtés.

Par ailleurs, une autre question qui est de plus en plus souvent posée est « de savoir s'il est préférable de réunir les structures de la population sous ses divers aspects dans une seule et même enquête ou s'il ne convient pas de les scinder plutôt entre plusieurs enquêtes » (CHARMES, 1992). Autrement dit, peut-on concevoir, pour un observatoire, un questionnaire composé suivant plusieurs thèmes (c'est-à-dire de groupes homogènes de questions quant à leur contenu) voire un questionnaire à objectifs multiples ? Il ne faut pas, bien sûr, tomber dans le piège d'un instrument qui tente de cibler trop d'objectifs ce qui risquerait de masquer l'élément important attribué aux observatoires : saisir le changement. En revanche, le changement peut être appréhendé sous divers aspects à partir d'une structure de référence de la population. Il faut sans doute mieux éviter, dans la mesure du possible, d'avoir plusieurs entrées sur cette population pour étudier une mesure prise, quitte à mettre en œuvre effectivement plusieurs enquêtes qui pourront être par la suite exploitées

ensemble à partir d'un niveau d'analyse adéquat. Dans ce cas, les enquêtes seront plus légères et donc plus facilement contrôlables et cibleront une population plus nombreuse.

On peut étendre cette vision à une autre conception de l'observatoire : disposer de plusieurs sources d'informations qui répondent aux objectifs posés et qui mettent en jeu plusieurs niveaux d'analyse d'échelles différentes. L'avantage est d'utiliser des bases de données déjà créées ou d'élaborer des enquêtes légères mais surtout d'appréhender un phénomène suivant différents angles de vues (par exemple, individus, ménages, quartiers, associations de voisinage).

La structure de l'information issue d'un observatoire engendre, par conséquent, un certain nombre de difficultés dans sa mise en place notamment parmi les réflexions et les choix à faire sur :

- les unités d'observation, les unités d'échantillonnage, l'emboîtement des différents types d'unités, les partitions d'une population ;
- le mode de renouvellement de l'échantillon, l'effectif de chaque type d'échantillon (indépendant et panel), la durée du renouvellement ;
- les variables répétitives ou non, les thèmes abordés ;
- les différentes échelles de perception du changement.

Elle nécessite une bonne identification des objectifs et une bonne définition des objets étudiés.

L'observatoire intègre, alors, deux dimensions fondamentales :

- le *temps* avec les notions de permanence, de représentativité et de pertinence qu'il induit, et qui est à la base même de la structure de l'information ;
- les différentes *échelles de perception*, au sein même de l'enquête ou par la mise en relation d'enquêtes ou de sources diverses, qui ne sont pas spécifiques d'une structure de données d'observatoire mais qui sont souvent implicitement considérées.

Par conséquent, un observatoire ne peut pas être improvisé. Compte tenu des investissements humains et financiers qu'il nécessite, il serait dommage, lors de son exploitation, de ne retenir qu'un faible ensemble des informations ou de ne pas fournir les éléments d'évolution, de changement et d'évaluation attendus. De plus, compte tenu des structures complexes de base d'un observatoire, une réflexion sur la méthodologie ne peut que s'imposer. Et ceci d'autant plus qu'il faut résoudre la contradiction qui existe entre la complexité du changement et les objectifs de pertinence et d'efficacité attribués aux observatoires.

Exploitation statistique

Face à ces vastes ensembles de données conçues sur plusieurs périodes, sur plusieurs thèmes mais aussi sur plusieurs niveaux d'analyse, il est nécessaire de disposer d'outils qui permettent de structurer la base d'information et son exploitation d'une manière générale, et qui répondent aux objectifs des observatoires tels que :

- décrire des comportements économiques et sociaux des individus ;
- définir des typologies et des indicateurs ;
- analyser une évolution, un changement ;
- comprendre les mécanismes du changement social ;
- relier des informations provenant de sources ou de niveaux d'observation différents ;
- mettre en jeu des échelles différentes (passage du micro au macro).

Nous envisageons ici l'exploitation statistique des données d'observatoire d'une manière descriptive et exploratoire. Il s'agit des techniques traditionnelles de traitement des données d'enquête et des méthodes d'analyse des données (les analyses multivariées).

Le dépouillement de données d'enquête classique met en œuvre des techniques simples, éprouvées, faciles à interpréter : calcul de pourcentages, tableaux croisés pour des variables qualitatives ; calcul de moyennes, de dispersion ou de corrélations pour des variables quantitatives. Des tests d'hypothèses, tels que régressions, analyses de la variance ou encore modèles log-linéaires, viennent en appui aux résultats.

Ces techniques élémentaires ne sont, cependant, pas toujours suffisantes pour répondre aux objectifs des observatoires. En effet, parce qu'elles font intervenir un nombre restreint de variables, elles ne peuvent pas déceler certaines relations déterminantes ; il leur arrive d'accorder trop d'importance aux hypothèses de base (presque toujours implicites).

En revanche, les techniques d'analyse des données sont tout à fait adaptées pour répondre à ces objectifs et deviennent des passages quasi obligés pour exploiter efficacement des systèmes d'information complexes tels que ceux produits par les observatoires.

Elles imposent une réflexion sur la nature des données (unités d'observation, unités d'analyse, population statistique, jeu de variables identiques ou non d'une période à l'autre), un mode de constitution du tableau de données, un mode de traitement, une façon d'interpréter les résultats, un retour sur la donnée (*cf.* annexes). Bref, l'analyse des données impose une réflexion sur la structure de l'information pour appréhender la mise en place de l'observatoire et l'exploitation de ses données.

Basées sur les méthodes factorielles et de classification, elles prennent encore plus d'intérêt dans des techniques dérivées reposant sur la construction du tableau de données adéquat tenant compte des effets temporels et d'échelles.

Ainsi, les problèmes méthodologiques posés par la structure complexe d'un observatoire peuvent être en partie résolus par les techniques d'analyse des données (*cf.* annexes), à savoir :

- *organiser la base d'information* : il est nécessaire, dans un premier temps, de disposer, période par période, d'une vue d'ensemble du recueil des données trop nombreuses et trop dépendantes les unes des autres pour être interprétables en première lecture ;
- *analyser l'évolution et le changement* : les éléments d'une évolution apparaîtront dans l'analyse de jeux de variables identiques quelque soit l'échantillon considéré et les éléments du changement seront mis en évidence par les jeux de variables différents entre les périodes portant sur les individus de l'échantillon de suivi ;
- *relier des échelles différentes* : il s'agira de mettre en relation des informations provenant de sources ou de niveaux d'observation différents ce qui revient à faire l'analyse de changements d'échelles.

Conclusion

Ces méthodes répondent bien à une problématique de description des comportements économiques et sociaux des individus. Elles fournissent des facteurs et des typologies qui pourront décrire des modifications de comportements d'individus ou de groupes d'individus, des processus d'évolution. Elles permettent de saisir ainsi la manière dont ces individus s'adaptent aux situations de crise et déterminent les effets des mesures prises à un niveau plus global.

Par ailleurs, elles mettent en évidence des variables ou des combinaisons de variables à partir desquelles pourront être construits des indicateurs pertinents dont il s'agira de tester la stabilité.

Bibliographie

- | | |
|--|--|
| <p>CHARMES (J.), 1992 —
<i>Les observatoires
comme alternative ou complément
aux enquêtes lourdes.</i>
Projet BEN/87/023.</p> <p>BLAIZEAU (D.), DUBOIS (J.-L.), 1989 —
<i>Connaître les conditions de vie
des ménages
dans les pays en développement.</i>
Ministère de la coopération
et du développement, 1, 2, 3.</p> <p>LEBART (L.), 1989 —
<i>Stratégies du traitement
des données d'enquête.</i>
<i>Revue de Modulad</i>, 3, 9 p.</p> | <p>PIRON (M.), 1990 —
<i>Structuration de l'information
à plusieurs niveaux
et analyse des données :
application aux données
d'une enquête socio-économique
à Ouagadougou (Burkina Faso).</i>
Thèse, univ. Paris-VI, 265 p.</p> <p>PIRON (M.), 1992 —
<i>Analyse statistique
d'un système d'échelles.</i>
Orstom, Réseau Adoc,
doc. de travail 4, 222 p.</p> |
|--|--|

■ Annexes : aspects techniques d'une démarche de l'analyse des données d'observatoires

Annexe 1 : identification des éléments de base

Il est indispensable, pour structurer la base d'information, pour la construire et pour l'exploiter, d'identifier et de différencier les éléments de la base d'informations et les éléments de base de l'analyse. Les premiers permettent de constituer le tableau d'information et les derniers permettent de constituer le tableau de données soumis à l'analyse.

Éléments de la base d'information

Par définition, la structure des données d'un observatoire est une succession d'événements, de données dans le temps. Pour exploiter cette base d'information, il faut la mettre, dans un premier temps, sous forme d'un tableau d'information à trois « dimensions » où l'on distingue :

- les *individus statistiques* qui correspondent aux *unités statistiques* d'observation sur lesquelles on relève des caractéristiques. Il s'agit des unités enquêtées, renouvelées ou non dans le dispositif ;
- les *variables* qui sont les *caractéristiques* de la population d'étude. Elles peuvent être regroupées par thème. On distinguera, pour le traitement, les variables qualitatives et les variables quantitatives ;
- le *temps* qui correspond aux *périodes* d'enquêtes portant sur les mêmes individus ou non et sur les mêmes variables ou non.

On a donc un triplet (*individus, variables, temps*) caractéristique de chaque donnée élémentaire : mais d'une part les individus ne sont pas tous les mêmes d'un passage à l'autre (l'échantillon est partiellement renouvelé) et d'autre part on dispose de variables à la fois identiques et différentes entre les passages ; par conséquent, les données de l'observatoire expérimental considéré se structurent de la manière suivante (il ne s'agit pas exactement d'un tableau ternaire) :

	1 ^{er} passage J ou J ₁	2 ^e passage J ou J ₂	3 ^e passage J ou J ₃	
I _{1,1}				J ₁ , J ₂ , J ₃ : ensemble des variables descriptives de chaque passage J ₁ = J ₂ = J ₃ = J : ensemble des variables I _{1,1} : individu ou ménage ayant fait l'objet du 1 ^{er} passage I _{1,2} : individu ou ménage ayant fait l'objet du 1 ^{er} et du 2 ^e passage I _{2,3} : individu ou ménage ayant fait l'objet du 2 ^e et du 3 ^e passage I _{3,3} : individu ou ménage ayant fait l'objet du 3 ^e passage
I _{1,2}				
I _{2,3}				
I _{3,3}				

(Ici, l'échantillon est renouvelé toutes les deux périodes). On envisagera le cas où l'on dispose des mêmes variables aux différentes dates ($J_1 = J_2 = \dots = J$) et celui où elles sont différentes (J_1, J_2). Les méthodes peuvent être différentes suivant si les variables sont identiques ou non.

De plus, la structure des données peut être en sorte que l'on ne considère pas un seul type d'unités d'observation mais plusieurs types d'unités emboîtés possédant chacun leurs propres caractéristiques.

Éléments de base de l'analyse

À partir des éléments de la base d'information, on identifie ensuite les éléments de base de l'analyse. Nous distinguons schématiquement trois notions fondamentales (les définitions s'affinent avec la complexité de la structure d'information) (BLAIZEAU, DUBOIS, 1989) :

– les *unités statistiques d'observation* qui font l'objet d'une description et dont l'ensemble constitue la *population d'étude*. Elles définissent le niveau d'observation. Si l'on exploite des variables qualitatives, il s'agit de ce que l'on compte : des individus ou des ménages, mais aussi des migrations, des changements de situation. Pour une même enquête, il peut y avoir plusieurs types d'unités d'observation et donc plusieurs populations d'étude potentielles ;

– les *unités statistiques d'analyse* qui définissent le niveau d'analyse c'est-à-dire l'échelle de perception du phénomène. Ce peut être des unités d'observation. Mais d'une manière générale, ce sont des entités non directement observables, recouvrant des unités d'observation et dont les caractéristiques sont déduites de celles des unités d'observation. Ces niveaux vont du plus fin aux plus englobants ;

– les *variables* où l'on distingue les *variables de base* qui correspondent aux caractéristiques directement observées et les *variables dérivées* (CHARMES, 1992) qui ne peuvent être directement appréhendés par le questionnaire. Ce sont des indicateurs définis au préalable à partir d'un petit nombre de variables de base (ratios, taux, combinaisons de variables).

Ce n'est qu'à partir de l'identification de ces éléments que l'on peut envisager la construction du tableau de données cohérent puis l'exploitation statistique des données.

Annexe 2 : approche des problèmes méthodologiques par l'analyse des données

Nous proposons une démarche pour exploiter des données d'observatoire à partir des méthodes d'analyses de données.

On se référera essentiellement au traitement de variables qualitatives.

Organiser la base d'information

D'une manière générale, les méthodes d'analyses de données permettent de :

- dégager les structures profondes de l'ensemble de données quel que soit le degré de complexité de la base ;
- piloter l'exploitation de données d'enquête en élargissant le champ des hypothèses : « Les hypothèses de base pourront souvent être critiquées, d'autres hypothèses pourront être suggérées » (LEBART, 1989) ;
- contrôler la qualité de l'enquête en validant : l'échantillon d'une manière empirique (ce qui n'est pas négligeable pour des enquêtes de suivi de population), le questionnaire (lacune dans le choix des variables), la codification (pertinence des codes des variables qualitatives et découpage en classes des variables quantitatives).

En ce qui concerne la présentation des résultats, il est souvent préférable de revenir aux techniques statistiques usuelles pour une lecture plus accessible. L'apport de ces méthodes multivariées aura pourtant été essentiel pour la formulation et la mise en place des résultats.

Ainsi, il est utile, avant même d'envisager l'exploitation des données prises dans le temps ou sur des sources ou niveaux d'observation différents, de procéder à une critique de la base d'information et à une identification de la population d'étude à chaque période.

Analyser l'évolution et le changement

Intéressons-nous à la particularité des enquêtes répétées dans le temps qui visent à caractériser des processus d'évolution et de changement.

À partir de quels critères définit-on un changement ? Y a-t-il modification du comportement des individus ? Si oui, y a-t-il stabilité de groupes d'individus ? Comment définir un profil d'évolution ? Dans quel sens va-t-il ?

Nous souhaitons ici montrer que les techniques d'analyse de données appliquées aux données des observatoires, en dégagant des tendances et des typologies, peuvent apporter des éléments de réponses à ces questions.

Suivant les caractéristiques de l'échantillon et des variables, les problématiques ne sont plus les mêmes et l'exploitation des données est envisagée différemment. Nous résumons dans le schéma ci-dessous les différentes situations (présentées pour deux passages d'enquête mais qui sont facilement transposables à plusieurs passages) :

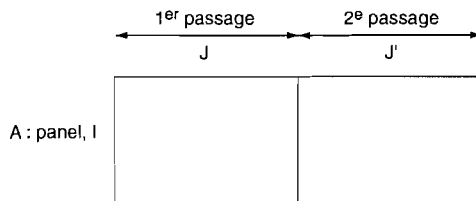
	Jeu de variables différents J, J'	Même jeu de variables J
Panel, I	A	D
Indépendant I, I'	B	C

Nous allons regarder indépendamment les quatre situations A, B, C et D. Mais on peut parfois considérer ensemble certaines de ces situations qui caractérisent l'observatoire étudié.

Les propositions suivantes sont des trames possibles pour aborder le problème de l'analyse de données dans le temps.

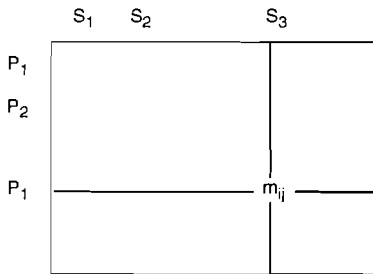
*L'échantillon est un panel
et les jeux de variables sont différents*

La situation A se ramène à une analyse d'enquête classique où nous avons au moins deux thèmes, le premier défini par les variables du premier passage et le deuxième par celles du deuxième passage.



Cependant, il faut pouvoir observer les modifications de comportement des individus intervenues entre les périodes et savoir si, à partir des typologies obtenues aux différentes périodes, il y a stabilité des groupes d'individus.

Une proposition est alors de constituer, à chaque date, des types d'individus (P_1, P_2, \dots, P_j) au premier passage, puis des types d'individus (S_1, S_2, \dots, S_j) au second passage et l'on construit ainsi une matrice d'échange :

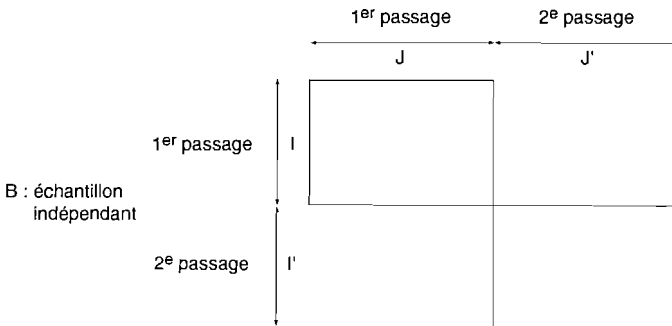


m_{ij} est le nombre d'individus appartenant à la fois au type P_i définie au premier passage et au type S_j définie au second passage. À partir de là, on peut connaître la répartition, au deuxième passage, des individus appartenant à un même groupe déterminé au premier passage.

Cette situation permet surtout de saisir les modifications du comportement de tel individu ou de tel groupe d'individus et de définir ainsi un changement.

*Les échantillons sont indépendants
et les jeux de variables sont différents*

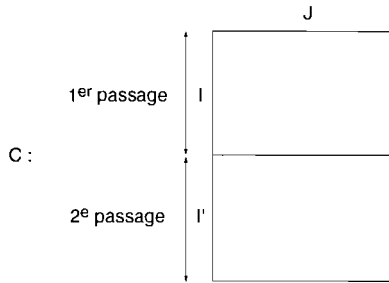
Cette situation qui correspond au schéma suivant, est *a priori* sans grand intérêt.



(On peut cependant intégrer ce cas de figure au précédent — panel et jeu de variables différents — par les techniques d'éléments supplémentaires ou d'analyse discriminante).

*Les échantillons sont indépendants
et les jeux de variables sont identiques*

On dispose du tableau de données suivant :

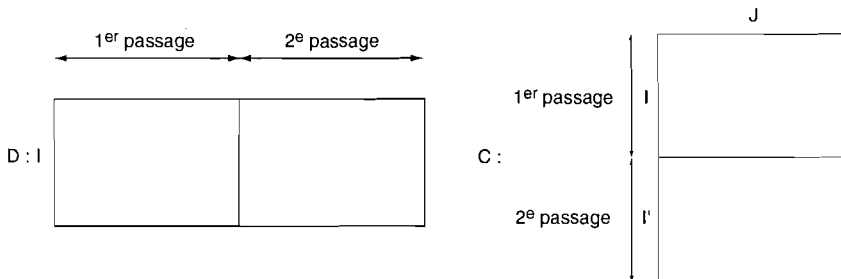


L'analyse d'un tel tableau se ramène à l'analyse d'enquête classique. On saisit l'évolution par une étude sur les variables. Cependant dans ce type d'échantillon, nous perdons la notion de suivi de la population et donc la notion d'individu.

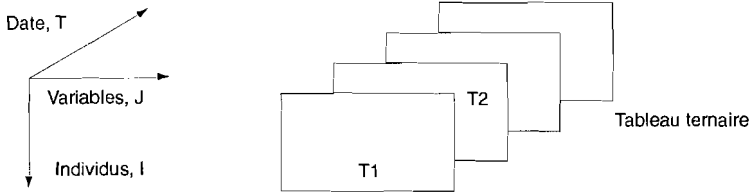
Cette situation permet surtout l'émergence d'une tendance globale de l'évolution et d'en saisir les caractéristiques.

*Il s'agit d'un panel
et les jeux de variables sont identiques*

Le tableau de données est construit de la manière suivante :



Dans ce cas, il s'agit d'un tableau ternaire classique (individus, variables, temps) que l'on représente de la manière suivante :



et qui permet d'effectuer le suivi de la population.

Cette situation permet surtout de saisir le sens et l'intensité de l'évolution.

Conclusion

La première situation (mêmes individus et jeux de variables différents) permet de saisir les modifications du comportement des individus, et les éléments d'un changement.

Les deux dernières situations (même jeu de variables) offrent l'avantage de saisir les éléments significatifs d'une évolution surtout si l'on dispose de plusieurs périodes. Cependant, entre deux passages, les caractéristiques se modifient peu et une évolution est difficilement perceptible. Il est alors possible, par des techniques d'analyses factorielles (analyse intra-classes) appliquées à un même jeu de variables, de saisir les évolutions minimales mais intéressantes. On peut ainsi mettre en évidence les modifications qui sont intervenues entre les différents passages sur des groupes d'individus.

Relier les différentes échelles de perception

On envisage deux situations quasi similaires dans l'approche méthodologique. Elles consistent à mettre en relation des informations provenant de sources ou niveaux d'observation différents ce qui revient à effectuer des changements d'échelles (PIRON, 1992) au moment de l'exploitation.

Mise en relation de niveaux d'observation emboîtés

Comment croiser des groupes de variables relatifs à des niveaux d'observation différents ?

Deux choix sont à faire : l'un sur la population statistique et l'autre sur les niveaux d'analyse. En effet, l'existence de plusieurs types d'unités d'observation entraîne l'existence de plusieurs populations statistiques potentielles et induisent diverses unités d'analyse.

Par un système de pondération cohérent, il est possible de déterminer une population statistique de référence. Puis, en identifiant la nature de la relation d'inclusion entre le niveau de définition des variables et le niveau d'analyse (relations d'affectation ou d'agrégation), on peut envisager de croiser des groupes de variables issues de niveaux d'observation différents (PIRON, 1992).

Mise en relation de sources diverses

On cherche, ici, à mettre en relation des informations provenant de sources différentes. Elles peuvent porter soit sur les mêmes entités (deux enquêtes ménages par exemple), soit sur des entités différentes emboîtées (ménages et quartier par exemple) ou non (quartiers et zones d'aménagements par exemple).

Ce cas de figure est une généralisation du premier (niveaux d'observation emboîtés). On peut alors relier ces informations si l'on dispose cependant d'éléments de liaison qu'il faut identifier :

- si les informations à mettre en relation appartiennent à un même type d'unités statistiques (le ménage), les éléments de liaison seront des groupes de variables identiques qui permettront ainsi de constituer des typologies de référence ;
- si les informations sont issues de niveaux d'analyse différents et emboîtés (ménage et quartier), il faudra affecter à chaque unité statistique du niveau le plus fin (le ménage) un identificateur le reliant à une unité statistique du niveau plus global (le quartier) ;
- si les niveaux d'analyse ne sont pas emboîtés (quartier et zone d'aménagement), leur mise en relation ne pourra se faire qu'à un niveau d'analyse commun et donc plus global (le secteur).

On peut sur le même principe envisager de comparer des données d'observatoires semblables relatifs à des sites différents.

Bien évidemment si, techniquement, il est possible de relier différentes sources d'information à la condition de disposer d'éléments de liaison, d'autres conditions spécifiques à l'étude devront impérativement être respectées (par exemple même définition du ménage).