

NOUVELLES APPROCHES MÉTHODOLOGIQUES EN SCIENCES SOCIALES

Une vue d'ensemble

Daniel COURGEAU*

Si, dans le passé, les sciences sociales se sont constituées indépendamment les unes des autres, avec leur propre objet d'étude et leurs méthodes particulières, on observe actuellement une convergence croissante de leurs approches et une extension de leur domaine. Ainsi les méthodes issues du même type de modèle statistique, sont aussi bien utilisées en biométrie, en démographie, en économie, en géographie humaine, en sciences de l'éducation, en sociologie, etc. Bien sûr, ce n'est pas un hasard si ces modèles s'adaptent correctement aux divers champs, mais c'est parce qu'ils reflètent de mieux en mieux la complexité des phénomènes sociaux, qui y sont à l'œuvre.

Par exemple, l'analyse des biographies est issue des travaux de probabilistes et de statisticiens, qui ont réussi à formuler correctement le déroulement au cours du temps de processus très généraux, avec l'effet de diverses caractéristiques, pouvant, elles-mêmes, varier selon la durée (Cox, 1972; Kalbfleisch et Prentice, 1980; Cox et Oakes, 1984; Andersen *et al.*, 1993). Ainsi des phénomènes sociaux aussi différents que les changements d'emploi, les migrations, le mariage, la poursuite d'études, les conflits ethniques, le passage d'un régime politique à un autre, etc., peuvent être abordés avec ces méthodes. Il en est résulté une généralisation et une adaptation de cette approche aux problèmes posés par la sociologie (Coleman, 1980; Tuma et Hannan, 1984; Blossfeld *et al.*, 1989), par l'économétrie (Lancaster, 1990), par la démographie (Courgeau et Lelièvre, 1989; 1992) et par bien d'autres sciences sociales.

Bien plus, une telle innovation peut remettre en question la validité de certains paradigmes. Ainsi, le paradigme classique en démographie permettait d'analyser un phénomène et un seul, considéré comme indépendant

* Institut national d'études démographiques.

des autres, dans une sous-population qui devait rester en permanence homogène (Henry, 1959). Ce paradigme, utilisable tant que l'on analysait des données simples (état civil, par exemple), était de moins en moins tenable avec le développement d'enquêtes détaillées, telles que l'enquête mondiale de fécondité. Il n'était plus possible de laisser de côté la dépendance entre phénomènes et l'hétérogénéité des populations. L'analyse des biographies permet, dans ce cas, de s'affranchir de ces contraintes, car elle considère qu'«un individu parcourt, tout au long de sa vie, une trajectoire complexe, qui dépend à un instant donné, de sa trajectoire antérieure et des informations qu'il a pu acquérir dans son passé» (Courgeau et Lelièvre, 1996). Ce nouveau paradigme ouvre la voie à l'analyse des interactions entre divers événements, survenus dans une existence. Il permet également de saisir les différences de comportements individuels qui interviennent dans une population hétérogène.

Il n'est pas dans notre propos d'aller plus avant dans la présentation de cette approche, maintenant parfaitement intégrée dans les diverses sciences sociales (Keilman, 1993), mais de présenter, dans ce numéro spécial, de nouvelles avancées méthodologiques, dont l'avenir nous dira si elles tiennent leurs promesses.

De l'individu au groupe

En premier lieu, l'émergence de comportements propres à des groupes réunissant des entités semblables, est un problème récurrent et important en sciences sociales. Il s'agit ici de dépasser les deux approches opposées qui, pour l'une, considère que le comportement agrégé est la somme ou la moyenne des comportements individuels et qu'il n'y a pas lieu d'introduire un effet propre des groupes, et pour l'autre, considère que l'effet de groupe est prépondérant et que les comportements individuels sont les manifestations, sans intérêt propre pour les sciences sociales, des causes globales sous-jacentes à étudier (Alexander *et al.*, 1987). Il faut dès lors, montrer l'existence de comportements agrégés et les relier à ceux des membres du groupe. Il faut chercher à déceler quels processus conduisent à l'apparition de ces agrégats, quels types de frontière les délimitent, comment les interactions entre membres du groupe amènent des comportements qui transcendent ceux des individus (Holland, 1995).

Deux articles abordent ici ces problèmes avec des approches différentes mais non opposées, car elles sont liées à la taille des groupes observés.

E. Lelièvre, C. Bonvalet et X. Bry envisagent des groupes de taille restreinte (famille, ménage ou entourage). Dans ce cas l'analyse peut faire intervenir de façon détaillée la structure du groupe et les interactions entre ses différents membres. Si une famille ou un ménage peuvent être bien définis à un moment donné, en transversal, ces notions ne sont plus guère

utiles dès que l'on envisage une étude longitudinale ou biographique (Keilman et Keyfitz, 1985). Il est, dans ce cas, préférable *d'envisager ce que les auteurs appellent l'entourage*. Pour le définir, on va suivre un individu clef tout au long de sa vie. Son entourage est constitué des membres des différents ménages auxquels il a appartenu et d'individus de sa famille ou de son réseau de relations, avec lesquels il n'a pas cohabité, mais entretient des relations étroites. On pourra alors suivre cet entourage, à partir du premier contact de chacun de ses membres avec l'individu clef. Des enquêtes originales doivent dès lors être mises en place, pour suivre correctement cet entourage. Il faudra ensuite modéliser son comportement, en faisant des hypothèses sur les liens entre ses divers membres. Les modèles biographiques à mettre en place seront alors bien plus complexes que les modèles habituels, mais devraient permettre une meilleure compréhension des comportements, non plus des individus isolés mais de groupes tels que l'entourage.

Une autre approche est suivie par D. Courgeau et B. Baccaïni, qui envisagent des groupes moins structurés (élèves d'une classe, par exemple) ou des groupes de taille beaucoup plus importante, tels que les habitants d'une région, sur lesquels il n'est guère possible d'introduire une structure. Cela conduit à la mise en place de modèles, qui ne personnalisent pas les membres d'un niveau d'agrégation donné, mais qui permettent de considérer de nombreux niveaux correspondant soit à des hiérarchies, soit à des classifications croisées. *Les modèles multi-niveaux* permettent de répondre à ces hypothèses. Les individus y ont à la fois des comportements qui dépendent de leurs propres caractéristiques, des caractéristiques agrégées des diverses zones dans lesquelles ils vivent et d'aléas correspondant à ces zones. Ces modèles cependant, posent encore des problèmes liés à l'hétérogénéité non observée, même lorsque celle-ci est indépendante des caractéristiques observées, et à la mise en place d'une analyse biographique qui soit pleinement multi-niveaux. En revanche, ils peuvent être utilisés dans de très nombreuses sciences sociales, et permettent de répondre à des questions importantes (Goldstein, 1995).

Des événements et des mots pour les dire

Un domaine en pleine extension, *l'analyse de données textuelles*, trouve de nombreuses applications en sciences sociales et ouvre des voies nouvelles pour l'analyse de trajectoires biographiques complexes. Ces méthodes, initialement développées à des fins stylistiques (Lebart et Salem, 1994), se révèlent en fait avoir des rapports étroits avec l'analyse biographique : elles envisagent, en effet, des trajectoires constituées de suites de mots ou d'étapes, ordonnées dans le temps, comme des biographies. Mais ces trajectoires sont ici beaucoup plus complexes, par la diversité des mots d'une langue, par la grande variété des suites envisagées. Il s'agit donc

d'abord d'une analyse essentiellement descriptive, qui cherche à retrouver la trace des états successifs investis par l'individu. Elle permet cependant, dans un deuxième temps, de mieux comprendre la trajectoire suivie, en introduisant diverses caractéristiques des répondants et en classant les trajectoires suivant diverses typologies.

F. Guérin-Pace présente ici ces méthodes d'analyse, en indiquant à nouveau combien elles peuvent s'appliquer à des domaines très divers des sciences sociales. Bien entendu, au départ, elles n'étaient pas du tout orientées vers la démographie. Ce n'est que récemment que trois grands domaines d'application se sont ouverts aux sciences sociales : l'analyse des réponses à des questions ouvertes, l'analyse de récits ou de discours et surtout l'analyse d'itinéraires complexes. Les significations données au mot « environnement » dans l'enquête *Population, espace de vie et environnement* de l'Ined, éclairent les différences d'interprétation données à un concept, apparemment simple, mais que chacun habille d'un vêtement différent. L'introduction des diverses caractéristiques des répondants permet de relier cet habillage aux différents profils individuels. L'étude de récits, par des méthodes de classification en particulier, met en évidence et analyse le cheminement du discours. Enfin, le suivi d'itinéraires professionnels compliqués, à partir des données de l'*Enquête sur l'emploi* de l'Insee, montre comment une analyse textuelle peut prendre le relais d'une analyse biographique, lorsque le nombre de trajectoires types observées devient très élevé (près de 300 dans le cas étudié). Elle permet cependant une approche plus descriptive qu'explicative, contrairement à l'analyse biographique.

Modèles d'évolution des populations

Ces extensions de l'analyse biographique touchent également le domaine des projections de population. E. Van Imhoff et W. Post nous présentent ici l'utilisation de *méthodes de micro-simulation* dans ce domaine. Ces méthodes permettent d'utiliser avec une très grande efficacité les résultats d'une analyse biographique, en introduisant les probabilités de connaître les divers événements selon l'état dans lequel se trouve l'individu et les durées écoulées depuis d'autres événements d'intérêt : par exemple, fécondité en fonction de l'âge, de la durée de mariage et de la parité de la naissance. Le processus simulé est alors aléatoire et permet, non seulement d'estimer les valeurs moyennes des diverses caractéristiques des populations au cours du temps, mais donne aussi une indication de leurs variations autour de ces valeurs moyennes. À nouveau, ces simulations peuvent conduire à des modèles plus complets que de simples projections démographiques, en introduisant un large spectre d'événements socio-économiques. Cependant, ces projections restent toujours dans le cadre classique des modèles linéaires.

Voyons donc maintenant, plus en détail, comment ces modèles linéaires ont été formulés dans le passé, en travaillant à un niveau plus agrégé,

tel qu'un pays ou un pays divisé en régions. Si l'on suppose que les comportements démographiques sont représentés par des quotients perspectifs par âge, supposés invariants au cours du temps, il est possible en multipliant ces quotients par les populations initiales correspondantes, de modéliser de façon linéaire l'évolution de leurs effectifs. Ainsi, en faisant intervenir la fécondité et la mortalité par âge, on met en place des modèles qui aboutissent à des populations stables ou stationnaires (Keyfitz, 1968). Si l'on introduit en plus l'émigration par âge des diverses régions vers les autres, on aboutira toujours à des populations stables, tous âges réunis d'abord, pour chaque groupe d'âges par la suite (Rogers, 1975). Tous les autres apports à ces modèles, leur conservant une forme linéaire, conduisent toujours à des solutions d'équilibre stables ou stationnaires.

On peut dès lors se demander si de tels modèles sont réalistes et s'il n'est pas nécessaire d'y introduire des non-linéarités. Ainsi, par exemple, le système proie-prédateur n'est correctement modélisé, qu'en faisant intervenir un terme d'interaction, correspondant au produit des populations en compétition à chaque instant. De même, si l'on fait intervenir, dans un modèle multirégional, un indice d'intensité migratoire, rapport du flux entre deux régions au produit de leurs populations respectives, on obtient à nouveau un système non linéaire, plus proche des observations.

D. Blanchet présente ici les nouvelles propriétés introduites par de tels *modèles non linéaires* et discute de l'opportunité de certaines d'entre elles en démo-économie.

En premier lieu, ces modèles ne conduisent généralement plus à des populations stables. Les populations peuvent varier de façon cyclique au cours du temps sans jamais atteindre un état d'équilibre. Certaines de ces populations peuvent s'éteindre, d'autres varier de façon acyclique. Mais le plus intéressant est que dans certains cas cette dynamique peut devenir chaotique, sans aucune composante cyclique. La trajectoire de ces populations, au cours du temps, devient alors difficile à prévoir, car de légers écarts sur les états initiaux peuvent aboutir à des évolutions entièrement différentes. Si cette propriété s'avérait correctement vérifiée par les populations humaines, cela entraînerait une imprévisibilité de certains phénomènes sociaux, en rupture avec les dynamiques classiques (Keilman, 1993). D. Blanchet apporte ici des éléments de réflexion critique sur cet important problème. Notons, cependant, que la complexité du comportement est telle que celui-ci semble chaotique, sans que cette apparence puisse être attribuée à des fluctuations erratiques des paramètres du processus, puisque le modèle lui-même est déterministe.

Il nous faut donc maintenant introduire l'aléatoire dans ces modèles, pour en voir les conséquences. *La théorie de la viabilité* (Aubin, 1990) permet de franchir ce pas, en introduisant l'absence de déterminisme dans l'évolution de macro-systèmes non linéaires. Elle suppose de plus la présence de contraintes auxquelles le système doit en permanence obéir, sous peine de s'éteindre. Ce système doit dès lors disposer de contrôles sur les-

quels il peut jouer, pour se maintenir. Cette théorie ne permet donc plus de prévoir l'évolution d'un système initial, mais donne les limites à partir desquelles il est nécessaire de changer certains comportements (les contrôles) pour maintenir le système viable. La grande généralité de ces principes permet d'appliquer cette théorie à de nombreuses sciences sociales : anthropologie, démographie, économie, sociologie, etc.

N. Bonneuil présente ici les principes de base de cette théorie et ses principaux acquis, tout en l'appliquant de façon très éclairante à des problèmes, aussi bien sociaux qu'économiques, qui se posaient à des populations observées par l'anthropologue R. Barth. Une structure sociale soumise à des contraintes pour son existence, telles que celle des nomades persans Basseri ou des pêcheurs norvégiens, considérés ici, peut arriver à se maintenir, non pas en équilibre, mais sur une trajectoire viable, en utilisant un certain nombre de stratégies, qui constituent ses contrôles. Il en est ainsi de la sédentarisation pour les nomades, ou de la décision de partir en exploration hors des sites déjà exploités, pour les pêcheurs. Ces stratégies permettent le maintien au cours du temps de ces groupes sociaux. En revanche, une multitude de trajectoires peuvent être utilisées par ces groupes, du fait que cette théorie n'est pas déterministe, mais cela sous la seule condition qu'elles les maintiennent dans le noyau de viabilité du système. Il est ainsi possible d'extraire, d'une grande diversité de situations concrètes, les régularités qui sous-tendent l'existence de ces groupes sociaux, sans les déterminer. L'absence de déterminisme au niveau macro, dans la théorie de la viabilité, vient répondre à l'absence de déterminisme au niveau individuel dans l'analyse des biographies.

Espace, réseaux, diffusion

Il est nécessaire, pour compléter ce numéro spécial, d'introduire maintenant de façon explicite, l'espace dans lequel les phénomènes sociaux se produisent. Il était déjà intervenu, de façon indirecte, dans les analyses multi-niveaux où les niveaux d'agrégation pouvaient être constitués de communes, de régions ou de villes. Il faut maintenant l'introduire explicitement, avec ses structures et sa morphologie propre, engendrées par les sociétés et les réseaux sociaux qui l'habitent, avec ses hiérarchies et ses ruptures.

Il faut bien voir ici, qu'en dehors de la géographie humaine et de l'épidémiologie, les autres sciences sociales ont longtemps considéré l'espace comme un élément secondaire et sans intérêt dans l'étude des sociétés. Les démographes, les économistes et les sociologues, s'ils ont parfaitement intégré la dimension temporelle, raisonnent le plus souvent sans faire intervenir l'espace. P. Gould (1993) a récemment montré que même l'épidémiologie pouvait revenir à une approche uniquement temporelle de la

diffusion du SIDA, laissant de côté ses aspects spatiaux, ce qui constitue à son avis une erreur fondamentale. Il paraît dès lors urgent de reconsidérer l'effet de l'espace, non seulement pour des phénomènes qui l'impliquent directement comme les migrations internes et internationales, la diffusion des épidémies, mais pour l'ensemble des phénomènes sociaux.

Ainsi, J.-P. Bocquet-Appel et L. Jakobi s'interrogent ici sur *la diffusion de la transition démographique*, que la plupart des auteurs ont considérée comme essentiellement sociale, allant des élites vers les travailleurs ruraux non qualifiés. Cependant, si cette diffusion s'était produite de proche en proche dans l'espace européen, il devrait être possible de suivre son cheminement sur des cartes géographiques. Le projet de Princeton, qui a recueilli les données à un niveau géographique fin, permet cette analyse. Celle-ci n'est cependant pas si évidente qu'il y paraît, car il faut déjouer de nombreux pièges que posent de telles données spatiales, dont on veut suivre le déplacement au cours du temps. Les auteurs présentent des méthodes issues, pour une part de la géostatistique, et d'autre part, de l'épidémiologie qui permettent de traiter correctement ce problème. Ils vérifient que la contraception semble bien s'être diffusée comme un processus de contagion en Grande-Bretagne, bien que cette diffusion se soit déroulée sur un temps beaucoup plus court qu'en France. Il reste à tester cette hypothèse au niveau de l'Europe toute entière et à faire intervenir les caractéristiques qui permettent de l'expliquer, sur un si vaste espace. Ce projet, qui remet la friction de l'espace au cœur de la diffusion des processus sociaux, montre clairement le danger d'écarter l'espace de l'étude des sociétés.

Il nous faut donc maintenant aborder les problèmes liés à la répartition des hommes sur un territoire donné. Cette répartition suit des règles dont il faut déterminer les origines et les formes prises selon les cultures. Ainsi les villes sont des structures très intéressantes à étudier. En dépit d'un flux continu et important d'hommes, d'entreprises, d'administrations, et bien qu'elles soient soumises à des changements permanents dans leurs bâtiments, dans leurs voies et moyens de transport et dans leur superficie, de telles structures peuvent persister pendant de très longues durées (Derycke *et al.* 1996). Qui plus est, bien que construites et habitées par des civilisations et des cultures très différentes de par le monde, elles constituent une forme d'organisation géographique universellement répandue à la surface de la terre.

P. Frankhauser s'interroge ici sur cette forme et sur les conditions de sa persistance au cours du temps. Il paraît nécessaire de dépasser l'analogie entre les systèmes urbains et les systèmes physiques, qui a longtemps prévalu, avec l'utilisation de modèles gravitaires ou de modèles de maximisation de l'entropie (Wilson, 1967). Ces modèles trop déterministes ont maintenant fait place à de nouveaux modèles, qui mettent en relief la façon dont les décisions locales, peu planifiées et non coordonnées entre elles, peuvent faire apparaître des structures globales ordonnées, définissant la taille et la forme des villes observées. Ces villes apparaissent alors comme des *structures auto-organisées* qui émergent d'actions locales très diverses.

Cela conduit à des *formations fractales*, dont l'évolution au cours du temps peut être modélisée. Cette modélisation permet alors de comprendre, aussi bien la formation de villes à centre unique de la révolution industrielle, que la formation d'agglomérations urbaines multi-centrées qui existent actuellement. On peut alors, en bouclant la boucle, revenir aux problèmes de l'agrégation posés en début de ce numéro : est-il possible qu'une structure agrégée, comme l'est une ville, puisse être considérée indépendamment des individus qui la composent et lui donnent vie ? L'extension des modèles urbains présentés ici, devrait permettre d'y répondre.

*

* *

Ce tour d'horizon rapide des nouvelles approches en sciences sociales, ne prétend pas être exhaustif. Il montre cependant l'extrême vitalité de la recherche dans ce domaine et le foisonnement des méthodes mises en place. Ces approches, cependant, ne se développent pas de manière aléatoire et nous avons déjà essayé, dans cette présentation, de montrer les liens qui existent entre elles. Nous avons ainsi dégagé un certain nombre d'hypothèses de base, qu'il nous semble maintenant utile de présenter.

En premier lieu, les individus ou les groupes sur lesquels on travaille sont *hétérogènes*, tant par leur histoire passée, que par leurs possibilités d'action. Cette diversité se retrouve également dans les trajectoires individuelles, dans les suites de mots d'un texte, etc.

En second lieu, l'action de ces individus ou de ces groupes ne peut se comprendre que comme une *interaction* entre les événements connus dans le passé ou connus par d'autres membres du groupe et les événements que l'on veut étudier.

En troisième lieu, il est nécessaire de considérer les effets d'*agrégation*, qui conduisent à des comportements qui ne sont pas la somme des comportements des unités agrégées. Si l'appartenance d'un individu à un groupe peut avoir des effets sur ses propres comportements, on peut se demander si le groupe en tant que tel, n'acquiert pas une certaine indépendance par rapport à ses membres.

En quatrième lieu, il importe de considérer des comportements de groupes qui conduisent à des modèles *non-linéaires* ou *non-déterministes*, ou les deux à la fois. Certains de ces modèles peuvent conduire à des comportements chaotiques, dont il faut vérifier la validité. Il paraît également utile de chercher à mieux relier le comportement de ces groupes à ceux des individus qui les composent.

En cinquième lieu, il est nécessaire d'introduire, tant dans les modèles individuels que dans les modèles agrégés, des *flux* de toutes sortes, qui se produisent dans l'espace physique ou dans un espace plus complexe.

Il reste maintenant à poursuivre la recherche dans ces différentes voies ou dans d'autres, non spécifiées ici, pour mettre en évidence les hypothèses qui seront assez solides pour résister à l'épreuve des faits et celles qu'il faudra abandonner, car insuffisamment établies et contredites par certaines observations. Ce numéro spécial aura permis de montrer l'important renouveau qui a pris place dans les sciences sociales au cours des dernières années, l'avenir de la recherche permettra de séparer le bon grain de l'ivraie.

Daniel COURGEAU

BIBLIOGRAPHIE

- ALEXANDER J., GIESEN B., MÜNCH R., SMELSER N., (1987), *The micro-macro link*, University of California Press, Berkeley.
- ANDERSEN P., BORGAN O., GILL R., KEIDING N., (1993), *Statistical models based on counting processes*, Springer-Verlag, New York.
- AUBIN J.-P., (1990), « A survey of viability theory », *SIAM Journal of Control and Optimization*, vol. 28, n° 4, pp. 749-788.
- BLOSSFELD H.-P., HAMERLE A., MAYER K.-U., (1989), *Event history analysis*, Lawrence Erlbaum Associates, Hillsdale NJ, Londres.
- COURGEAU D., LELIÈVRE É., (1989), *Analyse démographique des biographies*, Édition de l'Ined, Paris.
- COURGEAU D., LELIÈVRE É., (1992), *Event history analysis in demography*, Clarendon Press, Oxford.
- COURGEAU D., LELIÈVRE É., (1996), « Changement de paradigme en démographie », *Population*, 3, pp. 645-654.
- COX D.-R., (1972), « Regression models and life tables », *Journal of the Royal Statistical Society*, B.34, pp. 187-220.
- COX D.-R., OAKES D., (1984), *Analysis of survival data*, Chapman and Hall, London.
- DERYCKE P.-H., HURIOT J.-M., PUMAIN D., (1996), *Penser la ville. Théories et modèles*, Anthropos, Éditions Economica, Paris.
- GOLDSTEIN H., (1995), *Multilevel Statistical models*, Edward Arnold, London.
- GOULD P., (1993), *The slow plague*, Backwell Publishers, Cambridge, Massachusetts.
- HENRY L., (1959), « D'un problème fondamental de l'analyse démographique », *Population*, 1, pp. 9-32.
- HOLLAND J., (1995), *Hidden order : how adaptation builds complexity*, Addison Wesley Publishing Company, Reading, Massachusetts.
- KALBFLEISCH J.-D., PRENTICE R.-L., (1980), *The Statistical analysis of failure time data*, John Wiley and sons, New York.
- KEILMAN N., (1993), « Emerging issues in demographic methodology », *European population, II. Demographic dynamics*, Congresses and Colloquia, John Libbey/Ined, Paris.
- KEILMAN N., KEYFITZ N., (1988), « Recurrent issues in dynamic household modelling », *Modelling household formation and dissolution*, Clarendon Press, Oxford.
- KEYFITZ N., (1968), *Introduction to the mathematics of population*, Addison Wesley, Reading, Massachusetts.
- LANCASTER T., (1990), *The econometric analysis of transition data*, Cambridge University Press, Cambridge.
- LEBART L., SALEM A., (1994), *Statistique textuelle*, Dunod, Paris.
- ROGERS A., (1975), *Introduction to multi-regional mathematical demography*, John Wiley and sons, New York.
- TUMA N., HANNAN M., (1984), *Social dynamics. Models and methods*, Academic Press, Orlando, Florida.
- WILSON A.-G., (1967), « A statistical theory of spatial distribution models », *Transportation research*, 1, pp. 253-269.