

ATHIS STATISTIQUES. CONCLUSIONS

Benjamin Deruelle et Alain Dallo

En introduction Jean-Philippe Genet traçait un bilan en demi-teinte des relations entre recherche historique et statistiques. Tout en notant un recul des méthodes quantitatives chez les historiens, il relevait un frémissement récent indiquant une reprise. Il met ainsi en lumière un paradoxe fondamental des relations entre l'historien et l'informatique dans ce début de XXI^e siècle. Alors que les méthodes d'exploration et de modélisation sont de plus en plus nombreuses à être à la disposition des historiens ; que l'offre de formation en direction des étudiants et des enseignants est de plus en plus importante et de plus en plus performante ; et les machines et logiciels de plus en plus simples à utiliser ; il semble qu'il y ait une raréfaction de l'usage des traitements statistiques dans les travaux d'historiens ?

MÉTHODES EXPLORATOIRES

Un premier aspect de notre rencontre nous a permis de renouer avec des pratiques qui ont fait leurs preuves pour l'observation et la structuration des données. Vincent Loonis a montré par le biais de l'étude de la toponymie que l'usage de la statistique descriptive et des probabilités apportait toujours des résultats. Séverine Lepape a prouvé une nouvelle fois au travers de l'exemple de l'arbre de Jessé que les tableaux croisés, l'AFC (Analyse Factorielle de correspondances), l'ACM (Analyse en composantes multiples) et les PEM (pourcentage de l'écart maximum) permettaient de décrire et de croiser les variables, d'une base de données et de proposer des modèles. Brigitte Leroux a présenté l'intérêt des ellipses de concentration pour compléter l'interprétation des AFC.

Des méthodes novatrices pour l'historien, basées sur la classification des données et la production de types, apparaissent aussi pleines de promesses. L'algorithme de Kohonen, utilisé par Marie Cottrell et Patrice Gaubert, fournit des méthodes de classification et d'extraction d'individus représentatifs. Patrice Gaubert adjoint à l'algorithme de Kohonen, celui de Markov et les tests de ruptures pour mesurer les changements de régime des trois étalons monétaires du XIX^e. Laurent Lesnard, quant à lui, use des méthodes d'appariement optimal pour obtenir des types de journée de travail et analyser leurs rythmes et séquence. Sont présentés aussi de nombreuses autres méthodes de visualisation graphique des données : grilles, arbres, réseaux sociaux, cartes de migrations, et des calculs d'indice comme celui de la similarité pour la comparaison en (anthroponomie ?) historique présenté par Pascal Chareille.

Toutes ces pratiques rejoignent la formalisation chère à Alain Guerreau. Ce qu'ont traduit les différentes communications et nos travaux respectifs est un retour aux structures, aux classifications, et aux typologies. L'appropriation de ces méthodes et le renouvellement de la bibliographie ne semblent-ils pas indiquer l'émergence d'un « néo-structuralisme » ?

MODÉLISATION

Comme François Djindjian nous le montrait, il n'y pas de rupture radicale entre statistique descriptive et analyse de données d'une part et modélisation et simulation d'autre part. Ces dernières ayant pour objectif « la reconstitution des sociétés ». Le travail de Patrice Gaubert intègre la notion de modèle (modèle à changement de régimes), celui de Guillaume Daudin fait appel à la régression. Nous retrouvons celle-ci dans la présentation d'Anne-Sophie Bruno avec trois modèles de régression multi-

niveaux pour l'analyse des salaires ; ainsi que dans celle de Claire Lemerrier qui utilise la régression logistique et les réseaux afin d'étudier les migrations et les frontières linguistiques. Philippe Cibois, quant à lui, s'attache au lien entre catégories socioprofessionnelles des parents, position idéologique et choix d'une langue morte (grec ou latin) ou non pour leurs enfants au collège. Il met à contribution régression linéaire, logistique et analyse tabulaire multivariée, dans une démarche très pédagogique.

Noël Bonneuil nous présente une approche totalement différente de tout ce qui précède. En reprenant les travaux de Poincaré pour définir des états et des phases, il les applique à la démographie historique. Les cycles sont remplacés par des phases faisant apparaître une véritable circulation entre celles-ci. Le facteur temps prend ici une dimension complètement nouvelle nous permettant d'envisager toutes les trajectoires possibles entre elles.

Exploration et modélisation ne semblent pas impliquer la nécessité de disposer d'un grand nombre de données. Il est possible de faire un traitement avec des données absentes ou mal définies. Claire Lemerrier insiste sur la présence de lacunes et de nombreux biais, même avec des données très contemporaines. Parmi ces biais, elle souligne la différence d'enregistrement et de perception de la profession ainsi que de la catégorie sociale en fonction du département ou de l'administrateur.

Beaucoup de ces « nouvelles méthodes » nous ont été présentées par des non historiens : des sociologues, Philippe Cibois et Laurent Lesnard, des économètres, Marie Cottrell et Patrice Gaubert, un informaticien, Jean-Daniel Fekete, des mathématiciens, Brigitte Leroux, et Noël Bonneuil. Les historiens doivent-ils se les approprier ? Faut-il connaître et utiliser l'ensemble des méthodes sur « un jeu de données », en risquant un « effet boîte noire » par la méconnaissance des concepts mathématiques ? Ou au contraire restreindre les analyses à celles qui sont maîtrisées ?

ASPECTS PÉDAGOGIQUES : « UN ESSOUFLEMENT » ?

Aujourd'hui si certains historiens s'emparent de nombreuses méthodes statistiques, imitant en cela les sciences dures ; c'est le doute qui domine la pédagogie. De nombreuses questions se posent concernant la diffusion de ces méthodes auprès des chercheurs, des enseignants et des étudiants.

À l'échelle des institutions et des chercheurs, un premier constat peut apparaître. C'est celui de la différence, parfois extrêmement marquée, entre les institutions d'enseignement et de recherche. Les communications qui ont été faites ont montré l'étendue des situations que l'on pouvait rencontrer : du chercheur isolé à l'équipe de recherche. L'intervention de Jean-Daniel Fekete et de Nicole Dufournaud représentant certainement le plus manifestement les avantages incomparables d'une collaboration entre l'historien et le statisticien. Cette réalité est certainement imputable, d'abord, à des réalités financières, qui manifestent l'existence de pôles d'excellences au sein desquels sont réunis historiens, informaticiens, mathématiciens... En dehors de ces institutions, les initiatives sont le plus souvent liées aux compétences d'un chercheur au sein d'une équipe ou d'un laboratoire. Pourtant, ils semblent que les mentalités aient également leur part de responsabilité. C'est sur les pratiques même de l'historien que nous invite à réfléchir la communication de Serge Heiden. Elle expose la nécessité d'une collaboration non seulement entre les universités, les disciplines, les compétences, et les logiciels mais aussi d'un partage des sources dont les historiens sont si jaloux.

En tant qu'enseignants, nous avons aussi une responsabilité face aux étudiants. Celle de nous tenir informés des outils proposés par les sciences dures qui nous entourent et qui peuvent offrir à l'analyse historique de nouvelles voies exploratoires. Dans ce domaine, il faut reconnaître notre incapacité, parfois notre incompetence, à maîtriser l'ensemble de ces méthodes. Cet état de fait nécessite pour chacun de nous, afin que l'interdisciplinarité puisse s'exprimer sans retenue, de participer au développement des réseaux de formateurs potentiels pour orienter au mieux les

étudiants. Les méthodes présentées par Marie Cottrell, Patrice Gaubert, Guillaume Daudin, et bien d'autres ouvrent des perspectives de recherche tout à fait novatrices en histoire.

Finalement, sommes-nous, comme se le demandait Jean-Philippe Genet en introduction, à la fin d'une « période d'essoufflement » ? Ou au temps du paradoxe soulevé par Alain Guerreau et François Djindjian où aux ordinateurs de plus en plus performants, aux logiciels de plus en plus accessibles, aux capacités de formation sans précédents, et aux publications de recherche nombreuses répond le manque d'intérêt des étudiants qui arrivent à l'université. Parallèlement, alors que les possibilités d'utilisation prolifèrent, l'usage heuristique de l'outil informatique paraît se raréfier chez un grand nombre de chercheurs. Il est aujourd'hui nécessaire de s'arrêter un moment sur nos méthodes et de s'interroger sur ce phénomène. Ce désintérêt est-il générationnel, pédagogique, le manque de conviction ou de pratique de la majorité des enseignants, ou une question de maturité de nos étudiants ? Certains échappent à cette « apathie » méthodologique et la dynamique qui ressort des présentations des jeunes chercheurs que nous avons entendus laisse supposer l'existence d'un fossé entre l'étudiant et le chercheur en devenir. Toutefois, nous ne disposons pas d'indicateurs satisfaisants pour mesurer ces phénomènes. Un bilan historiographique semble s'imposer.