

MICHEL-HENRI GENSBITTEL

BENOÎT RIANDEY

Rencontres de la SFdS avec la statistique publique

Journal de la société française de statistique, tome 146, n° 3 (2005),
p. 5-7

http://www.numdam.org/item?id=JSFS_2005__146_3_5_0

© Société française de statistique, 2005, tous droits réservés.

L'accès aux archives de la revue « Journal de la société française de statistique » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

Rencontres de la SFdS avec la statistique publique

Ce numéro du *Journal de la SFdS*, coordonné avec le numéro 113-114 du *Courrier des statistiques*, nous amène à évoquer quelques-unes des rencontres passées entre associations de statisticiens et système statistique public.

Viennent immédiatement à l'esprit les initiatives internationales d'Adolphe Quetelet, et notamment son premier congrès international de statistique à Bruxelles en 1853, après lequel fut créée la Société de Statistique de Paris (SSP) en 1860. Cette dernière fut le lieu de débat des différents secteurs de la statistique française, tant au cours de ses dîners-débats que par les publications du *Journal de la SSP*, repris en 1998 par la présente revue. Le groupe Statistiques économiques et sociales (SES) de la SFdS s'inscrit dans cette lignée et le présent dossier, construit en commun avec l'Insee, en est le plus récent rejeton.

La relation dialectique entre la production statistique et les règles de protection de la vie privée s'est esquissée en de multiples occasions : loi du 7 juin 1951 et loi du 23 décembre 1986 sur la statistique publique, loi Informatique et libertés du 6 janvier 1978, loi Bioéthique de 1994 sur l'épidémiologie et, pour les instituts de recherche, débats aboutissant au vote du 6 août 2004.

Dès l'année 1992, Isabelle de Lamberterie et la cellule Science et Droit du Ministère de la recherche réunissaient des membres du CNRS, de l'Université et de l'INED pour évaluer les contraintes que la loi Informatique et libertés imposait à la recherche française. Les 23 et 24 octobre 1995, Paris et Strasbourg connaissaient deux événements simultanés : le vote de la directive européenne relative à la protection de la vie privée, transcrite en droit français le 6 août 2004, et la tenue du colloque du cinquantenaire de l'INED « Démographie, Statistique et Vie privée »¹. Christian Engelage, représentant d'Eurostat à ce colloque, soulignait l'intérêt des statisticiens français à voir le Parlement français approuver certaines dispositions optionnelles contenues dans la directive.

Telle fut l'origine de la Commission de déontologie de la SFdS présidée par René Padiou. Élargie en un *Intergroupe Statistique et Recherche*, cette commission a travaillé avec les instances ministérielles, le Conseil national de l'information statistique (CNIS) et des représentants de la Commission nationale de l'informatique et des libertés (CNIL). Ces réunions ont abouti à la publication du rapport du CNIS sur « Les attentes des statisticiens en matière de transcription de la directive du 24 octobre 1995 ». Gérard Lang, représentant l'Insee lors des arbitrages interministériels, a su faire entendre le point de vue des statisticiens comme en atteste le texte voté par le parlement le 6 août 2004. Il n'est donc pas excessif de dire que la mobilisation de la SFdS a permis une reconnaissance de la statistique et de la recherche dans le cadre de la nouvelle loi Informatique et Libertés.

1. B. Riandey, « Démographie, Statistique et Vie privée », *Population* n° 2, 1996, p. 439-458.

Un autre chantier s'était développé autour de la notion d'appariements sécurisés. Alors même que les statisticiens n'osaient pas prononcer le mot *appariement* devant un membre de la CNIL, les épidémiologistes avaient développé une technique d'appariements sécurisés par hachage des identifiants. Elle leur permettait de «dédoublonner», de suivre ou d'apparier des dossiers portant sur des sujets aussi sensibles que le sida ou le cancer. Ces techniques ont été appliquées au système d'information des hôpitaux, le PMSI, et aux fichiers de sécurité sociale. C'est la CNIL elle-même qui a suggéré au CREDOC d'y recourir pour son Observatoire parisien du RMI. La lecture très instructive du rapport de la CNIL (année 1999) a conduit l'INED et le groupe SES à organiser un séminaire « Appariements sécurisés et statistique publique » en février 2001, suivi d'un cours sur les appariements sécurisés en janvier 2003. La participation à ce cours d'Alain Goy (responsable à l'époque du service statistique du ministère de l'Éducation nationale) et de son équipe a permis l'avancée considérable présentée ici. La SFdS a joué là son rôle de pont entre des collègues aussi éloignés que les services statistiques ministériels et les épidémiologistes, notamment Catherine Quantin qu'on ne saurait suffisamment remercier. Cette rencontre méthodologique des fichiers d'étudiants et d'assurés n'est pas fortuite : en juin 1996, les convergences du fichier Scolarité des élèves et du futur SNIIR-AM² faisaient déjà l'objet d'un débat du séminaire de méthodes d'enquêtes. Mais les projets mûrissent moins vite que les paroles ne volent.

Citons un troisième chantier autour des échantillons permanents de l'assurance maladie, du RNIAM³ et d'un futur échantillon de bénéficiaires. Son intérêt pour la démographie est vite apparu : dès mars 1989, le séminaire de méthodes d'enquêtes INED-ASU notait la fécondité de l'appariement de l'échantillon permanent des assurés sociaux, l'EPAS, et de l'enquête ESPS sur la santé et la protection sociale du CREDES, devenu IRDES. L'intérêt d'appuyer les enquêtes sur des fichiers administratifs permanents en ressortait de façon convaincante. L'apport démographique du RNIAM, discuté en septembre 1996 au séminaire Démodynamiques de l'INED relatif à la dématérialisation des données de l'assurance maladie a suscité beaucoup d'espairs scientifiques pour un avenir certes lointain. Le futur panel de bénéficiaires de l'assurance maladie, base de sondage d'une constellation d'enquêtes spécifiques, apparaissait comme un second pôle très efficace de la statistique d'enquêtes à côté du recensement rénové.

Telle fut l'idée de la table ronde organisée par le groupe SES aux Journées de statistique de Montpellier en 2004. L'analyse n'était pas mauvaise car, à la même époque, le CNIS avait confié à Mylène Chaleix et Stéfan Lollivier un rapport sur les trajectoires sociales et d'emploi conduisant à des recommandations de même nature : structurer les enquêtes ménages de l'Insee autour de l'échantillon démographique permanent. De là à actualiser le débat auprès des parisiens en rassemblant ces deux courants d'inspiration, il n'y avait qu'un pas à faire. Les auteurs de ce petit historique le firent aussitôt pour le groupe

2. Système national d'information inter régimes de l'assurance maladie.

3. Répertoire national inter régimes de l'assurance maladie.

SES auprès de Jean-Pierre Le Gléau, Stéfan Lollivier et Carla Saglietti. Cet échange s'est concrétisé par le séminaire du 1^{er} février 2005 « Nouveaux enjeux, nouveaux atouts de la statistique sociale ». Le dossier commun au *Courrier des statistiques* et au Journal de la SFdS en est le compte rendu élargi.

En définitive, la réflexion sur la protection des personnes lors d'opérations statistiques, celle des techniques d'appariements anonymes dans la statistique administrative et la gestion efficace d'échantillons ciblés dans un univers porté à la non-réponse se consolident mutuellement. Cela peut être vu comme un nouvel âge de la statistique publique, celui où l'évolution de la loi et celle des techniques permettent une statistique plus assurée, libérée d'une certaine auto-censure que le choc de la loi Informatique et Libertés avait induite dans les années 1980. Souhaitons que ce débat aboutisse à de nouveaux projets statistiques. Le renouvellement du cours de la SFdS sur les appariements sécurisés en 2006 devrait en être un moteur. En favorisant ce débat avec l'Insee et le CNIS, la SFdS aura joué son rôle.

D'autres rencontres de la SFdS avec la statistique publique pourraient être citées, qu'il s'agisse des séminaires de méthode d'enquêtes, des séminaires SFdS-Insee sur la rénovation du recensement⁴ ou des colloques francophones sur les sondages. Mais il n'en sera pas davantage question ici car ils concernent peu la présente livraison du Journal de la SFdS dont la matière est déjà abondante. Souhaitons la innovatrice et d'agréable lecture.

Michel-Henri GENSBITTEL, Benoît RIANDEY

Et parlons d'avenir...

LA PRATIQUE DES APPARIEMENTS SÉCURISÉS

En janvier 2003, le groupe Statistiques économiques et sociales de la Société Française de Statistique (SFdS) a organisé un premier atelier de formation pratique aux méthodes d'appariement sécurisés : aspects juridiques, description des méthodes, présentation de logiciels disponibles, étude d'exemples et débat avec la CNIL.

Cet atelier a par exemple permis à la Direction de l'évaluation et de la prospective (DEP) du Ministère de l'éducation nationale de s'approprier ces méthodes et, en accord avec la CNIL, de les mettre en application rapidement.

Le groupe Statistiques économiques et sociales se propose d'organiser un nouvel atelier de formation sur ce thème au premier semestre 2006.

Contact : Michel-Henri Gensbittel (mhgens@univ-paris1.fr)

4. Ces séminaires ont reçu un très large écho dans le Journal de la SFdS (volumes 140, n°4, 142, n°3, 143, n°3-4, respectivement publiés en 1999, 2001, 2002).