Techniques d'enquête, juin 1990
Vol. 16, n° 1, pp. 55–66
Statistique Canada

Rôle de l'administration fédérale dans le développement des méthodes statistiques aux États-Unis

BARBARA A. BAILAR[1]

RÉSUMÉ

Dans cet article, nous montrons brièvement comment les statisticiens du U.S. Bureau of the Census ont pu, par leurs recherches, contribuer à l'avancement de la théorie et de la pratique des recensements et des sondages. Nous essayons aussi de voir ce que nous réserve l'avenir à ce chapitre.

MOTS CLÉS: Échantillonnage; erreur non due à l'échantillonnnage; estimation; confidentialité; désaisonnalisation.

1. INTRODUCTION

Aux États-Unis, l'administration fédérale a été un chef de file dans l'élaboration de méthodes statistiques pour les recensements et les sondages. Nous nous limiterons ici aux réalisations du U.S. Bureau of the Census et concentrerons notre attention sur quatre grands sujets de recherche – l'élaboration de méthodes d'échantillonnage, l'erreur non due à l'échantillonnage, la désaisonnalisation et l'élaboration de méthodes visant à protéger le caractère confidentiel des données fournies par les répondants (communément appelées méthodes de protection du secret statistique). Enfin, nous tenterons de voir comment pourrait évoluer la recherche dans les années à venir.

2. ÉCHANTILLONNAGE

L'histoire des méthodes de sondage au sein de l'administration fédérale des É.-U. est surtout celle d'un groupe de personnes remarquables qui ont oeuvré au U.S. Bureau of the Census sous la direction de Morris Hansen et de William Hurwitz. Lorsqu'on apprend que le U.S. Bureau of the Census était acquis à l'échantillonnage probabiliste dès le début des années 1940, on se demande comment une institution aussi conformiste a pu faire pour adopter si tôt une telle position? En règle générale, les organismes hésitent très longuement avant d'adopter de nouvelles méthodes et probablement que le Census Bureau est beaucoup moins empressé aujourd'hui à adopter de nouvelles méthodes et à en promouvoir l'utilisation. Hansen donne trois raisons pour expliquer que les divisions spécialisées du Census Bureau aient accepté assez rapidement l'idée de l'échantillonnage (Causey, Cox et Lawrence 1985) : attitude favorable de la part de la direction du Census Bureau (Bailar 1975), établissement de liens de coopération avec les divisions spécialisées (Bell et Hillmer 1984) et formation d'un groupe d'experts en sondages (appelés ultérieurement méthodologistes) au sein des divisions spécialisées, ce groupe étant chargé de conseiller la Statistical Research Division (SRD) sur les questions d'ordre technique. Hansen omet de mentionner un autre facteur important, soit la vigueur et le caractère du duo dynamique qu'il forme avec Hurvitz et de ses partisans.

En 1936, le Census Bureau a commencé à s'intéresser à l'échantillonnage et aux applications possibles. À cette époque, on pratiquait déjà l'échantillonnage mais il ne s'agissait pas d'échantillonnage probabiliste. Il y avait l'échantillonnage par choix raisonné et l'échantillonnage de

[1] Barbara A. Bailar, American Statistical Association, 1429 rue Duke, Alexandria, VA 22314-3402.

gros établissements. Toutefois, les méthodes d'échantillonnage n'avaient à peu près pas de fondements théoriques. En 1937, le Congrès a autorisé un recensement facultatif des personnes en chômage et des personnes employées à temps partiel dans tout le pays. Un questionnaire devait être expédié par la poste à chaque ménage. Comme on craignait que ce recensement facultatif renferme une erreur systématique quelconque, on a décidé de procéder à un recensement de contrôle dans un certain nombre de secteurs. Ce recensement consistait à former un échantillon probabiliste de routes postales et à interviewer tous les ménages dont le logement se trouvait sur ces routes. Les facteurs étaient chargés de faire les interviews et de classer les questionnaires qui étaient retournés par la poste. Ils produisaient ensuite des chiffres pour chaque route postale, y compris les routes échantillonnées. On disposait ainsi d'une variable indépendante pour l'estimation; c'était le début de l'estimation par quotient. Les résultats du recensement de contrôle ont confirmé l'utilité de l'échantillonnage. Cependant, l'expérience a été remarquable à plusieurs égards:

- évaluation précise des effets de la non-réponse pour un recensement facultatif;
- utilisation de l'estimation par quotient;
- production rapide des résultats.

Au cours d'une interview reproduite dans la revue *Statistical Science* (Olkin), Hansen raconte que le recensement facultatif s'est fait durant la semaine du 20 novembre 1937, que les interviews ont eu lieu dans la semaine du 4 décembre et que les résultats préliminaires ont été connus le 31 décembre. Il est difficile de croire que le Census Bureau puisse agir encore plus rapidement aujourd'hui.

Selon Hansen, les résultats convaincants du recensement de contrôle de 1937 auront eu le mérite de faire avancer la cause de l'échantillonnage probabiliste au sein du Census Bureau. Auparavant, on croyait au Census Bureau qu'il fallait s'en tenir à des recensements et que les sondages n'étaient pas un exercice sérieux. Grâce au succès de l'expérience de 1937, on pouvait envisager de recourir à l'échantillonnage à l'occasion du recensement de 1940; celui-ci fut d'ailleurs le premier recensement où certaines questions étaient adressées uniquement à un sous-ensemble de la population. Malheureusement, quelques membres du Census Bureau ont repris depuis quelques mois l'idée de faire une vérification complète des logements inoccupés sous prétexte qu'un recensement comporte moins d'erreurs qu'une enquête. Espérons qu'il ne s'agit là que d'un moment d'aberration causé par un différend quelconque.

La théorie des sondages a évolué au même rythme que l'enquête sur la population active. La Works Progress Administration (WPA) avait la responsabilité d'une enquête visant à mesurer le niveau de chômage. Lorsque cet organisme a cessé d'exister en 1942, le Census Bureau s'est vu confier la responsabilité de l'enquête. Il a alors fait une évaluation des méthodes de sondage et leur a apporté de nombreuses améliorations. Cet exercice de révision a contribué largement à l'avancement de la théorie des sondages. Mentionnons au passage quelques-uns des principes qui ont été définis lors de cet exercice de révision: unités primaires d'échantillonnage plus grandes, échantillonnage avec probabilité proportionnelle à la taille et sous-stratification de régions ou de secteurs. Hansen et Hurwitz ont analysé ces principes dans un article publié en 1943 dans la revue *Annals of Mathematical Statistics*. Lorsqu'on relit "On The Theory Of Sampling From Finite Populations", on découvre toujours quelque chose de nouveau. Ce serait le premier article que des employés d'un organisme fédéral auraient publié sur l'échantillonnage appliqué à des populations finies. Bien que les notions avaient déjà été traitées par d'autres auteurs, ce que Hansen et Hurwitz apportaient à la théorie était tout à fait inédit. En outre, Hansen et Hurwitz ont été les seuls à faire une analyse des résultats à l'aide d'une série de comparaisons qui faisaient ressortir les avantages des méthodes proposées.

Par la suite, on a continué d'améliorer l'enquête sur la population active. Ainsi est arrivée l'estimation composite, qui applique le principe du renouvellement de l'échantillon afin d'obtenir de meilleures estimations. Nul doute que la Current Population Survey (c'est le nom

qu'on donne à l'enquête sur la population active) a servi de modèle dans le monde entier pour l'élaboration d'enquêtes sur la population active.

La Statistical Research Division s'est aussi intéressée aux problèmes d'échantillonnage que posaient les enquêtes menées auprès d'établissements commerciaux. L'idée la plus courante était que l'échantillonnage était une opération appropriée pour des populations relativement homogènes comme les populations de personnes mais convenait peu pour des populations fortement asymétriques comme les populations d'entreprises. Tenant compte de l'asymétrie de la population étudiée, le groupe de recherche a stratifié les magasins de détail selon leur taille. Les plus gros magasins étaient nécessairement inclus dans l'échantillon et ceux de moindre importance étaient échantillonnés avec une probabilité proportionnelle à la taille.

On a aussi observé qu'il y avait de nombreuses créations et disparitions d'entreprises. Comme on croyait qu'un échantillon statique ne pourrait pas refléter cette rotation, on a créé un échantillon aréolaire qui allait produire des estimations pour les nouveaux magasins. Bien que l'enquête mensuelle sur le commerce de détail ait subi de nombreux autres perfectionnements, sa structure fondamentale est demeurée la même. L'estimation composite est aussi utilisée dans l'enquête sur le commerce de détail afin d'obtenir des estimations plus précises.

Nous pourrions citer beaucoup d'autres exemples de perfectionnement des méthodes de sondage. L'ouvrage en deux volumes de Hansen, Hurwitz et Madow intitulé *Sample Survey Methods and Theory* contient de nombreux exposés accompagnés de la théorie et des applications correspondantes. Bien que les exemples qui y sont présentés soient largement périmés, nous ne connaissons pas d'ouvrage qui contienne autant d'applications relatives aux sondages. Le seul inconvénient, c'est que ces applications n'aient été jamais mises à jour.

3. ERREUR NON DUE À L'ÉCHANTILLONNAGE

Dans cet effort soutenu pour améliorer les méthodes de recensement et d'enquête, il ne fallait pas s'arrêter uniquement aux erreurs d'échantillonnage. Il fallait en effet tenter de limiter les erreurs provenant d'autres sources, comme l'interview, le traitement, le questionnaire, *etc.* Hansen et Hurwitz ont commencé à s'intéresser à la question avant le recensement de 1950; ils ont intégré à ce recensement de nombreuses études expérimentales qui visaient à estimer l'effet des erreurs de mesure. L'erreur totale dans les enquêtes devint donc un sujet de recherche important au Census Bureau. L'évaluation et le contrôle des erreurs non dues à l'échantillonnage figuraient régulièrement au programme de recherche du Census Bureau.

L'assertion voulant que les erreurs de mesure pourraient avoir une influence beaucoup plus marquée sur les données que les erreurs d'échantillonnage, surtout aux niveaux d'agrégation supérieurs, a donné un nouvel élan à la recherche sur les erreurs non dues à l'échantillonnage. Hansen, Hurwitz et Bershad (1961) ont élaboré un modèle intégré pour les recensements et les enquêtes, qui tenait compte à la fois de l'erreur d'échantillonnage, de l'erreur de réponse et du biais. L'erreur de réponse était composée de ce qu'on appelle aujourd'hui la variance de réponse simple et la variance de réponse corrélée. La première représente la variabilité fondamentale observée d'une fois à l'autre à cause des différences entre les réponses, les répondants, les interviewers, *etc.* La variance de réponse simple représente aussi la variabilité observée d'une fois à l'autre dans le codage. La variance de réponse corrélée représente la variance engendrée par un facteur qui tend à orienter les réponses. Le facteur auquel on s'intéresse le plus souvent est l'interviewer. Parce qu'il peut avoir une certaine idée de la réponse que fournira la personne interviewée ou parce qu'il a acquis de l'expérience dans l'interview des ménages, l'interviewer peut orienter les réponses vers certaines catégories. Pour une région donnée, on observe une forte variabilité selon les interviewers des taux de non-réponse et des réponses concernant le niveau de scolarité et beaucoup d'autres caractéristiques.

Le modèle de Hansen, Hurwitz et Bershad a été testé pour la première fois lors du recensement de 1950 et il a eu une influence déterminante dans la décision de remplacer le "recensement sur place", où un interviewer visitait chaque ménage pour lui poser des questions et

enregistrer ses réponses, par un "recensement par la poste", où chaque ménage reçoit un questionnaire par la poste, le remplit et le retourne par la poste. L'expérience a été tentée à l'occasion des recensements de 1960 et de 1970 et on a observé une forte diminution de la variance lorsqu'il y avait auto-dénombrement (U.S. Bureau of the Census 1960, 1970).

Par ailleurs, Hansen et Hurwitz ont encouragé la recherche sur l'erreur de couverture. Le Census Bureau a passé beaucoup de temps à analyser les effets de l'erreur de couverture dans les recensements comme dans les enquêtes. Après le recensement de 1950, le Census Bureau a réussi à mesurer le niveau de sous-dénombrement au niveau national selon l'âge, l'origine raciale et le sexe en se servant d'un modèle qu'avait élaboré Ansley Coale à l'université Princeton. Cette méthode, connue sous le nom d'analyse démographique, a permis de constater que le niveau de sous-dénombrement était beaucoup plus élevé chez les noirs que chez les blancs (Citro et Cohen, 1985). De plus, le Census Bureau a entrepris l'élaboration d'une enquête postcensitaire dans le but de mieux connaître la population non dénombrée. À l'origine, le Bureau cherchait surtout à perfectionner ses méthodes de recensement; depuis quelques années, il met plutôt l'accent sur la répétition d'enquêtes ou de recensements et c'est justement l'approche qu'il entend adopter pour le recensement de 1990. De même, le sous-dénombrement observé dans des enquêtes a poussé les spécialistes à approfondir les méthodes d'estimation par quotient dans l'espoir de réduire l'impact du sous-dénombrement. Ces méthodes sont utilisées dans la plupart des enquêtes-ménages du Census Bureau.

Le U.S. Bureau of the Census est maintenant reconnu pour ses recherches sur les erreurs de mesure. En plus de ses recherches sur les erreurs de réponse et la couverture, il a favorisé la réalisation d'études sur les biais de renouvellement, qui ont un effet sur les estimations tirées d'enquêtes où les répondants sont sollicités plus d'une fois. L'enquête sur la population active, où les répondants demeurent dans l'échantillon quatre mois consécutifs, puis en sont exclus pendant huit mois, et enfin y sont inclus de nouveau pendant quatre mois, a été analysée soigneusement. Bailar (1975) a montré que les estimations du niveau d'emploi et du niveau de chômage étaient en règle générale plus élevées pour les personnes qui en étaient à leur premier mois dans l'échantillon que pour celles qui n'en étaient pas à leur première présence. Ces écarts influent sur les niveaux d'emploi et de chômage, mais n'ont probablement aucun effet sur les estimations de la variation d'un mois à l'autre.

Ceci n'est qu'un aperçu des recherches qui ont été entreprises au Census Bureau sur les erreurs de mesure. À l'heure actuelle, tous les organismes statistiques s'intéressent à la question.

4. DÉSAISONNALISATION

On a commencé à parler de désaisonnalisation au sein de l'administration publique lorsque Julius Shiskin était membre du Census Bureau. Il était alors chargé d'informatiser les opérations de désaisonnalisation. Aujourd'hui, la méthode X-11 est utilisée partout.

Shiskin raconte que dans les années 1950, le Council of Economic Advisors pressait les organismes fédéraux de produire des séries chronologiques désaisonnalisées. En 1953, le Census Bureau acquiert le premier ordinateur spécialisé dans le traitement des données, le UNIVAC I, et Éli Marks, qui se rend au travail dans la même voiture que Shiskin, l'entretient souvent de la difficulté qu'il éprouve à programmer cet ordinateur. Shiskin se demande s'il ne serait pas possible de se servir de l'ordinateur pour exécuter la désaisonnalisation des séries; après avoir consulté un technicien en informatique, il découvre qu'une série de 10 ans pourrait être désaisonnalisée en une minute. Aujourd'hui, une opération de ce genre prend évidemment beaucoup moins de temps.

La désaisonnalisation est en quelque sorte une opération qui exige de la maîtrise étant donné le très grand nombre d'options qu'offre le programme X-11 à l'analyste. Cependant, lorsque Shiskin a entrepris d'automatiser les opérations de désaisonnalisation, on se demandait si une machine pouvait vraiment effectuer le travail d'un technicien expérimenté. Shiskin décida donc

de prendre au mot le Federal Reserve Board. Sa proposition allait dans le sens suivant: le Federal Reserve Board choisirait une série quelconque et prendrait tout le temps voulu pour la corriger; de son côté, Shiskin traiterait la même série par ordinateur. Les deux versions ainsi corrigées seraient reproduites graphiquement, sans que l'on précise la façon dont elles ont été corrigées, et soumises à l'attention d'un petit groupe de spécialistes du Federal Reserve Board pour qu'il se prononce sur la qualité des résultats. Les membres de ce groupe ont été unanimes pour dire que la série corrigée par ordinateur était supérieure à l'autre.

Aujourd'hui, les organismes fédéraux désaisonnalisent des milliers de séries chronologiques à chaque année. Pendant de nombreuses années, on a cru que les méthodes fondées sur un modèle étaient difficilement applicables à cause des limites des ordinateurs. Par ailleurs, à chaque année on élaborait de nouveaux facteurs de désaisonnalisation à partir de données chronologiques. Par exemple, un facteur qui devait servir à calculer les données désaisonnalisées pour juillet était élaboré en décembre de l'année précédente. Dans les circonstances, on ne pouvait utiliser de données fondées sur des événements plus récents. Cette restriction était tout à fait logique compte tenu de ce que la préparation des cartes perforées et le traitement de la série à l'ordinateur étaient des opérations qui s'étendaient sur plusieurs jours. Cependant, au cours des dix dernières années cette méthode a fait l'objet de plus en plus de critiques au profit de la méthode de désaisonnalisation courante. Les employés du Census Bureau qui étaient affectés aux séries chronologiques, avec David Findley en tête, étudièrent en profondeur les avantages de cette méthode de désaisonnalisation pour les séries du Census Bureau et prônèrent son utilisation au sein du Bureau.

Ces mêmes employés se posèrent aussi des questions très fondamentales en ce qui a trait à la désaisonnalisation. Premièrement, sur quel critère doit-on se fonder pour déterminer si une série doit être désaisonnalisée ou non? Deuxièmement, étant donné qu'il existe plusieurs méthodes pour corriger les séries chronologiques, comment évalue-t-on ces diverses méthodes? Dans un article de fond, Bell et Hillmer (1984) s'interrogent sur les nécessités d'une désaisonnalisation lorsque la série en question peut être modélisée convenablement. Ils définissent également des critères permettant d'évaluer les opérations de désaisonnalisation. Il convient de souligner que le Census Bureau n'est pas le seul organisme d'État à avoir fait des recherches inédites dans ce domaine. Soit dit en passant, une des grandes réussites de l'équipe du Census Bureau affectée aux séries chronologiques est la série de conférences qu'elle organise régulièrement à l'intention des spécialistes oeuvrant au sein de l'administration publique. Ainsi, des membres du Federal Reserve Board, du Bureau of Labor Statistics, de l'Energy Information Administration et du Bureau of Economic Analysis, pour ne nommer que ceux là, assistent régulièrement à ces conférences pour se tenir au courant des progrès dans le domaine. Estella Dagum, de Statistique Canada, a mené à bien de nombreux projets de recherche, notamment l'élaboration de la méthode X-11 ARMMI.

5. PROTECTION DU SECRET STATISTIQUE

Que l'on soit d'accord ou non avec les lignes directrices du Census Bureau en matière de confidentialité des données, il faut reconnaître que le Bureau a été l'un des premiers à prôner l'utilisation de méthodes visant à protéger la confidentialité des données d'enquête. La protection du secret statistique est une mesure qui vise à garder confidentiels les renseignements fournis par un répondant. La question de la protection du secret statistique a toujours été un problème dans les recensements, mais elle l'est aussi dans les enquêtes, particulièrement les enquêtes de nature longitudinale ou celles où on risquerait d'associer des enregistrements à des données de l'enquête.

La principale préoccupation des responsables de recensements démographiques est d'éliminer les risques de divulgation liés à la publication de très petites fréquences. De trop petites fréquences risquent de dévoiler l'identité de répondants ou de petits groupes de répondants. En outre, la présence de zéros dans certaines cases est aussi une source potentielle de divulgation.

On considère qu'il y a divulgation dans un tableau de fréquences lorsqu'il est possible de conclure que le nombre de répondants visés par une case X quelconque est inférieur à un seuil préétabli. À l'occasion du recensement décennal de 1980, on avait défini des seuils différents pour les ménages et pour les individus.

Les méthodes de protection du secret statistique pour les tableaux de fréquences se divisent en trois catégories: la suppression de cases, la modification de fréquences de case et le remplacement de fréquences de case par des intervalles. La suppression de cases empêche le dévoilement de valeurs numériques et rend impossible toute inférence qui pourrait découler d'une tentative pour établir des relations linéaires entre les fréquences de cases publiées et les fréquences non publiées. La modification de fréquences de case consiste à modifier légèrement à la hausse ou à la baisse la plupart des fréquences de manière que l'on ne puisse tirer de conclusions précises sur les valeurs figurant dans le tableau. La troisième méthode, qui consiste à remplacer les estimations ponctuelles par des intervalles, est peu utile dans le cas de classifications croisées.

La suppression de cases a été la principale méthode utilisée par le Census Bureau au cours des années 1980. Les restrictions appliquées aux sommes de ligne ou de colonne d'un tableau de fréquences produisent une série de contraintes linéaires. Une fois que l'on a supprimé les cases les plus susceptibles de divulguer des renseignements confidentiels, on utilise la programmation mathématique pour vérifier s'il n'y aurait pas d'autres sources de divulgation dans le tableau. Bien que cette méthode ait été utilisée de façon empirique pendant de nombreuses années, Cox et ses collègues du Census Bureau ont défini les fondements mathématiques de la méthode (Causey, Cox et Ernst, 1985) et ont montré jusqu'à quel point la suppression de cases était une opération complexe.

Les méthodes de modification de fréquences de case, y compris l'arrondissement aléatoire, ont été élaborées et appliquées au Royaume-Uni, en Suède et au Canada. Toutes ces méthodes consistent à ajouter à des fréquences de case ou à soustraire de ces fréquences, avec une probabilité définie, une faible valeur pouvant parfois être zéro.

En ce qui concerne des données comme les chiffres de vente, la valeur nette, les stocks et les données financières d'établissements manufacturiers et d'établissements de vente au détail, le danger, selon le Census Bureau, est que l'on réussisse à découvrir les montants déclarés par un répondant. Si les dirigeants d'une entreprise venaient à déduire des chiffres d'un tableau ceux qui se rapportent à leur entreprise, ils pourraient découvrir par soustraction les chiffres fournis par un concurrent. Pour éviter cela, le Census Bureau a recours à la suppression de cases. À cette fin, il applique la règle dite de (n, k), selon laquelle X est considérée comme une case prêtant à divulgation si un nombre n de répondants représentent plus de k pour cent de la fréquence totale de la case. Cette règle fait partie d'une série de règles touchant la prédominance des cases et qui sont toutes additives.

La protection du secret statistique est maintenant un sujet de préoccupation dans tous les pays et plus particulièrement au sein des administrations publiques. En effet, celles-ci doivent chercher à résoudre les problèmes très délicats que pose depuis quelque temps la demande de microdonnées.

6. PERSPECTIVES D'AVENIR

L'élaboration de modèles mathématiques joue un rôle important dans les quatre cas que nous venons de traiter. L'échantillonnage repose bien sûr sur des méthodes de randomisation mais la volonté de limiter l'erreur totale dans les enquêtes a favorisé l'élaboration d'un modèle d'erreurs d'enquête énoncé pour la première fois par Hansen, Hurwitz et Bershad (1961). Ce modèle ainsi que les tests ayant servi à estimer les paramètres ont été à l'origine de nombreuses décisions concernant les modalités d'exécution des recensements et des enquêtes.

L'usage des modèles de séries chronologiques est très répandu dans le monde; ceux-ci tendent d'ailleurs à remplacer les méthodes empiriques telle la X-11. À l'heure actuelle, les spécialistes

conseillent vivement d'intégrer les méthodes chronologiques aux méthodes d'estimation afin d'obtenir des résultats plus précis. Il sera intéressant de voir si cette intégration se fera et comment elle se fera.

Les organismes statistiques tentent par la modélisation de produire des données pour petites régions. Alors que les données sont le plus souvent recueillies pour des unités géographiques relativement étendues comme les États, on aurait besoin de données pour des unités géographiques plus petites comme les comtés. Des conférences ont permis d'évaluer et de comparer diverses méthodes de production de données régionales. Au cours de la dernière décennie, le Census Bureau a pu établir des estimations de la population à l'aide de méthodes empiriques. Les recherches faites sur le sous-dénombrement au Census Bureau ont permis d'examiner plusieurs modèles et ont apporté beaucoup d'éléments pour mieux comprendre le problème.

On est en train d'examiner minutieusement des méthodes spéciales de contrôle et d'imputation et d'élaborer des modèles mathématiques. La tendance à la modélisation s'accentuera sûrement dans l'avenir.

En conclusion, nous croyons que l'avenir sera propice à la modélisation. Non pas que nous cherchions à dénigrer les méthodes empiriques qui sont en usage actuellement; les statisticiens ont toujours affirmé que théorie et pratique vont de pair. Les méthodes empiriques qui semblent donner des résultats satisfaisants mènent à l'établissement de modèles et à la définition de nouveaux principes théoriques qui sont nuancés par la pratique. Du fait que les organismes d'État sont appelés à résoudre de nombreux problèmes statistiques qui présentent le plus grand intérêt, depuis toujours ils font figure de pionnier dans le domaine de la méthodologie statistique.

BIBLIOGRAPHIE

BAILAR, B.A. (1975). The effects of rotation group bias on estimates from panel surveys. *Journal of the American Statistical Association*, 70, 23-30.

BELL, W.R., et HILLMER, S.C. (1984). Issues involved with the seasonal adjustment of economic time series. *Journal of Business and Economic Statistics*, 2, 291-317.

CAUSEY, B.E., COX, L.H., et ERNST, L.R. (1985). Applications of transportation theory to statistical problems. *Journal of the American Statistical Association*, 80, 903-909.

CITRO, C.F., et COHEN, M.L. (éds.) (1985). *The Bicentennial Census*. Washington, D.C.: National Academy Press.

DUNCAN, J., et SHELTON, W. (1978). *Revolution in United States Government Statistics*. Washington D.C.: U.S. Government Printing Office.

HANSEN, M.H. (1987). Some history and reminiscences on survey sampling. *Statistical Science*, 2, 180-190.

HANSEN, M.H., et HURWITZ, W.N. (1943). On the theory of sampling from finite populations. *Annals of Mathematical Statistics*, 14, 333-362.

HANSEN, M.H., HURWITZ, W.N., et BERSHAD, M.A. (1961). Measurement errors in censuses and surveys. *Proceeding of the International Statistical Institute*, 38, 358-374.

HANSEN, M.H., HURWITZ, W.N., et MADOW, W. (1953). *Sample Survey Methods and Theory*, Vols. 1 et 2. New York: John Wiley and Sons.

OLKIN, I. (1987). A Conversation with Morris Hansen. *Statistical Science*, 2, 191-210.

U.S. BUREAU OF THE CENSUS (1968). *Evaluation and Research Program of the U.S. Censuses of Population and Housing, 1960: The Effect of Interviewers and Crew Leaders*. Séries ER 60 n° 7.

U.S. BUREAU OF THE CENSUS (1979). *Evaluation and Research Program of the U.S. Censuses of Population and Housing, 1970: Enumerator Variance in the 1970 Census*. PHC(E) n° 13.

COMMENTAIRES

G.J. BRACKSTONE[1]

1. Introduction

Cet article montre quelle contribution importante a été celle du U.S. Bureau of the Census, depuis 50 ans, au développement de la méthodologie statistique. Les quatre exemples choisis par Bailar pour illustrer cet apport sont remarquables tant par leur importance intrinsèque que par leur diversité. Ces exemples ne sont pas des variantes d'une percée méthodologique unique, mais des contributions fondamentales dans quatre domaines différents. Peut-être ces exemples illustrent-ils la grande diversité et la difficulté des problèmes méthodologiques que doivent résoudre les organismes statistiques – diversité et difficulté qui feraient mentir ceux qui pourraient penser que dans le secteur public la statistique n'a rien que de routinier et de banal.

Dans la description de ces exemples, l'aperçu sur les circonstances dans lesquelles ont eu lieu ces développements offre un intérêt particulier. Les perfectionnements méthodologiques obtenus ont engendré des solutions dont la généralité dépasse de beaucoup l'envergure des problèmes qui les ont suscités au départ, mais les processus même qui les sous-tendaient méritent d'être examinés de près car ils permettent de déterminer quelles circonstances sont nécessaires pour que de telles percées soient possibles. Je reviendrai sur ce point.

Pendant cette même période de cinquante ans, le U.S. Bureau of the Census a également été très actif dans le domaine de l'automatisation des opérations statistiques. Premier à avoir mis au point des trieuses et des tabulatrices de cartes perforées au début du siècle, le Bureau of the Census a aussi été le premier organisme statistique à utiliser un ordinateur, dans les années 50. Dans les années 60, c'est encore lui qui a innové dans l'automatisation de la saisie des données en mettant au point le dispositif FOSDIC permettant de lire une version sur micro-fiche d'un questionnaire porteur de marques. Les innovations du Bureau of the Census sont donc manifestement présentes dans beaucoup d'aspects des tâches d'un organisme gouvernemental de statistique.

2. Diffusion des progrès méthodologiques

Chacune des contributions à la méthodologie décrites par Bailar est née de la nécessité pour un organisme statistique de résoudre un problème pratique réel. La nécessité de recueillir des données supplémentaires à un coût raisonnable et dans un délai acceptable a motivé le développement des méthodes d'échantillonnage probabiliste; la nécessité d'améliorer la qualité des données en tâchant de comprendre, de mesurer et de réduire les erreurs non dues à l'échantillonnage a suscité des travaux dans ce domaine; les progrès au chapitre de la désaisonnalisation semblent avoir été déterminés par le besoin d'accélérer et de normaliser une méthode manuelle complexe; le problème de la définition d'un processus rationnel et efficace pour assurer la confidentialité des renseignements individuels dans les résultats statistiques a inspiré la recherche dans le domaine de la protection du secret statistique. Les nombreux autres exemples qui auraient pu être cités ont pour caractéristique commune d'avoir eu leur point de départ dans un problème pratique réel.

L'organisme statistique qui met au point des méthodes statistiques conçues pour résoudre des problèmes comme ceux qui viennent d'être énumérés en tire évidemment un avantage immédiat. Mais des contributions de ce genre ont-elles eu un intérêt plus large? Ont-elles fait progresser l'ensemble de connaissances et de méthodes qu'on appelle la statistique? On peut dire que ces développements ont eu des avantages significatifs et globaux pour les organismes statistiques chargés de la production de données sociales et économiques, mais que leur importance pour la statistique comme discipline universitaire, bien qu'elle augmente, n'a pas été aussi grande qu'elle aurait pu l'être.

[1] G.J. Brackstone, statisticien en chef adjoint, Statistique Canada, Ottawa, Ontario.

Premièrement, considérons les organismes statistiques ailleurs qu'aux États-Unis. Dans la plupart des pays, l'organisme statistique d'État est une institution unique chargée de faire à intervalles réguliers d'importantes enquêtes sur les ménages et les entreprises, d'intégrer les données provenant de diverses sources, de tenir à jour et d'analyser des séries chronologiques et de diffuser de grandes quantités de données dans le public. (Sous ce rapport, les États-Unis sont un cas d'exception du fait qu'ils ont plusieurs grands organismes qui se partagent ces activités selon les domaines concernés.) Dans la plupart des pays, donc, les organismes statistiques doivent regarder à l'étranger pour observer des expériences semblables aux leurs ou pour avoir des échanges entre experts. Le réseau d'interaction entre les organismes statistiques des pays industrialisés est très développé. Les rapports peuvent être bilatéraux ou multilatéraux. L'ancienne et vivante tradition d'échanges d'information et de résultats d'expériences entre Statistique Canada et le U.S. Bureau of the Census est un exemple de rapports bilatéraux. Statistique Canada a tiré de grands avantages du fait de pouvoir utiliser et, dans certains cas, développer des méthodes statistiques mises au point par le Census Bureau, dont celles que décrit Bailar; et, je pense, le Census Bureau n'a pas moins profité des progrès méthodologiques réalisés à Statistique Canada.

En ce qui concerne les relations multilatérales, plusieurs organisations offrent une tribune où les organismes statistiques d'État peuvent échanger leurs vues: l'Organisation des Nations-Unies et ses organismes régionaux et spécialisés, l'Institut international de statistique, particulièrement sa section des statisticiens d'enquêtes et sa section des statistiques officielles, et les associations professionnelles de statisticiens de divers pays. En outre, le Census Bureau et Statistique Canada ont institué des conférences de recherche ou symposiums annuels où peuvent être communiqués les progrès et les résultats d'expériences. En somme, ces rapports bi et multilatéraux remplissent bien leur fonction, qui est de faire en sorte que les contributions d'un organisme à la méthodologie statistique – et beaucoup d'organismes apportent des contributions importantes – sont librement partagées et utilisées par d'autres organismes.

Mais quelle a été l'influence de ces développements sur la discipline de la statistique à l'extérieur des organismes d'État? Ici, nous allons nous arrêter aux exemples cités par Bailar pour illustrer son propos, encore qu'il y ait beaucoup d'autres domaines (dont certains sont énumérés à la section 4) auxquels des arguments semblables pourraient s'appliquer. Dans le cas de l'échantillonnage, l'influence sur la statistique comme discipline a été d'une grande portée. L'échantillonnage appliqué à une population finie est aujourd'hui une matière qui fait partie du programme d'enseignement de la statistique dans beaucoup d'universités et à laquelle sont consacrés de nombreux manuels. Les progrès méthodologiques réalisés dans les organismes d'État ont été assimilés et développés par la profession statistique. En fait, quelques-uns pourraient considérer qu'à certains égards ces développements ont été raffinés bien au-delà des besoins pratiques des responsables d'enquêtes. Dans le cas des erreurs non dues à l'échantillonnage, le problème est différent. Les contributions dans ce domaine n'ont pas encore conduit à un ensemble cohérent de théories et de méthodes. Cela ne veut pas dire qu'il n'y ait pas eu de progrès. Au contraire, beaucoup de travaux ont été faits sur cette question. Mais la plupart concernent une enquête en particulier. Ils ont, on peut l'espérer, amélioré de nombreuses enquêtes, servi à documenter beaucoup d'expériences et donné aux praticiens d'utiles enseignements. Mais cette question des erreurs non dues à l'échantillonnage n'a pas encore trouvé sa place dans la statistique comme discipline. En fait, les moyens accrus que nous ont donnés les travaux s'y rapportant trouvent souvent leur application dans des domaines particuliers (sociologie, démographie, *etc.*) plutôt que dans la statistique proprement dite.

Le problème de la désaisonnalisation est encore un autre cas. Technique à caractère plutôt empirique utilisée par les organismes statistiques, elle est de plus en plus étudiée depuis quelques années par des chercheurs qui tentent de lui donner un fondement solide dans la théorie statistique. Bailar mentionne un certain nombre de questions fondamentales actuellement étudiées à propos d'objectifs et de normes pour la désaisonnalisation. On étudie

également des modèles qui pourraient remplacer les techniques traditionnelles du modèle X-11 ARMMI. C'est là un domaine de la recherche en statistique auquel s'intéressent les universitaires spécialisés dans les séries chronologiques. Les méthodes de désaisonnalisation ont manifestement des applications qui dépassent l'utilisation que peuvent en faire les organismes statistiques d'État.

Enfin, le domaine dont parle Bailar où les contributions sont le plus récentes est celui de la protection du secret statistique. Il s'agit là d'un problème qui intéresse presque exclusivement les organismes assujettis à des règles de confidentialité interdisant la divulgation de tout renseignement qui pourrait permettre d'identifier une entité enquêtée. Le gros de la recherche dans ce domaine se fait dans les organismes statistiques d'État. Toutefois, les techniques utilisées viennent surtout de l'informatique, de l'analyse numérique et des mathématiques. Ce domaine est assez nouveau et n'a pas encore suscité beaucoup de travaux à l'extérieur des organismes d'État.

Ces exemples montrent que les contributions méthodologiques des organismes statistiques d'État non seulement servent à résoudre les problèmes qui se posent à ces organismes mais peuvent aussi faire faire des progrès importants à la statistique en général. Naturellement, ce ne sont pas toutes ces contributions qui ont un champs d'application très étendu, et certaines pourront rester sans grande utilité à l'extérieur des organismes d'État. Un des défis permanents qui se posent aux statisticiens du secteur public est de susciter l'intérêt de leurs collègues des autres secteurs, en particulier des universitaires, pour les recherches qui se font dans les organismes d'État.

3. Climat favorable à l'innovation

Les innovations sont rarement le fruit du hasard. Il faut un climat qui favorise l'éclosion des idées et le progrès de la recherche. Ce climat n'est pas toujours facile à créer dans une organisation dont la principale tâche est de diffuser régulièrement des données selon un calendrier préétabli. Bailar mentionne trois raisons données par Hansen pour expliquer pourquoi l'enquête par échantillonnage probabiliste a été adoptée assez rapidement au Census Bureau. Ces trois raisons définissent en substance quelles conditions préalables doivent exister pour qu'un organisme statistique soit un milieu favorable à l'innovation:

a) le soutien de la direction, c'est-à-dire la volonté d'investir dans la recherche;

b) la collaboration de clients, c'est-à-dire qu'il faut à la recherche, pour être fructueuse, une application particulière correspondant au problème qui se posait au départ et déterminant un échéancier; le gestionnaire d'un programme de recherche de ce genre doit être quelqu'un qui est à l'aise dans un climat d'expérimentation;

c) des chercheurs compétents, non seulement par leur connaissance d'un domaine en particulier, mais aussi par leur capacité de reconnaître quels problèmes peuvent être généralisés et résolus au moyen des méthodes statistiques.

Ces trois conditions peuvent créer un bon climat pour la recherche, mais d'autres efforts pourront être nécessaires pour faire en sorte que les résultats des travaux soient utilisés, et utilisés efficacement. Cela suppose que le statisticien possède des dons pour persuader et communiquer et que l'organisme où s'effectue la recherche offre le soutien voulu pour la nouvelle méthodologie.

4. Autres contributions

Bailar ne cherchait pas, dans le choix de ses exemples, à couvrir tous les domaines où des contributions ont été faites à la méthodologie statistique. Nous pouvons énumérer ici d'autres domaines de la méthodologie où des organismes d'État on apporté des contributions importantes. Certains sont mentionnés par Bailar parmi les domaines où elle envisage

des progrès pour l'avenir; mais des travaux importants y ont déjà été faits. Voici une liste de domaines dans lesquels Statistique Canada a fait des recherches.

a) **Méthodes d'analyse des données d'enquêtes complexes** Très utiles pour les utilisateurs de la plupart des statistiques gouvernementales, ces méthodes sont conçues pour permettre d'adapter ou pour remplacer les méthodes traditionnelles d'analyse statistique fondées sur l'échantillonnage aléatoire simple. Ce domaine de recherche a suscité l'intérêt des universitaires, qui ont également fait de nombreuses contributions sous ce rapport.

b) **Couplage d'enregistrements** Cette technique est utilisée pour calculer des valeurs statistiques à partir de données administratives, pour évaluer la qualité des données, dans le cas de couplages de micro-données, et pour la tenue à jour de listes. L'élaboration d'une théorie générale du couplage d'enregistrements a permis de mettre au point des logiciels de couplage. Sur cette question, la plupart des travaux ont été effectués par des organismes statistiques d'État.

c) **Contrôle et imputation** Largement utilisée dans de nombreuses enquêtes, cette technique a manqué de bases solides jusqu'au moment où l'on a élaboré une théorie, dans les années 70. Depuis, on a mis au point des méthodes et des systèmes d'application générale pour effectuer le contrôle et l'imputation dans diverses enquêtes. Ce sujet a suscité beaucoup d'intérêt et de nouveaux travaux à l'extérieur des organismes statistiques.

d) **Estimations régionales** Ces dernières années, on s'est de plus en plus intéressé à la production d'estimations pour des régions plus petites que la taille permise par l'estimation directe à partir d'enquêtes par échantillonnage. Les organismes statistiques, avec la participation d'universitaires, ont mis au point diverses méthodes pour résoudre ce problème. Mais l'utilisation de ces méthodes pour la production d'estimations est encore assez limitée, en partie parce que la question demeure de savoir s'il convient pour un organisme gouvernemental de produire des estimations au moyen d'un modèle.

e) **Utilisation statistique de données administratives** Les dossiers administratifs ont été utilisés comme un moyen parmi d'autres de réduire le coût de la collecte de renseignements. Cette source présente, relativement à la couverture et à la qualité des données, un ensemble de problèmes qui diffèrent de ceux que posent les enquêtes. Bien que les données administratives puissent être utilisées seules pour produire des données statistiques, on les utilisera plus efficacement en les conjuguant à des données d'enquête ou de recensement dans des systèmes d'estimation qui tirent parti des avantages relatifs de chaque type de données. La plupart des travaux dans ce domaine ont été effectués par des organismes statistiques d'État.

5. L'avenir

Considérant l'avenir, Bailar prévoit une utilisation accrue des modèles. Il est presque impossible de douter de la justesse de ce pronostic, à une époque où les organismes statistiques cherchent de plus en plus à tirer le maximum d'information de données existantes et à contenir l'augmentation du coût de la collecte de données. Bailar mentionne en particulier l'intégration des méthodes des séries chronologiques aux méthodes d'estimation, question actuellement étudiée dans plusieurs organismes statistiques. J'ajouterai trois autres domaines dans lesquels on peut espérer des progrès importants à long terme, ces progrès supposant dans chaque cas une interaction entre la statistique et d'autres disciplines.

Le premier domaine est l'application de systèmes experts à certaines activités des organismes statistiques d'État. Pour prendre un exemple dont nous avons déjà parlé, le choix des options ou des modèles qu'il convient d'utiliser pour la désaisonnalisation des séries chronologiques pourrait se prêter à une approche de ce genre. Le deuxième domaine est l'utilisation des méthodes cognitives pour comprendre et améliorer le processus de réponse. Plusieurs

organismes statistiques travaillent sur cette question. Avec l'aide de la psychologie, les statisticiens pourraient s'inspirer de ces méthodes pour mettre au point un meilleur modèle du processus de réponse – processus qui est sans doute la moins bien comprise des composantes de l'enquête. Le troisième domaine est le développement de systèmes intégrés d'information statistique alliant des modèles de structures sociales et économiques à des bases de données et permettant de simuler les effets de diverses politiques. De tels systèmes faciliteraient l'utilisation des données d'un organisme pour analyser des politiques et aideraient cet organisme à déceler les lacunes de données dans ses programmes.

Pour faire écho à la conclusion de Bailar, je dirai que les problèmes sont passionnants et qu'il n'en manque pas.