

# Les besoins en matière de formation des statisticiens d'enquête dans les pays développés et dans les pays en développement

Graham Kalton

Westat, États-Unis

La présente communication a été faite lors du colloque de 2002 de la Section des méthodes de recherche-sondage de l'ASA, durant une session invitée, dédiée à la mémoire de Leslie Kish

## 1. Introduction

La session au cours de laquelle nous faisons la présente communication est intitulée "La formation des statisticiens d'enquête: une session dédiée à la mémoire de Leslie Kish". Le thème de la session est tout à fait approprié, compte tenu du dévouement pour la formation des statisticiens d'enquête, et en particulier ceux des pays en développement, dont a fait preuve Kish tout au long de sa vie. Il est également approprié que la session soit coparrainée par l'Association internationale des statisticiens d'enquêtes (AISE) et par quatre sections de l'ASA: Statistiques relatives aux politiques en matière de santé, Méthodes de recherche-sondage, Statistiques gouvernementales et Statistiques sociales. Kish était un membre très important du comité qui a fondé l'AISE; il avait été président tant de l'AISE que de l'ASA.

Kish avait créé, en 1961, au Survey Research Center de l'Université of Michigan, un important programme de formation à l'intention des statisticiens d'enquête issus de pays en développement. Ce programme d'été très réussi d'une durée de deux mois qui porte le nom de Sampling Program for Survey Statisticians—SPSS (Programme d'échantillonnage pour statisticiens d'enquête) existe toujours et a permis de former jusqu'à maintenant plus de 500 statisticiens d'enquête provenant de plus de 100 pays. Le thème que Kish avait choisi pour la quatrième conférence commémorant Morris H. Hansen (*Morris H. Hansen Memorial Lecture*), soit "Former des spécialistes de l'échantillonnage pour les pays en développement" indique l'importance qu'il attachait à cette formation (Kish, 1996). Sa communication contient de nombreuses observations pertinentes sur les besoins en matière de formation des statisticiens d'enquête dans les pays en développement.

Dans la présente communication, nous nous penchons sur les besoins en matière de formation qu'ont les statisticiens d'enquête, tant dans les pays développés que dans les pays en développement. Le terme "statisticien d'enquête" désigne ici les personnes qui s'occupent de la conception de plans d'enquête et de l'analyse

statistique d'enquête, principalement les spécialistes de l'échantillonnage, et n'a pas le sens plus général qui englobe toutes les personnes participants à un quelconque aspect de la méthodologie d'enquête. Dans le cas des pays développés, nous nous concentrons sur les études qu'il faut suivre en vue de devenir statisticien d'enquête, principalement des études au niveau de la maîtrise. Dans le cas des pays en développement, nous mettons en évidence des programmes à court terme qui offrent une formation pratique dans le domaine de l'échantillonnage, à savoir des programmes comme le SPSS précité ainsi que des programmes de plus courte durée. Tout au long de la communication, nous mettons l'accent sur les aspects pratiques de l'échantillonnage, plutôt que sur les aspects théoriques.

Les besoins en éducation des statisticiens d'enquête dans les pays développés sont traités dans la section 3, tandis que les besoins en matière de formation des statisticiens d'enquête dans les pays en développement sont abordés dans la section 4. Comme introduction aux sujets traités dans ces deux sections, la section 2 contient des observations sur les besoins généraux en matière de formation qui existent dans le domaine de la recherche-sondage.

## 2. La formation relative aux méthodes de la recherche en sondage

De par sa nature, la recherche-sondage est une activité interdisciplinaire. Celle-ci est menée généralement dans le cadre d'une équipe dont les membres apportent leur expertise en ce qui a trait aux domaines visés par les enquêtes, à la conception de questionnaires, à l'échantillonnage, aux opérations des enquêtes, à la collecte des données, à l'informatique, au traitement des données et à l'analyse statistique. Au sein d'organismes d'enquêtes de grandes envergures, l'équipe d'enquête peut être composée de spécialistes de chacune des composantes des méthodes d'enquête. Dans le cas d'organismes d'enquêtes de petites tailles, les membres de l'équipe d'enquête peuvent avoir à s'occuper de plus d'une composante à la fois. Par exemple, il se peut que le membre chargé de la conception du

questionnaire gère également la collecte des données, et que le membre qui effectue la sélection de l'échantillon dirige aussi le traitement des données et l'analyse d'enquête. Mais quel que soit le nombre de composantes de l'enquête dont sont chargés les membres de l'équipe, ceux-ci doivent posséder une compréhension de base de toutes les composantes. La conception et la mise en œuvre efficaces d'une enquête nécessitent une intégration attentive des diverses composantes de celle-ci. Les membres de l'équipe d'enquête doivent donc être en mesure de travailler de concert pour concevoir un plan d'enquête qui peut être mis en œuvre avec succès. Chacun des membres de l'équipe devrait concevoir sa composante pour qu'elle soit la meilleure possible dans le contexte du plan d'enquête global, plutôt qu'optimale dans une optique plus restreinte. Afin de pouvoir interagir efficacement avec les autres membres de l'équipe d'enquête, chaque membre doit connaître les aspects que comporte le plan d'enquête dans son ensemble.

Une façon alternative de démontrer la nécessité, pour les membres de l'équipe d'enquête, de comprendre l'ensemble du processus d'enquête est à l'aide de la notion, largement utilisée, d'erreur totale dans l'enquête. Kish (1965, chapitre 13) décrit les différentes sources d'erreurs relatives aux estimations d'enquêtes, y compris l'erreur d'échantillonnage et les erreurs non liées à l'échantillonnage relatives à la non-observation (non-couverture et non-réponse) et à l'observation (erreurs liées à la collecte et au traitement des données). La conception de l'enquête nécessite souvent que l'on fasse des compromis entre les diverses sources d'erreurs. Ainsi, par exemple, une taille d'échantillon plus importante permet de réduire l'erreur d'échantillonnage, mais peut-être au prix d'un plus grand taux d'erreur liée à la non-réponse et à la collecte des données; ou alors, un protocole de collecte de données plus exigeant peut réduire l'erreur de mesure, mais au prix d'un taux d'erreur de non-réponse plus élevé. Tous les membres d'une équipe d'enquête doivent connaître ces interdépendances. Une partie importante de la formation que reçoivent les spécialistes de la recherche-sondage devrait donc porter sur la façon dont le domaine de spécialisation de chacun s'intègre à l'ensemble de l'activité d'enquête.

Il existe un manque manifeste de programmes d'études et de formation en recherche-sondage de par le monde. Peu d'universités offrent de tels programmes ou sont en mesure de les offrir. Les universités disposent rarement de tout le personnel enseignant nécessaire, et la plupart des établissements ne comptent pas dans leur corps professoral des personnes possédant la vaste expérience qu'il faut avoir pour enseigner

efficacement les aspects pratiques de la recherche-sondage. Reconnaisant la pénurie de programmes de ce genre aux États-Unis, les chefs de l'organisme fédéral de statistique, le chef du *Statistical Policy Office* de l'organisme américain *Office of Management and Budget* ainsi que le président du *Council of Economic Advisors* ont décidé, en 1992, de créer un centre pour les méthodes d'enquêtes qui offrirait des programmes d'études et de recherche du deuxième cycle universitaire, afin de répondre aux besoins des organismes fédéraux. Dans le cadre de l'élaboration de ce projet, le *Committee on National Statistics* (CNSTAT) du *National Research Council* avait organisé un atelier portant sur le centre proposé. Le sommaire des travaux de cet atelier (Wolf, 1992) contient un grand nombre d'idées importantes en ce qui a trait aux besoins en matière de formation des spécialistes de la recherche-sondage; ces idées ont été exprimées tant par des personnes travaillant dans ce domaine au sein d'organismes fédéraux que par des membres du corps professoral chargé de l'enseignement de la recherche-sondage. Quelques-unes de ces considérations sont mentionnées ci-après.

- ◆ Il faudrait mettre l'accent sur la nature interdisciplinaire de la recherche-sondage et sur la formation requise pour cette activité, comme nous l'avons mentionné précédemment.
- ◆ Le centre devrait offrir aussi bien des programmes menant à un diplôme que des programmes ne menant pas à un diplôme.
- ◆ La formation pourrait comprendre des cours obligatoires, des cours de spécialisation et des cours au choix.
- ◆ Une formation en apprentissage ou des stages pratiques sont importants pour acquérir une expérience de la réalité.
- ◆ Il faut un mécanisme permettant de répondre aux nouveaux besoins en matière de formation.
- ◆ Le programme offert par le centre peut comprendre un large éventail de cours.

Le centre en question a été créé à l'Université du Maryland en 1992, à l'issue d'un concours public. Le programme qui y est offert est le fruit d'une collaboration de l'Université du Maryland, de l'Université du Michigan et de Westat et porte le nom de *Joint Program in Survey Methodology* (JPSM). La façon dont le JPSM a mis en œuvre les idées mentionnées précédemment est décrite sur le site Web du programme, à l'adresse suivante: <http://www.jpsm.umd.edu>. Le JPSM a mis sur pied des programmes d'études de maîtrise et de doctorat, des programmes ne menant pas à un diplôme (mention et certificat de deuxième cycle)

ainsi que des cours de brève durée sans crédits. Le programme de maîtrise du JPSM exige 46 heures-crédits. Il comprend deux concentrations (statistique et sciences sociales) et tous les étudiants inscrits doivent suivre certains cours obligatoires.

Il est utile de mentionner la collaboration sur laquelle repose le JPSM. Cette collaboration tripartite rassemble la gamme d'enseignants et d'expérience nécessaire à la mise en œuvre du programme; en outre, le programme bénéficie de contributions en matière d'enseignement apportées par du personnel d'organismes fédéraux. Au Royaume-Uni, on trouve également une collaboration, pour la formation dans le domaine de la recherche-sondage, entre une université, un organisme de recherche-sondage et le gouvernement. Ainsi, le Département de la statistique sociale de l'Université de Southampton offre des programmes de diplôme et de maîtrise en statistique sociale et en statistiques officielles; le programme de maîtrise est offert en collaboration avec l'*Office for National Statistics*. Le département collabore également avec un organisme de recherche-sondage—le *National Centre for Social Research*—ainsi qu'avec le *Centre for Applied Social Surveys* (CASS) de l'Université de Surrey, qui offre des cours de brève durée portant sur les méthodes d'enquête. Pour plus de renseignements, on se reportera à l'adresse Internet <http://www.socstats.soton.ac.uk>.

### **3. Programmes d'études pour les statisticiens d'enquêtes**

La présente section porte sur les études universitaires du deuxième cycle pour les statisticiens d'enquête (principalement, les spécialistes de l'échantillonnage), et en particulier sur les questions qui ont trait à l'établissement d'un programme d'études de maîtrise. À l'instar de Kish (1996), nous mettons l'accent sur l'échantillonnage et sur le plan d'échantillonnage, et non sur les fondements de la théorie de l'échantillonnage.

Comme Kish également, nous privilégions en général un programme flexible, afin de répondre aux besoins d'un large éventail d'étudiants. Certains étudiants peuvent vouloir se concentrer sur l'échantillonnage, d'autres peuvent vouloir suivre un programme plus général sur les méthodes d'enquête qui met l'accent sur l'échantillonnage, et d'autres encore peuvent souhaiter combiner l'échantillonnage avec un domaine en particulier (par exemple, l'agriculture, l'éducation, la santé ou l'économie). Certains étudiants peuvent viser des postes de cadre au cours de leur carrière, alors que d'autres peuvent rechercher une éducation qui les prépare à une carrière consacrée entièrement au domaine de

l'échantillonnage. Le programme de maîtrise destiné à ce dernier groupe d'étudiants doit transmettre des bases solides sur lesquelles on peut s'appuyer à mesure que de nouveaux développements surviennent dans le domaine de spécialisation.

Nos critères pour la mise sur pied d'un programme de maîtrise à l'intention des statisticiens d'enquêtes sont les suivants:

- ◆ une base solide en ce qui a trait à la théorie de la statistique et aux méthodes statistiques;
- ◆ une formation spécialisée dans la théorie et la pratique de l'échantillonnage;
- ◆ une connaissance étendue de l'ensemble du processus d'enquête;
- ◆ l'acquisition de solides compétences en informatique; et
- ◆ une formation en communication orale et écrite.

Nous accordons la priorité à la théorie de la statistique et aux méthodes statistiques parce que ces éléments forment les assises pour l'échantillonnage. Les étudiants ont besoin d'une base solide dans ce domaine, afin de pouvoir comprendre les textes spécialisés d'actualité qui portent sur l'échantillonnage. Ils en auront sans doute besoin encore davantage pour mettre à niveau leurs compétences à mesure que l'on fera des progrès dans le domaine de l'échantillonnage dans l'avenir. À titre d'illustration, on peut mentionner les développements observés récemment en matière d'estimations régionales et d'estimation de la variance dans le cas d'ensembles de données contenant des valeurs imputées. Ces deux domaines montrent comment on a recours à des méthodes statistiques plus puissantes dans le cadre de nouvelles méthodes d'échantillonnage. Aujourd'hui, pour acquérir une solide base de connaissances en ce qui a trait à la théorie de la statistique et aux méthodes statistiques, il faut au moins quatre cours (deux consacrés au calcul des probabilités et à la théorie de la statistique, et deux portant sur les méthodes statistiques), et même cette exigence semble déjà inadéquate pour constituer le savoir fondamental nécessaire pour aborder les méthodes récentes utilisées dans le domaine de l'échantillonnage. Afin de se préparer pour l'avenir, il faudra probablement davantage de cours.

L'échantillonnage est maintenant un domaine très développé où l'on trouve une documentation riche et un large éventail de méthodes. Un statisticien d'enquête qui exerce son métier a besoin d'au moins trois cours de base portant sur ce domaine (par exemple, méthodes d'échantillonnage

appliquées, théorie fondamentale et estimation de la variance); il y a aussi un grand nombre de sujets particuliers qui, idéalement, devraient être traités d'une façon ou d'une autre. La liste qui suit contient une partie des sujets importants qui ne peuvent être traités adéquatement, en supposant qu'ils soient même mentionnés, dans les cours fondamentaux:

- ◆ non-réponse et ajustements de poids par collage
- ◆ méthodes de vérification et d'imputation
- ◆ conception et estimation relatives aux enquêtes au fil du temps
- ◆ modèles d'erreurs de mesure
- ◆ échantillonnage de populations rares
- ◆ estimations régionales
- ◆ échantillonnage pour enquêtes-établissements
- ◆ conception et analyse d'études d'observation
- ◆ prévention de la divulgation
- ◆ traitement des valeurs aberrantes
- ◆ coordination d'échantillons au fil du temps.

Un certain nombre de ces sujets peuvent être traités dans le cadre d'un quatrième cours sur l'échantillonnage, mais on ne peut les aborder tous, et ceux qui le sont ne peuvent être étudiés en profondeur. Plusieurs de ces sujets justifient un cours de trois crédits (où même plus d'un cours) pour pouvoir être traités de manière exhaustive, tandis que d'autres sujets justifient des cours d'une ou de deux crédits. On pourrait envisager des cours supplémentaires, mais il est probable que l'on ferait face alors aux contraintes liées au nombre maximum d'heures-crédits prévus pour un programme de maîtrise. Avec un seul cours portant sur ces sujets, il faut faire un choix difficile: essayer de sensibiliser les étudiants à tous les aspects des méthodes d'échantillonnage qu'ils pourraient devoir connaître dans leur travail, ou laisser les étudiants choisir un nombre restreint de sujets à étudier plus en profondeur.

Nous avons justifié dans la section précédente le besoin pour tous les spécialistes de la recherche-sondage d'acquérir une connaissance étendue de l'ensemble du processus d'enquête. Cette justification s'applique clairement aux statisticiens d'enquête, qui doivent concevoir des échantillons qui s'intègrent efficacement aux autres caractéristiques des plans d'enquête. L'acquisition d'une connaissance étendue du processus d'enquête requiert au moins deux ou trois cours de trois crédits chacun (par exemple, méthodes de collecte de données, sources d'erreurs dans les enquêtes) ainsi qu'un ou deux cours supplémentaires, si les études comprennent un

stage pratique. Des formations en apprentissage choisies avec soin, qui permettent de se familiariser avec l'ensemble du processus d'enquête, pourraient également répondre au besoin important de donner aux étudiants un aperçu des considérations liées à l'ensemble de la conception d'enquête dont il faut tenir compte dans la pratique.

La nécessité de posséder des compétences en informatique, et notamment en ce qui a trait à l'utilisation de logiciels de statistique, est évidente. L'informatique est présente dans la plupart des cours modernes portant sur les méthodes statistiques, et elle peut également être incluse dans des cours pertinents sur l'échantillonnage. Cela devrait normalement être suffisant, pourvu que l'on tienne compte de l'informatique dans l'élaboration des cours.

Une lacune largement reconnue que présente une grande partie de la formation en statistique est l'attention insuffisante accordée à la communication orale et écrite. De solides compétences en communication sont essentielles pour tous les statisticiens en statistique appliquée, et à plus forte raison pour les statisticiens d'enquête. Dans une communication portant sur la préparation des statisticiens à une carrière au sein du gouvernement fédéral, un comité de l'ASA a mentionné le besoin de posséder des compétences en communication: "On ne saurait accorder trop d'importance aux capacités de communication écrite. Lorsque des statisticiens accèdent à des postes de gestionnaires, la communication, tant orale qu'écrite, devient un élément essentiel de la réussite professionnelle." (Eldridge et coll., 1982, p. 76). Une communication intitulée "Préparer les statisticiens se destinant à une carrière à Statistique Canada", de J. Denis et coll. (2002), contient des recommandations à l'intention des universités, concernant la préparation des étudiants au travail effectué au sein d'un organisme de statistique, qui concordent avec les critères proposés précédemment, y compris une recommandation concernant la communication écrite:

- ◆ "Offrir des bases solides en statistique";
- ◆ "Donner aux étudiants l'opportunité de travailler sur des problèmes 'pratiques'";
- ◆ "Suscite l'intérêt pour l'apprentissage de logiciels statistiques"; et
- ◆ "Accorder une plus grande importance à la communication écrite".

La meilleure façon d'améliorer les compétences en communication orale et écrite des étudiants consiste à exiger de ceux-ci la présentation d'exposés et la rédaction de rapports dans le cadre

d'un grand nombre de cours. La réalisation d'exposés d'une durée de quinze minutes accompagnés de la présentation de diapositives préparées avec soin, et ce, dans plusieurs cours, peut aider les étudiants à perfectionner leurs capacités de communication orale et à s'exercer dans la présentation d'exposés lors de réunions professionnelles. On peut former les étudiants à la rédaction de rapports dans le cadre de travaux de cours, mais pour que cette formation soit efficace, des instructeurs ou des spécialistes de la rédaction doivent consacrer suffisamment de temps à l'écriture plutôt qu'aux aspects techniques des rapports. Dans certains cas, un cours distinct sur les éléments de la rédaction de rapports pourrait être à conseiller.

Dans l'élaboration d'un programme de maîtrise à l'intention des statisticiens d'enquête, le défi consiste à réaliser tous les objectifs précités tout en respectant une charge de cours raisonnable. Compte tenu des contraintes inévitables liées à la durée d'un programme de maîtrise, il faut établir des priorités. L'ensemble de priorités initiales que nous avons mentionné précédemment a trait à l'acquisition de connaissances de base solides en ce qui a trait à la théorie statistique et aux méthodes statistiques (au moins quatre à six cours), à une étude approfondie des méthodes d'échantillonnage (au moins quatre cours) et à un survol de l'ensemble du processus d'enquête (au moins deux cours). Pour la suite, nous privilégions de la flexibilité, afin de permettre aux étudiants de choisir des sujets supplémentaires dans n'importe lequel de ces domaines, ou dans tout autre domaine pertinent, comme un domaine d'application important. De façon générale, les étudiants qui souhaitent faire carrière comme spécialistes de l'échantillonnage et ceux qui veulent obtenir un doctorat en échantillonnage auraient intérêt à choisir davantage de cours portant sur la théorie de la statistique et sur les méthodes statistiques, et davantage de sujets liés à l'échantillonnage.

Quelle que soit la structure d'un cours de maîtrise, on ne peut s'attendre à ce qu'un tel cours traite de manière détaillée de tous les sujets relatifs à l'échantillonnage sur lesquels auront à se pencher les statisticiens d'enquête durant leur travail. En outre, à mesure que le domaine de l'échantillonnage progressera, de nouvelles méthodes verront le jour. Par conséquent, nous considérons qu'il y a un grand besoin de programmes de formation continue pour les statisticiens d'enquête, afin que ceux-ci puissent perfectionner toutes leurs compétences. Les cours supérieurs de brève durée aident les participants à se familiariser avec l'utilisation de méthodes spécialisées, mais ils ne comportent pas de

moyens permettant de s'assurer que les participants entreprennent les études supplémentaires nécessaires dont ils ont besoin pour maîtriser entièrement la matière. Nous sommes d'avis qu'il faut des cours structurés plus longs comportant des devoirs et des examens ainsi que la délivrance d'une attestation reconnaissant la réussite. Ce genre d'éducation permanente, qui est requise dans certaines autres disciplines, est maintenant nécessaire également dans le domaine de l'échantillonnage. Espérons que les employeurs sauront reconnaître les avantages que comporte l'éducation permanente, qu'ils encourageront leur personnel à participer et qu'ils récompenseront ceux qui terminent les cours avec succès. Il est utile de noter que Statistique Canada investit beaucoup dans la formation professionnelle. Ainsi, cet organisme a créé son propre Institut de formation afin d'offrir des cours internes et s'attend à ce que les surveillants et les employés discutent d'un plan de formation pour l'année à venir au cours de la réunion annuelle portant sur le rendement (Denis et coll., 2002).

#### **4. Cours de formation à l'intention des statisticiens d'enquêtes des pays en développement**

Dans la présente section, nous nous penchons sur les besoins en matière de formation des statisticiens d'enquêtes dans les pays en développement. Les programmes de maîtrise décrits précédemment permettent de bien répondre aux besoins des statisticiens d'enquêtes des pays en développement. Cependant, l'envoi d'étudiants à l'étranger depuis ces pays pour suivre des programmes de maîtrise coûte cher et prend du temps. Il y a donc un grand besoin de programmes de formation de plus courte durée.

Parmi les modèles pour l'établissement de tels programmes de formation, il y a le SPSS (Programme d'échantillonnage pour statisticiens d'enquête) que Kish a élaboré il y a plus de quarante ans et que nous avons mentionné au début. Le programme offert par l'*International Program Center* du *U.S. Census Bureau* est un autre programme à long terme pour la formation de personnel d'organismes de statistique nationaux. Ce centre offre maintenant des ateliers d'une durée de deux à six semaines sur divers sujets, dont l'échantillonnage (<http://www.census.gov/ipc>).

L'AISE propose avec succès depuis de nombreuses années des cours de brève durée sur l'échantillonnage et d'autres aspects de la recherche-sondage, en association avec les sessions biennales de l'Institut international de statistique (IIS). En 1973, Kish avait lancé la tenue de réunions d'une demi-journée qui ont été transformées en ateliers de plus longue durée en

1987. Lors de la session de l'IIS qui eu lieu à Séoul en 2001, l'AISE avait proposé les cours suivants: Atelier sur l'échantillonnage, Estimation de la variance dans les enquêtes complexes, Introduction aux estimations régionales, La recherche sur les erreurs non liées à l'échantillonnage et Vérification et imputation de données d'enquête. L'atelier sur l'échantillonnage est conçu spécialement pour les statisticiens d'enquêtes issus de pays en développement. Avec le cours sur l'estimation de la variance, cet atelier offre une introduction générale au sujet.

Les cours comme ceux offerts par l'AISE sont utiles, non seulement parce qu'ils proposent une formation en matière d'échantillonnage, mais également parce qu'ils permettent à des statisticiens d'enquête de différents pays de partager leurs expériences. De tels cours devraient mettre l'accent sur les aspects pratiques de l'échantillonnage et aborder les notions théoriques de base uniquement dans la mesure nécessaire. Ils devraient traiter des méthodes d'échantillonnage de base (par exemple, l'échantillonnage systématique, la stratification, l'échantillonnage à plusieurs degrés, la sélection avec probabilité proportionnelle à la taille) et également de la non-réponse, de la non-couverture, de la pondération et de l'utilisation de logiciels pour l'estimation de la variance. Ils devraient aussi comprendre des exemples relatifs à des pays en développement. Des exercices simples peuvent être utilisés pour souligner les principales observations. Lorsque cela est possible, il faudrait inclure l'utilisation pratique de logiciels.

La principale difficulté que comporte la proposition de cours de formation à l'intention des statisticiens d'enquêtes des pays en développement, que ce soit des cours longs comme le SPSS ou des cours de brève durée comme ceux de l'AISE, a trait au financement. Le SPSS a été lancé avec des fonds provenant de la *Ford Foundation* et est maintenant financé en partie par le *Leslie Kish International Fellows Fund*. De nombreux statisticiens issus de pays en développement ont pu participer aux cours de l'AISE grâce à des bourses de la Division de statistique des Nations Unies. Dans les deux cas, ces sources assurent le financement de base pour le maintien de ces cours, avec le soutien de diverses autres sources.

Lorsqu'on forme des personnes issues de différents pays, la langue représente un facteur important. Les cours précités se donnent généralement en anglais, de sorte que les participants doivent être en mesure de comprendre

cette langue. Toutefois, il peut être utile de disposer d'une traduction du contenu des diapositives présentées lors des cours, ou du moins des principaux termes, pour le moins lorsqu'il y a seulement une seule autre langue importante outre la langue d'enseignement. L'utilisation de deux projecteurs (un pour chaque langue) peut aider les participants qui ne connaissent pas très bien la langue d'enseignement.

Il semble particulièrement souhaitable d'avoir un programme de formation offert dans un pays en développement, à l'intention des statisticiens de ce pays et des pays voisins. Un programme régional d'une durée d'une ou deux semaines pourrait être une utilisation très efficace de ressources limitées. Dans l'analyse qu'il a été invité à faire de la présente communication, Hermann Habermann a proposé une amélioration de ce programme pour en faire une série d'activités de formation s'échelonnant sur une période de plusieurs années. Nous sommes tout à fait en faveur de cette extension, que nous considérons comme une façon de développer les ressources de formation internes de la région, où les instructeurs locaux pourront prendre la relève après quelques années.

Il y a un certain nombre d'années, l'Enquête mondiale sur la fécondité, dirigée par Sir Maurice Kendall pour le compte de l'Institut international de statistique, ainsi que le Programme de mise en place de dispositifs nationaux d'enquêtes sur les ménages (de la Division de statistique des Nations Unies) ont fait des contributions importantes à l'établissement d'infrastructures d'enquêtes dans de nombreux pays en développement. Depuis lors, un nombre important de programmes d'enquêtes internationales a été mis sur pied, mais malheureusement, aucun de ces programmes ne comporte une mission de formation en statistique d'enquêtes ou en méthodes d'enquêtes. On y gagnerait beaucoup, si on pouvait convaincre les responsables de ces programmes de consacrer une faible partie de leur budget à la formation et au renforcement des capacités.

Les programmes mentionnés dans la présente section jouent un rôle très utile pour la formation des statisticiens d'enquêtes dans les pays en développement. Cependant, il reste encore beaucoup à faire. L'ASA et l'AISE pourraient-elles travailler de concert pour soutenir de tels programmes de formation? Ces deux associations ont à cœur l'amélioration de l'utilisation des statistiques, et des efforts dans le domaine de la formation contribueraient à cette amélioration.

## 5. Références

Denis, J., Dolson, D., Dufour, J., and Whitridge, P. (2002), Preparing statisticians for a career at Statistics Canada, (Working Paper No. HSMD-2001-006E/F), Ottawa: Statistics Canada.

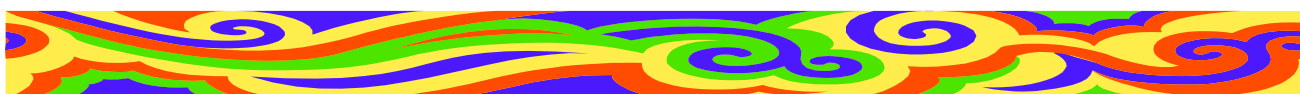
Eldridge, M.D., Wallman, K.K., Wulfsberg, R.M., Bailar, B.A., Bishop, Y.M., Kibler, W.E., Orleans, B.S., Rice, D.P., Schaible, W., Selig, S.M., and Sirken, M.G. (1982), Preparing statisticians for careers in the Federal Government: Report of the ASA Section on Statistical Education Committee on

Training of Statisticians for Government. *American Statistician*, 36, 2, 69-89.

Kish, L. (1965), *Survey sampling*, New York: Wiley.

Kish, L. (1996), Developing samplers for developing countries. *International Statistical Review*, 64, 143-162.

Wolf, F.E. (Ed.). (1992), *Center for Survey Methods: Summary of a workshop*, Washington D.C.: National Academy Press.



## Articles spéciaux: Recensements effectués dans divers pays

### **Incidence potentielle de la modification des questions sur les données et leur utilisation: étude de cas de deux recensements sud-africains (1996 et 2001)**

Avertissement: Les opinions exprimées dans le présent document sont celles de l'auteur et ne représentent pas nécessairement la position du Bureau de la statistique de l'Afrique du Sud.

La production de données exactes est très importante dans toute société en mutation rapide. Les décideurs, planificateurs, gens d'affaires et bien d'autres ont quotidiennement recours aux données qu'on leur fournit. L'utilisation de ces données a une incidence directe sur notre vie de tous les jours. Il importe donc que toutes les personnes qui utilisent des données sous une forme ou une autre comprennent bien ces renseignements et leurs limites. Ce processus commence au tout début, soit au moment de la conception de l'instrument de collecte de données.

Selon Bradburn (2000), "depuis le début des enquêtes scientifiques, les chercheurs savent que la modification du libellé des questions peut produire des changements considérables dans la répartition des réponses" [traduction libre]. Le présent article vise à examiner brièvement l'incidence des différences dans la formulation, l'ordre et le contenu des questions sur les données définitives qui sont produites. Le document examine également d'autres facteurs susceptibles d'entraîner des réponses différentes dans le cadre d'une enquête. L'analyse de ces questions est fondée sur la comparaison entre le dernier recensement sud-africain de 2001 et celui de 1996.

#### **Contexte**

Les données recueillies au fil du temps servent à établir des comparaisons permettant de cerner les changements, de dégager des tendances et de faire des prévisions. Par conséquent, la compatibilité des différents ensembles de données est fort importante. Les chercheurs doivent être conscients des différences entre deux ensembles de données et de l'incidence potentielle de ces différences sur la qualité des données. Les différences entre les méthodologies employées pour recueillir les données peuvent avoir une

incidence sur la mesure de certaines variables ou sur l'application des données à des secteurs de petite taille. Les écarts de mesure des variables peuvent également être attribuables aux différences dans la conception des questionnaires. Il est donc essentiel de mener des études mettant en relief ces différences et leurs répercussions. De telles recherches permettent d'établir le fondement des travaux d'autres chercheurs qui exploitent les données. Si les données comportent des problèmes majeurs, les utilisateurs éventuels doivent en être informés pour qu'ils puissent décider d'avoir recours ou non à ces renseignements (Jaffe, 1951).

Le Bureau de la statistique de l'Afrique du Sud (Stat AS) est l'organisme officiel chargé de la réalisation des recensements nationaux de la population en Afrique du Sud. En vertu de la *Loi sur la statistique*—n. 6 de 1999, Stat AS est tenu de réaliser un recensement tous les cinq ans. En octobre 2001, Stat AS a mené le deuxième recensement de la population depuis les premières élections démocratiques de 1994. Le premier recensement remonte à octobre 1996. Avant de commencer à utiliser les deux ensembles de données, il convient d'effectuer une analyse rigoureuse des différences entre ces bases d'information. Plusieurs facteurs peuvent avoir des effets sur cette comparaison. Cependant, le présent article portera principalement sur les enjeux relatifs aux questionnaires utilisés.

#### **Libellé et ordre des questions**

McLaughlin (1999) explique de façon détaillée comment le contexte des questions peut conditionner les réponses. Elle soutient que le contexte d'une question, notamment le format, le libellé, les catégories de réponse ou l'ordre des questions, peut avoir des effets fortuits (ou intentionnels) sur les réponses obtenues.

L'un des principaux facteurs permettant de recueillir des réponses convenables auprès des répondants est le libellé des questions. Par exemple, les questions "Quel âge avez-vous?" et "En quelle année êtes-vous né?" n'ont qu'un seul mot en commun, mais elles permettent de recueillir les mêmes renseignements. Les questions "Quand êtes-vous né?" et "Où êtes-vous né?" ne se distinguent que par un seul mot. Pourtant, elles donneront lieu à des réponses complètement

différentes. La formulation des questions représente ce que le répondant doit traiter pour comprendre quelle information on cherche à obtenir. En règle générale, le libellé doit être simple, clair et éviter la forme négative.

Il importe d'établir que les répondants comprendront la formulation utilisée. Il arrive que les expressions argotiques, propres à une culture particulière ou les termes techniques ne permettent pas au chercheur de mesurer ce qu'il entendait mesurer au début de son étude. La formulation peut amener le répondant à répondre d'une certaine manière. On doit éviter l'emploi du mot "et" dans la formulation pour ne pas poser de questions doubles. Il importe également d'éviter les formulations négatives, particulièrement si le répondant est appelé à répondre par "oui" ou par "non".

Les questions fermées, soit celles qui proposent aux répondants des catégories de réponse, facilitent le traitement des données d'enquête. Si on peut répondre plus rapidement à de telles questions, les répondants peuvent, en revanche, y répondre trop vite et faire des erreurs. En outre, les catégories de réponse définies ne correspondent pas toujours à ce qu'aurait choisi le répondant. Dans le cas des questions fermées, les options doivent couvrir toute la gamme des réponses possibles et s'exclure mutuellement. Les questions ouvertes ne proposent aucune réponse prédéfinie au répondant. Les répondants doivent donc fournir leur propre réponse, exprimée dans leurs propres mots, ce qui peut se traduire par des renseignements plus probants et par l'expression d'opinions différentes mais intelligentes. Le traitement des données découlant de ces questions est souvent plus difficile, le codage des réponses présentant de nombreux problèmes.

Outre le libellé du questionnaire, l'ordre des questions peut conditionner les réponses obtenues. Les répondants sont souvent influencés par les stimuli visuels ou autres, et il importe donc d'établir l'ordre des questions de manière à obtenir des réponses positives. Il semble également que les gens tendent à se souvenir davantage du premier et du dernier élément d'une liste que des éléments intermédiaires, ce qui peut entraîner de mauvaises perceptions lorsqu'il s'agit de répondre selon des critères d'accord-désaccord ou d'appétence-aversion. Bradburn (2000) analyse le modèle de Tourangeau et Rasinski qui conceptualise le processus de questions et de réponses en tant qu'une des étapes nécessitant un traitement différent et susceptible de subir divers types d'effets de contexte. Ce modèle s'articule autour de l'interprétation, du repérage, du jugement et de la réponse. La question "Combien d'enfants avez-

vous?" posée à des enseignants illustre bien ce phénomène. Dans le premier cas, la question était précédée de questions sur la famille des répondants, et les enseignants l'ont interprétée en relation avec leur famille. Dans le second cas, la question était précédée de questions liées à l'école, et les enseignants l'ont interprétée comme portant sur le nombre d'enfants dont ils avaient la charge en classe. Cette question, bien qu'elle ait été formulée exactement de la même façon, a donné lieu à des réponses variant de 0 à 4 dans le premier cas et de 25 à 40 dans le second.

Il est possible de distinguer deux types d'effets de l'ordre: les effets d'assimilation et les effets de contraste. Les premiers se produisent lorsque les répondants ajustent leurs réponses à des questions ultérieures pour que les réponses fournies à un ensemble de questions soient cohérentes. Les effets de contraste se produisent lorsque les répondants omettent ou soustraient des renseignements relatifs à une question antérieure dans leur réponse à une question subséquente (Gendall, Carmichael et Hoek, 1997).

Traditionnellement, la conception et la présentation étaient établies suivant que le questionnaire devait être rempli par le répondant lui-même ou par le truchement d'un intervieweur. Bien que la conception continue d'être orientée par ces facteurs, la conception de même que les questions sont maintenant conditionnées aussi par les technologies de traitement des résultats, notamment les technologies de balayage et de reconnaissance optique de caractères, de lecture optique de marques et de reconnaissance d'images. Ces techniques modernes guident parfois la conception du questionnaire pour ce qui est de la présentation et de la séquence. Il convient d'accorder une attention rigoureuse à la formulation et à l'ordre des questions lorsqu'on adopte des technologies modernes, pour s'assurer que les questions mesurent bien ce que les chercheurs entendaient mesurer au début de leurs travaux.

### **Sujets couverts**

Il arrive que les sujets couverts aient des effets sur les réponses. Dans les recensements, par exemple, les questions portant sur le revenu sont généralement fort controversées. Les données relatives à ces questions sont aussi généralement suspectes. Ces questions sont insérées vers la fin du questionnaire pour ne pas qu'elles influent sur l'état d'esprit du répondant. Le bien-fondé de l'inclusion de questions de cette nature suscite des débats périodiques. Les questions se rapportant à l'ethnicité ou aux groupes de population font aussi l'objet de débats houleux. Hurst (2000) décrit de façon détaillée le débat dans lequel s'est engagé Statistique Canada relativement à la question de

l'ethnicité. Des problèmes semblables se sont manifestés aux États-Unis et ont amené les autorités à modifier la question après le recensement de 1990; en raison de ces changements, il n'est plus possible d'établir de comparaison directe entre les données des différents recensements.

On se sert de plusieurs types de questions pour couvrir les sujets retenus. Notons à cet égard les questions filtres (également désignées par le terme "questions de sélection") qui dirigent les répondants vers les questions pertinentes, les questions binaires qui proposent aux répondants deux réponses possibles et les questions par catégories qui proposent aux répondants une liste de réponses possibles. Les questions ordinales sont associées à des réponses à échelle graduée comme l'échelle à cinq points de Likert, tandis que les questions liées sont présentées sous une même forme et donnent lieu à des réponses consignées dans une matrice.

#### **Autres facteurs ayant un effet sur les réponses**

Bradburn (2000), en collaboration avec Sudman, a mis au point un modèle permettant de cerner les sources des effets relatifs aux réponses. Le modèle est fondé sur la prémisse selon laquelle l'interview représente une forme particulière de conversation entre des personnes parlant un langage commun. L'interview est considérée comme un système microsocial dans lequel trois rôles sont réunis par la tâche commune représentée par le questionnaire. L'une des personnes (l'intervieweur) a le rôle de poser les questions, la deuxième (le répondant) a pour rôle de répondre aux questions, la troisième (le chercheur) définit les tâches. Ces trois acteurs ont un effet sur les réponses obtenues.

Le processus qui consiste à répondre à une question d'enquête s'appuie sur un certain nombre de processus cognitifs (compréhension, extraction de l'information, jugement et évaluation de l'information extraite, sélection d'une réponse pertinente), chacun de ceux-ci pouvant conduire à des variations et à des réponses erronées de la part du répondant. Le jeu de ces divers processus dépend de la nature du processus de réponse, du mode d'enquête (questionnaire rempli par le répondant ou par le truchement d'un intervieweur) ainsi que des caractéristiques du répondant.

Les variations et les réponses erronées à des questions d'enquête sont attribuables à divers facteurs, notamment la confusion, l'ignorance, la négligence ou le besoin de désirabilité sociale de la part du répondant; une formulation ambiguë des questions et des catégories de réponse; l'ordre ou le contexte de présentation des questions; le choix

et l'ordre de présentation des catégories de réponse; la méthode de collecte des données (Meadows, Abrams et Greene, 1998).

Le contexte social dans lequel se déroule l'interview joue aussi un rôle de premier plan en ce qui a trait aux résultats de l'interview et à l'exactitude des données. Le contexte social est lié, entre autres, au milieu physique dans lequel l'interview est réalisée. Par exemple, les répondants pourraient se sentir plus à l'aise de remplir un questionnaire à la maison que dans leur milieu de travail.

Les caractéristiques du répondant et de l'intervieweur constituent un autre facteur susceptible d'influer sur les réponses. Divers facteurs tels que la race, le sexe et l'âge peuvent avoir une incidence sur la propension à exprimer des attitudes. Le mode de réalisation de l'interview peut aussi avoir des répercussions. Par exemple, l'interview peut être menée par téléphone, par un questionnaire rempli par le répondant ou au moyen de diverses formes d'assistance informatique (Bradburn, 2000).

#### **Perspective sud-africaine questionnaires**

Il n'est pas inhabituel de changer, d'un recensement à l'autre, le nombre et le type de questions incluses dans le questionnaire, à la lumière de nouvelles initiatives stratégiques, de l'examen des questions redondantes, des nouveaux besoins des utilisateurs, de la situation sociale générale et des perceptions du public. La présentation du questionnaire de recensement s'est également modifiée en raison des techniques de conception.

Dans le cas des deux recensements sud-africains, l'élaboration du questionnaire s'est appuyée sur la participation active des utilisateurs et du comité consultatif du recensement. Les divers intervenants ont suggéré des questions devant être prises en considération. Des consultants internationaux ont aussi fourni des conseils spécialisés quant à la conception et à la présentation des questionnaires. Dans les deux cas, on a effectué des essais préparatoires et sur le terrain pour assurer la mise au point d'un questionnaire convivial. Les résultats des essais ont facilité l'élaboration de libellés adéquats pour les questions et ont fourni de l'information qualitative et quantitative sur la fiabilité des réponses ainsi que sur l'acceptation des questions.

Les décisions relatives aux méthodologies de traitement comptent parmi les autres facteurs exerçant une influence sur l'élaboration des questions. Ainsi, les progrès technologiques en matière de balayage et de reconnaissance

d'images de même que le codage assisté par ordinateur ont orienté la présentation et la conception du questionnaire.

Le questionnaire des deux recensements a été traduit dans les 11 langues officielles. Cependant, on relève des différences au chapitre de l'application et de la disponibilité des traductions d'un recensement à l'autre. Dans le cas du recensement de 1996, le matériel dans les 11 langues était mis à la disposition des répondants sur le terrain. Pour le recensement de 2001, les recenseurs avaient un livret de traduction que le répondant pouvait consulter pour lire les questions dans la langue de son choix. Les réponses étaient ensuite transcrites sur la version anglaise du questionnaire. Néanmoins, lors du recensement de 2001, le questionnaire traduit était aussi disponible sur demande. Pour le recensement de 1996, chaque langue était associée à une couleur distincte, tandis qu'en 2001, toutes les traductions étaient présentées dans la même couleur.

Les questionnaires suivants ont été utilisés pour le recensement de 1996:

**Questionnaire 1 (questionnaire des ménages):** Ce questionnaire s'adressait aux répondants vivant dans des ménages privés et dans des établissements offrant des services d'hébergement aux familles. Le questionnaire comportait 50 questions se rapportant aux personnes et 15 questions se rapportant aux ménages. Le questionnaire permettait de recueillir les renseignements visant neuf personnes au maximum.

**Questionnaire 2 (livret sommaire des établissements d'hébergement):** Ce questionnaire servait à dénombrer l'ensemble des personnes et des ménages vivant dans les établissements d'hébergement et comprenait neuf questions se rapportant à l'établissement.

**Questionnaire 3 (questionnaire des particuliers):** Ce questionnaire visait les personnes vivant dans les établissements d'hébergement et comprenait 49 questions (soit les questions posées aux ménages, sauf celle relative aux liens entre les personnes).

**Questionnaire 4 (livret du recenseur pour le dénombrement spécial):** Ce questionnaire permettait de recueillir des renseignements très élémentaires sur les personnes vivant dans des établissements tels que les hôtels, les prisons et les hôpitaux, de même que sur les sans-abri. Seulement six questions étaient posées à ces personnes. Le questionnaire comportait également neuf questions relatives à l'établissement.

Trois questionnaires distincts ont été employés dans le cadre du recensement de 2001. Le premier (questionnaire A) visait tous les ménages, y compris les ménages composés d'une personne seule, ainsi que les résidents des centres d'hébergement pour les travailleurs, des résidences pour les étudiants universitaires et des hôtels-résidences. Ce questionnaire permettait de recueillir des renseignements sur chacun des membres du ménage (41 questions), sur l'ensemble du ménage (1 question) de même que sur les services (13 questions).

Les deux autres questionnaires s'adressaient aux résidents d'établissements, d'hôtels pour touristes, de même qu'aux sans-abri. Les établissements regroupaient les hôpitaux, les établissements de soins pour enfants, les pensionnats, les couvents, les casernes des forces armées, les prisons, les salles communautaires et paroissiales de même que les camps de réfugiés. Le questionnaire B servait à recueillir des renseignements personnels sur les répondants (40 questions), tandis que le questionnaire C comportait des questions relatives à l'établissement (8 questions). Là aussi, le questionnaire relatif aux personnes ne comportait pas de question sur les liens et sur le ménage.

L'introduction des codes à barres constitue l'un des principaux changements apportés au recensement de 2001 par rapport à celui de 1996. Tous les questionnaires du recensement de 2001 portaient un code à barres unique imprimé sur chacune des pages.

### **Sujets couverts**

Les deux questionnaires couvraient des questions de base d'ordre démographique, social et économique. Certaines questions s'adressaient à chacun des membres du ménage. Par exemple, les répondants devaient indiquer leur âge, leur sexe, le groupe de la population auquel ils appartenaient, la langue parlée à la maison, leur niveau de scolarité (dans le cas des personnes âgées de plus de 5 ans), leur profession ou leur occupation (pour les personnes âgées de 5 ans et plus) ainsi que le revenu de chacune des personnes présentes dans le ménage le jour du recensement. Les autres questions se rapportaient à la situation du ménage dans son ensemble, par exemple, l'accès ou non du ménage à des services d'électricité et d'aqueduc.

Dans le cadre du recensement de 2001, des efforts concertés ont été déployés pour harmoniser les questions du recensement dans les États membres de la *Southern African Development Community*. On a tenu des réunions et des ateliers pour discuter un ensemble de questions de base. Cette démarche avait pour principal objectif de recueillir

des données comparables dans la région. Stat AS a inclus ces questions de base dans le questionnaire et les a complétées par des questions permettant de répondre aux besoins particuliers de l'Afrique du Sud en matière de données.

Le questionnaire de 2001 comportait, au total, 14 questions qui n'avaient pas été posées dans le cadre du recensement de 1996. Par ailleurs, 9 questions figurant sur le questionnaire de 1996 n'ont pas été reprises dans le recensement de 2001. Cela signifie que 23 questions ne peuvent être utilisées pour établir des comparaisons entre les données des deux recensements. Il conviendra de déterminer si l'inclusion ou l'exclusion de certaines questions accroît ou réduit la valeur de l'ensemble des données du recensement. On pourrait aussi examiner les raisons motivant ces exclusions. Il arrive que des questions soient redondantes; parfois, de nouvelles recherches mettent en relief la nécessité d'un changement d'orientation. L'exploitation des données peut également se traduire par l'abandon d'une question ou l'inclusion d'une nouvelle question dans le recensement suivant.

### **Perspectives d'avenir**

Bien des recherches devront être menées pour déterminer l'incidence des différences décrites précédemment sur les données. Certains des problèmes relevés pourraient être réglés par une vérification adéquate des données et ce, dès que les ensembles de données seront disponibles. Il faudra également réaliser des études sur l'effet d'autres facteurs, comme les différences culturelles et les réponses, de même que sur les données.

Toutes les questions discutées dans le présent article convergent vers l'adoption d'une démarche interdisciplinaire intégrale en vue de la sélection des sujets traités et de la conception du questionnaire, démarche axée sur l'apport de domaines aussi variés que la psychologie, la linguistique, la technologie de l'information, le graphisme et la statistique.

Il importe tout particulièrement que les sociétés caractérisées par une grande diversité culturelle accordent une attention spéciale à certains aspects de la conception du questionnaire et du libellé des questions pour veiller à ce que les données se rapportant aux divers groupes culturels soient cohérentes et comparables.

### **Références**

Bradburn, N.M. 2000. *Questionnaire Design: From Art to Science*. Paper delivered at the Fifth International Conference on Social Science Methodology, Cologne, Germany.

Gendall, P., Carmichael, V., and Hoek, J. 1997. A Test of the Conversational Logic Analysis Model of Question Order Effects. *Marketing Bulletin*, vol. 8, pp. 41-52.

Hurst, L. 2000. The Politics of People-Counting. Question on Ethnicity Haunts Census-Takers. *Toronto Star*. January 22.

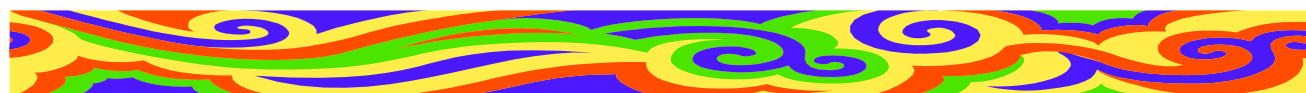
Jaffe, A.J. 1951. *Handbook of Statistical Methods for Demographers*. U.S. Government Printing Office, Washington, DC.

McLaughlin, M.E. 1999. Controlling Method Effects in Self-Report Instruments. *Research Methods Forum*, Vol. 4.

Meadows, K.A., Abrams, C., and Greene, T. 1998. The Application of Cognitive Aspects of Survey Methodology (CASM) in the Development of HRQOL Measures. *Quality of Life Newsletter*, <http://www.mapi-research-inst.com/index02.htm>.

### **Pour contacter l'auteur:**

Marius Cronjé  
Bureau de la statistique de l'Afrique du Sud  
[mariusc@statssa.gov.za](mailto:mariusc@statssa.gov.za)



## Forum de discussion

**Note de la rédaction:** Deux participants à une discussion sur le thème de substitution en ont donné leur point de vue. Nous invitons les autres membres à participer aussi à cette discussion. Envoyez-nous une brève description de la stratégie que vous préconisez pour traiter la substitution des unités de sondage. Nous aimerions savoir comment divers organismes des quatre coins du monde surmontent les difficultés que posent la non-réponse sur les enquêtes. Dans la première discussion, on utilise le terme “remplacement,” dans la traduction. Dans ce sens, il est utilisé exactement de la même manière que “substitution.”

### Remplacer ou ne pas remplacer—Voilà la question

par David W. Chapman

La non-réponse aux enquêtes entre généralement dans l'une de deux catégories: la non-réponse totale ou la non-réponse partielle. Dans le cas de la non-réponse totale, on n'obtient aucune donnée d'enquête utilisable de l'unité d'échantillonnage admissible. Dans le cas de la non-réponse partielle, on n'obtient pas de réponse à certaines questions de l'enquête (ce qui peut comprendre des rejets au contrôle), mais on obtient des réponses à suffisamment de questions pour inclure le participant comme répondant.

À ma connaissance, pour tenir compte de la non-réponse totale les chargés d'enquête ont recours à l'une des trois méthodes suivantes:

- ◆ Correction des poids des répondants;
- ◆ Imputation d'enregistrements complets; et
- ◆ Remplacement (ou remplacement d'une zone).

D'après mon expérience, la correction des poids est de loin la méthode utilisée le plus souvent pour tenir compte de la non-réponse totale, du moins dans le cas des enquêtes menées aux États-Unis. Selon cette méthode, les répondants dans une catégorie de pondération (qui peut être une strate ou une grappe) font l'objet d'une “pondération à la hausse” selon l'échantillon complet d'unités admissibles dans la catégorie de pondération.

Je ne connais pas beaucoup de cas de recours à la méthode de l'imputation bien qu'on l'utilise, je crois, dans le cas du formulaire abrégé du questionnaire du recensement décennal exécuté aux États-Unis (essentiellement, on applique la méthode “hot deck” à des enregistrements complets). Il s'agit peut-être d'une méthode acceptable pour corriger le biais de non-réponse, mais elle augmente la variance des estimations d'enquête puisque les poids des répondants “donneurs” sont au moins doublés.

Par remplacement ou remplacement d'une zone, j'entends la méthode qui consiste à remplacer chaque non-répondant dans un échantillon par une unité de population qui n'a pas été sélectionnée initialement dans l'échantillon et à traiter cette unité remplaçante comme si elle avait été sélectionnée dans l'échantillon original. L'objectif est de sélectionner des unités remplaçantes qui possèdent des caractéristiques d'enquête similaires à celles des non-répondants. D'après mon expérience, la plupart des statisticiens considèrent le remplacement comme une méthode inférieure (ou même inacceptable) utilisée pour tenir compte de la non-réponse.

Le présent article vise à décrire les trois types de méthodes de remplacement de base et à examiner les avantages et les inconvénients de chacun. Une bonne partie des observations formulées ici sont tirées d'un chapitre de mon livre intitulé *Incomplete Data in Sample Surveys* (Chapman, 1983). J'espère que le présent article suscitera de l'intérêt et une discussion du recours au remplacement pour tenir compte de la non-réponse totale.

À mon avis, le remplacement est une méthode sous-estimée parce qu'elle a été mal utilisée dans le passé. Je ne prétends pas qu'il s'agit de la meilleure méthode pour toutes les enquêtes, mais seulement qu'elle peut être appropriée dans de nombreux cas. J'ai l'impression qu'on n'a pas souvent recours au remplacement pour tenir compte de la non-réponse totale pour les enquêtes menées aux États-Unis. Toutefois, on a peut-être recours à cette méthode plus souvent dans d'autres pays. Dans son excellent article sur le remplacement, Vehovar (1999, p. 377) signale que, même si cette méthode est rarement recommandée dans les manuels son application est relativement répandue dans de nombreuses enquêtes sur échantillon probabiliste.

Les méthodes de remplacement utilisées entrent dans l'une de deux catégories de base:

- ◆ Sélection d'une unité de remplacement désignée spécialement
- ◆ Sélection d'une unité de remplacement aléatoire

Dans le premier cas, il s'agit d'une sélection raisonnée d'une unité de remplacement dont les caractéristiques correspondent le mieux à celles du non-répondant selon la base de sondage. Par exemple, dans le cas d'une enquête auprès des établissements, l'unité de remplacement pourrait être définie comme l'établissement (parmi ceux non sélectionnés dans l'échantillon) dont les caractéristiques correspondent le mieux à celles de l'établissement non-répondant dans la base de sondage, comme le type d'entreprise, le lieu géographique, le degré d'urbanisation, s'il s'agit d'un établissement du secteur public ou privé, à but lucratif ou non lucratif, le nombre d'employés et les revenus ou les bénéfices.

Dans le deuxième cas, il s'agit d'une sélection aléatoire afin d'obtenir une unité de remplacement parmi les établissements non sélectionnés dans l'échantillon. Habituellement, l'unité de remplacement est sélectionnée aléatoirement à partir d'une sous-population (par exemple, la même strate ou grappe que celle du non-répondant) afin d'accroître les chances de sélectionner une unité de remplacement possédant des caractéristiques similaires à celles du non-répondant. Pour les unités de remplacement sélectionnées aléatoirement, Vehovar (1999) calcule des expressions de l'effet sur le biais et la variance de l'utilisation d'unités de remplacement.

De ces deux méthodes, la sélection aléatoire d'unités de remplacement est probablement celle utilisée le plus souvent, du moins pour l'échantillonnage probabiliste (qui est le contexte de cet article). Cela tient à ce que la sélection aléatoire est une méthode plus scientifique, semble-t-il, et permet d'éviter d'autres biais qui pourraient résulter de la sélection raisonnée. Toutefois, pour ce qui est d'obtenir des données de remplacement pour un non-répondant, une unité de remplacement sélectionnée spécialement qui possède toutes les mêmes caractéristiques pertinentes de la base de sondage que le non-répondant peut être supérieure à une unité de remplacement sélectionnée aléatoirement à partir d'une sous-population définie par certaines des caractéristiques de la base de sondage.

On ne peut (sans modélisation créative) établir des expressions statistiques qui caractérisent l'effet sur le biais et la variance du remplacement que pour la méthode de sélection aléatoire des unités de remplacement. En ce qui concerne le remplacement aléatoire, Vehovar (1999, p. 238)

signale que tout biais additionnel associé au remplacement découlerait de différences entre les procédures utilisées pour communiquer avec les répondants et les recruter. Par exemple, on peut faire un plus grand nombre d'appels pour joindre les unités d'échantillonnage initiales que pour joindre les unités de remplacement. Il conclut que, même si cette différence existe probablement dans les procédures de collecte des données, le biais qui en découle est probablement "relativement faible".

Pour les comparaisons de la variance, Vehovar (1999, section 2.2) élabore une expression de l'augmentation de la variance pour la correction des poids comparativement au remplacement dans le cas de l'échantillonnage en grappes à deux degrés. Il conclut que le remplacement procure des avantages seulement dans des conditions très particulières qui ne se produisent pas très souvent. Il en arrive à des conclusions semblables pour les plans d'échantillonnage en grappes à plusieurs degrés et pour l'échantillonnage stratifié.

Pour ce qui est de l'effet du remplacement, j'ai résumé les résultats de quatre études empiriques que j'ai mentionnées dans mon étude antérieure (Chapman, 1983, section 3). En outre, lorsque j'étais à l'U.S. Census Bureau, j'ai participé à une étude empirique du remplacement pour une enquête à composition aléatoire de numéros de téléphone (Chapman et Roman, 1985). De plus, Vehovar (1999, section 2.3) a décrit une étude empirique qu'il a menée. Ces études empiriques n'ont pas produit de résultats décisifs quant à la valeur de l'utilisation du remplacement pour tenir compte de la non-réponse totale.

Par conséquent, pour décider s'il y a lieu d'avoir recours au remplacement dans le cas d'une enquête donnée, on doit sans doute surtout user de jugement et tenir compte de considérations d'ordre pratique. Par exemple, quand je travaillais à Westat en 1971-1972, j'ai participé à une enquête auprès des écoles élémentaires et secondaires dont l'un des objectifs importants était d'élaborer des liens d'équivalence entre plusieurs tests de rendement nationaux. À cette fin, les analystes avaient besoin qu'un nombre précis d'écoles participent à l'étude et exactement deux écoles de chaque strate. Le remplacement était la seule façon de procéder pour obtenir le type d'échantillon nécessaire aux fins d'analyse. Dans mon étude antérieure portant sur le remplacement (Chapman, 1983, p. 47), j'ai décrit la méthode de remplacement utilisée pour cette étude, qui reposait sur une combinaison d'unités de remplacement désignées spécialement (première priorité) et d'unités de remplacement sélectionnées aléatoirement.

Le recours au remplacement présente deux grands avantages. En premier lieu, dans la mesure où l'on peut trouver des unités de remplacement de tous les non-répondants, on préserve le plan d'échantillonnage et la taille de l'échantillon originaux. Ainsi, si le plan d'échantillonnage comprend des probabilités de sélection uniformes (cette à dire s'il est autopondéré), l'échantillon final aura les mêmes propriétés. Ainsi, l'analyse des résultats de l'enquête sera simplifiée. En outre, les variances des estimations de l'enquête seront vraisemblablement les mêmes que celles qui auraient été obtenues si le taux de réponse avait été de 100%.

Le deuxième avantage tient à la réduction éventuelle du biais lorsque le plan de sondage comprend de nombreuses caractéristiques, comme dans le cas de nombre d'enquêtes auprès des entreprises. Si on peut alors identifier des unités de remplacement dont plusieurs caractéristiques correspondent (ou presque) aux caractéristiques du non-répondant dans la base de sondage, le biais de non-réponse pourrait être réduit, comparativement à celui associé à la correction des poids, lorsque les valeurs de l'enquête manquantes sont imputées effectivement par une combinaison des valeurs des répondants dans la même catégorie de pondération. Bien entendu, les variables de la base de sondage doivent être en corrélation étroite avec les variables de l'enquête pour procurer pareils avantages.

Le recours au remplacement présente toutefois plusieurs désavantages éventuels qu'il faut prendre en compte. En premier lieu, la façon dont les unités de remplacement sont sélectionnées pourrait introduire un biais dans les estimations de l'enquête, particulièrement si l'on permet aux intervieweurs de choisir les unités de remplacement. Pour éviter pareils biais, il faut élaborer des critères bien définis pour identifier les unités de remplacement. Idéalement, les unités de remplacement seraient sélectionnées au moyen d'une procédure administrative et l'intervieweur ne saurait pas s'il tâche d'interviewer une unité faisant partie de l'échantillon initial ou une unité de remplacement. Il devrait être facile d'établir ce type de contrôle dans le cas d'enquêtes à composition aléatoire de numéros de téléphone, au moyen de l'ITAO (interview téléphonique assistée par ordinateur).

En deuxième lieu, on pourrait avoir tendance à déployer moins d'efforts pour communiquer avec les unités d'échantillonnage initiales et pour les recruter lorsque des unités de remplacement sont disponibles. Dans une certaine mesure, les intervieweurs et, peut-être, certains chargés d'enquête pourraient considérer les unités de

remplacement comme presque aussi bonnes que les unités faisant partie de l'échantillon original. Cela aurait pour effet de réduire le taux de réponse et d'accroître la possibilité d'un biais de non-réponse.

Le recours au remplacement présente une difficulté connexe qui tient à la limitation associée à cette méthode en raison de l'aspect séquentiel du processus de mise en œuvre (cette à dire, on ne téléphone à une unité de remplacement que lorsque les efforts faits pour recruter l'unité d'échantillonnage comprise dans l'échantillon original se sont révélés infructueux). Pour une période de collecte de données fixe, il faudrait réduire le délai prévu pour communiquer avec les unités d'échantillonnage comprises dans l'échantillon initial et pour les recruter afin de prévoir suffisamment de temps pour recruter des unités de remplacement. Plus probablement, une période de collecte de données plus longue serait prévue pour permettre de tâcher de recruter les unités d'échantillonnage comprises dans l'échantillon original ainsi que les unités de remplacement. Cela aurait pour effet vraisemblablement d'accroître les coûts de la collecte des données (comparativement à l'approche fondée sur la correction des poids), même dans le cas d'un nombre comparable d'interviews complètes réalisées.

Une façon de résoudre partiellement le problème de la nature séquentielle de l'utilisation d'unités de remplacement consiste à tâcher de communiquer avec une unité de remplacement avant que toutes les tentatives prévues soient faites pour recruter l'unité d'échantillonnage comprise dans l'échantillon original. On procède alors à des tentatives simultanées de communiquer avec l'unité originale et l'unité de remplacement. Bien entendu, cette façon de procéder présente certains désavantages: l'intervieweur qui a communiqué avec l'unité de remplacement peut avoir gaspillé son temps et on peut même, en fin de compte, avoir interviewé et l'unité d'échantillonnage comprise dans l'échantillon original et l'unité de remplacement, auquel cas les données de cette dernière ne devraient pas être utilisées.

Un autre désavantage du recours au remplacement tient à la difficulté d'identifier de façon suffisamment précise les répondants de remplacement dans la base de données finale pour permettre de calculer le véritable taux de réponse. On pourrait avoir tendance à calculer le taux de réponse comme si les unités de remplacement étaient des unités d'échantillonnage comprises dans l'échantillon initial.

J'aimerais mentionner une dernière lacune du recours au remplacement: l'impossibilité éventuelle de recruter une unité de remplacement pour chaque non-répondant. Ce sera certainement le cas si une seule unité de remplacement est sélectionnée pour chaque non-répondant. (C'est pourquoi je recommande la sélection de plusieurs unités de remplacement pour chaque non-répondant, de préférence dans le cadre du même processus de tirage de l'échantillon.) Si on n'a pas obtenu d'unités de remplacement pour chaque non-répondant, il faudra appliquer une autre méthode pour tenir compte de la non-réponse résiduelle (comme la correction des poids).

En conclusion, la question fondamentale est de savoir s'il est préférable d'utiliser une unité de remplacement d'un non-répondant à une enquête que d'imputer les données du non-répondant à partir d'un mélange des données déclarées par les répondants dans la même catégorie de pondération. Dans de nombreux cas, j'estime qu'il serait préférable d'utiliser une unité de remplacement, à la condition de pouvoir contourner suffisamment les désavantages mentionnés plus haut du recours à une méthode de remplacement.

J'apprécierais tout commentaire que veulent bien me faire parvenir les lecteurs qui ont eu recours au remplacement aux fins d'enquête, y compris sur les méthodes particulières utilisées et le taux de succès en matière d'obtention d'unités de remplacement des non-répondants disposées à coopérer, ainsi que toute observation concernant la mesure dans laquelle la méthode a ou n'a pas donné de bons résultats. Bien entendu, j'invite également tous les intéressés à me faire part de leurs opinions sur la question de savoir s'il y a lieu d'avoir recours au remplacement pour tenir compte de la non-réponse totale aux enquêtes.

#### Références

Chapman, D. (1983). The impact of substitution on survey estimates. In W. Madow, I. Olkin, and D. Rubin (Eds.). *Incomplete data in sample surveys, vol. II, Theories and bibliographies*, (pp. 45-61), Academic Press: New York.

Chapman, D. and Roman, A. (1985). An investigation of substitution for an RDD survey. *Proceedings of the Survey Research Methods Section, American Statistical Association*, 269-274.

Vehovar V. (1999). Field substitution and unit nonresponse. *Journal of Official Statistics*, 15, (2), 335-350.

#### Pour contacter l'auteur:

David W. Chapman, statisticien en chef  
Federal Deposit Insurance Corporation

Room 2022, 550 17<sup>th</sup> St. N.W.  
Washington, DC 20429  
USA  
(202) 898-7280  
dachapman@fdic.gov



## L'expression "substitutions sur le terrain" redéfinie

par Vasja Vehovar

### Introduction

L'expression "substitutions" sur le terrain fait habituellement référence à la pratique courante de remplacer une unité non-répondante par une autre unité à l'étape du travail sur le terrain de l'enquête. Comme le remarque Chapman (2003), cette pratique a été "mal utilisée" dans les décennies passées et s'est retrouvée avec une mauvaise réputation. L'aperçu dans Vehovar (1999) démontre que, en général, les substitutions n'apportent que de petites améliorations quant à la variance d'échantillonnage, sans compenser les nombreux inconvénients qu'elles causent, particulièrement le manque de contrôle sur les intervieweurs et les retards dus à la nature séquentielle de la substitution, (car il faut clarifier le statut des unités avant de pouvoir utiliser les unités de remplacement). Il est également vrai qu'aucun des exemples décrits dans la documentation ne démontre convenablement que cette pratique était justifiée. Nous pourrions conclure que le recours aux substitutions ne résulte souvent que de l'ignorance des intervenants. Le rejet relativement systématique de cette pratique dans le cas des échantillons probabilistes dans bon nombre d'enquêtes face à face aux États-Unis semble donc être la bonne attitude à prendre.

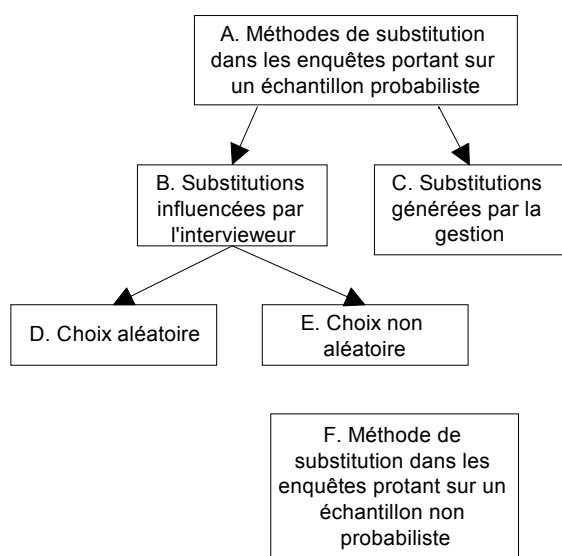
On pense habituellement (à tort) que cette démarche est préférable au simple fait de ne rien faire à propos des non-réponses. Plus subtilement, on s'appuie sur la croyance (jamais démontrée) que les substitutions sont préférables à d'autres rajustements de pondération. Ceci est, en fait, une idée fautive intuitive très trompeuse, car il est plutôt étrange qu'en obtenant la structure optimale de l'échantillon on n'y gagne à peu près rien en précision par rapport à l'échantillon non équilibrée des unités répondantes sans substitutions. Dans le cas de l'échantillon par grappes à deux degrés avec  $n = 100 \times 10$  unités, il semble donc préférable d'obtenir la structure désignée de répondants (cette-à-dire, 100 grappes de 10 unités chacune) avec des substitutions plutôt que les mêmes 1000 unités avec 100 grappes de tailles différentes (par exemple, 5-15 unités) sans substitutions. Si la

différence est plutôt petite, c'est parce que la formule de Kish bien connue pour l'augmentation de la variance attribuable à la pondération ne joue pas pour le poids des non-réponses dans des échantillons en grappes. Ceci est aussi assez surprenant puisque cela remet en question une autre certitude intuitive des intervenants—la conviction que la formule Kish est extrêmement solide.

### Définition de "substitutions"

Toutefois, indépendamment des objections théoriques énoncées ci-dessus, un reproche beaucoup plus fondamental s'adresse à l'influence de l'intervieweur sur le choix des substitutions. Ceci comprend aussi la situation où l'intervieweur peut déclarer une certaine unité difficile à traiter comme étant une unité non-répondante et obtenir de ses supérieurs une substitution aléatoire. Pourtant, l'influence de l'intervieweur—l'un des plus graves inconvénients—n'est pas une partie nécessaire de la démarche. Dans Vehovar (1999), cette distinction de base n'était pas reconnue, ce qui laisse planer sur l'ensemble du document une certaine confusion quant à la nature exacte des substitutions étudiées. Faisons la distinction ici. Si l'on veut améliorer la définition, il faut séparer deux composantes, a) l'influence de l'intervieweur sur le choix et b) l'effet séquentiel du remplacement de l'unité manquante par l'unité de substitution. La dernière composante est la seule qui est essentielle aux substitutions. Il n'existe non plus aucune raison de limiter les substitutions aux non-réponses, car on peut aussi faire séquentiellement des substitutions d'unités non admissibles (par exemple, unités non admissibles).

Figure 1: Types de méthode de substitution



Définition: Les substitutions se rapportent au remplacement séquentiel de l'unité manquante par une autre unité au cours de l'étape du travail sur le terrain de l'enquête. Les deux types qu'on doit distinguer sont les suivants:

- ◆ Substitutions influencées par l'intervieweur. Ici, l'intervieweur peut décider si l'unité de substitution sera utilisée. Il est vrai que dans le cadre d'enquêtes face à face il est impossible de complètement éliminer cet impact; nous ne pouvons que l'estimer avec des contrôles stricts sur le terrain. En ce qui concerne la façon dont la substitution est choisie, nous avons deux autres options:
  - le choix aléatoire où les unités de substitution sont choisies dans une liste préparée d'avance, basée sur un mécanisme d'échantillonnage probabiliste strictement contrôlé;
  - le choix non aléatoire, où les unités de substitutions sont choisies soit à partir de critères de sélection intentionnels ou soit avec un certain montant de liberté de la part de l'intervieweur.
- ◆ Substitutions générées par les gestionnaires. Ici, l'intervieweur n'a aucun impact ni sur la décision de substituer ni sur la méthode utilisée. Même lorsqu'il déclare une unité comme étant non-répondante, il ne sait pas si une substitution sera faite. Cette situation se rapproche beaucoup de ce que Kish (1965) appelle un "échantillon complémentaire"—une liste d'unités préparée d'avance appliquée lorsque les gestionnaires chargés de l'enquête en décident. Toutefois, avec des unités complémentaires, aucun remplacement séquentiel individuel n'est effectué, ce qui constitue l'essence même des substitutions. L'exemple le plus typique pour ce type de substitution est l'environnement complet de l'ITAO où l'intervieweur par téléphone n'a aucune façon de savoir que la gestion applique effectivement des substitutions. De l'autre côté, lorsque l'intervieweur connaît la situation dans sa grappe (cette-à-dire, ce qu'il a accompli pour répondre aux critères de réalisation de la "prise b" d'unités admissibles ou répondantes) ou même s'il décide de faire une substitution et comment choisir l'unité de substitution, cette situation devient une substitution influencée par l'intervieweur. Il est évident que nous substituons aussi dans la méthode courante Waksberg-Mitofsky, toutefois, seules les unités non admissibles sont substituées dans ce cas, car il a été plus ou moins entendu, pour cette méthode, de ne pas substituer dans les cas de non-réponse, même si certains ont essayé.

La figure 2 ci-dessous illustre des exemples typiques de substitutions dans des échantillons probabilistes. Si nous ne faisons pas de distinction explicite entre elles, nous pouvons aisément nous retrouver avec une communication discordante, qui existe aussi dans Vehovar (1999), où la majorité du traitement est de type D (figure 1), avec un commentaire occasionnel qui renvoie vraiment aux types E et C. Étant donné que ce sont tous des types différents, des critères très différents s'adressent à chacun d'eux.

### Exemple classique de substitution

Il existe un exemple de choix d'une situation d'enquête face à face avec deux unités (par exemple, écoles et ménages) par strates ou par grappes où la substitution est jugée de loin préférable. Dans ce cas, Chapman (1983) et Vehovar (1999) s'entendent pour dire que la substitution est l'option préférable. Devrions-nous vraiment permettre à l'intervieweur de tirer avantage du fait qu'il est sur place et le laisser prendre un substitut afin qu'il ramène de l'information nécessaire de ce segment géographique spécifique inclus dans l'échantillon? Ou devrions-nous appliquer l'option suivante:

- ◆ laisser la taille initiale de l'échantillon se gonfler de non-réponses anticipées. Si elle augmente de 20%, laisser toutes ou une portion des grappes ou strates à la taille accrue (3 au lieu de 2),
- ◆ avoir des échantillons supplémentaires au cas où la non-réponse dépasse nos attentes. Dans le cas de grappes d'échantillon, il peut y avoir des unités supplémentaires à l'intérieur des grappes existantes ou, dans le cas de structures de coûts spécifiques, on peut préparer les nouvelles grappes,
- ◆ le problème d'une ou d'aucune unité par strate ou grappe peut être résolu par des techniques d'échantillonnage regroupé, ou par imputation ou pondération.

Dans le cas ci-dessus, le bon sens et l'intuition statistique semblent dicter que les substitutions sont très "naturelles" et avantageuses. Toutefois, seule une recherche approfondie (et compliquée) peut permettre de trouver la bonne réponse à ce problème. C'est aussi vrai pour d'autres situations apparemment avantageuses, particulièrement pour la méthode Waksberg-Mitofsky, pour laquelle les recherches antérieures ne sont pas encore exhaustives (Chapman et Roman, 1985). Cette méthode de substitution est encore légèrement surestimée par rapport à la méthode modifiée sans substitutions (Brick et Waksberg, 1992). Bien sûr, il est aussi vrai que la méthode Waksberg-Mitofsky a fréquemment été abandonnée aux États-Unis au cours des dernières années, vu les bons

échantillons assistés d'une liste et les inconvénients de sa nature séquentielle. Toutefois, avec les enquêtes par téléphones cellulaires, du moins en Europe, ceci pourrait changer du tout au tout.

Figure 2: Exemples de méthodes de substitution dans des échantillons probabilistes

Choix	Mode	
	Téléphone	Face à face
Substitutions générées par la gestion (C)	Waksberg-Mitofsky dans une méthode complètement automatisée d'ITAO	Compléments (seulement conditionnelle-ment)
Substitutions influencées par l'intervieweur (B→D)	Méthodes de substitution aléatoire courantes dans des enquêtes par téléphone	Méthodes de substitutions aléatoires courantes dans des enquêtes face à face

### Substitutions dans des échantillons non probabilistes

Jusqu'à maintenant, la discussion n'a porté que sur les échantillons probabilistes. Toutefois, il est de plus en plus important, de nos jours, de reconnaître les problèmes des échantillons non probabilistes. Nous pouvons continuer de dire qu'il ne peut pas y avoir de modification de la notion d'échantillon scientifique, pourtant, nous sommes confrontés à une sérieuse augmentation de composantes de non probabilité dans tous les échantillons. Peut-on vraiment décider qu'on ne peut rien faire avec ces échantillons, sauf de laisser chaque intervenant élaborer ses propres trucs maison pour extraire et ajuster ses échantillons non probabilistes?

Il vaut la peine de mentionner ici deux directions. La première est l'utilisation de substitutions avec l'addition d'imputations multiples (Rubin&Zanutto, 2002). L'idée de base est d'obtenir de l'information des substituts à certains endroits (grappe, strate, famille) pendant que l'intervieweur est sur place, et de l'utiliser dans la stratégie d'imputation multiple. Les règles strictes des échantillons probabilistes sont quelque peu mises de côté, mais on obtient en échange, des avantages de coût-erreur.

La deuxième direction est reliée aux échantillons déterminés par quota et au débit actuel d'échantillons non probabilistes sur le Web (Vehovar et al, 2002), où la demande d'un certain encadrement théorique est particulièrement forte. La question est de savoir si rien ne peut être ajouté

à Kish (1965, 1999), qui déclare que ces échantillons sont simplement des objets d'art sans aucune substance scientifique. D'un autre côté, pour la majorité des analystes de marché, particulièrement ceux qui travaillent au niveau des ménages ou Internet, la notion que "l'échantillonnage n'a plus aucune valeur" est déjà acquise depuis des décennies. Une présélection minutieuse et des ajustements savants après enquête fournissent souvent des données satisfaisantes pour la majorité de leurs besoins. On a toutefois le sentiment que nous ne sommes toujours pas à même d'extraire tout le potentiel d'information caché dans ces échantillons non probabilistes. La pondération des codes de propension est un exemple de tentative d'obtenir une meilleure inférence de ce genre de données. Ici, les substitutions (type F et E) font l'objet de la même demande que pour les échantillons non probabilistes de plus d'efforts de recherche élaborant les stratégies d'erreurs de coûts correspondantes (Vehovar et al, 2001).

### Conclusions

L'essence même des substitutions est constituée par les remplacements séquentiels d'unités manquantes. D'un autre côté, l'influence de l'intervieweur sur le choix des substituts n'existe que pour certains types de substitutions. Dans ces cas, les substitutions ne peuvent être justifiées. La seule exception possible est le cas extrême ne comportant que quelques unités par strate ou grappe, où la situation n'est pas encore très claire et où plus de recherches sont nécessaires. Toutefois, aucune de ces critiques ne s'adresse aux substitutions générées par la gestion, y compris la méthode Waksberg-Mitofsky dans les enquêtes par téléphone. Dans ces cas, les substitutions sont confrontées—outre l'inconvénient général de leur nature séquentielle—à un autre problème: elles n'offrent qu'un gain minime en matière de précision, ce qui les rend beaucoup moins favorables qu'elles ne paraissent. Lorsque les substitutions sont appliquées dans des échantillons non probabilistes, elles partagent le destin général de cette famille grandissante d'enquêtes non scientifiques: elles surclassent très souvent les échantillons probabilistes (si l'on en juge par les considérations globales de coût-erreur), mais peu de recherches ont été effectuées pour mieux comprendre cet avantage et pour aider les intervenants.

### Références

Brick & Waksberg (1991). Méthode pour éviter l'échantillonnage progressif dans une enquête téléphonique à composition aléatoire. *Techniques d'enquête*, 17(1), pp. 27-41.

Chapman, D. W. (2003). To substitute or not to substitute—that is the question. *Survey Statistician*, 48, 32-34.

Chapman, D. W. et Roman, A. (1985). An investigation of substitution for an RDD survey. *Proceedings of the Survey Research Methodology Section, ASA*, 269-274.

Chapman, D. W. (1983). The impact of substitutions on survey estimates. In Meadow et al.: *Incomplete data in sample surveys, vol. II*, National Academy of Sciences and Academic Press.

Kish (1965). *Survey sampling*. New York:Wiley.

Kish (1999): *Quota sampling: Old plus new thought*. Available on <http://ris.org/vasja/kish.doc>.

Rubin, D. B. et Zanutto, E. (2002): Using matched substitutes to adjust for nonignorable nonresponse through multiple imputation. In Groves et al (Eds): *Survey nonresponse*. New York:Wiley.

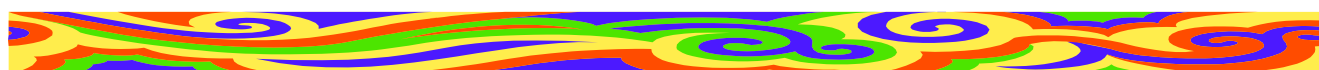
Vehovar, V. (1999): Field substitution and unit nonresponse. *Journal of Official Statistics*.

Vehovar, V., Lozar, K., Batagelj, Z. et Zaletel (2002): Nonresponse in web surveys. In Groves et al (Eds). *Survey nonresponse*. New York:Wiley.

Vehovar, V., Lozar, K., et Batagelj, Z. (2001). Sensitivity of e-commerce measurement to survey instrument. *International Journal of Electronic Commerce*, 2001, 6/1.

### Pour contacter l'auteur:

Vasja Vehovar  
Université de Ljubljana, Slovénie  
[vasja@ris.org](mailto:vasja@ris.org)



## Annonces

### Assemblée générale de l'AISE

La prochaine assemblée générale des membres de l'AISE aura lieu pendant les réunions de l'IIS à Berlin, le **jeudi 14 août de 11 h 15 à 13 h**. Tous les membres à titre personnel de l'AISE, ainsi que les représentants des membres-organismes sont chaleureusement invités à participer à cette réunion. Vous trouverez d'autres renseignements à l'adresse <http://www.isi-2003.de/>.



### Élections au sein de l'AISE

Des bulletins de vote ont été remis aux membres de l'AISE en février en prévision de l'élection d'un président désigné, de deux vice-présidents, d'un secrétaire scientifique et de six membres du conseil. On annoncera les résultats de l'élection et le mandat des candidats débutera à la fin de l'assemblée générale.



### Mini-réunion conjointe de l'Institut de statistiques mathématiques et de la *Survey Research Methods Section (SRMS)* sur les tendances actuelles de l'échantillonnage et sur les statistiques officielles, à Calcutta, Inde, les 2 et 3 janvier 2004

L'Institut de statistiques mathématiques et la Survey Research Methods Section (SRMS) organiseront conjointement une mini-réunion qui portera sur les tendances actuelles de l'échantillonnage et sur les statistiques officielles (Current Trends in Survey Sampling and Official Statistics) et qui aura lieu à Calcutta, en Inde, les 2 et 3 janvier 2004. Le tout sera coparrainé par le U.S. Census Bureau, le Gallup Research Center de l'University Nebraska-Lincoln et le Department of Statistics de l'Université de Calcutta. La mini-réunion a pour but de tisser des liens entre les spécialistes de la statistique mathématique et les spécialistes de la mathématique qui travaillent à des enquêtes-échantillon et aux statistiques officielles, soit à des organismes gouvernementaux

ou privés. Les chercheurs suivants ont accepté (sous toute réserve) de présenter une communication:

Arijit Chaudhuri (Inde), William Bell (É.-U.), Heleno Bolfarine (Brésil), Jay Breidt (É.-U.), Daniela Cocchi (Italie), Stephen E. Fienberg (É.-U.), David Findley (É.-U.), Marco Fortini (Italie), Wayne Fuller (É.-U.), Malay Ghosh (É.-U.), J.K. Ghosh (Inde/É.-U.), Pilar Iglesias (Chili), Michael Larsen (É.-U.), Brunero Liseo (Italie), Viviana Lencina (Argentine), Michel Mouchart (Belgique), S. James Press (É.-U.), J.N.K. Rao (Canada), Stuart Scott (É.-U.).

Vous trouverez à l'adresse <http://www.jpsm.umd.edu/ims> des renseignements pertinents sur la réunion, notamment l'inscription et le programme. Pour y présenter une communication, veuillez contacter Partha Lahiri à l'adresse [plahiri@survey.umd.edu](mailto:plahiri@survey.umd.edu) et pour obtenir des renseignements de nature locale, veuillez contacter Tathagata Banerjee à l'adresse [btathaga@yahoo.com](mailto:btathaga@yahoo.com).



### Comité du programme de l'AISE pour la session de l'IIS à Sydney en 2005

Lorsque vous lirez ce message, notre comité aura commencé à dresser la liste des sujets admissibles des personnes invitées à présenter une communication à la partie de l'AISE de la session de l'IIS à Sydney. Les discussions menant à la décision finale auront lieu à la session de l'IIS à Berlin en août. Si vous désirez proposer un sujet, veuillez envoyer le titre et un court résumé (d'un maximum de cent mots) ainsi que vos commentaires et suggestions au président du comité dont l'adresse figure ci-dessous. Les suggestions de dernière minute sont acceptées, à condition qu'elles parviennent au comité avant le 31 juillet.

Pedro Silva  
[pedrosilva@ibge.gov.br](mailto:pedrosilva@ibge.gov.br)  
Tél.: 55-21-25144548



## Exercices corrigés de méthodes de sondage, Editions Ellipses, 2003

Les ouvrages d'exercices corrigés de sondage sont extrêmement rares. Aussi, pour contribuer à la bonne compréhension de la discipline, Pascal Ardilly (INSEE, France) et Yves Tillé (Université de Neuchâtel, Suisse) proposent une sélection de 116 exercices systématiquement accompagnés d'une correction approfondie. Ces exercices sont regroupés par chapitre et couvrent les principaux aspects de la théorie des sondages. Les premiers chapitres sont consacrés aux méthodes d'échantillonnage: sondage aléatoire simple, sondage stratifié, tirage à probabilités inégales, tirage à plusieurs degrés, tirage équilibré, méthode des quotas. Les chapitres suivants traitent des méthodes de redressement: estimateur par le ratio, estimateur par la régression, estimateur post-stratifié, calage généralisé. Un chapitre est consacré à l'estimation de variance. Enfin, le livre s'achève par des exercices sur la non-réponse.

Le lecteur trouvera dans cet ouvrage (300 pages, 25 euros) des exercices de tous niveaux, certains développant des aspects théoriques, d'autres

traitant de problèmes plus appliqués avec des applications numériques. À noter que chaque chapitre est précédé d'un bref rappel de cours précisant les notations et les principaux résultats à connaître.

Pascal Ardilly est chef du service Échantillonnage et traitement statistique des données, à l'Institut National de la Statistique et des Études Économiques (INSEE), France. Yves Tillé est docteur en statistique de l'Université libre de Bruxelles, et actuellement professeur à l'Université de Neuchâtel, Suisse. Les deux auteurs ont déjà publié des ouvrages de cours en français:

- ◆ Les techniques de sondage, P. Ardilly, (1994).Technip;
- ◆ Théorie des sondages, échantillonnage et estimation en population finie, Y. Tillé, (2001) Dunod.



1992.....1996.....1999.....

# ASC2003

## THE IMPACT OF TECHNOLOGY ON THE SURVEY PROCESS

The Association for Survey Computing  
announces its  
**Fourth International  
3 - day Conference**

### ASC2003 – What is it ?

The Fourth International Conference on Survey and Statistical Computing, hosted by **The Association for Survey Computing**. It follows on from similar events presented by the ASC in 1992 (Bristol University, England), in 1996 (Imperial College, London, England) and in 1999 (Edinburgh University, Scotland).

### ASC2003 – When is it ?

Wednesday 17th to Friday 19th September 2003.

### ASC2003 – Where is it ?

Ideally located for a residential conference, **The University of Warwick, England, UK**, is situated in the heart of England on a large and pleasant campus, close to the City of Coventry and the historic towns of Warwick, Kenilworth, Stratford-upon-Avon and Royal Leamington Spa. Delegates will be accommodated close to the University's Central Campus and main Conference areas.

### ASC2003 – Who is it for ?

The Conference targets practitioners across all aspects of the survey research process. These are likely to be drawn from across the UK, Europe, North and South America and Australasia. The Conference's central theme will be **The Impact of Technology on the Survey Process**, and with contributions from commissioners, producers, and consumers of survey research alike. It will explore the often complex relationships between the push and pull of technological change and the expectations and demands created by them. It will also seek to examine how outcomes feed back to affect the processes which initially gave rise to them. Although a conference about technology, the topics to be discussed will appeal to everyone with an interest in survey design, data collection, analysis, reporting, or statistical computing.

### ASC2003 – Conference Proceedings

A bound volume of **Proceedings**, incorporating all the papers presented to the Conference, will be published prior to the Conference and distributed to all delegates.

### ASC2003 – What does it include ?

#### THE CONFERENCE IS BASED ON FIVE ELEMENTS

##### Plenary Sessions

The following Invited Speakers will address the Conference:

##### Norman Glass

(Chief Executive, National Centre for Social Research)

##### Denise Lievesley

(Director, UNESCO Institute for Statistics)

##### Peter Mouncey

(Cranfield University School of Management)

##### David Pullinger

(Head of National Statistics Online, ONS)

##### Themed Streams of Contributed Papers

Some 40 Papers will be presented within six streams:

Data Dissemination  
Organisational Behaviour  
Survey Development  
Survey Modalities  
Survey Software  
The Effectiveness of the Internet

##### Poster Sessions

Providing an additional way of demonstrating the variety of developments in the field covered by the Conference.

##### The Conference Exhibition

An essential element at all ASC Conferences, providing opportunities for companies to meet the delegates.

##### The Social Programme

A number of social events have been arranged to enable delegates to meet together in a more informal setting.

### ASC2003 – What is the ASC ?

**The Association for Survey Computing** is a not-for-profit organisation based in the United Kingdom but drawing significant international membership and support. It provides a valuable forum for those concerned with survey research and statistical computing. Its many activities and links to other organisations are publicised through its web site ([www.asc.org.uk](http://www.asc.org.uk)) including the many conferences and workshops it runs, the register of available software it maintains and its support of other initiatives in the survey and statistical field. Over the past thirty years, the ASC has organised many one-day conferences and workshops, mainly in London but also occasionally in Edinburgh. ASC membership is wide-ranging at both individual and corporate level. It aims to keep its members up to date with developments in the application of computers to all aspects of the survey process. It has links with the British Computer Society, the International Association for Statistical Computing, the Market Research Society, The Royal Statistical Society and the Social Research Association.

### ASC2003 – More information ?

##### Contact The Administrator

ASC, PO Box 60, Chesham, Bucks, HP5 3QH, England.  
Tel/Fax: 01494 793033 (International +44 (0)1494 793033);  
e-mail: [admin@asc.org.uk](mailto:admin@asc.org.uk)

##### Visit The Web Site at [www.asc.org.uk](http://www.asc.org.uk)

The ASC web site provides the latest Conference updates, including titles of Papers and delegate fee structures.

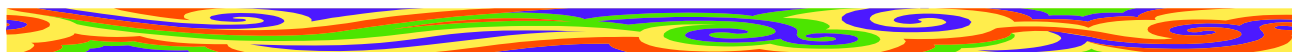
## Visitez le nouveau site Internet de l'AISE et lisez *Le Statisticien d'enquêtes* en ligne!

[www.isi-iass.org](http://www.isi-iass.org)

	<p><b>International Association of Survey Statisticians (IASS)</b> The Premier World Organization representing who's who in Sample Survey and Census Methodologies</p> <hr/>
<p><b>More about the IASS</b></p> <p><b>Becoming a member</b></p> <p><b>Services for members</b></p> <p><b>Survey Statistician</b></p> <p><b>Publications</b></p> <p><b>IASS Conferences</b></p> <p><b>Cochran-Hansen Prize</b></p> <p><b>Ask the experts</b></p> <p><b>Links</b></p> <p><b>What's new</b></p>	<p><b>FOUNDATION AND OBJECTIVES:</b></p> <p>Founded in 1973, the International Association of Survey Statisticians (IASS) has as its charge and mandate to promote the study and development of the theory and practice of sample surveys and censuses. It also aims to increase interest in surveys and censuses among statisticians, governments, and the public the world over.</p> <p><b>MEMBERSHIP AND EXECUTIVE:</b></p> <p>At present the IASS has approximately 1.200 members from 130 countries and 38 institutional members.</p> <p>IASS headquarters are situated in Libourne (FRANCE) and operate, to a large extent, under the auspices of the French statistical agency INSEE, which lends its expertise and status to aid and promote the association's work.</p> <p>The society is run by an Executive Committee, elected for a period of 2 year and a Council, elected for a four year period.</p> <p>Anyone interested in learning more about IASS should contact Christophe Lefranc IASS Executive Director INSEE 18, Bd. Adolphe Pinard 75675 PARIS, France e-mail: <a href="mailto:christophe.lefranc@insee.fr">christophe.lefranc@insee.fr</a></p> <hr/>
	<p><b>Association Internationale des Statisticiens d'Enquête (AISE)</b></p>

### Avis important

- ◆ Un fichier PDF du bulletin est disponible sur le site Internet de l'AISE. Très peu de membres ne souhaitent plus recevoir la version papier du bulletin mais souhaitent simplement être informés de la mise en ligne d'un nouveau bulletin. En ce moment, nous n'avons pas de procédé établi afin d'appuyer cette option. Un procédé sera développé quand un nombre suffisant de membres auront choisi l'option ci-dessus. D'ici à ce moment là, tous les membres continueront à recevoir la version papier du bulletin. Si vous voulez bénéficier de l'option susmentionnée, veuillez envoyer un courriel en ce sens à [LeylaMohadjer@Westat.com](mailto:LeylaMohadjer@Westat.com)
- ◆ Les membres sont invités à visiter le site web de l'AISE ([www.isi-iass.org](http://www.isi-iass.org)) et à soumettre leurs commentaires et suggestions à l'adresse suivante: [eric.rancourt@statcan.ca](mailto:eric.rancourt@statcan.ca).



## Autre revues



### Techniques d'enquête

#### Une revue éditée par Statistique Canada

Volume 28, numéro 2, décembre 2002  
Table des matières

Dans ce numéro.....	123
GORDON J. BRACKSTONE Stratégies et approches relatives aux statistiques régionales.....	125
DENNIS TREWIN L'importance d'une culture de la qualité .....	135
YVES THIBAUDEAU Imputation pour la non-réponse partielle basée sur un modèle explicite pour les catégories démographiques .....	147
BALGOBIN NANDRAM, GEUNSHIK HAN et JAI WON CHOI Un modèle bayésien hiérarchique de non-réponse non-ignorable pour les données multinomiales des petites régions .....	157
JAY STEWART Évaluation du biais lié à diverses stratégies de prise de contact dans les enquêtes téléphoniques sur l'emploi du temps .....	171
ROBERT M. BELL et DANIEL F. MCCAFFREY Réduction du biais des erreurs-types pour la régression linéaire dans le cas d'échantillons à plusieurs degrés.....	185
MONROE G. SIRKEN Effets de plan de sondage dus aux bases de sondage dans les enquêtes auprès des établissements .....	199
LOUIS-PAUL RIVEST Une généralisation de l'algorithme de Lavallée et Hidiroglou pour la stratification dans les enquêtes auprès des entreprises.....	207
WILSON LU et RANDY R. SITTER Méthode pratique de stratification multiple par programmation linéaire .....	215
ROBBERT H. RENSSSEN et GERARD H. MARTINUS De l'utilisation des matrices inverses généralisées dans la théorie de l'échantillonnage.....	225
Remerciements .....	231



**Journal of Official Statistics**  
**Une revue internationale publiée**  
**par Statistique Suède**

**Volume 18, numéro 4, 2002**  
**Table des matières**

Le JOS est une revue érudite trimestrielle consacrée à la méthodologie d'enquête et sur d'autres questions en rapport avec la production de statistiques par les bureaux nationaux et d'autres organismes statistiques. Tous les manuscrits sont examinés rigoureusement par des évaluateurs indépendants et des membres du Comité de rédaction.

Measures to Evaluate the Discrepancy Between Direct and Indirect Model-based Seasonal Adjustment Eduardo Otranto and Umberto Triacca.....	511
Satisfying Disclosure Restrictions with Synthetic Data Sets Jerome P. Reiter .....	531
Item Nonresponse as a Predictor of Unit Nonresponse in a Panel Survey Geert Loosveldt, Jan Pickery, and Jaak Billiet.....	545
Using Administrative Records to Improve Small Area Estimation: An Example from the U.S. Decennial Census Elaine Zanutto and Alan M. Zaslavsky.....	559
Asymptotically Efficient Generalised Regression Estimators Giorgio E. Montanari and M. G. Ranalli .....	577
Editorial Collaborators.....	591
Index to Volume 18, 2002 .....	595

**Journal of Official Statistics**  
**An International Review Published**  
**by Statistics Sweden**

**Volume 19, numéro 1, 2003**  
**Table des matières**

Multiple Imputation for Statistical Disclosure Limitation T. E. Raghunathan, J.P. Reiter, and D.B. Rubin .....	1
Confidence Interval Coverage Properties for Regression Estimators in Uni-Phase and Two-Phase Sampling J. N. K. Rao, W. Jocelyn, and M. A. Hidiroglou .....	17
Sequential Weight Adjustments for Location and Cooperation Propensity for the 1995 National Survey of Family Growth Vincent G. Iannaccione .....	31
Constrained Inverse Adaptive Cluster Sampling Emilia Rocco .....	45
Poisson Mixture Sampling Combined with Order Sampling Hannu Kröger, Carl-Erik Särndal, and Ismo Teikari .....	59
Ischemic Heart Disease and Acute Myocardial Infarction as Cause of Death in a Cause-of-Death Register Boo Svartbo, Gösta Bucht, Lars Olov Bygren, Anders Eriksson, and Lars Age Johansson .....	71

Toutes les demandes de renseignements concernant les présentations d'articles et les abonnements doivent être adressées au [jos@scb.se](mailto:jos@scb.se).



## Statistics in Transition

### Revue de l'Association polonaise de statistique

Volume 5, numéro 5, octobre 2002  
Table des matières

From the Editor ..... 721

#### Classification and Data Analysis

K. JAJUGA  
Foreword ..... 723

H. H. BOCK  
Clustering Methods: From Classical Models to New Approaches ..... 725

W. J. KRZANOWSKI  
Orthogonal Components for Grouped Data: Review and Applications ..... 759

Cz. DOMAŃSKI  
Some Remarks on the Tasks of Statistics on the Verge of the Twenty First Century ..... 779

A. ZELIAŚ  
Some Notes on the Selection of Normalisation of Diagnostic Variables ..... 787

D. BANKS and R. T. OLSZEWSKI  
Combinatorial Search in Multivariate Statistics ..... 803

M. GREENACRE and J. G. CLAVEL  
Simultaneous Visualization of Two Transition Tables ..... 823

E. GATNAR  
What is Data Mining? ..... 837

M. D. LARSEN  
Impact of Latent Class Clustering of NSF Doctoral Survey Data on Adjusted Rand Index Values ..... 843

#### Other Articles

A. MŁODAK  
An Approach to the Problem of Spatial Differentiation of Multi-feature Objects Using  
Methods of Game Theory ..... 857

H. P. SINGH and N. MATHUR  
An Alternative to an Improved Randomized Response Strategy ..... 873

#### Reports

International Federation of Classification Societies Conference—IFCS-2002, Cracov, July 16-19, 2002  
(K. Jajuga) ..... 887

The Ninetieth Anniversary of the Foundation of the Polish Statistical Association—the Scientific  
Conference, Krakow, Poland, July 14-15, 2002 (A. Zeliaś) ..... 891

# Statistics in Transition

## Revue de l'Association polonaise de statistique

Volume 5, numéro 6, décembre 2002  
Table des matières

From the Editor .....	897
<b>Survey Sampling (1)</b>	
D. THORBURN Foreword .....	899
M. CARLSON Assessing Microdata Disclosure Risk Using the Poisson-Inverse Gaussian Distribution .....	901
M. DAVIDSEN and M. KJØLLER The Danish Health and Morbidity Survey 2000—Design and Analysis .....	927
D. HEDLIN Estimating Totals in Some U.K. Business Surveys .....	943
A. HOLMBERG A Multiparameter Perspective on the Choice of Sampling Design in Surveys .....	969
V. KIVINIEMI and R. LEHTONEN Web Tools in Teaching and Learning of Survey Sampling: The ViSS Application .....	995
S. LAAKSONEN Traditional and New Techniques for Imputation .....	1013
K. MEISTER Asymptotic Considerations Concerning Real Time Sampling Methods .....	1037
<b>Other Articles</b>	
B. PRASAD, R. S. SINGH, and H. P. SINGH Modified Chain Ratio Estimators for Finite Population Mean Using Two Auxiliary Variables in Double Sampling .....	1051
G. N. SINGH Estimation of Population Ratio in Two Phase Sampling .....	1067
T. P. TRIPATHI, H. P. SINGH, and L. N. UPADHYAYA A General Method of Estimation and Its Application to the Estimation of Coefficient of Variation .....	1081
<b>Book Reviews</b>	
W. Charemza and K. Strzala (Eds), East European Transition and EU Enlargement: A Quantitative Approach, (by Subrata Ghatak) .....	1103

**Statistics in Transition**

**Revue de l'Association polonaise de statistique**

**Volume 5, numéro 6, décembre 2002 (suite)  
Table des matières**

**Economic and Social Trends in Transition Countries**

Basic Economic Trends in Countries of Central and Eastern Europe and CIS Countries  
(by M. Bieńkowska, E. Czumaj, J. Gniadzik) ..... 1103

**Obituary**

Tore E. Dalenius (1917-2002)..... 1117

**Acknowledgements**

Referees of Volume 5 ..... 1121



The Allgemeines Statistisches Archiv (ASTA)

[www.uni-koeln.de/wiso-fak/wisostatsem/asta/](http://www.uni-koeln.de/wiso-fak/wisostatsem/asta/)

Le *Allgemeines Statistisches Archiv* (ASTA) est le journal de la Société statistique d'Allemagne. Il est rédigé par le professeur Karl Mosler, Université de Cologne. Il offre une tribune internationale pour les chercheurs et les utilisateurs de tous les domaines de la statistique. La première partie (articles) présente les contributions à la théorie, aux méthodes et aux applications statistiques. On y met surtout l'accent sur les problèmes statistiques qui se posent dans l'analyse des phénomènes économiques et sociaux. La deuxième partie (nouvelles et rapports) se compose de nouvelles, rapports et autres documents qui portent sur les activités de la Société. Les recensions d'ouvrages paraissent dans la troisième partie (ouvrages). Les communications sollicitées qui ont été présentées au congrès annuel de la Société sont publiées régulièrement dans le deuxième numéro d'un volume. Toutes les communications dans la première partie sont examinées par des vérificateurs anonymes. Pour être acceptée, une communication doit présenter une nouvelle approche méthodologique ou un résultat, obtenu par une utilisation poussée d'une méthode statistique, qui a des répercussions scientifiques ou sociétales importantes. On encourage aussi les auteurs à publier des examens critiques hors série. Les auteurs de documents doivent envoyer trois copies de leur contribution au rédacteur à des fins d'examen. Le journal publie les contributions originales en anglais et en allemand, et les articles, de préférence en anglais. Une fois un document accepté, l'auteur doit fournir un fichier LaTeX ou ASCII du manuscrit final tout en respectant le style du *Allgemeines Statistisches Archiv* dans les citations et la disposition. Pour les envois, prière de communiquer avec:

Petra Sablotny  
Allgemeines Statistisches Archiv (Prof. Mosler)  
Universität zu Köln  
Albertus-Magnus-Platz  
D-50923 Köln  
e-mail: [sablotny@marge.statistik.uni-koeln.de](mailto:sablotny@marge.statistik.uni-koeln.de)

**Volume 87, numéro 1, 2003**  
**Table des matières**

**Articles**

Core Inflation in the Euro Area: Evidence from the Structural Linebreak VAR Approach Elke Hahn.....	1-23
On the Optimal Design in Stratified Regression Estimation Ralf Münnich .....	25-38
Systematische Ausfälle im Mikrozensus-Panel: Ausmass und Auswirkungen auf die Qualität von Arbeitsmarktanalysen Sylvia Zühlke.....	39-58
Heaping and Its Consequences for Duration Analysis: A Simulation Study Joachim Wolff, Thomas Augustin .....	59-86

**News and Reports**

Die Statistische Erfassung von Gründungen in Deutschland—Ein Vergleich von Beschäftigtenstatistik, Gewerbeanzeigenstatistik und den Mannheimer Gründungspanels Michael Fritsch, Reinhold Grotz, Udo Brix, Michael Niese, Anne Otto .....	87-96
Rolf Krengel In Memoriam Reiner Stäglin.....	97-99

## The Allgemeines Statistisches Archiv (AStA)

### Volume 87, numéro 1, 2003 (suite) Table des matières

#### Books

Buchbesprechungen ..... 101-105

Kucerá, T., Kucerá, O., Opara, O., Schaich, E.  
New Demographic Faces of Europe. The Changing Population Dynamics in Countries of Central  
and Eastern Europe  
Reviewed by G. Uebe

Shumway, R. H., Stoffer, D. S.  
Time Series Analysis and Its Applications  
Reviewed by W. Schmid

Wilcox, R. R.  
Fundamentals of Modern Statistical Methods, Substantially Improving Power and Accuracy  
Reviewed by G. Uebe

Härdle, W., Rönz, B.  
MM\* Stat, Eine Interaktive Einführung in die Welt der Statistik  
Reviewed by P. von der Lippe

Buchinformation ..... 106

### Volume 87, numéro 2, 2003 Table des matières

#### Articles

Volkswirtschaftliche Gesamtrechnungen im Zentrum der Wirtschaftspolitischen  
Diskussion—Zur Einführung  
Ruth Meier, Reiner Stäglin ..... 107-112

Stability Criteria and Convergence: The Role of the System of National Accounts for Fiscal  
Policy in Europe  
Tilman Brück, Andreas Cors, Klaus F. Zimmermann, Rudolf Zwiener ..... 113-131

Das System der Volkswirtschaftlichen Gesamtrechnungen als Informationsquelle für die  
Geldpolitik in der EWU  
Hermann Remsperger ..... 133-146

Gemeinschaftliches Europa—Nationale Volkswirtschaftliche Gesamtrechnungen  
Josef Richter ..... 147-163

Wachstumsrückstand in Deutschland?—Probleme der Deflationierung  
Heinrich Lützel ..... 165-175

Volkswirtschaftliche Gesamtrechnungen als die Kunst des Möglichen  
Wolfgang Strohm ..... 177-182

Qualität und Genauigkeit der Volkswirtschaftlichen Gesamtrechnungen  
Albert Braakmann ..... 183-199

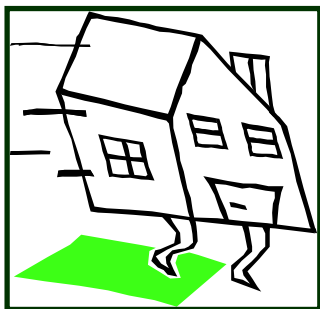
**The Allgemeines Statistisches Archiv (AStA)**

**Volume 87, numéro 2, 2003 (suite)  
Table des matières**

An Approach for Timely Estimations of the German GDP Andreas Cors, Vladimir Kouzine .....	201-220
Gibt es Eine Theorie der VGR? Utz-Peter Reich .....	221-237

# ASSOCIATION INTERNATIONALE DES STATISTICIENS D'ENQUÊTES

## FORMULAIRE DE CHANGEMENT D'ADRESSE



*Si votre adresse personnelle ou professionnelle a changé,  
veuillez faire une copie de ce formulaire et envoyer le  
complété à:*

Secrétariat de l'AISE  
c/o INSEE-CEFIL  
Att. Ms. Claude Olivier  
3, rue de la Cité  
33500 Libourne — France

Nom: M./Mme./Mlle. \_\_\_\_\_ Prénom: \_\_\_\_\_

Adresse électronique (prière d'indiquer une seule adresse):  
\_\_\_\_\_

*J'autorise l'AISE à publier mon adresse électronique sur la page web de l'AISE:*

Oui  Non

### Adresse personnelle

Rue: \_\_\_\_\_

Localité: \_\_\_\_\_

État/Province: \_\_\_\_\_ Code postal: \_\_\_\_\_

Pays: \_\_\_\_\_

Téléphone: \_\_\_\_\_

Fax: \_\_\_\_\_

### Adresse professionnelle

Nom de l'établissement: \_\_\_\_\_

Rue: \_\_\_\_\_

Localité: \_\_\_\_\_

État/Province: \_\_\_\_\_ Code postal: \_\_\_\_\_

Pays: \_\_\_\_\_

Téléphone: \_\_\_\_\_

Fax: \_\_\_\_\_

*Veuillez préciser l'adresse à laquelle votre correspondance de l'AISE devra être envoyée:*

Adresse personnelle  Adresse professionnelle

## Représentants de l'AISE et membres du conseil

Président (2001-2003):	Xavier Charoy (France), <a href="mailto:xavier.charoy@libertysurf.fr">xavier.charoy@libertysurf.fr</a>
Président-élu (2003-2005):	Luigi Biggeri (Italie), <a href="mailto:pres@istat.it">pres@istat.it</a>
Vice-présidents (2001-2003):	David Binder (Canada), <a href="mailto:binddav@statcan.ca">binddav@statcan.ca</a> Anders Christianson (Suède), <a href="mailto:anders.christianson@telia.com">anders.christianson@telia.com</a>
Secrétaire scientifique (2001-2003):	Seppo Laaksonen (Finlande), <a href="mailto:seppo.laaksonen@stat.fi">seppo.laaksonen@stat.fi</a>
Membres du conseil (1999-2003):	Florentina Alvarez (Espagne), <a href="mailto:falvarez@ine.es">falvarez@ine.es</a> Cynthia Clark (États-Unis), <a href="mailto:cclark@census.gov">cclark@census.gov</a> Pedro Do Nascimento Silva (Brésil), <a href="mailto:pedrosilva@ibge.gov.br">pedrosilva@ibge.gov.br</a> Jean Martin (Royaume-Uni), <a href="mailto:jean.martin@ons.gov.uk">jean.martin@ons.gov.uk</a> Vitalis Muba (Tanzanie), <a href="mailto:eastc@ud.co.tz">eastc@ud.co.tz</a> Bhagwan Tikkiwal (Inde), <a href="mailto:bdtikkiwal@rediffmail.com">bdtikkiwal@rediffmail.com</a>
Membres du conseil (2001-2005):	Kari Djerf (Finlande), <a href="mailto:kari.djerf@stat.fi">kari.djerf@stat.fi</a> David Fitch (États-Unis-Guatemala), <a href="mailto:dfitch@uvg.edu.gt">dfitch@uvg.edu.gt</a> Lidija Gligorova (Croatie), <a href="mailto:lgligor@dzs.hr">lgligor@dzs.hr</a> Langhui Huang (Chine), <a href="mailto:huanglh@stats.gov.cn">huanglh@stats.gov.cn</a> Moshe Sicron (Israël), <a href="mailto:sicron@internet-zahav.net">sicron@internet-zahav.net</a> Awa Thiongane (Sénégal-Éthiopie), <a href="mailto:athiongane@uneca.org">athiongane@uneca.org</a>
Présidents des comités:	
Comité de programme de Berlin 2003:	Danny Pfeffermann (Israël), <a href="mailto:msdanny@mscc.huji.ac.il">msdanny@mscc.huji.ac.il</a>
Comité de programme de Sydney 2005:	Pedro Do Nascimento Silva (Brésil), <a href="mailto:pedrosilva@ibge.gov.br">pedrosilva@ibge.gov.br</a>
Comité de nomination de 2003:	Oladejo Ajayi (Nigeria), <a href="mailto:ooajayi@infoweb.abs.net">ooajayi@infoweb.abs.net</a>
Jury du Prix Cochran-Hansen 2003:	Chris Skinner (Royaume-Uni), <a href="mailto:c.j.skinner@soton.ac.uk">c.j.skinner@soton.ac.uk</a>
Le secrétariat:	
Directeur exécutif:	Alain Charraud (France), <a href="mailto:alain.charraud@insee.fr">alain.charraud@insee.fr</a>
Trésorier:	François Fabre (France), <a href="mailto:francois.fabre@insee.fr">francois.fabre@insee.fr</a>
Secrétaire exécutive:	Anna Maria Vespa-Leyder (France), <a href="mailto:vespa@cepii.fr">vespa@cepii.fr</a>
Secrétaire:	Claude Olivier (France), <a href="mailto:claudelivier@insee.fr">claudelivier@insee.fr</a>



## Membres institutionnels

4 organismes internationaux

AFRISTAT  
CICRED  
EUROSTAT  
UNITED NATIONS STATISTICAL DIVISION

31 bureaux de statistiques

**ALLEMAGNE** - STATISTICHE BUNDESAMT  
**ARGENTINE** - INSTITUTO NACIONAL DE ESTADISTICA Y CENSOS - INDEC  
**AUSTRALIE** - AUSTRALIAN BUREAU OF STATISTICS  
**BELGIQUE** - INSTITUT NATIONAL DE STATISTIQUE  
**BRÉSIL** - INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATISTICA - IBGE  
**CANADA** - STATISTIQUE CANADA  
**CHINE** - GOVERNO DE MACAU  
**CÔTE D'IVOIRE** - INSTITUT NATIONAL DE LA STATISTIQUE - INS  
**DANEMARK** - DANMARKS STATISTIK  
**ESPAGNE** - INSTITUTO NACIONAL DE ESTADISTICA  
**ÉTATS-UNIS** - BUREAU OF THE CENSUS  
**ÉTATS-UNIS** - DEPARTMENT OF EDUCATION  
**ÉTATS-UNIS** - DEPARTMENT OF HEALTH AND HUMAN SERVICES  
**FINLANDE** - STATISTICS FINLAND  
**FRANCE** - INSTITUT NATIONAL DE STATISTIQUE ET D'ÉTUDES ÉCONOMIQUES - INSEE  
**GAMBIE** - CENTRAL STATISTICS DEPARTMENT  
**GRÈCE** - NATIONAL STATISTICAL SERVICE OF GREECE  
**IRAN** - STATISTICAL CENTER OF IRAN  
**ITALIE** - INSTITUTO CENTRALE DI STATISTICA - ISTAT  
**MEXIQUE** - INSTITUTO NACIONAL DE ESTADISTICA, GEOGRAFIA E INFORMATICA - INEGI  
**NIGÉRIA** - FEDERAL OFFICE OF STATISTICS  
**NORVÈGE** - CENTRAL BUREAU OF STATISTICS  
**NOUVELLE-ZÉLANDE** - STATISTICS NEW ZEALAND  
**PAYS-BAS** - CENTRAL BUREAU OF STATISTICS  
**PORTUGAL** - INSTITUTO NACIONAL DE ESTATISTICA - INE  
**RÉPUBLIQUE DE CORÉE** - NATIONAL STATISTICAL OFFICE - NSO  
**RÉPUBLIQUE TCHÈQUE** - CZECH STATISTICAL OFFICE  
**ROYAUME-UNI** - OFFICE FOR NATIONAL STATISTICS  
**SUÈDE** - STATISTIQUE SUÈDE  
**SUISSE** - OFFICE FÉDÉRAL DE STATISTIQUE  
**TANZANIE** - BUREAU OF STATISTICS

6 universités, centres de recherche, compagnies privées de statistiques

**ARGENTINE** - UNIVERSIDAD NACIONAL DE TRES DE FEBRERO  
**DANEMARK** - SFI  
**ÉTATS-UNIS** - RESEARCH TRIANGLE INSTITUTE  
**ÉTATS-UNIS** - WESTAT  
**EUROPE** - A.C. NIELSEN MANAGEMENT SERVICE  
**FRANCE** - Institut National d'Études Démographiques - INED