

Éditorial du numéro spécial Sondages

Title: Editorial to the special issue on Survey Sampling

Guillaume Chauvet¹ et David Haziza²

La théorie des Sondages s'intéresse à la façon d'inférer sur une population constituée par exemple de ménages, d'entreprises, ou de consommateurs d'électricité, en se basant sur un échantillon de quelques centaines ou de quelques milliers d'unités seulement. La procédure d'échantillonnage dépend de la connaissance a priori que l'on a de la population. En présence d'une base de sondage unique, l'échantillon peut être obtenu par tirage direct. Dans certains cas, il est nécessaire de recourir à plusieurs bases pour couvrir l'ensemble de la population. Un échantillon est alors sélectionné dans chacune, et l'objectif est de les combiner afin de produire une estimation aussi précise que possible. Quand on ne dispose d'aucune base de sondage, des méthodes d'échantillonnage indirect sont généralement privilégiées. Les techniques de sondage constituent également une approche intéressante dans le cas où de très grands volumes de données sont disponibles, et qu'il est nécessaire d'en réduire la dimension afin de les rendre exploitables.

Au cours des dernières années, les taux de réponse pour les enquêtes auprès des ménages n'ont cessé de baisser. Des efforts particuliers ont été faits sur la relance des non-répondants afin d'optimiser les taux de retours. L'objectif est maintenant de mieux cibler les non-répondants visés par une relance, de façon à équilibrer les profils des répondants finaux et ainsi à limiter les risques de biais. Une fois la phase de relance terminée, la non-réponse résiduelle est prise en compte et modélisée afin d'être finalement corrigée.

Ce numéro spécial se propose de considérer ces problèmes, et d'apporter des solutions pour les traiter. Dans le premier article, Carl-Erik Särndal et Peter Lundquist examinent des méthodes permettant de réduire le biais de non-réponse à l'étape de la collecte des données et à celle de l'estimation. Les auteurs proposent un indicateur de déséquilibre pour le sous-ensemble de répondants. Ils montrent l'importance de cibler la relance des non-répondants afin que les estimations basées sur les répondants soient proches de celles basées sur l'échantillon complet pour des variables auxiliaires disponibles pour toutes les unités échantillonnées. Les méthodes proposées sont illustrées au moyen de données provenant de l'enquête sur les conditions de vie en Suède. L'article de Fulvia Mecatti et Avinash Singh propose une revue unifiée de l'estimation dans un contexte de bases multiples, où l'estimateur ponctuel est de type Horvitz-Thompson. Il est construit à partir des facteurs de multiplicité qui représentent, pour une unité donnée, le nombre

¹ ENSAI (CREST), Campus de Ker Lann, 35170 Bruz, France.

E-mail : guillaume.chauvet@ensai.fr

² Département de mathématiques et de statistique, Université de Montréal, QC, Canada H3C 3J7.

E-mail : haziza@DMS.UMontreal.CA

de bases de sondage sur lesquelles cette unité est présente. Cette approche est particulièrement attrayante lorsque le nombre de bases est plus grand que 2. Les auteurs présentent également un estimateur de la variance obtenu au moyen des facteurs de multiplicité.

Pauline Lardin-Puech, Hervé Cardot et Camelia Goga proposent dans leur article une revue des méthodes d'échantillonnage et d'estimation pour des données fonctionnelles. Ils présentent les estimateurs de Horvitz-Thompson pour des paramètres fonctionnels tels que la courbe moyenne ou la fonction de covariance, ainsi que pour des paramètres non linéaires. La convergence uniforme de ces estimateurs ainsi que des bandes de confiance asymptotiques sont obtenues sous des conditions générales. L'estimation de variance et l'utilisation d'estimateurs calés est également discutée, et les auteurs proposent une illustration dans le cas de courbes de charges individuelles de consommation d'électricité.

L'article de Lionel Qualité et Eric Graf fait le point sur les méthodes utilisées pour les enquêtes auprès des personnes et des ménages par l'Office Fédéral suisse de Statistique. Après avoir rappelé les principales enquêtes concernées, les auteurs présentent les techniques de coordination d'échantillon utilisées dans le cas d'enquêtes successives. Le cheminement suivi pour obtenir des estimateurs pondérés est ensuite détaillé. Ce cheminement inclut la prise en compte du plan de sondage, la correction de la non-réponse totale et le calage final des estimateurs, ainsi que l'utilisation éventuelle de partage des poids ou d'estimation composite. La prise en compte des valeurs aberrantes et de la non-réponse partielle, ainsi que l'estimation de variance, sont ensuite discutés. Les auteurs présentent finalement des pistes d'amélioration des méthodes actuellement utilisées.

L'article d'Elvire Demoly, Arnaud Fizzala et Emmanuel Gros a pour but de passer en revue les différentes phases du processus de production d'une enquête auprès des entreprises à l'Institut national de la statistique et des études économiques. Ainsi, la constitution de la base de sondage, la sélection de l'échantillon dont les problèmes d'allocation et de coordination, le traitement de la non-réponse partielle et totale, le traitement des valeurs atypiques, la diffusion ainsi que la gestion de la confidentialité, sont discutés en détails.

Ces contributions invitées sont complétées par l'article de Karim Claudio, Vincent Couallier, Yves Le Gat et Jérôme Saracco qui s'intéresse à l'estimation de la consommation d'eau sur la base d'un échantillon d'utilisateurs dont la consommation est télérelevée. L'article décrit de façon détaillée la mise en place d'un plan de sondage stratifié, et notamment la construction des strates et le choix de leur nombre, ainsi que la répartition de l'échantillon dans les strates. Plusieurs stratégies sont alors comparées, selon que l'estimation est obtenue à l'aide de l'estimateur de Horvitz-Thompson ou à l'aide d'un estimateur par calage.

Nous remercions l'ensemble des auteurs et des rapporteurs qui ont permis la publication de ce numéro spécial. Nous remercions également l'éditeur en chef Gilles Celeux de nous avoir donné l'opportunité de le préparer.