

## Dans ce numéro

Ce numéro de la revue *Techniques d'enquête* débute par un article spécial à la mémoire de M.P. Singh, rédacteur en chef fondateur, qui a dirigé la revue pendant 30 ans et en a fait la source internationalement reconnue d'information sur les derniers progrès concernant les techniques d'enquête et les méthodes de production de statistiques officielles qu'elle est aujourd'hui. Un grand nombre de collègues et amis proches qu'a comptés M.P. au fil des ans y partagent leurs souvenirs et évoquent sa carrière et ses contributions.

Dans le premier article ordinaire du présent numéro, Thompson discute de l'utilisation des plans de sondage à marche aléatoire pour l'échantillonnage d'une population réseautée. Il montre comment cette approche peut mener à des échantillons en réseau où les probabilités d'inclusion peuvent être estimées indépendamment de la façon dont l'échantillon initial de noeuds est choisi, ce qui donne des méthodes valides d'inférence fondée sur le plan de sondage. La sélection préférentielle de certains types de noeuds ou caractéristiques des graphes est possible grâce au choix du mécanisme de marche aléatoire. Il décrit des plans de sondage à marche aléatoire uniforme ainsi que ciblée, et présente certains exemples.

Durrant et Skinner examinent le recours à l'imputation et à la pondération pour corriger l'erreur de mesure dans l'estimation d'une fonction de répartition. Ils étudient diverses méthodes d'imputation par le plus proche voisin et d'imputation hot-deck, ainsi que la pondération par le score de propension à répondre sous divers modèles de réponse. Ils discutent des propriétés théoriques de ces méthodes et les comparent au moyen de simulations afin d'estimer la distribution de la rémunération horaire au Royaume-Uni d'après des données provenant de l'Enquête sur la population active. Ils concluent qu'une approche fondée sur l'imputation fractionnaire semble être celle qui, dans l'ensemble, est la plus efficace et la plus robuste.

Harms et Duchesne étudient le problème de l'estimation des quantiles en utilisant des données d'enquête. Ils calent une estimation interpolée d'une fonction de répartition sur des quantiles donnés d'une variable auxiliaire, puis inversent l'estimateur interpolé calé résultant de la fonction de répartition de la variable d'intérêt. Enfin, ils réalisent une étude par simulation afin de comparer leur approche à d'autres méthodes.

Dans leur article, Haziza et Rao proposent une nouvelle méthode d'imputation par la régression avec utilisation des probabilités de réponse. Cette nouvelle méthode mène à des estimateurs valides sous l'approche du modèle de non-réponse ou sous celle du modèle d'imputation. Sous la première approche, le mécanisme de réponse est modélisé paramétriquement et n'est pas limité au modèle de non-réponse uniforme, tandis que sous la seconde, les variables d'intérêt sont modélisées et la non-réponse est considérée comme étant ignorable. Les auteurs fournissent aussi des estimateurs de la variance sous leur méthode d'imputation. Ils présentent, pour l'estimation ponctuelle ainsi que l'estimation de la variance, des résultats de simulation qui témoignent des bonnes propriétés de la méthode proposée d'imputation par la régression.

L'article de Zanutto et Zaslavsky traite du problème de l'estimation dans le cas du recensement décennal de la population des États-Unis sous échantillonnage pour le suivi des non-répondants. Au lieu d'essayer d'obtenir l'information auprès de tous les non-répondants, un échantillon est tiré pour le suivi, ce qui pose un problème d'estimation pour petits domaines. La stratégie proposée consiste à prédire le nombre de ménages non répondants dans diverses catégories au moyen d'un modèle hiérarchique loglinéaire, puis à imputer des renseignements détaillés sur les personnes et les ménages selon la méthode d'imputation par donneur. L'idée, à la première étape, est de modéliser les caractéristiques du ménage en utilisant des covariables peu détaillées à des niveaux détaillés de géographie et des covariables plus détaillées à des niveaux plus élevés d'agrégation géographique. Une étude par simulation indique que les propriétés du modèle proposé se comparent favorablement à celles d'autres modèles.

Dans l'article de Théberge, on propose une nouvelle approche pour répartir l'échantillon de la Contre-vérification des dossiers (CVD) de 2006 qui vise à mesurer le sous-dénombrement du recensement et une partie du sur-dénombrement. Les estimations de la CVD sont utilisées conjointement avec les chiffres du recensement pour produire des estimations démographiques, lesquelles servent à établir les paiements de péréquation du gouvernement fédéral canadien aux provinces. L'approche proposée permet d'établir une répartition qui fournit un équilibre entre quatre objectifs. Elle consiste d'abord à calculer une répartition distincte pour chaque objectif. On prend ensuite pour chaque province la taille d'échantillon maximale sur chacune des répartitions. La répartition infraprovinciale de l'échantillon de la CVD est obtenue en utilisant la technique du calage pour effectuer un lissage de paramètres définis au niveau des strates.

Dans son article, Longford discute de la façon de concevoir une enquête lorsque l'on doit produire des estimations pour plusieurs petits domaines, pour lesquels les priorités varient éventuellement, par minimisation d'une somme pondérée des variances espérées. Il commence par développer ses idées dans le contexte de l'estimation directe, puis les étend à l'estimation composite qui combine l'estimateur direct à un estimateur synthétique. Pour illustrer les méthodes, il présente les résultats, sous diverses hypothèses, de la répartition de l'échantillon d'une enquête auprès des ménages entre les divers cantons suisses.

You et Chapman proposent une approche hiérarchique bayésienne de l'estimation pour petits domaines lorsque les erreurs d'échantillonnage des estimateurs directs sont estimées. Ils démontrent leur approche en produisant des estimations pour petits domaines à partir de deux ensembles de données et étudient sa sensibilité aux hypothèses de modélisation.

Khoshgooyanfar et Monazzah comparent des méthodes d'estimation pour petits domaines basées sur un estimateur synthétique, un estimateur composite et un estimateur empirique bayésien en vue de produire des estimations intercensitaires des taux provinciaux de chômage en Iran. Ils constatent que l'estimateur composite et l'estimateur empirique bayésien produisent l'un et l'autre des résultats satisfaisants.

La brève note de Gabler, Häder et Lynn, qui conclut ce numéro, constitue une extension intéressante de l'article publié antérieurement par Gabler, Häder et Lahiri dans *Techniques d'enquête* (1999). Elle offre une solution pratique au problème de la détermination de l'effet de plan lorsque des échantillons différents sont utilisés pour différents domaines exclusifs.

Enfin, nous tenons à souligner que *Techniques d'enquête* est maintenant disponible en ligne dans un format PDF entièrement interrogeable. Tous les articles publiés dans la revue peuvent désormais être consultés gratuitement en direct sur le site Web de Statistique Canada dès leur diffusion. Nous prévoyons également inclure les numéros antérieurs. Tous les articles parus dans les sept derniers numéros sont déjà mis en ligne et les travaux se poursuivent en vue d'ajouter ceux qui ont été publiés au cours des dix années antérieures. Une version imprimée de la revue continue d'être produite pour les abandonnés. Les anciens numéros peuvent être obtenus sur demande en version imprimée ou en format PDF scanné. La revue peut être consultée sur le site Web de Statistique Canada à l'adresse <http://www.statcan.ca/bsolc/francais/bsolc?catno=12-001-X>.

Harold Mantel, Rédacteur en chef délégué