Component of Statistics Canada
Catalogue no. 12-001-X Business Survey Methods Division

Article

In this issue

June 2000

Statistics Canada Statistique Canada

Canada

Survey Methodology, June 2000
Vol. 26, No. 1, pp. 1-2
Statistics Canada, Catalogue No. 12-001

In this issue

This issue of *Survey Methodology* continues the celebration of 25 successful years that was marked by the publication of the December 1999 issue. The first seven papers of this issue, by prominent statisticians working in survey methods, were invited to help mark this occasion but were not included in the December issue due to space limitations. I would like to extend a special word of thanks to all of the authors who helped to make these two celebration issues so special and memorable.

To start off this special issue Kalton reviews developments in survey research over the past 25 years since *Survey Methodology* first started publishing. He first describes developments in survey taking as a profession, specifically the rise of specialist journals and professional associations for survey methodologists, as well as the international and multidisciplinary aspects of the profession. He then reviews developments in survey methods including questionnaire design, data collection, non-sampling errors, sampling methods and estimation. Finally he discusses the rise in importance of panel surveys and international surveys, administrative data sources and analysis of survey data.

Bellhouse traces the parallel developments in survey taking and computing over the twentieth century. He first describes the interaction between census taking and the early development of computing machines and digital computers. Later, developments in scientific computation lead to the use of more sophisticated statistical methods and models. He concludes his story with discussions of the development of statistical software for surveys and of model-related methods.

Bailar discusses the role of statistics in census taking, with particular emphasis on errors in census counts due to census errors of various sorts and adjustment of census counts using sample based estimates of net undercount. The various sources of errors in censuses are described. Use of statistical methods for census evaluation, quality control in census processing, and imputation is also discussed. Using a model for census bias and variance, the potential efficacy of census adjustment procedures is illustrated.

Isaki, Tsay and Fuller consider estimation of census adjustment factors using data from the 1990 post enumeration survey. Their estimators are based on a components of variance model with a fixed linear predictor and a random effect describing the unknown true adjustment factor for each of 336 post-strata. They consider alternatives based on using an estimate of the full variance-covariance matrix of the direct survey errors of the post-stratum adjustment factors versus using only the diagonal elements. Use of the diagonal elements only can reduce the effects of instability in the estimate of the full variance-covariance matrix. In an empirical comparison they find that a compromise between these two extremes works best. They also restrict the model based adjustment factors so that the estimate of total population matches that obtained from the direct survey estimates of these adjustment factors.

Lachapelle and Kerr present an innovative use of a coverage study to examine the demographic estimates of population. Their approach decomposes the results from Statistics Canada's Reverse Record Check (RRC) to provide an additional source of data that can be compared to the more traditional administrative record based estimates of the components of growth. The objective of this comparison is to identify major sources of error in either the administrative record based or the RRC estimates. They also show how the error of closure can be decomposed into two parts: differences between the RRC and the Census estimates of enumerated population and differences between the RRC and administrative record based estimates of growth.

In their paper Feder, Nathan and Pfeffermann consider repeated sampling from a hierarchical population. At each fixed time point the population can be described by a two level model; first and second level random effects are then allowed to evolve stochastically over time. In particular, the case where second level units remain in the sample for only a few occasions, as for example in many labour force surveys, is considered. A two step estimation procedure is proposed. In the first step the two-level model is fit to each time point independently to obtain estimates of the fixed effects. Time series parameters are estimated in the second step. Sampling weights can be incorporated into both steps to account for possibly informative sampling.

Rivest and Belmonte propose measurement of the mean square error of small area estimators conditionally on the realized smoothing model. They propose a natural estimator for this MSE; however, the estimator can be quite unstable when there is a lot of smoothing. They also propose a correction for bias in the case that the distributions of the direct estimators are skewed. Finally

they investigate the properties of their estimator in an Empirical Bayesian context and illustrate their method using undercoverage data from the 1991 Canadian Census.

Shao addresses an important topic - the evaluation of cold deck imputation methods. Since computer technology continues to make it easier to store and access data from previous and related surveys, imputation methods that make use of this auxiliary data will become increasingly important. As a result, Shao takes the first steps in evaluating how various cold deck imputation methods will perform relative to other imputation methods.

Thompson and Frank discuss model based estimation for link-tracing designs. In link-tracing designs, links are followed from one respondent to another. Network sampling and snowball sampling are just two examples. After a general introduction to the area, they present several link-tracing designs. They then present a graphical model for the linked population. Finally they develop likelihood based inference procedures for such populations using data from link-tracing designs.

Théberge attempts to solve the problem of extreme weights due to the calibration estimator by relaxing somewhat the calibration equation requirements. In fact, the problem is one of minimization similar to that encountered in ridge regression. He also reviews other means of restricting weights. He discusses the asymptotic properties of calibrated weights, and provides necessary and sufficient conditions for the existence of restricted weights satisfying the calibration equation. He also outlines a way of formulating the estimation problem by controlling the significance given to the calibration equation, and describes various means of restricting weights that do not rely on the use of a specific distance. Finally, he suggests an estimator having restricted weights that is useful for small domains, and deals with outliers by developing a method similar to that used to handle extreme weights.

Two short notes conclude this issue. Losinger, Garber, Wagner and Hill present a case study in the care that must be taken when adjusting for non response in different waves of a survey. Finally, Shaffer looks at the estimation of regression coefficients using survey data when the assumption of fixed auxiliary variables is relaxed.

You may recall that the December issue of *Survey Methodology* was made available, on an experimental basis, in an electronic format on the Statistics Canada web site. There was also a web based survey to guage your reactions and preferences with respect to an electronic version of the journal. Although there was quite a bit of interest in an electronic version, it seems that the time is not yet ripe for publishing electronically on a regular basis. We will certainly be reconsidering this option in the near future, and your responses to the survey will help to improve any future electronic version. In the meantime, we will continue to publish a print version of the journal for the foreseeable future.

M.P. Singh