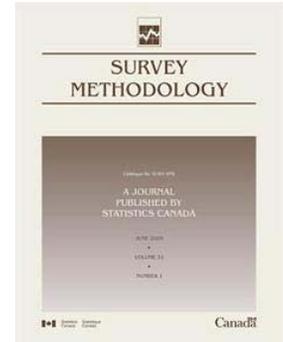


Article

In This Issue



June 2008

In This Issue

This issue of *Survey Methodology* includes papers on a variety of methodology topics.

In the first paper, Thompson and Wu consider the problem of obtaining inclusion probabilities, for the derivation of sampling weights, when modifications or compromises to the original sample design have been made due to practical constraints or limitations. The problem was motivated by the International Tobacco Control Policy Evaluation Survey of China which used a multi-stage unequal probability design for the selection of adult smokers and nonsmokers from seven cities. Due to refusal to participate by some districts, substitution units had to be selected after the original sample was selected. This substitution made it very difficult to calculate first order inclusion probabilities and practically impossible to calculate second order probabilities. In the paper the authors demonstrate, both theoretically and empirically, that the first and second order inclusions probabilities can be accurately estimated through Monte Carlo simulations.

Torabi and Rao derive the mean squared error (MSE) of a proposed new generalized regression estimator (GREG) of a small area mean under a two-level model and provide both theoretical and empirical comparisons between the new GREG and the best linear unbiased prediction estimator in terms of relative efficiency.

The paper by You discusses various cross-sectional and time series small area models for unemployment rate estimation for Canadian sub-provincial areas. In particular he considers an integrated non-linear mixed effects model under the hierarchical Bayes (HB) framework. An HB approach with the Gibbs sampling method is used to obtain estimates of posterior means and posterior variances of small area unemployment rates. The proposed HB model leads to reliable model-based estimates in terms of CV reduction. You also analyses the proposed model fitness and compares the model-based estimates to direct estimates.

The paper by Wang, Fuller and Qu studies small area estimation under a restriction. The authors study the impact of different augmented models in terms of MSE of the EBLUP. They consider small area models augmented with one additional explanatory variable for which the usual small area predictors achieve a self-calibrated property. They then consider small area models augmented with an added auxiliary variable that is a function of area size to reduce the bias when an incorrect model is used for prediction.

Nandram and Choi present an interesting approach to allocating undecided voters in surveys conducted prior to an election. Data from election polls are typically presented in two-way categorical tables with many polls taken before the actual election. They present the construction and analysis of a time-dependent nonignorable nonresponse model using Bayesian methods. They compare their model to extended versions (to include time) of ignorable and nonignorable nonresponse models introduced by Nandram, Cox and Choi (*Survey Methodology*, 2005). They also construct a new parameter to help predict the winner. The approach is illustrated using polling data from the 1998 race for governor of Ohio.

In their paper, Lazar, Meeden and Nelson develop a Bayesian approach to finite population sampling, through the use of a Polya posterior, when prior information is available in the form of partial knowledge about an auxiliary variable. The authors introduce the constrained Polya estimator and show that it has similarities with the generalized regression estimator under simple random sampling. However, their estimator does not require specification of a linear model. It is also related to empirical likelihood methods. Examples are used to illustrate the theory.

Zaslavsky, Zheng and Adams consider optimal sampling rates in element-sampling designs when the anticipated analysis is a survey-weighted linear regression and the estimands of interest are linear combinations of regression coefficients from one or more models. Methods are first developed assuming that exact design information is available in the sampling frame and then generalized to situations in which some design variables are available only as aggregates for groups of potential subjects, or from inaccurate or old data. Potential applications include estimation of means for several sets of overlapping domains, estimation for subpopulations such as minority races by disproportionate sampling of geographic areas, and studies in which characteristics available in sampling frames are measured with error.

The paper by Li explores the problem of estimating a finite population total using a nonlinear generalized regression estimator. The Box-Cox technique along with pseudo maximum likelihood estimation is used to obtain data-driven predictions. The author shows that the resulting regression estimator is design-consistent. Its performance is also evaluated through a simulation study.

Béguin and Hulliger extend the BACON algorithm to handle incomplete survey data. The BACON algorithm was developed to identify multivariate outliers using Mahalanobis distance. In the presence of missing values, the EM algorithm can be considered to estimate the covariance matrix at each iteration step of the BACON algorithm. The authors modify the EM algorithm to handle finite population sampling, which they call the EEM (Estimated Expectation Maximization) algorithm, and combine this algorithm with the BACON algorithm. This leads to the proposed BACON-EEM algorithm. It is then applied to two datasets and compared with alternative methods.

The paper by Jäckle and Lynn provides an empirical assessment of the effects of continued incentive payments on attrition, nonresponse bias and item nonresponse, and whether these effects change across waves of a multi-mode panel survey of young people in the UK. They test several hypotheses about the effects of incentives. They conclude that respondent incentives are an effective means of maintaining sample sizes of a panel, thus ensuring its value in terms of efficiency, especially for subgroup analyses. However, they also found that incentives had no effect on attrition bias.

Finally, Mohadjer and Curtin discuss challenges in designing and implementing a sample selection process that satisfies the goals of the National Health and Nutrition Examination Survey (NHANES). They describe how the sample design for NHANES must balance the requirement for efficient subdomain samples with the need for an efficient workload for the interview and examination staff at the Mobile Examination Centres (MEC), while keeping response rates as high as possible and costs down. The article elaborates on a number of unique features of the NHANES design and concludes with a brief summary of what has been achieved and some of the challenges facing future NHANES designs.

Harold Mantel, Deputy Editor