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

Article

In this issue

June 2009

Statistics Canada Statistique Canada

Canada

Survey Methodology, June 2009
Vol. 35, No. 1, pp. 1-2
Statistics Canada, Catalogue No. 12-001-X

In this issue

In the first paper of this issue of *Survey Methodology*, Estevao and Särndal consider the problem of calibration estimation in the context of two-phase sampling. The contributions of the paper include the choice of initial weights in the calibration procedure as well as the important problem of variance estimation. New variance estimators are proposed and results from a simulation study show that the proposed variance estimators are more efficient than the traditional ones.

Next, Li and Valliant investigate the problem of the detection of influential units in linear regression analysis of survey data. They first give an expression for the hat matrix and its associated leverages (diagonal of the hat matrix) when a weighted least squares technique is used to estimate model parameters. They then propose a decomposition of the leverages and highlight that the leverage for a given unit can be large when either its survey weight is large or its vector of explanatory variables is far from the center. They illustrate the effect of influential units on both ordinary and weighted least squares using a numerical example.

Beaumont and Bocci propose a bootstrap methodology for testing hypotheses about a vector of unknown model parameters when the sample has been drawn from a finite population. The technique uses model-based test statistics that incorporate the survey weights and can usually be obtained easily using standard software packages. Using a simulation study the authors show that the proposed method performs similarly to the Rao-Scott procedure, and better than the Wald and Bonferroni procedures when testing hypotheses about a vector of linear regression model parameters.

The paper by Park, Choi and Choi present an interesting approach to nonresponse. Studies have shown that the voting behaviour of the undecided voters can have a significant impact on the final result of an election and that by considering these undecided voters, the accuracy of election forecasting can be improved. The authors present two Bayesian models whose priors depend on information from both respondents and undecided. They analyze an incomplete two-way contingency table using four sets of data from the 1998 Ohio state polls to illustrate how to use and interpret estimation results for the elections.

Ghosh, Kim, Sinha, Maiti, Katzoff and Parsons develop hierarchical and empirical Bayes methods for estimation of proportions in small domains using unit-level models. They propose a hierarchical Bayes analogue of the generalized linear mixed model to obtain posterior means and posterior standard errors of the population small domain proportions. Using an approach based on the theory of optimal estimating functions, they also obtain emprical Bayes estimators and corresponding asymptotic mean square error estimators. The methods are illustrated using data from the National Health Interview Survey (NHIS) to obtain small domain estimates of the proportions of Asians without health insurance.

In the McElroy and Holan paper, the problem of testing for residual seasonality in seasonally adjusted data is investigated. The authors propose a statistical significance test for peaks in the spectral density of the time series under consideration that is indicative of seasonality. The theory of the proposed method developed and is illustrated and compared with existing methods through both simulation and empirical studies.

Gabler and Lahiri provide a model-assisted justification of the traditional interviewer variance formula for equal probability sampling with no spatial clustering. They then obtain, in the context of a complex sampling design, a definition of interviewer variability that appropriately accounts for unequal probabilities of selection and spatial clustering. They also propose a decomposition of total effects into effects due to weighting, spatial clustering and interviewers. Their results can help to more effectively understand and control sources of variability.

In their paper, Schouten, Cobben and Bethlehem investigate the problem of assessing the similarity between the response to a survey and the sample or population under investigation. They propose a representativeness indicator to replace response rates as a quality indicator for the impact of nonresponse bias. This indicator, called the R-indicator, is shown to be somewhat related to Cramer's V measure for the association between response and auxiliary variables. In fact, the R-indicator is better viewed as a lack of association measure since a weaker association implies that there is no evidence that nonresponse has affected the composition of the observed data. The theoretical properties of the proposed indicator are developed and it is illustrated through empirical studies.

Finally, in his article, Chauvet addresses the issue of balanced sampling when sizes in each stratum are too small for exact balancing. The author proposes an algorithm adapted to the Cube method, which guarantees balancing at the population level. A simulation study confirmed that the proposed method performed well.

Harold Mantel, Deputy Editor