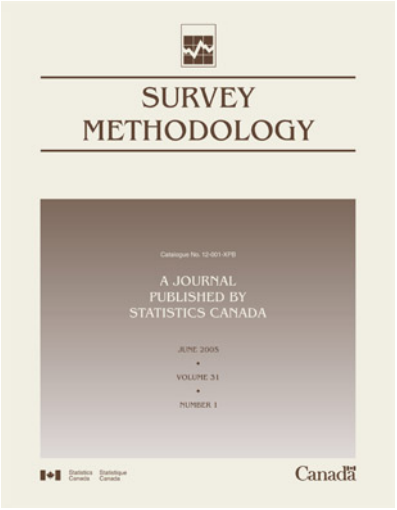
Catalogue no. 12-001-XIE

# Survey Methodology

December 2003




Statistics Canada Statistique Canada

Canada

## How to obtain more information

Specific inquiries about this product and related statistics or services should be directed to: Business Survey Methods Division, Statistics Canada, Ottawa, Ontario, K1A 0T6 (telephone: 1 800 263-1136).

For information on the wide range of data available from Statistics Canada, you can contact us by calling one of our toll-free numbers. You can also contact us by e-mail or by visiting our website.

| | |
|---|---|
| National inquiries line | 1 800 263-1136 |
| National telecommunications device for the hearing impaired | 1 800 363-7629 |
| Depository Services Program inquiries | 1 800 700-1033 |
| Fax line for Depository Services Program | 1 800 889-9734 |
| E-mail inquiries | infostats@statcan.ca |
| Website | www.statcan.ca |

## Information to access the product

This product, catalogue no. 12-001-XIE, is available for free. To obtain a single issue, visit our website at www.statcan.ca and select Our Products and Services.

## Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner and in the official language of their choice. To this end, the Agency has developed standards of service that its employees observe in serving its clients. To obtain a copy of these service standards, please contact Statistics Canada toll free at 1 800 263-1136. The service standards are also published on www.statcan.ca under About Statistics Canada > Providing services to Canadians.

Statistics Canada
Business Survey Methods Division

# Survey Methodology

December 2003

Published by authority of the Minister responsible for Statistics Canada



February 2006

Catalogue no. 12-001-XIE
ISSN 1492-0921

Frequency: semi-annual

Ottawa

Cette publication est disponible en français sur demande (nº 12-001-XIF au catalogue).

---

**Note of appreciation**

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued cooperation and goodwill.

Vol. 29, No. 2, pp. 105-106
Statistics Canada, Catalogue No. 12-001

# In This Issue

This issue of *Survey Methodology* includes a special section on Census Coverage Error which presents four papers, including two papers on the coverage survey used in the United States, one from Turkey, and one from Italy. The special section is preceded by a discussed paper, and followed by four papers on various topics.

In the first paper of this issue, Rao, Scott and Benhin study the repeated inverse sampling method proposed by Hinkins, Oh and Scheuren. In this approach, random subsamples are drawn from a complex sample in such a way that each subsample is unconditionally a simple random sample from the population. Rao, Scott and Benhin present some theoretical results for the expectation and variance of the repeated inverse sampling estimator. They then explore some conditions under which the repeated inverse sampling estimator converges to the original full sample estimator. They finally propose an approach based on estimating equations that avoids some of the potential bias of the repeated inverse sampling estimator for nonlinear parameters. The paper is followed by two fascinating discussions by Eltinge and Hinkins, and a rejoinder by the authors.

Hogan, in the first paper of the special section on Census Coverage Errors, presents a concise overview of the survey used to provide estimates for net undercoverage in the 2000 Census. He presents the Accuracy and Coverage Evaluation (ACE) study in the context of general post enumeration surveys and dual system estimators. He also presents the assumptions needed for these types of surveys to produce unbiased estimates and a detailed discussion where these assumptions failed in the 2000 ACE. The results are very interesting.

The next paper is also concerned with the 2000 ACE. Cantwell and Ikeda examine the crucial assumptions made when some data is missing. One of the points the authors note is that when a rare characteristic – persons missed by the Census in this case – is being estimated the methods used to adjust for missing data are very important. The authors point out the changes made from the methods used in previous post enumeration surveys for the 2000 ACE.

Ayhan and Ekni present the coverage procedures used in a different census context. While the basic post enumeration survey design is used in Turkey, there are some interesting differences between their experiences and those of the United States. Since Turkey uses a de facto approach to Census residence as opposed to the de jure approach used in the United States, there are some operational differences in the post enumeration surveys. These are clearly pointed out by the authors.

The final paper in the special section on Census Coverage Errors, by Cocchi, Fabrizi and Trivisano, describes the 1991 Italian Population Census and the the Post Enumeration Survey (PES) used to measure undercount. Since the census is administered by municipalities, data on the statistical quality of municipalities are used as auxiliary information for PES modelling and estimation. Poisson regression trees and hierarchical Poisson models are used to analyze the data. Results are summarized and discussed, and some recommendations are given.

Skinner and Carter extend estimation for Skinner and Elliot's measure of disclosure risk for survey microdata from the equal probability sampling case to the unequal probability sampling case under an assumption of Poisson sampling. Effects of possible departures from Poisson sampling are also considered.

The problem of inference for partially synthetic microdata sets is considered by Reiter. Statistical agencies may release microdata sets with completely synthetic data in order to protect respondent confidentiality. Methods for inference when the complete dataset is synthetic have been developed but most agencies release only partially synthetic datasets, that is, datasets for which only sensitive variables are imputed. There has been little reported in the literature under this situation. Reiter's proposed method is shown to be valid under a Bayesian framework and under a design-based framework and is illustrated by simulation studies.

In Brewer and Donadio, a variance estimator for the Horvitz-Thompson estimator that does not require the calculation of the second-order inclusions probabilities is obtained under high entropy situations. High entropy situations occur when there is the absence of any detectable pattern or ordering in the selected sample units. Under high entropy situations, an approximate variance formula is derived and verified through a model-assisted approach. A sample estimator for this approximate design-variance of the Horvitz-Thompson estimator is then developed. Finally, the proposed estimator is empirically compared with several other estimators using several populations.

Finally, Chow and Thompson present a Bayesian approach to designs where social links are exploited to obtain a sample of hidden or hard-to-access human populations. The authors provide an accessible introduction to the Bayesian approach in which the social links from one person to another are used to create the prior distribution. It is easy to adjust these priors when information is vague. The result is that from the resulting posterior distribution a large number of questions can be answered.

M.P. Singh