Editorial to the special issue on Survey Sampling

Titre: Éditorial du numéro spécial Sondages

Guillaume Chauvet[1] and David Haziza[2]

In survey sampling, we are interested in inferring on a finite population of, for example, households, businesses or electricity users, based on a sample of only few hundred or few thousands units. The sampling procedure depends on our a priori knowledge of the population. In the case of a single sampling frame, the sample can be obtained using a direct sampling procedure. In some cases, one must recourse to multiple sampling frames in order to cover the whole population. A sample is then selected within each frame and the goal is to combine them to obtain an accurate estimate. When no sampling frame is available, indirect sampling procedures are typically used. Also, sampling methods offer an interesting alternative in the context of large volumes of data when it is required to reduce the dimension, which in turns permits data exploitation.

Response rates have been steadily decreasing over time in household surveys. Efforts have been made for following up the nonrespondents in order to increase the response rates. Now, the objective consists of targeting the nonrespondents in order to balance the characteristics of the respondents at the end of the process, which may be useful for controlling the risks of bias. After data collection, nonresponse is treated at the estimation stage using some models.

This special issue considers these problems and discusses some solutions. In the first article, Carl-Erik Särndal and Peter Lundquist examine some methods for reducing the nonresponse bias at both the collection stage and the estimation stage. The authors propose a balance indicator for the set of respondents. They discuss the importance of targeting the nonrespondents so that estimates based on the full sample are close to estimates based on the respondents for a number of auxiliary variables available for all the sample units. The proposed methods are applied to the Swedish Living Condition survey.

The article written by Fulvia Mecatti and Avinash Singh gives an unified review of estimation in the context of multiple frames, where the point estimator is a Horvitz-Thompson type estimator. It is based on the multiplicity, which represents, for a given unit, the number of frames in which the unit is present. This approach is particularly attractive when the number of frames is larger than 2. The authors also present a variance estimator based on multiplicity.

In their article, Pauline Lardin-Puech, Hervé Cardot and Camelia Goga review sampling and estimation procedures for functional data. They present the Horvitz-Thompson estimator for

---
[1] ENSAI (CREST), Campus de Ker Lann, 35170 Bruz, France.
     E-mail: guillaume.chauvet@ensai.fr
[2] Département de mathématiques et de statistique, Université de Montréal, QC, Canada H3C 3J7.
     E-mail: haziza@DMS.UMontreal.CA

functional parameters such as the mean curve or the covariance function, as well as for non-linear parameters. The uniform consistency of these estimators as well as asymptotic confidence bands are obtained under general conditions. Variance estimation and calibration estimation are also discussed and the authors illustrate the proposed method using the case of a dataset of electricity load curves.

The article written by Lionel Qualité and Eric Graf describe methods used in household surveys at the Office Fédéral suisse de Statistique. After a brief description of the main concerned surveys, the authors present the methods of sample coordination used in repeated surveys. The steps involved in the derivation of weighted estimators are discussed in some detail. The steps include the sampling design, the treatment of unit nonresponse, calibration as well as the use of the weight sharing method or composite estimation. Then, the authors discuss the treatment of outliers and item nonresponse. Finally, they present some possible solutions for improving the methods currently used.

The article written by Elvire Demoly, Arnaud Fizzala and Emmanuel Gros reviews the different steps of a business survey at the Institut national de la statistique et des études économiques. The construction of the sampling frame, the sample selection from which stem the problems of allocation and sample coordination, the treatment of unit and item nonresponse, the treatment of atypical values, data dissemination and the problem of confidentiality, are discussed in some detail.

In addition to the invited contributions, the reader will find an article written by Karim Claudio, Vincent Couallier, Yves Le Gat and Jérôme Saracco, who are interested in estimating water consumption based on a sample of users, whose consumption is obtained by automatic water meter reading. The paper describes the stratified sampling design used in the survey, more specifically the construction of strata as well as the number of strata and the allocation within strata. Several estimation strategies are then compared, namely the Horvitz-Thompson estimator and calibration estimators.

We thank all the authors and referees who made this special issue possible. Our thanks go also to the Editor-in-Chief, Gilles Celeux, who gave us the opportunity to contribute to this special issue.