INTERNATIONAL STATISTICAL INSTITUTE

INTERNATIONAL ASSOCIATION OF SURVEY STATISTICIANS

THE SURVEY STATISTICIAN

n°16

SURVEY STATISTICIAN N° 16

CONTENTS

EDITORIAL

The present issue of the Survey Statistician doesn't come alone and I think we have to thank greatly Statistics Canada for so coming towards IASS members ; no doubt that the enclosed Survey Methodology issue will give rise to many subscriptions.

By another way, a problem is set in this issue of the Survey Statistician, which many members pointed out in the past. Several conferences or meeting are announced, the deadline of which will be in the next future, if not in the past when this issue will arrive. The bi-annual rythm of the Survey Statistician doesn't allow to fit the rythm with which announces come to the secretariat ; and the IASS cannot immediately dispatch information to members. In agreement with the Executive Director I thought however it was necessary to publish these announcements, at least to inform you about themes the statistical community takes interest in ; so, in case of need, you can ask further information to the organizers.

The Survey Statistician staff wishes to the IASS members a happy year 1987.

Gildas ROY

1. Election for the Council

The election for the next Council will take place during the first half of 1987. The voting ballots will be sent to all members, together with a notice about each candidate, aroud March, 1987.

2. Meeting of the Executive Committee

The meeting of the Executive Committee, which must be held at least once between two ordinary meetings of the General Assembly, took place on September 12 th, 1986, in Paris. I.P. Fellegi (President), D. Holt (Scientific Secretary), X. Charoy (Executive Director) and A.M. Vespa (Administrative Secretary) were present ; G. Théodore and E. Lunenberg (ISI Permanent Office) also attented a part of the meeting. The following item have been mainly dealt with :

- Survey Consultation by Correspondence : the problems arisen by V.K. Verma's departure from the I.S.I. Research Centre. A solution is sought for, but has not been found yet.

- "Survey Methodology" : a free copy will be sent to all members for advertising.

- Local Representatives : the President will write to them in order to

- Tokyo session : a seminar, for which funding is sought, will be organised before the session for some members from developping countries. This should unable them to take part to the session, for which they may have difficulty to get financing in their country. However, the number of participants will be limited .

- Various other seminars are planned :

  . seminar on censuses and surveys methodology in Nairobi (May 1987), sponsored by IUSSP (International Union for Scientific Study of Population) ;
  . seminar cosponsored by the U.S. Bureau of the Census and Statistics Canada on population censuses, for french-speaking countries ;
  . seminar co-sponsored by IASS and INSEE (Paris) in Lomé (Togo) or Abidjan (Côte d'Ivoire), at the beginning of 1988 ; the subjects are to be precised.

3. Call for dues

When this issue of the Survey Statistician will be published, the members of the Association will have received the call for dues 1987. Don't forget to answer quickly !

The Secretariate wishes to remind here two points :

a. The account of dues is expressed in French Francs. Previous accounts expressed in dollars should not be taken into account any more. However, it is possible to pay in other convertible currencies, using the exchange rate with the Fr. franc at the day of payment.

Amount of the dues

Fr. Francs 130

Fr. Francs 65 for members from developping countries.

b. It is possible, simultaneously, to subscribe at a preferential rate for the review of Statistics Canada : "Survey Methodology". On the call for dues, the amount of the subscription is expressed in French Francs (FF. 60 for one year, FF. 30 for members from developping countries). See below.

4. Survey Methodology Journal

Available at reduced rates

The Journal publishes articles dealing with various aspects of statistical development relevant to a statistical agency such as :

- design issues in the context of practical constraints
- use of different data sources and collection techniques
- total survey error
- survey evaluation
- research in survey methodology
- time series analysis
- seasonal adjustment
- demographic studies
- data integration
- estimation and data analysis methods
- general survey systems development.

Emphasis is placed on the development and evaluation of specific methodologies as applied to data collection or data themselves. All papers submitted for publication are refereed. The Journal is published twice a year in bilingual format (English and French). Contributions on the above mentioned topics are welcome. The editorial address is : Editor, Survey Methodology, Statistics Canada, Ottawa, Ontario, Canada KIA OT6.

The Editorial Board consists of R. Platek (Chairman), M.P. Singh (Editor), and, in alphabetical order, J. Armstrong, K.G. Basavarajappa, D.R. Belhouse, L. Biggeri, E.B. Dagum, W.A. Fuller, J.F. Gentleman, G.J.C. Hole, T.M. Jeays, G. Kalton, H. Lee, C. Patrick, J.N.K. Rao, C.E. Särndal, F. Scheuren, V. Tremblay, and K.M. Wolter.

According to an agreement between IASS and Statistics Canada the Journal is available to IASS members at a reduced rate. The price of an annual subscription is 60 French Francs for members in developed countries and 30 French Francs for members in developing countries.

The 1986 issues are special editions featuring selected papers presented at the Symposium on Missing Data in Surveys held in Ottawa, April 1986.

To subscribe to the Journal and/or to receive the special editions, you may send the following order form to the address indicated.

ORDER FORM

Mail to : IASS Secretariat, c/o INSEE, 18, boulevard Pinard,
75675 Paris Cedex 14, France

Name : _____

Address : _____ City : _____

Province : _____ Country : _____

Postal Code : _____ Telephone n° : _____

Enter my subscription to Survey Methodology, 2 issues at 60 Fr.F/30 Fr.F
Send me the special editions on Missing Data in Surveys, 2 issues at 60 Fr.F/30 Fr.F

Total Remittance : _____

Signature : _____ Date : _____

Modes of payment :

0 bank cheque here enclosed

0 bank transfer directly credited to the IASS bank account with the Banque Nationale de Paris in Paris (bank code : 30 004 ; agency code : 00874 ; account n° 004586-32)

it is necessary that your mention :

. the name of the establishment issuing the order : .....
.....

. its address : .....
.....

. the number of the order : .....

0 postal transfer sent directly to the account of the IASS : CCP Paris, Postal code 30041, agency code 00 001, account n° 12233182Z

0 UNESCO coupons here enclosed (for countries with inconvertible currencies)

Thanks.

The Secretariate heartfully thanks Statistics Canada who offers to all the members of the Association the issue of the journal "Survey Methodology" which is sent together with the present issue of the Survey Statistician. Let us hope that this will incite many of them to subscribe to this review and to read it carefully.

5. Tokyo session

a. Administrative meetings

The various sections of the I.S.I. will hold their administrative meetings during the Tokyo session. Here is the agenda of these which are planned for the I.A.S.S. :

Wednesday September 9th: Programme Committee
Thursday September 10th : Council meeting I
Monday September 14 th : General Assembly
Tuesday September 15 th : Council meeting II.

Other meetings (Local Representatives, other Committees) will take place, but the dates are not planned yet.

b. IASS WORKSHOP

A Workshop for survey statisticians from developing countries will take place in Tokyo, Japan, 4th-7th September, 1987, immediately preceeding the ISI Biennial Meeting. Funds are available for a number of participants to attend the Workshop and the ISI Meetings.

The Workshop will bring together statisticians from developing countries with a small number of other experienced people. The Workshop will be focussed on problem solving and each participant will be expected to provide a topic from which a selection will be made for use in the Workshop.

Statisticians who would like to be considered for inclusion should write as soon as possible to D. Holt, Department of Social Statistics, University of Southampton, Highfield, Southampton SO9 5NH, United Kingdom.
Applications should include :

(I) A brief outline (500 words) of a suitable topic for the Workshop.
(II) As much information as possible on travel costs from the applicants home to Tokyo.
(III) A statement describing any financial support already available to the applicant.
(IV) A brief description of education and work experience including current position.
(V) Address, telephone and telex numbers to enable fast communication.

Decisions on applications will be made in February/March of 1987 and applicants should respond without delay to this announcement.

c. IASS Question/Answer Session

As part of the ISI Conference in Tokyo, IASS will organise a question/answer session in which a panel of experienced survey statisticians will respond to questions put by participants. The emphasis will be on problems which arise in developing countries. Anyone planning to attend the Tokyo meeting will be welcome at the question/answer session. This notice is to encourage you to come prepared with a question which would be suitable for the panel.

For further details, or with suggested questions, please write to Professor D. Holt, Department of Social Statistics, University of Southampton, SO9 5 NH, United Kingdom.

6. Survey Consultation by Correspondence

Dr. V.K. Verma asked us to insert here a brief adress to all those who volunteered for this IASS service :

Dear Colleagues involved in the SCC :

I convey to you once again my appreciation for your volunteering to participate in the SCC service which the IASS has endeavoured to implement with the help of the ISI Research Centre. Your offer was communicated to IASS Executive Director and President.

However, it is with regret that I inform you that my direct personal involvement with execution of the SCC will conclude with my departure from the ISIRC as its Director on 15 November 1986. My involvement with IASS as its council member and as a SCC volunteer will of course continue.

Pending the nomination of my successor at the ISIRC, Mr. X. Charoy as Executive Director of IASS has kindly accepted, as an interim measure, to coordinate further development and implementation of the SCC.

I wish the SCC, this important intitiative of our association, success.

Vijay K. Verma

Besides that, additionnal members volunteered to give consultations by correspondence. They complete the list which was published in the Survey Statistician, n°15.

M. Ahmad (Pakistan)
Paul Burke (U.S.A.)
G. Büti (Switzerland)
G. Jones (Canada)
Carol Kindel (U.S.A.)
Guillermo Llanos (Peru)
Bakary Sacko (Mali)
A. Sen (Canada)
D.J. Trewin (Australia)
V.K. Verma (United Kingdom)

So, we now have 27 consultants. But no request for consultation has been received yet ! Until V.K. Verma is replaced in his function of responsible for Survey Consultation by Correspondence, requests can be mailed either to D. Holt (Scientific Secretary) or to X. Charoy (Executive Director) who will transmit to an appropriate consultant.

7. Publication of a handbook by Pr. T. Dalenius

Professor Tore Dalenius has recently published a handbook on survey methods entitled "Elements of Survey Sampling". The book is written for practising statisticians and students of statistics in developing countries. Professor Dalenius presents the most basic and important principles in survey sampling in an easy-to-grasp and pedagogical way. The book also serves as a guide to the vast literature on survey sampling. IASS members in developing countries will receive two copies as a gift from statistics Sweden.

LARS LYBERG
STATISTICS SWEDEN.

The President and all members of the Council heartfully thank Statistics Sweden for that generous

1. Announcement and call for contributed papers, international telephone survey methodology conference, November 1987.

An International conference on telephone survey methodology will be held in the U.S. in november 1987. The exact time and place of the conference will be announced at a later date. The organizing committee seeks abstracts for contributed papers. Topics of interest include coverage of the household population by telephone, sample designs for telephone surveys, administrative orgainization of telephone survey data collection, nonresponse in telephone surveys, measurement error properties of telephone surveys, and computer assisted telephone interviewing (CATI) systems. Research papers are most desirable but review papers or descriptive papers on experiences in conducting telephone surveys could also be appropriate if they illustrate widely applicable properties of the methodology.

The organizing committee will screen the abstracts obtained, and it is essential that abstracts are explicit about the content and the nature of the data analyzed. Participants in the conference will come from around the world. The conference will produce a monograph that includes the edited invited papers. Some contributed papers may also be included in the monograph. Other selected papers will be published as a special journal issue. Those interested in proposing a contributed paper should mail a one to two page abstract describing the paper to :

Robert Groves
Survey Research Center
The University of Michigan
Ann Arbor, MI 48109-1248, USA

To be considered for the conference, abstracts must be sent before april 1, 1987.

2. Census Bureau's Third Annual Research Conference (ARC III)

The program for the Census Bureau's Third Annual Research Conference has been set. Response to the Bureau's first "Call for Papers" was overwhelming. More than 115 abstracts were received from which only 30 could be selected. The choices were very difficult ; many fine papers and ideas had to be turned away.

ARC III is scheduled to begin Sunday evening, March 29, 1987, and run through 2 p.m. on Wednesday, April 1, 1987, at the Omni International Hotel in Baltimore, Maryland.

The overall theme is "Improving the Quality of Survey Results." The program will devote one session to each of 14 topics within the overall theme, and also include the Shirley Kallek Memorial Lecture. Session topics include historical perspectives on quality, quality issues in economic data, questionnaire design for automated interviewing, automating the survey process, designing automated data editing systems, effects of confidentiality protection methods on data use, issues in response error, quality improvements in statistical programs in Mexico, modeling survey estimates, quality of housing statistics, evaluation of imputation procedures, designing quality into censuses and surveys, management issues in surveys and censuses, and quality issues in income statistics.

For additional information contact Carolee Bush, Office of the Director, Bureau of the Census, Washington, DC 20233, 301/763-4900.

3. Eighth National Conference on Demography : Demographic Projections.
(Grenoble - France - 5-7 may 1987)

The purpose of that Conference is to analyse methods and tools used today in demographic projections, but also to develop a reflexion about stakes of projections : differences between forecasting, perspectives and projecting, basis and critical examination of hypothesis, links between national and regional levels, technics of projecting specific populations : total strength of company, an administration, etc.

The provisionnal agenda is as follows : history of projections and comparison with actual changes, projections of french population starting from 1982 census, regional and local projections, projecting subpopulations, demographic analysis and developments.

For further information, please contact :

Paul Boyer
Secrétariat du Colloque de Démographie
Université des Sciences Sociales
38040 - Grenoble 47X
FRANCE
Tél : 76.54.81.78

4. First International Conference on Statistical Data Analysis based on the L1-Norm and Related Methods. (Neuchatel, Switzerland, 31 august-4 september 1987)

The L1-norm criterion as a base for statistical data analysis which may include estimation or model fitting (constrained or unconstrained), analysis of variance, hypotheses testing, cluster analysis and nonparametric analysis, depending on the field of application has been studied in several contexts under a variety of names : minimum or least sums of absolute errors ; minimum or least absolute deviations, or errors, or residuals ; and L1-norm method (from minimizing the L1 norm of the vector of deviations). Historically, the L1-norm method in simple models was suggested and studied in the early work of Boscovich (1757) and Laplace (1793).

While the method of least squares, which is in fact the minimization of L2-norm, enjoys certain well know optimality properties within strictly Gaussian parametric models, the least absolute error is a widely recognized superior robust method specially well-suited to longer tailed error distributions. Increasingly, the minimum L1 estimator is recommended as a preliminary (consistent) estimator for onestep and iteratively reweighted least squares procedures.

Recently, considerable progress has been made in this field and a substantial amount of knowledge has been accumulated. But the know-how is too spread out. A collective effort is therefore needed to bring all of the past learning together and to help give birth to a coherent field. The preliminary purpose of this conference therefore is to bring the knowledge together in order to simplify the theory, from which tremendous possibilities for applications will result.

The conference takes place over five days and will include Invited talks and Contributed papers. The following sessions are scheduled : - Histcrical Development :
- Computational Aspects ; - L1 Estimation of Linear Models ; - Constrained L1 Estimation ; - Tests and Diagnostics ; - L1 Analysis of Variance ; - L1 and the Analysis of Tables ; - L1 Cluster Analysis ; - L1 Nonparametric Estimation ; - Statistical Packages ; - Applications.

5. European Population Conference
(Jyväskylä - Finland - 11-16 june 1987)

Sessions will be as follows :

. Societal values and population policies
. The demography and minorities
. The changing marriage system's effects on fertility and children
. Social and geographical contrasts in major chronic diseases
. Population forecasting and planning
. Changing patterns of household structures
. Family policy and population change
. The correlates of the regional fertility transition
. Fertility fluctuation in conception and births
. Utilization of data sources for new analytical demands
. The demographic implications of highly effective birth control methods
. Pre-natal and infant mortality
. Economic activity and fertility
. Social and economic consequences of ageing populations
. Internal migrations and regional development
. Urban historical demography
. 1990 round of census
. The dynamic of labour force participation
. The family support of the elderly
. Migration of the elderly
. International labour migration
. Early contribution to demography
. Population and technology
. Economic demographic models.

For further information, please contact :

Dr. Altti Majava
Chairman of the Organizing Committee
c/o FINNCO - TRAVEL EXPERTS
PL 722, SF-0101. Helsinki - Finland

6. Symposium on Statistical Uses of Administrative Data

Faced with an increasing demand for more frequent and detailed data for a wide range of social, economic and business variables on the one hand and with increasing response burden and budget constraints on the other, the statistical agencies in different countries are engaged in investigating and evaluating the uses of administrative records to meet needs for data. This symposium sponsored by Statistics Canada will focus on the methodological researches being carried out on the subject and will include sessions on topics such as uses of administrative data in the development of sampling frames, in estimation and in evaluation of the quality of survey data. Sessions concerning conceptual integration of data, record linkage, uses of longitudinal administrative data and policy and public relations issues will also be included to provide a fuller perspective.

Speakers from various government departments and institutions within Canada and other countries will be invited to participate. The symposium will be held in Ottawa during the last week of November 1987. If you have a paper to contribute or need further information, please contact Dr. M.P. Singh, Methodology Branch, 4th Floor C-2, Jean Talon Building, Statistics Canada, Ottawa, Ontario, K1A 0T6, Tel: (613) 990-9894.

7. First International Conference on Statistical Computing
(Cesme, Izmir-Turkey, 30 March-2 April 1987)

The First International Conference on Statistical Computing has as its dual purposes, interaction among experts at the forefront of statistical computing broadly conceived and exposition of its recent advances to those to whom they are important. Topics of importance for the conference include (but are not limited to) : - Expert Systems ; - Simulation and Modeling ; - Industrial Applications ; - Density Estimation ; - Artificial Intelligence ; - Statistical Computation ; - Scientific Applications ; - Statistical Consulting.

Eminent researchers at the forefront on statistical computing have agreed to present talks with the needs of the novice as well as the expert in the mind.

The deadline for contributed papers was 1st november 1986, but it seemed to us interesting to inform members about this first conference.

Further details can be obtained from :

Professor Aydin Oztürk, Conference Secretary
First International Conference on Statistical Computing
Ege Universitesi, Bilgisayar Arastirma ve Uygulama Merkezi
35100 Bornova, Izmir, TURKEY
Telephone : (51) 181080-(51) 181802

8. Arab Demographers Association

The Arab Demographers Association (ADA) was established in 1984, as a nonprofit professional association which aims to promote population research and policies in the 21 memberstates of the League of Arab States.

The First General Conference of ADA is scheluded to take place in Hammamet, Tunisia from 9-13 March 1987. The conference will contain 10 scientific and several informal sessions.

Although only Arab nationals can be full members, ADA accepts in principle non-Arab individuals and organizations actively engaged in population studies and policies relating to the Arab world, as associate (non-voting) members.

Associate members will be listed in the forthcoming ADA directory, receive information on all ADA activities, be eligible to participate in conferences and will receive the ADA newsletter.

For further information and application write : Arab Demographers Association, Université de Tunis, B.P. 237, Le Bardo 2000, Tunisia.

1. A list of the Local Representatives has been published in the Survey Statistician n°14 and completed in n°15. Here are two modifications :

BOTSWANA : Mr. F. MODISE
Government Statistician
Private Bag 0024. GABORONE
BOTSWANA
(replacing R.J. PEMBER)

BANGLADESH : Dr. MOSLEHUDDIN
Associate Professor
Dept of Statistics
Dhaka University
DHAKA
BANGLADESH
(replacing Dr. Siddiquer RAHMAN)

2. Country report
Federal Republic of Germany

In N°. 15 - June 1986 - of "The Survey Statistician" on pages 12 and 13 a brief report was given on a preliminary meeting of the section "Methods of Statistical Surveys" (Deutsche Statistische Gesellschaft and IASS) in Bonn on 24 September 1985.

In the following a brief report on the development and activities of the section is given :

A second meeting took place with Professor Dr. Heinrich Strecker in the chair on 23 September 1986 in Frankfurt am Main during the Annual Session of the Deutsche Statistische Gesellschaft. Three papers were presented and followed by discussions :

a) Non-response in official statistics in Sweden - development and extent of non-response and methods to meet the problems involved (Director E. Rapaport, Statistics Sweden; Stockholm)

b) Micro-Census in Austria, open questions, obligatory and voluntary information, experience in conducting these surveys (Diplom-Ingenieur A. Haslinger, Osterreichisches Statistisches Zentralamt, Wien)

c) On the method of surveys by telephone calls in the Federal Republic of Germany (Diplom-Volkswirt C. von der Heide, Infratest, München and Diplom-Mathematiker G. Rösch, München).

The General Assembly of the Deutsche Statistische Gesellschaft with Professor Dr. H. Grohmann, University of Frankfurt am Main, in the chair decided on 25 September 1986 to establish the section "Methods of Ststistical Surveys" as a permanent one of the Society in co-operation with IASS. Professor Dr. H. Strecker, University of Tübingen, was elected as chairman and Professor L. Herberger, Statistisches Bundesamt, Wiesbaden, as vice chairman.

The tasks of the section will be to deal with issues concerning planning, data collection and quality of data, methods of processing and keeping data available. Special attention will be given to the theory and application of sampling. Besides, contacts to other societies such as IASS (International Association of Survey Statisticians) and IAOS (International Association for Official Statistics) will be fostered.

H. Strecker, Local Representative of IASS, Federal Republic of Germany.

This column is conducted by Pr. D. Holt; Department of Social Statistics, University of Southampton, SO 95 NH, U.K.

1. The linear non-homogeneous estimator in sample surveys

Author : Stephen Haslett, Institute of Statistics and Operation Research.Victoria University of Wellington. Private Bag. Wellington

Journal : Sankhya, 1985, Ser. B, Pt.1, Vol. 47, pp 101-117.

Traditional sample survey estimators, their variance, and estimated variance are the central topic of many standard texts on sample survey design. Such texts derive the relevant formulae as a series of particular results for non-informative sample designs. All common estimators of means and totals are, however, linear in the values of the responses for the sampled population units, and can be written in the form of a linear non-homogeneous estimator, namely

$$\hat{Y} = W_{0(s)} + \sum_{i=1}^{N} W_{i(s)} Y_i$$

where $\hat{Y}$ is the estimator ; i denotes one of N possible respondents ; $W_{i(s)}$ are random variables (dependent on i and possibly the sample s), which are zero when $i \notin s$. (Y_i) is the set of random variables associated with the N population units and in the traditional sample survey case is (Y_i) a set of fixed population values. Interestingly then, for traditional sample survey estimators only the $(W_{i(s)} : i = 0, 1, 2, \ldots, N)$ are random quantities. For example, for simple random sampling with or without replacement $W_{0(s)} = 0$ for all s, and $W_{i(s)} = 1/n$ if $i \in s$ and zero otherwise, where n is the sample size.

In this paper, general formulae for expected value, bias, variance, mean square error and estimated variance of $\hat{Y}$ are derived within the traditional sample survey design framework. These design based formulae include as special cases, formulae for such traditional sampling estimators as the expansion estimator for simple random sampling (with or without replacement), the Horwitz-Thompson estimator, the expansion estimator for two stage cluster sampling or stratified random sampling with or without replacement and ratio estimation, among others. In a number of cases, these are derived explicitly. The general formulae have two advantages in the traditional survey context : they provide a coherent framework for the teaching of traditional sample design, and also a design tool in the form of a general method for variance estimation for surveys with known inclusion and joint inclusion probabilities.

While the $(W_{i(s)})$ have stochastic properties with respect to the sample design, p(.), in a superpopulation context $Y_1, Y_2, \ldots, Y_N$ are also random variables with respect to some superpopulation distribution, ξ. (Such a superpopulation need not involve Bayesian assumptions, and is particularly useful when considering regression coefficients calculated from survey data, for example.)

Formulae for expected value, bias, variance, mean square error, and estimated variance of $\hat{Y}$ with respect to the joint distribution ξ are also derived, together with formulae for the corresponding joint survey design, superpopulation quantities. Design and superpopulation expectation for non-informative sample designs ere shown to be conditional expectations.

The superpopulation formulae are of particular interest in optimisation contexts.

The central issue addressed in the paper, however, remains the systematisation of what were previously a series of disparate formulae. There are consequent benefits both to teaching sample survey techniques, and to variance estimation for a wide range of survey designs.

2. Efficiencies in Variance Estimation for Complex Survey Data

Authors : Steven B. Cohen, Vicki L. Burt, and Gretchen K. Jones

Data from complex survey designs require special consideration, with regard to variance estimation and analysis, because of design components that include unequal selection probabilities, stratification, and clustering. Using data from the National Medical Care Expenditure Survey (NMCES) (U.S.), which is characterized by a complex survey design, four variance estimation programs are compared : SESUDAAN/RATIOEST, SUPERCARP, PSALMS, and HESBRR. The comparisons concentrate on program capabilities, computational efficiency, and user facility. The study was also designed to measure the effect of alternative specifications for data base size and type of statistic on program performance.

This software comparison was directed to variance estimation tasks associated with health care estimates, presented in the same table format as the NMCES Data Preview Series. The table format consisted of 10 univariate demographic breakdowns in addition to a single cross-classification of 3 demographic measures. It should be recognized that the SESUDAAN/RATIOEST and the HESBRR procedures were specifically developed for the derivation of population estimates and associated standard errors in a cross-tabulated format, whereas the other two software packages were not primarily designed for this purpose. The programs, however, possess other unique capacities that distinguish them. As a consequence of the frequency of generating standard errors for such NMCES reports, the identification and subsequent use of the most efficient software procedure should yield substantial savings in survey costs. The SESUDAAN/RATIOEST procedure is the recommended program of choice for these operations. The SESUDAAN/RATIOEST procedure was consistently superior in terms of programming facility, whether implemented to obtain standard errors for statistics expressed in mean, total, or ratio form. The SESUDAAN/RATIOEST procedure was also the most efficient program in terms of CPU time used. This finding was consistent over all specifications of data base size and type of statistic. Since the evaluation of program performance has focused on a setting for which the SESUDAAN/RATIOEST was specifically designed, this would largely explain superior performance in terms of computational efficiency and user facility.

3. Measuring Gross Flows in the Labor Force : An Overview of a Special Conference

Journal : Journal of Business and Economic Statistics (January, 1986)

Authors : Carma R. Hogue and Paul O. Flaim

The measurement of gross flows in the American labor market has challenged researchers for years. The volume of these flows between labor force states cannot be determined from the monthly published labor force statistics because the flows are largely off-setting. Several problems, including response variability, rotation group bias, noninterview and mover effects, and matching and clerical errors, have been identified as possible reasons for the large discrepancies between the published net changes and the gross flows.

The Bureau of the Census and the Bureau of Labor Statistics jointly sponsored a conference to examine these problems, present current research on solutions to the problems, and submit recommendations for improving the data. Conference participants presented several alternative estimation procedures which are summarized in the article. One adjustment procedure used stochastic models to adjust the flow data for nonresponse. Several participants addressed the problem of response variability with several proposals for response error adjustments. After presentations of the adjustment procedures, some alternatives to using gross flow data were discussed. Then, researchers from other countries reported on the quality and availability of flow data for countries outside of the United States.

In addition to proposing alternative estimation procedures for the gross flows, researchers recommended several Current Population Survey (CPS) processing changes which would improve the quality of the flow data. Some of the recommendations (e.g. assigning unique identifiers to improve matching) would improve not only the gross flows but perhaps also the cross-sectional data (e.g. better imputation and editing procedures) that are estimated from the Current Population Survey. Some of the recommendations included following movers and asking retrospective questions of persons in the sample for the first time. This would enable researchers to use a large portion of the unmatched data which are currently omitted from the gross flow estimation. Alternative reinterview procedures which emphasize changes in status also were recommended in order to gain better estimates of response errors. The use of computer-assisted telephone interviewing was suggested in order to ease the burden of record keeping and provide better quality data.

The processing recommendations and the research findings provided at this conference established a solid foundation for further research and CPS procedural changes at the Bureau of the Census and the Bureau of Labor Statistics.

Key words : gross changes ; longitudinal data ; response variability ; dynamic flow data.

4. Sample Attrition and Labor Supply Response in Experimental Panel Data

Authors : Philip K. Robins and Richard W. West
Journal : Journal of Business and Economic Statistics -July 1986-

Sample attrition is a potentially serious problem for analysis of panel data. This is because if the attrition process is not random with respect to unobservable variables, estimated parameters of a behavioral model based on the sample of nonattriters will be biased. For experimental panel data, such as those collected in the U.S. income maintenance experiments, sample attrition can be an especially serious problem because the experimental transfer payment provides a differential incentive for experimental families to remain in the sample, relative to control families.

In this article, a variety of estimation procedures are used to assess the importance of attrition bias in the Seattle and Denver Income Maintenance Experiments (SIME/DIME), the largest of the U.S. income maintenance experiments. The unique feature of this study is the availability of data on both attriters and nonattriters. The data on attriters were obtained from Social Security earnings records. Because of privacy requirements, the Social Security Administration only provided the data on an aggregate basis in cells that would not permit identification of specific individuals. The availability of data on attriters allows an evaluation of the usefulness of alternative estimation procedures based on the sample of nonattriters. In this study, the focus is on the importance of attrition bias in the estimation of labor supply response to SIME/DIME. The methodology used in this study, however, is general enough to be applied to behavioral models in other panel data sets.

The empirical results suggest that standard procedures of correcting for attrition bias do not always yield the proper results. The use of these procedures, however, depends to a large extent on the degree of attrition in the sample. In the SIME/DIME panel, attrition was fairly modest (11% for married men and 7% for the other groups) and hence standard techniques simply do not have the power to precisely measure the biases caused by attrition. In panels in which attrition is much more extensive (perhaps 20% or more) such procedures are likely to serve a more useful purpose, although care must be exercised in modeling the attrition process. In general, the results presented in this article suggest that attrition bias is not a serious enough problem in the SIME/DIME panel to warrant extensive correction procedures. A similar conclusion was reached by Hausman and Wise in their analysis of attrition in the Gary Income Maintenance Experiment. Hence, standard analysis, using the sample of nonattriters, appears to be a reasonable approach to estimating behavioral response models in these particular data sets.

5. Statistical Matching Using File Concatenation with Adjusted Weights and Multiple Imputations

Author : Donald B. Rubin
Journal : Journal of Business and Economic Statistics -Jan. 1986-

Statistically matched files are created in an attempt to solve the practical problem that exists when no single file has the full set of variables needed for drawing important inferences. In the important and common case of two data bases, say A and B, one set of variables, say X, is observed in both A and B ; another set of variable, Y, is observed in file A but not file B ; and another set of variables, say Z , is observed in file B but not file A. Since no statistical file matching technique can actually recreate the true Y data for file B or the true Z data for file A, the created file has uncertainty that leads to uncertainty in resultant inferences based on the matched file. In particular, measures of partial association between Y and Z given X (e.g., partial correlations between Y variables and Z variables given the X variables) are inestimable by any procedure yet are usually assumed to be zero. Typically and incorrectly, such uncertainty is ignored when drawing inferences from statistically matched files.

In this article, previous methods of file matching are reviewed, and the method of file concatenation with adjusted weights and multiple imputations is described and illustrated on an artificial example. This method concatenates the files A and B and then multiply imputes values for each missing Y and Z value to reflect uncertainty in the correct value to impute.

Letting n_A and n_B be the number of units in the respective files, the result is a file of $n_A + n_B$ units with sampling weights and complete X data for all units, complete Y data for the A units, multiple versions (say, two) of the Z data for the A units, complete Z data for the B units, and multiple versions of the Y data for the B units. Each set of imputations creates a complete data set, which can be analyzed using standard complete data methods. A major benefit of this approach is the ability to display sensitivity of inference to untestable assumptions, which were made when creating the matched file, using only standard methods of analysis.

6. SOME ASYMTOTIC RESULTS FOR THE SYSTEMATIC AND STRATIFIED SAMPLING OF A FINITE POPULATION

Author : Philip S. Kott
Journal : BIOMETRIKA (1986). 73. Nº. 2 pp. 485-491.

This theoretical article analyzes the roles of size stratification and randomization in finite population sampling theory. When properly used, both are shown to have desirable asymptotic properties under broadly specified models.

The problem is to estimate the population mean of some variable, Y, using a sample of distinct units. The Y values are unknown for nonsampled units, but there exists a known auxiliary variable, X, which is somehow related to Y.

There is extensive literature on this subject when the functional relationship between Y and X is assumed to be linear with unknown parameters. Unfortunately, the exact relationship between the two variables is generally unknown and is likely to be much more complicated.

Consider a sample drawn systematically from a list ordered by X values. Under any model in which $E(Y_i)$ is a finite combination of bounded monotonic transformations of X ; the model bias of the sample mean (as an estimator of the population mean) is shown to be an asymptotically insignificant contributor to model mean squared error. The model bias and standard error of the sample mean under simple random sampling, by contrast, are of the same order (on average).

Given any form of systematic sampling, even with a randomly chosen start point, there is shown to be a model under which the sample mean fails to converge in probability to the population mean. By coupling a sample mean with a stratified random sample, however, one can retain the desirable model based property when Y and X are related in the broad manner discussed above, while assuring the design consistency of the joint estimation strategy under all models.

The analysis can be (and is) extended to more general and often more efficient estimation strategies involving probability proportional to size sampling and model based stratification schemes.

QUESTION/ANSWERS

Conducted by Leslie Kish. Please send Questions to him (ISR - The U of Michigan, Ann Arbor, 48106, USA), or, to IASS, Paris. Please indicate whether or not you want your name given with the question. This has become an open forum, and we shall gladly print (after refereeing) additions, modifications, discussions of past published answers. Contributors to answers will be acknowledged with their agreement.

Q. "Computing sampling errors" : what do you and others mean by that, beyond computing variances and standard errors? Why have these new terms for survey sampling, when we already have the well known variances?

A : The main reasons for introductiny the new concepts of sampling error arise from the usual needs for averaging, generalizing and inferring variances between statistics. Those needs are common, even typical, in sampling practice. The concepts and terms below are in common use, but not strictly and universally defined.

Variances for statistics have fairly clear definitions in many books, and for simplicity we shall focus on variances for means, $var(\bar{y})$, and on the standard errors, $ste(\bar{y}) = \sqrt{var(\bar{y})}$. These statistics estimate corresponding parameters, usually distinguished by capitals $STE(\bar{y}) = \sqrt{Var(\bar{y})}$, though greek letters are often used instead. We should avoid confusion with element variances s^2 and standard deviations $s = \sqrt{s^2}$. We note the relationship $var(\bar{y})=deft^2 s^2/n$, where $deft^2$ denotes the design effect, that includes here the finite population correction (1-f), and which comes to 1 for srs. There are problems about computing s^2 from complex samples, some fine (should n or n-1 be used) some important (what to do about weights), but we must leave those aside here.

Mean-square-errors should also be considered when large biases are suspected or possible, though computing them may not be feasible, thus $MSE(\bar{y})=Var(\bar{y})+Bias^2(\bar{y})$. (This is different from the "mean-squares" of the analyses of variance.)

Relvariances : (i.e., relative variances) are sometimes used for positive (or non-negative) variables, and they refer to $cv^2(\bar{y})=var(\bar{y})/\bar{y}^2$, the square of the coefficient of variation $cv(\bar{y})=ste(\bar{y})/\bar{y}$. They are more useful for averaging and inferring between statistics of a survey than variances because they remove the units of measurement, e.g., of income, of areas of cultivation, and body length, etc.

Furthermore multiplying with constants like n or N/n have the same effect on the numerator and the denominator of $cv(\bar{y})=ste(\bar{y})/\bar{y}$; the sample total y, the estimated mean $\bar{y}=y/n$ and the estimated total $\hat{Y}=N\bar{y}$ have the same cv and cv^2. Then the inference between the standard errors for $\bar{x}$ and $\bar{y}$ may take this path : $ste(\bar{y}) \rightarrow cv(\bar{y}) \rightarrow cv(\bar{x}) \rightarrow ste(\bar{x})$. But the cv^2 and the cv have two severe limitations : (1) They are unstable for statistics that can take negative values and center $(\bar{y})$ near zero, such as changes of income. (2) The $cv^2(\bar{y})$ are not only affected by sample design but also are inversely proportional to the sample size n (unlike the $deft^2$).

The coefficients of variation cv(x)=ste(x) of the sample size x are very useful for ratio mean r=y/x, and I believe they should be always computed and checked. Values of cv(x)> 0.1 or 0.2 are signals that the estimates r may have serious biases and that the estimates var(r) may be too unreliable (See textbook on sampling).

Design effects, $deft^2(\bar{y})=var(\bar{y})/var_0(\bar{y})$, are widely and commonly used, with the $var_0(\bar{y})$ base denoting simple random variance. For the mean this would be s^2/n or $(1-f)s^2/n$, depending on how the fpc=(i-f) is to be treated. $Deft^2$ have also been used under names like variance factors and efficiency factors, but those terms have had other connotations also. $Deft^2$ have broader utility than relvariances because : (1) They can be readily applied to statistics other than means and totals, e.g., coefficients of regression, correlation etc., also to test statistics. (2) The effects of sample size are removed by the denominator $var_0(\bar{y})$. $Deft^2$ are most useful for : a) checks on gross errors of computations, b) generalizing and inferring to other statistics of the same sample, and c) inferring to statistics if other similar samples. Inferring takes the path : $var(\bar{y})$-$deft^2(\bar{y})$-$deft^2(\bar{x})$-$var(\bar{x})$. However, generalizing and inferring with $deft^2$ must be done with caution because the value of $deft^2$ depend on the sample design. This may be seen most simply if we write $deft^2(\bar{y})=1=roh(\bar{b}-1)$, where $\bar{b}=n/a$ the average sample cluster size.

Roh, the ratio of homogeniety, is a synthetic value, computed as $(deft^2-1)/(\bar{b}-1)$. If clusters were equal it would equal (almost), rho, the coefficient of intraclass correlation. By removing the effect of the sample cluster size $\bar{b}$ we can generalize more readily to subclass statistics especially to crossclass means ; also to other samples that use different sizes of $\bar{b}$. By removing the effect of $\bar{b}$, the values of roh become more portable than the values of $deft^2$. Thus the path of inferrence may be sketched as : $var(\bar{y})$-$deft^2(\bar{y})$-$roh(\bar{y})$-$roh(\bar{x})$-$deft^2(\bar{x})$-$var(\bar{x})$.

Components of the variance would provide a theoretically sounder path of inferrence, by providing separate components for each stage of selection, and for the strata used for each. A four stage sample with stratification for each stage may need eight components. These are seldom computed because of the complexities of such computations and because of the instability of such residual statistics.

Averaging of sampling errors is often done to facilitate generalizing and inferring, also to reduce the instability of sampling errors, especially for those based on small degrees of freedom (independent replications), also for the sake of simplicity.

Tables of sampling errors are often used for presenting averaged sampling errors as appendices to survey reports.

Some recent references

. Kish, L. (1987) Statistical Design for Research, New-York, John Wiley and Sons, Section 7.1

. Kish, L. Groves, R.M. and Krotki, K. (1976), Sampling Errors for Fertility Surveys, WFS Occasional paper 17, The Hague : ISI.

. Verma, V., Scott, C. and O'Muircheartaigh, C. (1980), Sample designs and sampling errors for the World Fertility Survey, Jour. Roy. Stats. Soc. (A), 143, 431-473.

. Waterton, J. (1986), Software package for variance estimation in complex surveys, Survey Methods Newsletter, Autumn 86, London : SCPR, 11-13.

. Wolter, K.M. (1985), Introduction to Variance Estimation, New York : Springer-Verlag.