

INTERNATIONAL STATISTICAL INSTITUTE

N° 28

DECEMBER 1992

THE SURVEY STATISTICIAN

**INTERNATIONAL ASSOCIATION
OF SURVEY STATISTICIANS**

iasr

Editor : Anne Marie Vespa-Leyder

***Association Internationale
des Statisticiens d'Enquêtes
A.I.S.E. - I.A.S.S.
c/o INSEE - 18, Bd Adolphe Pinard
75675 PARIS CEDEX 14***

CONTENTS

1. NEWS FROM THE ASSOCIATION	p	3
1.1 Dues	P	3
1.2 Changes to the Survey Statistician	p	3
1.3 Local Representatives - Up-to-date list	p	4
2. FLORENCE SESSION - Workshops.....	p	9
3. BEIJING SESSION - Programme	p	10
4. COUNTRY REPORTS	p	11
4.1 Canada.....	p	11
4.2 United Kingdom	p	12
4.3 Pakistan.....	p	13
4.4 New Zealand.....	p	14
4.5 Malawi.....	p	15
4.6 Brazil.....	p	17
4.7 U.S.A.....	p	19
5. QUESTION/ANSWER	p	21

1. NEWS FROM THE ASSOCIATION

1.1 Dues :

For 1993, the dues remain fixed at F.F.130. (or F.F.65 for members from developing countries), or the equivalent in other currencies.

Members are reminded that they can use UNESCO coupons in some countries when the local currency is not convertible.

1.2 Changes to the Survey Statistician

Members of the IASS will probably have read of the planned changes to the Survey Statistician in the ISI Newsletter. These changes will take place beginning with the next issue of the Survey Statistician.

As the newly appointed editor of the Survey Statistician, I am seeking the views of fellow IASS members on their ideas for improvements. Offers of assistance or material for inclusion are especially welcome. In particular we are seeking :

Review articles - which highlight developments in a particular area. They could have a methodological, technological or substantive focus. Our current ideas for such articles are :

- disability surveys
- environmental monitoring
- PC software for sampling error calculation
- survey documentation
- omnibus surveys
- quality control in survey research.

If anyone has ideas for these articles or is willing to write on these topics I would be very pleased to hear from them.

Questions - the question and answer column will continue with a newly appointed "consultants team" assisting Leslie Kish. We will be examining ways of generating a wider range of questions from practising researchers with special focus upon issues of relevance in the developing world.

Meanwhile, if you have a question you would like this team to address, please send it to Leslie Kish directly or I can forward it for you.

Bibliographies - I would be interested in hearing from members who are willing to prepare annotated bibliographies on particular topics. Why not volunteer to do this for your area of specialism ?

The popular feature on country reports will continue. Gordon Brackstone will be the associate editor with responsibility for this section and he has ideas for making it more lively. Please give your country members items of news which you think will interest fellow IASS members.

My address is :

Professor Denise Lievesley,
ESRC Data Archive
University of Essex
Wivenhoe Park
Colchester, Essex CO4 3SQ, UK

Tél : 44 206 872009 ; Fax : 44 206 872003
E-mail : denise@uk.ac.essex

1.3 Up-to-date list of local representatives :

ALGERIE	Mr. Elyes BENTABET	5 rue Yahia Banaiche (ex Valentin) 16000 ALGER CENTRE
ARGENTINE	Mrs Clyde CHARRE DE TRABUCHI Mrs Clara BADLER	French 2740 1425 BUENOS AIRES Bd Orono 610 P2° 2000 ROSARIO
BELGIUM	Mr. Louis LOHLE-TART	Rue Americaine, 81 B. 1050 BRUXELLES
BOTSWANA	Mr. B.C. MOLOMO	Central Statistical Office Private Bag 0024 GABORONE
BRAZIL	Mrs E. PINHO DE CASTRO SILVA Mr. Wilton DE OLIVEIRA RUSSAB	Rua Tabapua 308 Apto 52 04533 SAO PAULO SP Universidade de Sao Paulo Instituto de Matematica e Estadistica CP 20570 AQ-IQUATEMI SAO PAULO 01451
CANADA	Mr. Gordon BRACKSTONE	78 Charing Road NEPEA ONTARIO K2G4C9
CHINA	Mr. Hui WU	State Statistical Bureau BEIJING

COTE D'IVOIRE	Mr. Bakary DAHO	Direction de la Statistique 01 BP V 55 ABIDJAN 01
CUBA	Mr. Luis SILVA	Instituto Superior de Ciencias Medicas de la Habana Vice Rectoria de Invest. e Posgrado Avenida 25 N 15005 PLAYA, HABANA 16
CZECHOSLOVAKIA	Mr. Vaclav CERMAK	Severozapadni 11 n° 18 Praha 4 14100
DENMARK	Mr. Finn MADSEN	Danish National Institute of Social Research Borgergade 28 DK 13000 KOBENHAVN-K
FRANCE	Mr. Jean-Michel REMPP	I.N.S.E.E. 18 Boulevard Adolphe Pinard 75675 PARIS CEDEX 14
GERMANY	Mr. Gotthard FORBRIG	Wilhelm Pieck Universität Klement Gottwald St. Nr: 6 ROSTOCK 1 2500
	Pr. Heinrich STRECKER	Rosenst. 11 D 8130 STARNBERG
GHANA	Mrs Elisabeth A. ALLOTEY	Industrial Census Coordinator Ghana Statistical Service P.O. BOX 1098 ACCRA
HUNGARY	Mr. Adam MARTON	Central Statistical Office Keleti K. N° 5-7 BUDAPEST 1525
INDIA	Mr. J. ROY	Indian Statistical Institute 203 Barrackpore Trunk Road CALCUTTA 35
INDONESIA	Mrs Sri Budianti SUKMADI	Central Bureau of Statistics J.L. DR Sutomo 8 P.O. Box 3 JAKARTA

ISRAEL	Mrs Malka KANTOROWITZ	Central Bureau of Statistics HAKIRYA - Romena JERUSALEM
ITALY	Mr. Claudio QUINTANO	Via Plinio il Vecchio n° 43 Castellamare di Stabia 80053
	Mr. Andrea GIOMMI	Dipartimento Statistico Universita di Firenze Via Curtatone 1 FIRENZE 50123
JAPAN	Dr. C. HAYASHI	Instituto of Statistical Matematics 4-6-7 Minami-Azabu TOKYO 106
KENYA	Mr. Philip GACHUKI	Central Bureau of Statistics P.O. Box 30266 NAIROBI
KOREA	Mr. Seung-Kon LIM	National Statistical Office 90 Kyeongun-Dong Jongro-Ku SEOUL
LEBANON	Mr. François FARAH	M.A.S.S. B.P. 208 HAZMIEH
LIBYA	Mr. Farouk EL BISHITI	Statistics Department Faculty of Science Al Fateh University P.O. Box 13574 TRIPOLI
MALAWI	Dr. Suresh Chandra BABU	UNICEF P.O. Box 30375 LILONGWE - 3
MALI	Mr. Bakary SACKO	ESPR C/O ODEM SEVARE MOPTI

MEXICO	Dr. Edmundo F. BERUMEN TORRES	INEGI Insurgentes Sur 795 12 Piso MEXICO D.F. C.C. 03810 28, Vuurbloem Str. 2317 LEIDEN
NETHERLANDS	Mr. S.J.M. DE REE	
NEW-ZEALAND	Mr. Léonard Warren COOK	Department of Statistics P.O. BOX 2922 WELLINGTON
NIGERIA	Mr. O.O. AJAYI	Federal Office of Statistics P.M.B. 12528 LAGOS
NORWAY	Mr. I.B. THOMSEN	Central Bureau of Statistics Dronningsgate 16 OSLO
PAKISTAN	Mr. Mohammad ISHAQ (S.O.)	HIES - Section Federal Bureau of Statistics RR. Ahmad Plaza, Plot N° 17 G-8 MARKAZ - ISLAMABAD
PANAMA and CENTRAL AMERICAN COUNTRIES	Mr. David FITCH	Instituto de Nutricion de Centro America y Panama Apartado Postal 1188 GUATEMALA, CA GUATEMALA
PAPUA-NEW GUINEA	Mr. John SHADOW	National Statistical Office P.O. Wards Strip
PERU	Mrs Leonora LAGUNA	581 Belisario Flores LIMA 14
PHILIPPINES	Mr. Tomas P. AFRICA	National Statistics Office (NSO) MANILA
RUSSIA	Mrs Galina N. SOTNIKOVA	Institute for Social-Political Research Krdzhidzhanovsky 24/35 b. 5 MOSCOU 117252

SENEGAL	Mme Awa THIONGANE	Direction de la Statistique B.P. 116 DAKAR
SPAIN	Mrs Carmen ARRIBAS	Instituto Nacional de Estadística Paseo de la Castellana 183 28016 MADRID
	Mrs Iztueta AZKUE ANJELES	Instituto Vasco de Estadística Servicio de Metodología y Diseño C/DATO 14-16 01005 VICTORIA-GASTEIZ (ALAVA)
SWEDEN	Mr. Anders CHRISTIANSON	Statistics Sweden S-115 81 STOCKHOLM
SYRIA	Mr. A. MALEK AL AKHRAS	Office of the Prime Minister-Central Bureau of Statistics DAMASCUS
TRINIDAD-TOBAGO	Mr. Carl ATHERLEY	Central Statistical Office 23 Park Street PORT OF SPAIN
TURKEY	Mr. Oztas AYHAN	Department of Statistics Middle Est Technical University ANKARA 06531
UNITED KINGDOM	Mr. Roger THOMAS	O.P.C.S. St. Catherines House 10 Kingsway LONDON WC 2B 6JP
U.S.A.	Dr. Daniel KASPRZYK	4906 Colonel Contee PL UPPER MARLBORO MD 20772-2825

2. FLORENCE SESSION - Workshops :

IASS WORKSHOPS, PADOVA, 1993

As previously announced, IASS plans to sponsor a program of short courses to be held at the University Degli Studi di Padova, immediately preceding the ISI Session in Firenze. Many IASS members and other survey statisticians, both from developed and developing countries, will want to attend one or more of these special courses.

The provisional program includes five separate courses :

1. Workshop on Survey Sampling in Developing countries

Instructors : Graham Kalton, Colm O'Muircheartaigh, Nanjamma Chinnappa and Anis Maitra

Language : English, Duration : 3 days

2. Variance Estimation in Complex surveys

Instructors : Wayne Fuller, Kirk Wolter and Michael Hidiroglou

Language : English, with assistance in French,
Duration : 1 day

3. Nonresponse in Surveys

Instructors : Donald Rubin and Roderick Little

Language : English, Duration : 1 day

4. Calibration Estimators and Generalized Raking

Instructors : Jean Claude Deville and Carl Eric Särndal

Language : French, Duration : 2 days

5. Survey Measurement : Methods for Error Evaluation and Reduction

Instructors : Paul Biemer, Barbara Forsyth and Judith Lessler

Language : English, Duration : 2 days

6. Environmental Sampling, Statistical Modeling and Observational Economy

Instructors : G.P. Patil et al.

Language : English, Duration : 3 days

3. BEIJING SESSION - Programme :

Members of IASS are invited to comment on possible items for inclusion as topics in the above session.

Recent developments are having, and will have, considerable impact on user needs and hence ways of meeting these, and on the traditional methods of data gathering and processing.

- * The changing role of the state in deregulated economies has altered the balance between administratively-sourced and survey-based information, with implications for survey frame construction, and choices about collecting information on transactors and transactions.
- * In many rapidly changing sectors, there is interest in measuring social and economic transitions taking place, and participation in sectors, rather than assuming a static population is available. This leads to an increased need for longitudinal data analysis for public policy formation.
- * Graphical editing techniques can take advantage of powerful workstations to improve outlier detection and management in very large and complex data sets.
- * Users seek the linking of data from a range of statistical surveys for model building and want access to unit record data files, while surveys must protect confidentiality of supplied information.
- * Cognitive reasoning studies assist the process of questionnaire development.
- * Key user frameworks, such as National Accounts, impact developing costing models for the design of complex statistical surveys.
- * Ongoing research and applications are still important in such areas as managing the response rates for household surveys. For example, public cooperation with surveys is affected by perceptions of increasing use of computer data bases, and consequent intrusion into matters previously difficult to monitor.
- * New developments in survey design to reflect theoretical developments.

These are a selection of issues that will be included in the range being considered by the Program Committee for the 1995 IASS sessions in Beijing. Comments are sought at this early stage from IASS members on the range, or suggestions for its extension. Please contact :

Len Cook, Department of Statistics, P.O. Box 2922
Wellington, New Zealand

Facsimile : 64 4 471-1183 ; E-mail : Lcook@stats.govt.NZ

4. COUNTRY REPORTS

4.1. CANADA by G. BRACKSTONE

Statistics Canada has launched a new six-week training course in survey-taking. Known as the Survey Skills Development Course, it takes participants through all aspects of survey-taking from design to dissemination by having them carry out a small survey. The students are required to meet with a client, design the questionnaire, conduct fieldwork, prepare and run a processing system, prepare an analytic report and discuss it with the client. The course teaches the different roles of the members of a survey team and builds teamwork. Originally designed for new professional and technical recruits, it is now also being offered to existing employees. For more information contact Hank Hofmann, Survey Skills Development Course, Statistics Canada, Tunney's Pasture, Ottawa, Ontario, Canada K1A 0T6 (Tel : (613) 951-0844).

Since 1987, the Canadian Centre for Health Information has produced current year estimates of cancer incidence and mortality counts and age-standardized rates, as part of a joint project with the Canadian Cancer Society, the National Cancer Institute of Canada, Health and Welfare Canada, and the provincial cancer registries. The results are published in the annual monograph Canadian Cancer Statistics. The purpose of this publication is to provide health professionals, and others who have an interest in cancer, with Canadian data describing the current incidence of, and mortality from, the more common types of cancer at the provincial and national levels. Such data are useful to plan, monitor and evaluate cancer control programs.

Information on cancer incidence and mortality comes from the provincial cancer registries and vital statistics offices. The process of collecting this information in each province and compiling it at the national level results in a considerable delay before complete data for a particular year is available for all of Canada. To overcome this delay, crude incidence and mortality rates are estimated by fitting maximum likelihood models (linear, quadratic, loglinear, logistic or other) to the available provincial and Canadian yearly values, the yearly counts being assumed to follow independent Poisson distributions. Rate estimates for the current year are then applied to population estimates to obtain estimated counts. The age-standardized incidence rates are estimated using weighted least squares regression, with weights taken as the inverse of the estimated variances of the actual age-standardized rates. The variances are calculated under the assumption that the age-specific counts used in the computation of the age-standardized rates follow independent Poisson distributions.

For more information on this project, please contact Leslie A. Gaudette, Canadian Centre for Health Information (Tel : 613-951-1740) or Carol Morin, Methodology Branch (Tel : 613-951-0659), Statistics Canada, Tunney's Pasture, Ottawa, Ontario, Canada K1A 0T6.

4.2 UNITED KINGDOM by Roger THO MAS

1991 Census Validation Survey and other census-related checks

As for previous UK censuses, a Census Validation Survey has been carried out to estimate the accuracy and reliability of the results of the 1991 Census of Population. Trained survey interviewers repeated the census enumeration procedures in a sample of areas to check on completeness and accuracy and interviewed a sample of households to check for errors in the answers given to census questions. First results on coverage are likely to be published shortly. The CVS will also be used to check the completeness of electoral registration and of the Postcode Address File used in survey sampling. Samples of households selected for voluntary OPCS continuous surveys will be matched with census returns to check for survey non-response bias. (P Heady, OPCS, tel : 071-396-2240)

Other 1991 Census Follow-up Surveys

The legislation under which the 1992 UK Census was conducted permits the Office of Population Censuses and Surveys (OPCS) to carry out specified sample surveys of population sub-groups identified by the census which are otherwise hard to sample and locate. These groups include persons and households in particular housing circumstances and persons with certain vocational qualifications who were not using those qualifications in employment at the time of the census. Several of these surveys will be reporting in the course of 1992. (R. Barnes, OPCS, tel : 071-242-0262 ext 2242)

Samples of anonymised census records (SARS)

For the first time in the UK samples of anonymised records from the 1992 Census are to be made available to the non-governmental research community. There will be two main samples, one of individuals (about one million records) and one of households (about half a million records). It is hoped that these samples will be accessible to researchers by mid-1993. (Dr Catherine Marsh, University of Manchester, tel : 061-275-2517)

Family Resources Survey

The Department of Social Security has launched a new continuous survey to cover around 25,000 households annually. Fieldwork and other operational stages are being carried out jointly by OPCS and SCPR, an independent survey research institute. The survey covers household structure, sources and amounts of income, incidence of state benefit payments and their effect on income maintenance and economic activity. Computer assisted personal interviewing (CAPI), based on the Blaise package developed by the Netherlands Central Bureau of Statistics, is used to collect the data. (R. Boothroyd, Department of Social Security, tel : 071-962-8000 ext 28242)

British Household Panel Survey

The ESRC Research Centre on Microsocial Change at the University of Essex has carried out, with funding from the Economic and Social Research Council, the first round of a major household panel survey. The intention is to run annual rounds of the survey, which will cover many aspects of the circumstances, attitudes, beliefs and behaviour of household members which are of interest to students of microsocial change. (Professor D. Rose, tel : 0206-872957)

National Child Development Survey

Fieldwork for a 1991 round of the NCDS has been completed, with funding from the ESRC and government departments. The NCDS is based on a national age cohort of persons born in 1958. The previous round of the survey was carried out in 1981, when the cohort members were aged 23, and the current round updates their personal histories with respect to relationships and family building, housing, employment, health etc. (Professor J. Bynner, tel : 071-253-4399 ext 4142)

National Survey of Sexual Attitudes and Lifestyles

This survey aims to investigate sexual attitudes and behaviour associated with the propagation of the HIV virus. It is funded by the Wellcome Trust and is being carried out by Social and Community Planning Research (SCPR). It parallels a similar survey carried out in France. Fieldwork and data processing are now complete and first results will be published shortly. (Mrs Julia Field, SCPR, tel : 071-250-1866).

4.3 PAKISTAN by Mohammad ISHAQ

Federal Bureau of Statistics (FBS) is the biggest organization in PAKISTAN in the field of Statistics, having head Office at ISLAMABAD and Regional/Field Offices in all over the country.

This organization Conduct the Surveys and Censuses Covering a wide range of Social and related topics and publish their reports regularly.

The following are some of the important surveys conducted by this organization regularly.

1. Household Income and Expenditure Survey.
2. Labour Force Survey.
3. Survey on use of Fertilizer in Rural Areas.
4. Survey of Small and Household Manufacturing Industries
5. Population Growth Survey.
6. Pakistan Demographic Survey.
7. National Health Survey.

In addition FBS also publish the following reports.

1. Monthly Statistical Bulletin
2. Statistical Pocket Book of Pakistan. (Annual)
3. News Letter (Monthly).
4. Social Indicators of PAKISTAN.
5. National Accounts of PAKISTAN.

Household Income and Expenditure Survey in one of the main surveys being conducted since 1984-85 to 1987-88. The Questionnaire has been revised and enlarged to cover maximum information and the name of the Survey has been changed as "House Hold Integrated Economic Survey" (HIES). The Survey with new Questionnaire was launched in 1990-91. Field enumeration, Editing/Coding and Data Processing has been completed. The Analysis of the Tabulation is in progress. The final report is expected to come out very soon. HIES 1992-93 has also been launched and Field enumeration for the 1st quarter (July-September 1992) has been started.

4.4 NEW ZEALAND by L.W. COOK

Details of Major New Surveys

Employment Contracts Survey

This survey will collect information on types of employment contracts currently in use and the number of employees covered by each type of contract. The output from the survey will be used for analysing changes under the Employment Contracts Act, new labour market legislation introduced in 1991.

Quarterly Local Authority Survey

This survey is collecting financial data from the local government sector which will be used in the calculation of quarterly estimates of New Zealand's Domestic Product.

Retirement Income Provision Survey

This survey will collect information on general issues pertaining to retirement provision. The resulting statistics will be used to inform public policy on state-provided superannuation.

Real Estate Investment Survey

The objective of this survey is to obtain information regarding :

1. Purchases, ownership and sales on New Zealand real estate by foreign principals ; and

2. Purchases, ownership and sales of real estate in countries outside New Zealand by Zealand-based organisations and individuals.

The data derived from the survey will be used to improve the coverage and accuracy of New Zealand's Balance of Payments statements.

Household Health Survey

This survey will provide health information not currently available. The subject areas of the survey are health status, health service users' satisfaction with health services, factors which influence the level of use of health services, the unmet need for health services and information on selected health behaviour.

4.5 MALAWI by Dr. S.C. BABU

Food Security and Nutrition Monitoring Surveys in Malawi

Food security and Nutrition Monitoring (FSNM) surveys conducted, regularly in Malawi provide information to the policy decisionmakers on the food security and nutrition situation in various parts of the country. The Food Security and nutrition Monitoring surveys involve use of four different modules : Food Security Module, Markets and Prices Module, Household Expenditure Module and Nutrition Monitoring Module.

The FSNM Surveys are conducted twice a year to capture the effects of seasonality on the food availability and nutritional status of the population. In a year, the first round of survey is conducted just before the harvest in June and the next round is implemented in December, during the planting period.

The food security module is used in both rounds. It generates information on the demographic characteristics, cultivation characteristics, livestock ownership, food production, marketing, and storage, time of running out of food stocks, food purchasing behaviour, food intake pattern and composition of meals, nature and availability of employment, wage levels, remittances, and the coping strategies followed by households to acquire food. The markets and prices module which is implemented at the enumeration area level, generates information on food market accessibility of organized markets, private trader operations, availability of food in the markets and prices of several food and non food commodities.

The expenditure module collects information on the monthly expenditures on fifteen food and twelve non-food commodities. Details on the total quantity of purchase, value of foods consumed from own production, and value of purchased using cash and kind are recorded. The nutrition module generates information on the anthropometry measurements of children under five years of age, child mortality, morbidity, drinking water availability, sanitation, availability and use of health facilities, breast feeding, immunization, and food eating pattern of children and their attendance of growth monitoring clinics.

The food security and market and prices modules are implemented during each round of survey. Additionally, the expenditure and nutrition modules are alternated between the two rounds.

Sampling Methodology : The sampling procedure followed for the FSNM surveys is multistage stratified random sampling. In the first stage the eight Agricultural Development Divisions (ADD) are stratified in to the Rural Development Projects (RDPS). In the second stage, the thirty RDPs are further stratified into Enumeration Areas (EA) which have similar agro-ecological characteristics such as soil type, cropping pattern and weather patterns. Documentation of these characteristics which follow closely the Extension Planning Areas (EPA) of the agricultural department, precede the stratification of enumeration areas. The 1987 population census was used to list the population in each of the RDP and the enumeration areas were randomly selected from each of the RDPs. The number of enumeration areas selected from each RDP depended on the total cost of conducting surveys and the population size of the RDP. About 240 enumeration areas are selected at this stage which ranged from 5 to 10 per RDP. The households in selected enumeration areas are listed and a random sample of twenty households are chosen systematically from the listing for the interviews. A total of 4800 sample households are chosen for the Food Security and Nutrition Monitoring Surveys.

Data collection and supervision : The data on the food security and nutritional status of the households are collected by the Enumerators stationed in the enumeration areas using the pretested modules described above. The enumerators are supervised by Field Supervisors placed in the EPA offices. Each Field Supervisor supervises 5 to 10 enumerators. The Field Supervisors report to the Field Officers placed in the Agricultural Development Divisions, who are responsible for the field level implementation of the surveys.

Data Entry and Verification : The time involved in preparing the data in a readily usable form, from the time of data collection has been seen as a constraint in developing countries in the use of information in decision making. To avoid such time lag, the data entry system for the FSNM surveys has been decentralized at the Agricultural Development Division level. Since, the same questionnaire is used throughout the country, the data entry program is prepared, field tested and given to the ADDs. The program has been written with error correction mechanisms so that simple data entry errors could be prevented. As the sample size at the ADD level ranges only from 400 to 800 the data is entered within two weeks immediately after data collection. The Statistical Clerks at the ADDs who are incharge of data entry have also been trained in data cleaning and verification. Once the data is entered, basic descriptive statistics are analyzed to verify the validity of the data. After the data is cleaned, they are sent to the headquarters for analysis. A copy of the data set is also retained at the ADD level where the data is also analyzed for the planning and policy making purposes.

Data analysis and Reporting : At the ADD level, the data is analyzed using Statistical Package for social Sciences (SPSS). A basic program to conduct tabular analysis of the data is prepared at the Ministry of Agriculture headquarters and sent to the ADDs along with the data entry program. The Statistical Clerks and the Evaluation Economists are trained in statistics and SPSS to use the program along with other statistical analysis that may be pertinent for the ADDs. Based on the data analysis, the ADDs prepare discussion

paper on the situation of food security and nutritional status in their area. The results are presented at the ADD level to the Agricultural planners and other Subject Matter Specialists to obtain their comments and suggestions on the situation and remedial measures in the form of policies and action plans. The persons to take follow-up actions are identified.

Incorporating the suggestions of Subject Matter Specialists, the discussion papers are revised. They are presented at a national level meeting in the presence of senior level government officials and appropriate policies are recommended. The data received at the headquarters are further analyzed by the Economists to prepare national level plans and policy suggestions. The reports are circulated among other sectoral ministries.

One of the accomplishments of the Food Security and Nutrition Monitoring surveys has been the quick turn over of results which has resulted in localized food and nutrition interventions. The Food Security and Nutrition Monitoring surveys have also been instrumental in initiating debate on the food and nutrition issues at various levels of decisionmaking in Malawi.

4.6 BRAZIL by P. de CASTRO SILVA

The Survey of Japanese descendants in Brazil

The survey was conducted as a special project for the 80th Anniversary of Japanese Immigration to Brazil in order to provide statistical information on the size and characteristics of Japanese descendants. The Survey was conducted by the Centro de Estudos Nipo-Brasileiros under the financial sponsorship of the Japanese International Cooperation Agency. The Survey was carried out by a Special Committee established in the CENB with the coordinating guidance of a statistician from Japan.

The basic problem of the Survey was the concept of Japanese descendants, While the meaning of a JD appeared to be self-evident to the society, various alternative notions were observed as to the details. Thus it was decided to use the widest concept, i. e., any person who had at least one Japanese in one's ancestral hierarchy, which enabled provision of data on any alternative definition with the addition of required characteristics in tabulation. For the survey purpose, the ancestral line was limited back to 1908, the year of the first Japanese immigration.

Unlike the situation 3 decades before, when the JD's were canvassed by a census, the contemporary JD's were not living closely together in limited areas and the local Associations of JD's, which engaged in the field work in the census, were vanishing and weak.

High accuracy of the estimate for the population was desired with limited human resources and financial allocation. The field force mobilizable were primarily university students in São Paulo area, who were to travel to the assigned areas after training. The proportion of JD's in the total Brazilian population was expected to be in the order of 1 % with concentrations in few

Provinces but with scattered residences in all the Provinces. It was therefore necessary to cover the entire area of Brazil with the only exception of the Island of Fernando de Noronã which had about 1000 military population in the 1980 Census.

For covering all the territory of a vast country like Brazil, it required full cooperation of the Instituto Brasileiro de Geografia e Estatística with the provision of information and materials on the sectors in particular, which were established for the 1980 Population Census of Brazil by the IBGE.

The Setors were the only areal breakdowns of the country below the Municipality level, but the number was about 140 000 each with a population of nearly 1000 on average. They were therefore too big to be canvassed within a day or so, it was decided accordingly to use sub-setors, which were the further breakdowns of setors into around 10 portions, as the survey for a field enumerators. The actual work of dividing the setor on the map required hundreds of mandays at the CENB. The IBGE sampled in a single stage selection, on the basis of stratification by the Survey Committee. Information utilized at the stage was limited to regional grouping of Provinces, existence of an Association of JD's, urban or rural defined by the IBGE, those with persons born in Japan, those without persons yellow-colored, those with persons yellow-coloured but none born in Japan, -the last 3 were from the 1980 Census. The number of strata was 103.

For sampling, equal probability was employed in each stratum. Allocation of sub-setors to respected stata was decided on the basis of judgement.

To make possible error calculation simple, the sample was composed of two independent sub-samples, with about 500 sub-setors each.

Items to be studied were numerous for a JD found, but no enquiring beyond identification of JD for a non-JD.

Reducing the burden on the field force, for the complete identification of JD, in particular, it was decided to complete the study in two field visits. It was expected in the order of hundred, the sampled sub-setors actually had no-JD's.

The first round of the enquiry, termed the Basic Survey, studied only sex, age beside the record of JD. The enumerators contacted all households in the area for identification of JD's.

About 100 enumerators were selected from among the 700 applicants. For about 10 enumerators one supervisor was assigned from among the staff of the Headquarters.

They were trained for 8 days and then despatched to the assigned places. For few remote areas, field workers of the IBGE participated as enumerators.

The Basic survey was conducted during the period of the 16th-31st of July 1987.

Excepting one area which was inaccessible because of rain, all the sampled areas were surveyed.

As the post enumeration check, a sub-sample of original sample areas were selected. At the same time, sample areas with many of no-entries in the cases of large scale Apartments which were unaccessible, were surveyed during the period of the 4th-12th September.

The survey materials were checked at the Headquarters and coded for computer processing. After entry, the data tapes were sent to Japan and processed.

For the computer processing, technical assistance was provided by the Statistics Bureau of Japan and the actual processing was done at the Japan Statistical Institute.

The basic results were released in time for the Anniversary. The population of JD's, of July 1987, was estimated to be 1.228,000 (± 3000).

The second round of the survey, termed the Socio-Economic Survey, was conducted in August 1988. This time the sample were limited only to those households with JD's, in the Basic Survey. To facilitate contacts in the subsequent round, a special item was included in the Basic Survey, recording the name, address and telephone number, if existed of the person most close, for the case of possible movement. Thus the enumerators chased after such removed cases. It was, however, found that some were not contacted even with such an arrangement.

The editing and entry was done in Brazil and further processing was done in Japan as in the previous year.

Main results were made available in early 1990. Detailed and special tabulation is well scheduled but to be done at the CENB, using a micro computer.

So far reports in published form appear in two separate versions of Brazilian (Portugese) and Japanese, first on the Basic Survey in June 1987, and the second on the Basic and Socio-Economic Surveys in 1990.

4.7 UNITED STATES by D. KASPRZYK

The National Center on Child Abuse and Neglect in the U.S. Department of Health and Human Services is sponsoring the third National Study of the Incidence of Child Abuse and Neglect ; Westat is the data collection agent for this study. The study is designed to measure the total number of children who are abused or neglected in the United States ; it is also designed to indicate the extent to which this number has changed since the data were collected for earlier incidence studies en 1979 and 1986. In addition, this study will be used to determine the incidence of substantiated child abuse cases that result in civil and criminal court proceedings and the outcomes of such cases.

The study has two components : the first component will be conducted in a national sample of country-level primary sampling units. Survey respondents include investigative caseworkers and other professionals who have direct contact with children and families. Data collection was scheduled for Fall 1992. The

second component of the study involves a retrospective exam of child protective services agency records. For more information, contact Ms. Judy Poston, National Center on child Abuse and Neglect, Room 2006, Switzer Bldg., 330 C Street, S.W., Washington, D.C. 20201 ; phone 202-345-6087.

The Continuing Survey of Food Intakes by Individuals (CSFII), conducted as three separate 1-year surveys in 1989, 1990, and 1991, is the most recent of many U.S. Department of Agriculture surveys designed to measure what Americans eat and drink. The Diet and Health Knowledge Survey (DHKS), conducted as a telephone follow-up to the CSFII, is designed to improve understanding of factors that affect food choices and to obtain information on people's knowledge and attitudes about the Dietary Guidelines for Americans. Together, the CSFII and the DHKS provide the first opportunity on a national scale to link an individual's knowledge and attitudes (from the DHKS) to his or her dietary behavior as indicated by food intake information collected from the same individual in the CSFII. Both surveys were conducted by USDA's Human Nutrition Information Service.

Individuals who took part in the CSFII were asked to provide 3 consecutive days of dietary data. The first day's data were collected in a personal in-home interview using a 1-day dietary recall. The second and third day's data were collected using a self-administrated 2-day dietary record. Individuals who were identified as the main meal planners/preparers in the CSFII were contacted by telephone, if possible, about 6 weeks after collection of the dietary data and were asked a series of questions about knowledge and attitudes toward diet, health, and food safety.

Results from the first year of data collection in 1989 are available on data tape. For more information contact Dr. Ellen Harris, Acting Division Director, Nutrition Monitoring Division, Human Nutrition Information Service, U.S. Department of Agriculture, 6505 Belcrest Road, Hyattsville, MD 20782 (301-436-8485).

The U.S. National Center for Education Statistics has begun to plan a new longitudinal studies program whose overall purpose is to describe children's 1) preparation for school ; 2) transition into school ; and 3) their progression through elementary school. The proposed program will consist of two cohorts--a birth cohort and a kindergarten cohort. Annual follow-up surveys of these cohorts are planned. The study will allow for the collection of data from the child, the child's parents/guardians, child's teachers and care providers, child care programs and schools. Field tests are planned for 1994 with studies being fielded in 1996. For more information contact Dr. Jerry West, National Center for Education Statistics, 555 New Jersey Avenue, N.W., Room 417B, Washington, D.C. 20208, phone 202-219-1574.

QUESTION/ANSWER

Conducted by Leslie Kish. Please send questions to him (ISR - The University of Michigan, Ann Arbor, MI 48106, USA, Internet leslie.kish@um.cc.umich.edu, Bitnet usergdqe@umichum, Fax 313-764-8263) 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 if they agree.

28.1 Question

How would one define and determine the best kind and the number of administrative units to use for primary sampling units for a national multi-stage sample of households?

Answer.

The question concerned a specific large country, but we may cover other countries also, without trying to generalize it to PSUs (primary sampling units) for all survey situations. We may begin with pointing out that the largest units in most countries (the 50 states in the USA, the 30 provinces of China, 23 states in Brazil, 10 in Canada, and 6 in Australia) are both too few and too large to serve as PSUs. They have been used (mistakenly, alas) in some plans, and even in some actual surveys that were designed without the benefit of sampling experts. They can serve as strata, and usually as domains, with all of them covered by the design. The roughly 3000 counties in the USA have often been used for PSUs and descriptions may be found [Hess 1985; Hansen, Hurwitz, and Madow 1953, I, 6.9, 9.11, 12.B.4; Kish 1965, 10.2, 10.4].

Defining the population elements is the first step in any survey; households and dwellings usually serve well the definitions of many kinds of persons. But next in the process and also in importance are generally the PSUs; other units (e.g. the second-stage and third-stage units in a four-stage sample) have less generality, because they depend on the availability of local resources and lists. I describe here a "county model" for PSUs that is frequently used for surveys, especially when they serve as a "master frame" for many diverse surveys (as they do in the references above). The counties are fairly large, heterogeneous units (urban, plus rural, etc.), often with urban centers where semi-permanent teams (1, 2, 3, 4 persons) of interviewers live, who can cover the PSUs as "primary survey areas."

A different model of PSUs may be preferred for a one-time survey covered by a team of interviewers travelling from a central office (or a few); also for a small, confined population, such as a city. Then one may select a larger number of smaller PSUs, such as "Census Enumeration Areas" or "blocks," if these are available. These may be used in some one-stage, or two-stage, or three-stage design. Even simpler designs may be useful for selections from national lists and registers for mail or telephone surveys. These models differ from the "county model" to which the

following remarks apply:

- a) The number of units (counties) in the population must be large, preferably over 1000, so that a sample of 1/10 (or 1/20) of them will still yield over 100 (or 50) PSUs for the sample. A much larger fraction would reduce the cost benefits of sampling them. A much smaller number of PSUs in the sample would 1) increase the variance, 2) increase the instability of variance estimates (Survey Statistician 1987, Q/A No 18), and 3) would greatly hinder regional estimates.
- b) Recognizable and stable boundaries for the PSUs are needed; also good and ample data for stratification and for measures of size for selections with PPS (probabilities proportional to size). In addition to samples of households, the PSUs may also serve other surveys: of agriculture, of schools, hospitals, businesses, etc.
- c) Heterogeneity of the populations within PSUs decreases the variances of estimates for fixed number of PSUs. Counties typically contain both urban and rural populations, different industries, occupations, social classes etc.
- d) The county's urban center is often the home of the interviewers, and roads radiate to the county's limits. When that radius is 15 miles (25 km), the average travel to households may be only 2 to 4 miles, because of greater density near the center. This is convenient for interviewers with autos, but interviewers on bicycles or foot and public transport may need smaller PSUs, or wider spread.

All of the above should be changed for the portion of the households in large urban concentrations, which are usually over 20 percent and may reach 100 percent in small and highly urbanized countries. These counties are often designated as "self-representing areas," because their sample sizes, $f_i N_i = n_i$ are expected to be sufficiently large to constitute an efficient work load. Each of these counties becomes a stratum within which smaller units, such as enumeration areas or blocks, become the first units the primary selection units. Thus the national sample is sharply divided into two large distinct strata, with very different primary selection units suited to the two strata.

Hansen MH, Hurwitz WA, and Madow WG (1953), *Sample Survey Methods and Theory*, Vol I, New York: John Wiley.

Hess I (1985) *Sampling for Social Research Surveys*, Ann Arbor MI: Institute for Social Research

Kish L (1965), *Survey Sampling*, New York: John Wiley.

28.2 Question

The bootstrap is a new and popular method that is being used in other areas of statistics to provide estimates of standard errors and confidence intervals for complex statistics. How can we use this method for computing variances in complex sample surveys?

Answer (by Dr. J. M. Brick of Westat, Inc., Rockville MD 20850-3129)

The bootstrap is a relatively new nonparametric method developed by Efron (1982) to estimate characteristics of the sampling distribution for statistics that are not easily addressed using ordinary distribution theory. The bootstrap has been examined for its applicability to sample surveys, but the results have been somewhat mixed, at best.

The new book by Sarndal, Swensson, and Wretman (1992) contains a brief section on the bootstrap which is a good introduction to the topic. In essence, the idea behind the method is to create a bootstrap population based on the sample by using the inverse of the selection probabilities. Then, bootstrap samples are repeatedly drawn from that artificial population using the same sample design used in the survey. The estimates from these bootstrap samples are used to estimate sampling errors and confidence intervals for the desired statistics. The bootstrap is closely related to other replication methods, like balanced half-samples.

Rao and Wu (1988) take a much more detailed look at the bootstrap in the survey sampling setting and discuss some of the modifications that must be made to make the inferences from this method valid. The procedure described above is biased, and Rao and Wu introduce corrections to scale the estimates of sampling errors appropriately. They indicate that the correctly scaled bootstrap does well for estimating confidence intervals, but the estimates of variances are less stable than jackknife and Taylor series methods.

At present, the use of the bootstrap as a general solution to computing variances is not very promising. A major problem is computational. Creating a bootstrap population and resampling from it is very computer intensive, even if some shortcuts are used to reduce the number of computations. In surveys with many estimates for which standard errors are needed, the bootstrap is not very practical. However, it may be more useful in circumstances where there are only a few estimates needed and the other methods of estimating sampling errors are not effective. Of course, if other methods of replication like the jackknife and half samples are not appropriate, the bootstrap may also run into difficulties.

References:

Efron B (1982). *The Jackknife, the Bootstrap, and Other Resampling Plans*. CBMS-NSF 38 SIAM.

Rao JNK and WU CFJ (1988). "Resampling Inferences with Complex Survey Data." *Journal of the American Statistical Association*, 83, 231-241.

Sarndal, Swensson, and Wretman (1992). *Model Assisted Survey Sampling*. New York: Springer-Verlag.

28.3 Question

In survey sampling we frequently must accept the presence of biases, due both to sampling and nonsampling causes. It is common then to rely on mean square errors that combine variances and biases. What justification have we for this substitution of mean square errors for variances in our confidence intervals?

Answer

This question concerns a common problem we often face. The answer I can find thus far is only partly satisfactory. I hope that readers can send us more complete answers in future references and studies. We often find these familiar symbols:

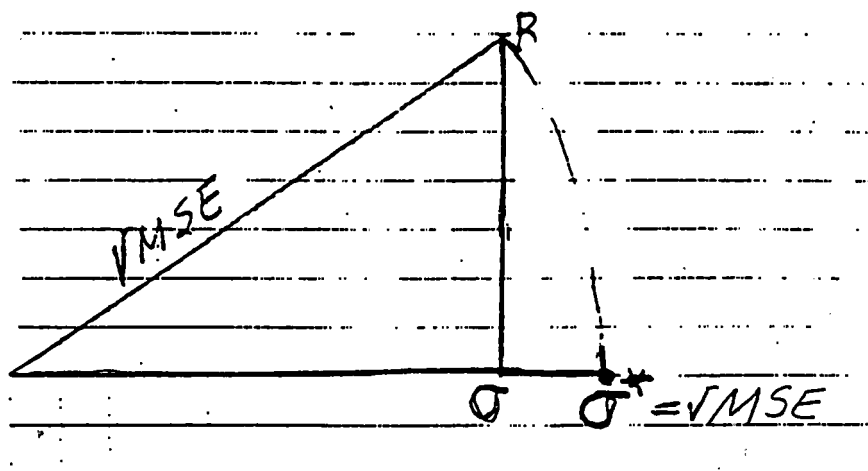
$$MSE = E(\bar{y} - \bar{Y})^2 = E[\bar{y} - E(\bar{y})]^2 + [E(\bar{y})]^2 - \bar{Y}^2 = \sigma^2 + B^2.$$

This is represented with a right triangle in several sampling books (Hansen, Hurwitz, Madow 1953, 2A; Cochran 1977, 1.8, 1.9; Kish 1965, 13.1, 13.2). With unbiased estimates B^2 would disappear and $MSE = \sigma^2$, but these are seldom obtained. Then, as Cochran notes briefly: "Use of the MSE as a criterion of the accuracy of an estimate amounts to regarding two estimates that have the same MSE as equivalent. This is not strictly correct because the frequency distributions of errors ($\hat{\mu} - \mu$) of different sizes will not be the same for the two estimates if they have different amounts of bias. It has been shown, however, by Hansen, Hurwitz, and Madow (1953) that if B/σ is less than about 1/2 the two frequency distributions are almost identical in regard to *absolute* errors $|\hat{\mu} - \mu|$ of different sizes. Table 1.2 illustrates this result."

Probability Of An Absolute Error $\geq 1\sqrt{MSE}$, $1.96\sqrt{MSE}$ And $2.576\sqrt{MSE}$

B/σ	$1\sqrt{MSE}$	$1.96\sqrt{MSE}$	$2.576\sqrt{MSE}$
0	0.317	0.0500	0.0100
0.2	0.317	0.0499	0.0100
0.4	0.319	0.0495	0.0095
0.6	0.324	0.0479	0.0083

The numbers of this table illustrate the approximation involved when the unbiased σ^* is substituted for \sqrt{MSE} in the triangle. Many examples of the nature of biases in surveys may be noted, and they all involve balancing B^2 against σ^2 for reducing the MSE.



- 1) Response errors and nonresponses are unavoidable as sources of potential biases; imputations, substitutions, callbacks always involve MSE reductions. Randomized response techniques increase variances to reduce biases.
- 2) Small area estimates must always balance biases with sampling errors (Prasad and Rao 1990; Platek, Rao, Sarndal, Singh 1987, see Stroud).
- 3) The above are closely related to "shrinkage" (James-Stein) estimates, also to "iterated proportional fitting."
- 4) Weighting of survey data must also balance potential biases against variance increases for MSE reductions (Kish, 1992).
- 5) Ratio estimators and post-stratifications are two examples of estimators that introduce technical biases that are usually small and with lower MSE; these are often described in the sampling literature.
- 6) The design of comparisons - of periodic samples, of domains, of multinational surveys - also must balance biases and sampling errors (Kish 1993).

Note: The famous triangle above must be adapted to the diverse estimates of multipurpose (all?) surveys. In particular the ratio B/σ tends to be reduced drastically for subclasses and especially for their comparisons (Kish 1987, 2.4).

a) And there are many other examples of biases in surveys. However we still may wish for future justification beyond those in our first three references. I just add a more elaborate statement (Kish 1965, 13.8) and some favorable empirical results (Kish and Frankel 1974). More empirical results should be welcomed and also feasible with today's computers. These both use the reasonable equivalence of "coverage errors" of probability statements for MSE as the criterion for justifying the $\sqrt{\text{MSE}}$ criterion.

b) For $t=1.96$ the error rates for MSE decrease for higher values of B/σ ; e.g. for $B/\sigma = 1.0$ the error rate is 0.038, instead of the 0.05 for $B/\sigma = 0$ and 0.048 for $B/\sigma = 0.6$. Decreases are even greater for higher values of B/σ and higher values of t (Kish, 1965, Fig 13.8 III).

c) For one-sided intervals the error rates diverge greatly, depending on which side of the estimate the bias lies (Kish 1965, Fig 13.8 II, Cochran 1977, Table 1.1).

d) The above begs for the use of Bayesian estimates for the biases, when these are feasible, but this is a deep topic.

e) In empirical results from complex surveys the coverage errors are increased (e.g. over 0.05) due to "losses in degrees of freedom" from unequal cluster sizes (Kish and Frankel 1974). These also show closer coverage errors for BRR methods than for JRR and Taylor).

f) The "quadratic loss functions" for both variances and mean square errors have been challenged, discussed, and justified since the time of Laplace and Gauss. Simplicity, convenience, and comprehensibility are also cited. I solicited comments and received them from JNK Rao and Bruce Spencer. (These are especially relevant in combination with the Central Limit Theorem, so basic for large samples of surveys.

References

Cochran WG (1977), *Sampling Techniques*, New York: John Wiley.

Hansen MH, Hurwitz WA, and Madow WG (1953), *Sample Survey Methods and Theory*, Vol I, New York: John Wiley.

- Kish L (1965), *Survey Sampling*, New York: John Wiley.
- Kish L (1987), *Statistical Design for Research*, New York: John Wiley.
- Kish L (1992), Weighting for unequal P_i , *Journal of Official Statistics*, 8,
- Kish L (1993), Comparative survey designs, *International Statistical Review* (submitted)
- Kish L and Frankel MR (1974), Inference from complex samples, *Journal of the Royal Statistical Society*, (B) 36, 1-37. Table 6.
- Platek R, Rao JNK, Sarndal CE, Singh MP (1987), *Small Area Statistics*, New York: John Wiley, Stroud TWF 124-37.
- Prasad NGN and Rao JNK (1990), The estimation of mean squared error of small-area estimators, *JASA*, 85, 163-71.

INTERNATIONAL ASSOCIATION OF SURVEY STATISTICIANS
18, boulevard Adolphe Pinard
75675 Paris, Cedex 14, France

MEMBERSHIP APPLICATION FORM

I wish to enroll as a member of the International Association of Survey Statisticians.

Name : Mr./Mrs./Miss

Date of birth :

Nationality :

Profession (if retired, please indicate) :

Title (professor, doctor) :

Business address :

Name of Office :

Street Address :

City and Postal Code :

Country :

Telephone and Extension :

Fax number :

Home address :

Street Address :

City and Postal Code :

Country :

Telephone N° :

I am a member of the ISI

My mother tongue is (Please specify) :

My IASS correspondence should be sent to :

- Business address

- Home address

The official language of my IASS correspondence should be :

- English

- French

Which field(s) of survey are you specially interested in ?

1. Theory of surveys
2. Collection of data
3. Data processing
4. Quality of data : errors of measurement, sampling, non-sampling errors
5. Analysis of survey data :
6. Other (please notify)

Which field(s) of application are you specially interested in ?

1. Demographic statistics :
2. Housing statistics :
3. Industrial statistics :
4. Statistics of commerce and services :
5. Agricultural statistics :
6. Socio-economic studies of households :
7. Market research and opinion surveys
8. Education :
9. Other (please specify) :

I am aware that the membership dues have been fixed at 130 F. Francs (or the equivalent in other currencies) and that these dues have been reduced to 65 F.F. for nationals of developing countries.

Modes of payment

- Cheque payable to the International Association Of Survey Statisticians*
- UNESCO coupons (for countries with nonconvertible currencies)
- Bank account of the IASS 4586-32 with the Banque Nationale de Paris-Agence Didot, 67, boulevard Brune, 75014 Paris, France
- Postal account : CCP Paris n° 12 233 18 Z
- Other
- *Sent directly to : Mrs. A.M. Vespa-Leyder, IASS Secretary
c/o INSEE - Room 1059 - Ms Anselmi -
18, boulevard Adolphe Pinard
75675 Paris Cedex 14; France

I have been recommended by :

Date :

Signature :



