Balancing Statistical Theory, Sampling Concepts, and Practicality in the Teaching of Survey Sampling

Colm O'Muircheartaigh
University of Chicago, NORC and Harris Graduate School of Public Policy,
1155 East 60th Street, Chicago IL 60637, USA

colm@norc.uchicago.edu

1. Introduction

The origins of survey sampling are in the tradition of official statistics developed from the 16th century onwards and culminating in the meetings of the International Statistical Institute (ISI) in the late 19th century. This tradition owes more to political economy (in particular political arithmetic) and to censuses and administrative records than it does to mathematics or mathematical statistics. Nevertheless, the teaching of sampling in universities (if it has been taught at all) has generally been taken over by departments of statistics (primarily mathematical statistics), and has been taught as a course in applied statistics. The texts devoted to sampling have required a fairly high level of technical training, and have discouraged the attention and education of social scientists and administrators. The practice of sampling has also grown up as an activity almost exclusively for statisticians.

This paper first outlines some of the history of sampling. Second, the contradiction between the logic of sampling textbooks and the logic of sample design is described. A couple of programs where sampling has been taught as a skill for non-statisticians are described. The neglect of practicality in the teaching of sampling is discussed, both in the application of sampling principles, and in the balancing of sampling and measurement errors.

2. History of Survey Sampling

The origins of survey sampling are in official government statistics in the late 19th and early 20th centuries; the provenance is non-mathematical, and the motivation is practical and operational. [A more extensive version of the material in this section can be found in O'Muircheartaigh and Soon (1981).]

One of the finest discussions of the philosophy of sampling took place at the International Statistical Institute (ISI) meetings in Berne in 1895. This was the first recorded occasion on which the statistics establishment (then primarily the collectors of data for government purposes) was confronted with a coherent plea for the use of samples in data collection. A N Kiaer, the director of the Norwegian Bureau of Statistics, presented a report on his experience with sampling surveys and advocated further investigation of the field. Kiaer's definition of a representative investigation was that it was a partial inquiry in which observation was made on a large number of units distributed throughout a country or territory. These units were chosen neither arbitrarily nor strictly randomly but according to a rational scheme based on the general results of some previous statistical investigations. The French geographer Levasseur pointed out that there were in fact three survey methods which were possible: (i) complete enumeration (les statistiques generales); (ii) monography, which he described as a detailed investigation of a unique phenomenon; and (iii) statistical exploration - the term he used to describe Kiaer's representative method. He expressed the view that the ISI would do well to promote discussions of the third method. The issue was put to a vote and the Institute decided by a very narrow margin to allow the topic to be placed on the agenda for further discussion at the next ISI meeting.

Kiaer defended his method four years later at the next session of the ISI held in St. Petersburg. He claimed that in so far as these investigations were representative, they constituted in the totality

of observed units a "photograph which reproduces the details of the original in its true relative proportions". He asserted that there were numerous ways of obtaining representativeness. He believed that if comparisons with census results showed that a sample was representative with respect to those variables on which information was collected in the census, it was probably representative with respect to other variables. Kiaer's arguments were compelling enough to induce the ISI to appoint a Sub-Committee to study the representative method.

Kiaer's efforts bore fruit at the Berlin session of the ISI in 1903. In this session, the ISI adopted a resolution which recommended the use of the representative method subject to the provision that in the publication of the results the conditions under which the selection of the observations was made were completely specified. The sample survey had become an acceptable method of data collection.

There are three important principles involved in Kiaer's approach. The first is the overwhelming importance of the 'representativeness' of the sample. It is worth emphasizing here that these representative samples were not simple random samples, even in intent; all the participants felt that it was appropriate that information about the population structure should be used in the design of the sample. The second principle, described in Kiaer (1897), was that the selection of elements for observation should, insofar as possible, be made objectively and that the subjective judgment of enumerators should not be allowed to influence the selection - enumerators were, for example, instructed to follow distinct routes in the execution of the fieldwork. The third principle was that for every survey the reliability of the results should be assessed. Each survey should be divided into a number of distinct parts, using for each a different (but apparently equally acceptable) representative method. The comparison of the results of these parts would provide evidence as to how much faith could be placed in the results of the survey (Kiaer, 1901); this procedure is the ancestor of later replicated (interpenetrated) methods of variance estimation, and is the only expression of concern for precision in the early discussions.

It may be worth noting that Kiaer's 1895 design was, in effect, a multi-stage stratified area sample with systematic sampling of households at the final stage.

3. The Logic of Survey Sampling

Survey sampling is the province of the practitioner, a person with knowledge of a particular subject matter and an appreciation of the importance of this subject matter relates to issues of policy or politics. Faced with the problem of obtaining sufficient information about a population without having the resources to measure all the elements in the population, the practitioner needs to make decisions about how to design/select a sample from the population so that the sample mirrors the structure of the population – in Kiaer's words "is a miniature of the population".

Consider the problem facing the sample designer in four exemplar surveys: (i) a national labor force (employment) survey; (ii) a survey of educational attainment among school-age students; (iii) a survey of consumer finances being carried out on behalf of the Ministry of Finance; and (iv) a telephone survey to ascertain and contrast inner-city and suburban health behavior. An economist carrying out a labor force survey will require (at least) stratification by region and urbanicity; cost and practicality considerations will require clustering of the sample; particular analyses may demand disproportional allocation of the sample; the estimation of population parameters will usually require weighting due to nonresponse and other adjustments. The sample of school-age students will probably be a heavily stratified sample of schools (possibly with a preceding stage of selection of administrative school districts); within schools a sample of teachers and or classes will be needed; for operational reasons associated with test administration, all students in the selected classes will usually be included in the sample. The sample of consumers will probably be a hybrid, a combination of a national probability sample of households (stratified by economic variables) and a list sample of high tax-payers from administrative records; both parts of the sample will be clustered for cost reasons. The telephone sample (at least in the USA) would be a list-assisted random digit dialing (RDD) sample with telephone exchanges stratified by proportion of listed

numbers within the designated inner-city area. Geocoding of addresses for listed numbers might also be used to provide further stratification for selected listed telephone numbers.

4. The Pedagogical Challenge

The challenge facing those of us who try to teach sampling, or devise training programs in sampling, is to combine an understanding of the subject matter objectives and problems with an adequate level of technical expertise to appreciate the implications of design decisions. There are two dimensions of sampling that require consideration: technical (mathematical and statistical) and practical (substantive and operational). There is the related issue of which academic discipline should take responsibility for the subject. So far, statisticians dominate the field, and there is therefore a tendency to seek and endorse textbooks that have both a sufficiently rigorous technical approach and impart a set of easily verifiable (or testable) skills. This means a tendency away from complex practical problems with no single optimal solution, and a resistance to narrative descriptions of practical trade-offs.

Imagine that the practitioner turns to a sampling textbook for help. There are two classes of sampling textbook in the literature, and they offer quite different views of the subject. As representatives of the two genres I will use Kish and Cochran, two fine books. Kish's book, published in 1965, belongs to the tradition established by Yates (1949) and Hansen, Hurwitz, and Madow (1953). This genre resembles sampling manuals rather than textbooks; they provide the kind of tool-kit that might be needed by a practical designer of samples, with theory used as support for the decisions. The main purpose of the theory is to permit the designer to attach standard errors to the estimates, and to evaluate the impact of design decisions on precision in arriving at a cost-effective design. The exercises in Kish have almost no algebraic content, but are numerical or design applications of the principles presented in the text. Cochran's book, first published in 1953, with later editions in 1963 and 1977, is in the tradition of statistics textbooks, and concentrates on presenting a logical technical development of the subject. Cochran was of course also a distinguished practitioner, both in experimental design and survey design, but his book focuses on the technical rather than the applied aspects of sampling. For courses in sampling in American universities, his has been for decades the dominant text. Though it is no longer quite as dominant, I will use it as an example; many of the more recent books mirror its general approach, at least in the features that I refer to here.

Cochran's text is a fine book, carefully and well written. It is systematic in its notation, careful in its derivations, and logical in its development. It starts with simple random sampling, for means and for proportions, considers determination of sample size, progresses through stratified element sampling, and arrives at ratio estimation in chapter 6. Regression estimation and systematic sampling follow. Selection of one-stage cluster samples, with first equal and then unequal cluster sizes follows in chapters 9 and 9A; multistage sampling appears is dealt with in chapters 10 and 11. A chapter on double (two-phase) sampling and a chapter on sources of error in surveys (nonresponse and measurement errors) complete the book.

This approach treats simple random sampling (SRS) as the foundation for all sample designs. The theory is presented for SRS, and is modified subsequently to allow for departures from SRS; often a good deal of attention is paid to the distinction between sampling with and without replacement. Stratification is introduced as a means of reducing the variance of the estimator. Systematic sampling is dealt with at some length. Cluster sampling is introduced relatively late (more than half-way through the book in the case of Cochran) and most of the attention is devoted to clusters of equal sizes. The exercises at the end of the chapters are primarily algebraic.

Kish's text on the other hand is a complex sprawling work, closer to the social science tradition than the mathematical. It does not follow a linear path through the material, topics occur and recur periodically throughout the text. An almost complete first reading of the book is essential to get the full benefit of the material in the first couple of chapters. Many of the exercises are practical design-related problems for which there is no single correct answer. Not ideal for the busy academic teacher with a large class and homework exercises and examinations requiring grading!

5. Possible Approaches

There will be diversity among those wishing and needing to learn sampling, and there should be a parallel diversity among the courses offered to them. A student wishing to become a survey sampler is faced with a difficult challenge, as is the teacher wishing to construct a curriculum. After training, the ideal sampler should have a technical understanding of sample design and estimation, an appreciation of the way in which the subject matter of the survey should inform sample design, and an understanding of the interdependence of sample design and survey operations. Candidates are unlikely to be prepared for all of these. It is not necessary that everyone should be equally skilled in all areas, but it is essential that everyone should have a foundation in all.

There is no single course, or type of course, that can deal adequately with all aspects of sample design; the goal therefore should be to construct a training portfolio that does so. The relevance of different parts of this portfolio will vary for different audiences; each aspect of design (technical, substantive, operational) needs treatment at different levels. An example of an effective program is the Sampling Program for Survey Statisticians (originally the Sampling Program for Foreign Statisticians) at the Institute for Social Research at the University of Michigan. The program was first offered by Leslie Kish in 1946 and has been offered every year since; it combines formal classes with applied workshops. In my presentation I will describe the components I would like to see in a training portfolio and offer some suggestions as to possible courses.

The teaching of survey sampling needs to be re-oriented; the hegemony of statisticians and of theory distracts attention from the purposes of sample design and prevents subject matter specialists from making their important contribution to tailored design of samples. The exclusion of subject matter specialists – such as economists, sociologists, researchers in education and health – from training in survey sampling has impoverished the field and restricted our horizons. The continuing failure to accumulate either cost data or empirical data on structural parameters of populations is an indication of our failure to integrate practical and theoretical considerations. An alliance between statisticians, survey operations staff, survey methodologists, and the users of the final data could provide designs that are both technically and operationally superior to those being produced by statisticians. Our teaching must be designed for all these audiences in order to bring such alliances into being.

REFERENCES

Cochran, W.G. (1953; 2nd edition 1963; 3rd edition 1977). *Sampling Techniques*. John Wiley and Sons, New York.

Hansen, M.H., Hurwitz, W.N., and Madow, W.G. (1953). *Sample Survey Methods and Theory*. John Wiley and Sons, New York, Vols. I and II.

Kiaer, A. (1895, 1897,1901). Proceedings. *Bulletin of the International Statistical Institute*.

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

O'Muircheartaigh, C. and Soon, T.W. [Soon Teck Wong] (1981). "The Impact of Sampling Theory on Survey Sampling Practice: A Review," *Bulletin of the International Statistical Institute*, XLIX, 1, pp 465-493.

Yates, F. (1949; 3rd edition 1960, 4th edition 1980). *Sampling Methods for Censuses and Surveys*. Charles Griffin and Co., London.