

# NEW PARADIGMS (MODELS) FOR PROBABILITY SAMPLING

Leslie Kish<sup>†</sup>  
*The University of Michigan*

## 1. Statistics as a New Paradigm

In several sections I discuss new concepts in diverse aspects of sampling, but I feel uncertain whether to call them new paradigms or new models or just new methods. Because of my uncertainty and lack of self-confidence, I ask the readers to choose that term with which they are most comfortable. I prefer to remove the choice of that term from becoming an obstacle to our mutual understanding.

Sampling is a branch of and a tool for statistics, and the field of statistics was founded as a new paradigm in 1810 by Quetelet (Porter, 1987; Stigler, 1986). This was later than the arrival of some sciences: of astronomy, of chemistry, of physics. “At the end of the seventeenth century the philosophical studies of cause and chance...began to move close together.... During the eighteenth and nineteenth centuries the realization grew continually stronger that aggregates of events may obey laws even when individuals do not.” (Kendall, 1968). The predictable, meaningful, and useful regularities in the behavior of population aggregates of unpredictable individuals were named “statistics” and were a great discovery.

Thus Quetelet and others computed national (and other) birth rates, death rates, suicide rates, homicide rates, insurance rates, etc. from individual events that are unpredictable. These statistics are basic to fields like demography and sociology. Furthermore, the ideas of statistics were taken later during the nineteenth century also into biology by Frances Galton and Karl Pearson, and into physics by Maxwell, and were developed greatly both in theory and applications.

Statistics and statisticians deal with the effects of chance events on empirical data. The mathematics of chance had been developed centuries earlier for gambling games and for errors of observation in astronomy. Also data have been compiled for commerce, banking, and government. But combining chance with real data needs a new theoretical view, a new paradigm. Thus statistical science and its various branches arrived late in history and in academia, and they are products of the maturity of human development (Kish, 1985).

---

<sup>†</sup> Leslie Kish passed away on October 7, 2000

The populations of random individuals comprise the most basic concept of statistics. It provides the foundation for distribution theories, inferences, sampling theory, experimental design, etc. And the statistics paradigm differs fundamentally from the deterministic outlook of cause and effect, and of precise relations in the other sciences and mathematics.

## 2. The Paradigm of Sampling

*The Representative Method* is the title of an important monograph that almost a century after the birth of statistics and over a century ago now, is generally accepted as the birth of modern sampling (Kiaer, 1895). That term has been used in several landmark papers since then (Jensen, 1926; Neyman, 1934; Kruskal & Mosteller, 1979). The last authors agree that the term “representative” has been used for so many specific methods and with so many meanings that it does not denote any single method. However, as Kiaer used it, and as it is still used generally, it refers to the aims of selecting a sample to represent a population specified in space, in time, and by other definitions, in order to make statistical inferences from the sample to that specified population. Thus a national representative sample demands careful operations for selecting the sample from all elements of the national population, not only from some arbitrary domain such as a “typical” city or province, or from some subset, either defined or undefined.

The scientifically accepted method for survey sampling is probability sampling, which assures known positive probabilities of selection for every element in the frame population. The frame provides the equivalent of listings for each stage of selection. The selection frame for the entire population is needed for mechanical operations of random selection. This is the basis for statistical inferences from the sample statistics to the corresponding population statistics (parameters) (Hansen, Hurwitz, & Madow, 1953, Vol. II). This insistence on inferences based on selections from frame populations is a different paradigm from the unspecified or model based approaches of most statistical analyses.

It took a half century for Kiaer’s paper to achieve the wide acceptance of survey sampling. In addition to neglect and passive resistance, there was a great deal of active opposition by national statistical offices who distrusted sampling methods that replaced the complete counts of censuses. Some even preferred the “monograph method,” which offered complete counts of a “typical” or “representative” province or district instead of randomly selected national sample (O’Muirheartaigh & Wong, 1981). In addition to political opposition, there were also many opponents among academic disciplines and among academic statisticians. The tide turned in favor of probability sampling with the report of the U.N. Statistical Commission led by Mahalanobis and Yates (U.N., 1950). Five influential textbooks between 1949 and 1954 started a flood of articles with both theory and wide applications.

The strength, the breadth, and the duration of resistance to the concepts and use of probability sampling of frame populations implies that this was a new paradigm that needed a new outlook both by the public and the professionals.

### 3. Complex Populations

The need for strict probability selection from a population frame for inferences from the sample to a *finite* population is but one distinction of survey sampling. But even more important and difficult problems are caused by the complex distributions of the elements in all the populations. These complexities present a great contrast with the simple model of independence that is assumed, explicitly or implicitly, by almost all statistical theory, all mathematical statistics.

The assumption of independent or uncorrelated observations of variables or elements underlies mathematical statistics and distribution theory. We need not distinguish here between independently and identically distributed (IID) random variables and “exchangeability,” and “superpopulations.” The simplicity underlying each of those models is necessary for the complexities of the mathematical developments.

Simple models are needed and used for early stages and introductions in all the sciences: for example, perfect circular paths for the planets or  $d=gt^2/2$  for freely dropping objects in frictionless situations. But those models fail to meet the complexities of the actual physical worlds. Similarly, independence of elements does not exist in any population whether human, animal, plant, physical, chemical, or biological. The simple independent models may serve well enough for small samples; and the Poisson distribution of deaths by horsekicks in the Prussian Army in 43 years has often served as an example (precious because rare) (Fisher, 1926).

There have also been attempts to construct theoretical populations of IID elements; perhaps the most famous was the classic “collective” of Von Mises (1939); but they do not correspond to actual populations. However, with great effort tables of random numbers have been constructed that have passed all tests. These have been widely used in modern designs of experiments and sample surveys. *Replication* and *randomization* are two of the most basic concepts of modern statistics following the concept of populations.

The simple concept of a population of independent elements does not describe adequately the complex distributions (in space, in time, in classes) of elements. Clustering and stratification are common names for ubiquitous complexities. Furthermore, it appears impossible to form models that would better describe actual populations. The distributions are much too complex and they are also different for every survey variable. These complexities and differences have been investigated and presented now in thousands of computations of “design effects.”

Survey sampling needed a new paradigm to deal with the complexities of all kinds of populations for many survey variables and a growing list of survey

statistics. This took the form of robust designs of selections and variance formulas that could use a multitude of sample designs, and gave rise to the new discipline of survey sampling. The computation of “design effects” demonstrated the existence, the magnitude, and the variability of effects due to the complexities of distributions not only for means but also for multivariate relations, such as regression coefficients. The long period of disagreements between survey samplers and econometricians testifies to the need for a new paradigm.

#### 4. Combining Population Samples

Samples of national populations always represent subpopulations (domains) which differ in their survey characteristics; sometimes they differ slightly, but at other times greatly. These subclasses in the sample can be distinguished with more or less effort. First, samples of provinces are easily separated when their selections are made separately. Second, subclasses by age, sex, occupation, and education can also be distinguished, and sometimes used for poststratified estimates. Third, however, are those subclasses by social, psychological, and attitudinal characteristics, which may be difficult to distinguish; yet they may be most related to the survey variables. Thus, we recognize that national samples are not simple aggregations of individuals from an IID population, but combinations of subclasses from subpopulations with diverse characteristics. The composition of national populations from diverse domains deserves attention, and it also serves as an example for the two types of combinations that follow. Furthermore, these remarks are pertinent to combinations not only of national samples but also of cities, institutions, establishments, etc.

In recent years two kinds of sample designs have emerged that demand efforts beyond those of simple national samples: (a) periodic samples and (b) multipopulation designs. Each of these has emerged only recently, because they had to await the emergence of three kinds of resources: (1.) effective demand in financial and political resources; (2.) adequate institutional technical resources in national statistical offices; (3.) new methods. In both types of designs we should distinguish the needs of the *survey methods* (definitions, variables, measurements), which must be harmonized, standardized, from *sample designs*, which can be designed freely to fit national (even provincial) situations, provided they are probability designs (Kish, 1994). Both types have been designed first and chiefly for comparisons: periodic comparisons and multinational comparisons, respectively. But new uses have also emerged: “rolling samples” and multinational cumulations, respectively. Each type of cumulation has encountered considerable opposition, and needs a new outlook, a new paradigm.

“Rolling samples” have been used a few times for local situations (Mooney, 1956; Kish, Lovejoy, & Rackow, 1961). Then they have been proposed several times for national annual samples and as a possible

replacement for decennial censuses (Kish, 1990). They are now being introduced for national sample censuses first and foremost by the U.S. Census Bureau (Alexander, 1999; Kish, 1990). Recommending this new method, I have usually experienced opposition to the concept of averaging periodic samples: "How can you average samples when these vary between periods?" In my contrary view, the greater the variability the less you should rely on a single period, whether the variation is monotonic, or cyclical, or haphazard. Hence I note two contrasting outlooks, or paradigms. Quite often, the opposition disappears after two days of discussion and cogitation. "For example, annual income is a readily accepted aggregation, and not only for steady incomes but also for occupations with high variations (seasonal or irregular). Averaging weekly samples for annual statistics will prove more easily acceptable than decennial averaging. Nevertheless, many investors in mutual stock funds prefer to rely more on their ten-year or five-year average earnings (despite their obsolescence) than on their up-to-date prior year's earnings (with their risky "random" variations). Most people planning a picnic would also prefer a 50-year average "normal" temperature to last year's exact temperature. There are many similar examples of sophisticated averaging over long periods by the "naïve" public. That public, and policy makers, would also learn fast about rolling samples, given a chance." (Kish, 1998)

Like rolling samples, combining multipopulation samples also encountered opposition: national boundaries denote different historical stages of development, different laws, languages, cultures, customs, religion, and behaviors. How then can you combine them? However, we often find uses and meanings for continental averages; such as European birth and death rates, or South American, or sub-Saharan, or West African rates. Sometimes even world birth, death, and growth rates. Because they have not been discussed, they all usually combined very poorly. But with more adequate theory, they can be combined better (Kish, 1999). But first the need must be recognized with a new paradigm for multinational combinations, followed by developing new and more appropriate methods.

## 5. Expectation Sampling

*Probability sampling* assures for each element in the population ( $i=1,2,\dots,N$ ) a known positive probability ( $P_i>0$ ) of selection. The assurance requires some mechanical procedure of chance selection, rather than only assumptions, beliefs, or models about probability distributions. The randomizing procedure requires a practical physical operation that is closely (or exactly) congruent with the probability model (Kish, 1965). Something like this statement appears in most textbooks on survey sampling, and I still believe it all. However, there are two questionable and bothersome objections to this definition and its requirements.

The more important of the two objections concerns the frequent practical situations when we face a choice between probability sampling and expectation

sampling. These occur often when the easy, practical selection rate for listing units of  $1/F$  yields not only the unique probability  $1/F$  for elements, but also some with variable  $k_i/F$  for the  $i$ th element ( $i=1,2,\dots,N$ ) and with  $k_i>0$ . Examples of  $k_i>1$ , usually a small integer, occur with duplicate or replicate lists, dual or multiple frames of selection, second homes for households, mobile populations and nomads, farm operators with multiple lots. Examples of  $k_i<1$  are: selecting a single adult from households, selecting single dwellings from buildings. In these examples often the  $k_i>1$  or the  $k_i<1$  is a small integer and can be easily ascertained, and it is cheaper, more convenient and economical to use weighting than attempting to obtain  $1/F$  for all the elements. These problems are described in books and articles.

In most cases, we find it more convenient and less expensive to accept the variable probabilities and to counter them with weighting the expected values  $1/k_i$  or  $k_i$ , than to operate another stage of selection. Thus, to paraphrase probability sampling: *expectation sampling* assures for each element in the population ( $i=1,2,\dots,N$ ) a known positive expectation ( $F_i>0$ ). These procedures are used in practice for descriptive (first order) statistics where the  $k_i$  or  $1/k_i$  are neither large nor frequent. The treatments for inferential – second order or higher – statistics are more difficult and diverse, and are treated separately in the literature. Note that probability sampling is the special (and often desired) situation when all  $k_i$  are 1.

The other objection to the term probability sampling is more theoretical and philosophical and concerns the word “known” in its definition. That word seems to imply belief. Authors from classics like John Venn (1888) and M.G. Kendall (1968) to modern Bayesians like Dennis Lindley – and beyond at both ends – have clearly assigned “probability” to states of belief and “chance” to frequencies generated by objective phenomena and mechanical operations. Thus, our insistence on operations, like random number generators, should imply the term “chance sampling.” However, I have not observed its use and it also could lead to a philosophical problem: the proper use of good tables of random numbers implies beliefs in their “known” probabilities. I have spent only a modest amount of time on these problems and agreeable discussions with only a few colleagues, who did agree. I would be grateful for further discussions, suggestions and corrections.

## 6. Some Related Topics

We called for recognition of new paradigms in six aspects of survey sampling, beginning with statistics itself. Finally, we note here the contrast of sampling to other related methods. Survey methods include the choice and definition of variables, methods of measurements or observations, control of quality (Kish, 1994; Groves, 1989).

Survey sampling has been viewed as a method that competes with censuses (annual or decennial), hence also with registers (Kish, 1990). In some other context, survey sampling competes with or supplements experiments and

controlled observations, and clinical trials. These contrasts also need broader comprehensive views (Kish, 1987, Section A.1). However, those discussions would take us well beyond our present limits.

## 7. References

- Alexander, C.H. (1999). A rolling sample survey for yearly and decennial uses. *Bulletin of the International Statistical Institute*, Helsinki, 52<sup>nd</sup> session.
- Fisher, R.A. (1926). *Statistical Methods for Research Workers*. London: Oliver & Boyd.
- Groves, R.M. (1989). *Survey Errors and Survey Costs*. New York: John Wiley.
- Hansen, M.H., Hurwitz, W.N., and Madow, W.G. (1953). *Sample Survey Methods and Theory*, Vol. I, New York: John Wiley.
- Jensen, A. (1926). The representative method in practice, *Bulletin of the International Statistical Institute*, Vol. 22, pt. 1, pp. 359-439.
- Kendall, M.G. (1968). Chance, in *Dictionary of History of Ideas*. New York: Chas Scribners.
- Kiaer, A.W. (1895). *The Representative Method of Statistical Surveys*, English translation, 1976. Oslo: Statistik Sentralbyro.
- Kish, L. (1965). *Survey Sampling*. New York: John Wiley.
- Kish, L. (1981). Using Cumulated Rolling Samples, Washington, Library of Congress.
- Kish, L. (1985). Sample Surveys Versus Experiments, Controlled Observations, Censuses, Registers, and Local Studies, *Australian Journal of Statistics*, 27, 111-122.
- Kish, L. (1987). *Statistical Design for Research*. New York: John Wiley.
- Kish, L. (1990). Rolling samples and censuses. *Survey Methodology*, Vol. 16, pp. 63-93.
- Kish, L. (1994). Multipopulation survey designs, *International Statistical Review*, Vol. 62, pp. 167-186.
- Kish, L. (1998). Space/time variations and rolling samples. *Journal of Official Statistics*, Vol. 14, pp. 31-46.
- Kish, L. (1999). Cumulating/combining population surveys. *Survey Methodology*, Vol. 25, pp. 129-138.
- Kish, L., Lovejoy, W., and Rackow, P. (1961). A multistage probability sample for continuous traffic surveys, *Proceedings of the American Statistical Association, Section on Social Statistics*, pp. 227-230.
- Kruskal, W.H. and Mosteller, F. (1979-80). Representative sampling, I, II, III, and IV, *International Statistical Review*, especially IV: The history of the concept in statistics, 1895-1939.
- Mooney, H.W. (1956). *Methodology of Two California Health Surveys*, OS Public Health Monograph 70.
- Neyman, J. (1934). On the different aspects of the representative method: the method of stratified sampling and the method of purposive selection. *JRSS*, Vol. 97, pp. 558-625.
- O'Muircheartaigh, C. and Wong, S.T. (1981). The impact of sampling theory on survey sampling practice: a review. *Bulletin of International Statistical Institute*, 43<sup>rd</sup> Session, Vol. 1, pp. 465-493.
- Porter, T.M. (1987). *The Rise of Statistical Thinking: 1820-1900*. Princeton, NJ: Princeton University Press.
- Stigler, S.M. (1986). *History of Statistics*. Cambridge: Harvard University Press.
- United Nations Statistical Office. (1950). *The Preparation of Sample Survey Reports*. New York: UN Series C No. 1; also Revision 2 in 1964.
- Venn, J. (1888). *The Logic of Chance*. London: Macmillan.
- Von Mises, R. (1939). *Probability, Statistics, and Truth*. London: Wm. Hodge and Co.

(Received: March 2000).

This is the last paper of Leslie Kish. It has also appeared in *Survey Methodology*, June 2002, No1, 31-34, by the permission of Rhea Kish, Ann Arbor, Michigan.