ESSENTIAL
Methods for Design Based Sample Surveys

Edited by
D. Pfeffermann
C.R. Rao

N·H

Essential Methods for Design Based Sample Surveys

A derivative of Handbook of Statistics: Sample Surveys: Design, Methods and Applications, Vol. 29a

Edited by

Danny Pfeffermann

Department of Statistics, Hebrew University, Israel

C.R. Rao

The Pennsylvania State University,
University Park, USA

ELSEVIER

Amsterdam • Boston • Heidelberg • London • New York • Oxford
Paris • San Diego • San Francisco • Singapore • Sydney • Tokyo
North-Holland is an imprint of Elsevier

N·H

North-Holland is an imprint of Elsevier
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA
Linacre House, Jordan Hill, Oxford OX2 8DP, UK
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands

First edition 2011

Material in the text originally appeared in the Handbook of Statistics, Vol 29 edited by D. Pfeffermann and C.R. Rao (Elsevier B.V. 2009)

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

ISBN: 978-0-444-53734-8

Typeset by: diacriTech, India

Printed and bound in Great Britain
10 11 12 13 10 9 8 7 6 5 4 3 2 1

ESSENTIAL METHODS FOR DESIGN BASED
SAMPLE SURVEYS

List of Contributors

Brewer, Kenneth, *School of Finance and Applied Statistics, College of Business and Economics, L.F. Crisp Building (Building 26), Australian National University, A.C.T. 0200, Australia; e-mail: ken.brewer@anu.edu.au* (Ch. 1).

Chowdhury, Sadeq, *NORC, University of Chicago, 4350 East-West Highway, Suite 800, Bethesda, MD 20814; e-mail: sadeqc@yahoo.com* (Ch. 3).

Frankovic, Kathleen A., *Survey and Election Consultant, 3162 Kaiwiki Rd., Hilo, HI 96720; e-mail: kaf@cbsnews.com* (Ch. 11).

Gambino, Jack G., *Household Survey Methods Division, Statistics Canada, Ottawa, Canada K1A 0T6; e-mail: jack.gambino@statcan.gc.ca* (Ch. 5).

Glickman, Hagit, *National Authority of Measurement and Evaluation in Education (RAMA), Ministry of Education, Kiryat Hamemshala, Tel Aviv 67012, Israel; e-mail: hglickman.rama@education.gov.il* (Ch. 10).

Gregoire, Timothy, Weyerhaeuser, J.P. Jr., *Professor of Forest Management, School of Forestry and Environmental Studies, Yale University, 360 Prospect Street, New Haven, CT 06511-2189; e-mail: timothy.gregoire@yale.edu* (Ch. 1).

Hidiroglou, Michael A., *Statistical Research and Innovation Division, Statistics Canada, Canada, K1A 0T6; e-mail: Mike.Hidiroglou@statcan.gc.ca* (Ch. 6).

House, Carol C., *National Agricultural Statistics Service, U.S. Department of Agriculture, Washington, DC, USA; e-mail: Carol.House@usda.gov* (Ch. 7).

Kalton, Graham, *Westat, 1600 Research Blvd., Rockville, MD 20850; e-mail: grahamkalton@westat.com* (Ch. 2).

Kelly, Jenny, *NORC, University of Chicago, 1 North State Street, Suite 1600, Chicago, IL 60602; e-mail: Kelly-Jenny@norc.org* (Ch. 3).

Lavallée, Pierre, *Social Survey Methods Division, Statistics Canada, Canada, K1A 0T6; e-mail: pierre.lavallee@statcan.gc.ca* (Ch. 6).

Marker, David A., *Westat, 1650 Research Blvd., Rockville Maryland 20850; e-mail: DavidMarker@Westat.com* (Ch. 8).

Naidu, Gurramkonda M., *Professor Emeritus, College of Business & Economics, University of Wisconsin-Whitewater, Whitewater, WI 53190; e-mail: naidug@uww.edu* (Ch. 9).

Nirel, Ronit, *Department of Statistics, The Hebrew University of Jerusalem, Mount Scopus, Jerusalem 91905, Israel; e-mail: nirelr@cc.huji.ac.il* (Ch. 10).

Nusser, S. M., *Department of Statistics, Iowa State University, Ames, IA, USA; e-mail: nusser@iastate.edu* (Ch. 7).

Panagopoulos, Costas, *Department of Political Science, Fordham University, 441 E. Fordham Rd., Bronx, NY 10458; e-mail: costas@post.harvard.edu* (Ch. 11).

Shapiro, Robert Y., *Department of Political Science and Institute for Social and Economic Research and Policy, Columbia University, 420 West 118th Street, New York, NY 10027; e-mail: rys3@columbia.edu* (Ch. 11).

Silva, Pedro Luis do Nascimento, *Southampton Statistical Sciences Research Institute, University of Southampton, UK; e-mail: pedrolns@soton.ac.uk* (Ch. 5).

Skinner, Chris, *Southampton Statistical Sciences Research Institute, University of Southampton, Southampton SO17 1BJ, United Kingdom; e-mail: C.J.Skinner@soton.ac.uk* (Ch. 4).

Stevens, Don L. Jr., *Statistics Department, Oregon State University, 44 Kidder Hall, Corvallis, Oregon, 97331; e-mail: stevens@stat.oregonstate.edu* (Ch. 8).

Velu, Raja, *Irwin and Marjorie Guttag Professor, Department of Finance, Martin J. Whitman School of Management, Syracuse University, Syracuse, NY 13244-2450; e-mail:rpvelu@syr.edu* (Ch. 9).

Wolter, Kirk, *NORC at the University of Chicago, and Department of Statistics, University of Chicago, 55 East Monroe Street, Suite 3000, Chicago, IL 60603; e-mail: wolter-kirk@norc.uchicago.edu* (Ch. 3).

Table of Contents

Essential Methods for Design Based Sample Surveys
ISSN: 0169-7161

DOI: 10.1016/B978-0-444-53734-8.00001-7

1

Introduction to Survey Sampling

Ken Brewer and Timothy G. Gregoire

1. Two alternative approaches to survey sampling inference

1.1. Laplace and his ratio estimator

At some time in the mid-1780s (the exact date is difficult to establish), the eminent mathematician Pierre Laplace started to press the ailing French government to conduct an enumeration of the population in about 700 communes scattered over the Kingdom (Bru, 1988), with a view to estimating the total population of France. He intended to use for this purpose the fact that there was already a substantially complete registration of births in all communes, of which there would then have been of the order of 10,000. He reasoned that if he also knew the populations of those sample communes, he could estimate the ratio of population to annual births, and apply that ratio to the known number of births in a given year, to arrive at what we would now describe as a ratio estimate of the total French population (Laplace, 1783[1], 1814a and 1814b). For various reasons, however, notably the ever-expanding borders of the French empire during Napoleon's early years, events militated against him obtaining a suitable total of births for the entire French population, so his estimated ratio was never used for its original purpose (Bru, 1988; Cochran, 1978; Hald, 1998; Laplace, 1814a and 1814b, p. 762). He did, however, devise an ingenious way for estimating the precision with which that ratio was measured. This was less straightforward than the manner in which it would be estimated today but, at the time, it was a very considerable contribution to the theory of survey sampling.

1.2. A prediction model frequently used in survey sampling

The method used by Laplace to estimate the precision of his estimated ratio was not dependent on the knowledge of results for the individual sample communes, which

[1] This paper is the text of an address given to the Academy on 30 October 1785, but appears to have been incorrectly dated back to 1783 while the Memoirs were being compiled. A virtually identical version of this address also appears in Laplace's *Oeuvres Complètes* 11 pp. 35–46. This version also contains three tables of vital statistics not provided in the Memoirs' version. They should, however, be treated with caution, as they contain several arithmetical inconsistencies.

would normally be required these days for survey sampling inference. The reason why it was not required there is chiefly that a particular model was invoked, namely one of drawing balls from an urn, each black ball representing a French citizen counted in Laplace's sample, and each white ball representing a birth within those sample communes in the average of the three preceding years. As it happens, there is another model frequently used in survey sampling these days, which leads to the same ratio estimator. That model is

$$Y_i = \beta X_i + U_i, \tag{1a}$$

which together with

$$E(U_i) = 0, \tag{1b}$$

$$E\left(U_i^2\right) = \sigma^2 X_i \tag{1c}$$

and

$$E(U_i U_j) = 0 \tag{1d}$$

for all $j \neq i$ can also be used for the same purpose.

Equation (1a) describes a survey variable value Y_i (for instance the population of commune i) as generated by a survey parameter, β, times an auxiliary value, X_i, (that commune's average annual births) plus a random variable, U_i. Equation (1b) stipulates that this random variable has zero mean, Eq. (1c) that its variance is proportional to the auxiliary variable (in this case, annual births), and Eq. (1d) that there is no correlation between any pair of those random variables.

Given this model, the minimum variance unbiased estimator of β is given by

$$\hat{\beta} = \frac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} X_i}, \tag{2}$$

which in this instance is simply the ratio of black to white balls in Laplace's urn.

1.3. *The prediction model approach to survey sampling inference*

While, given the model of Eqns. (1), the logic behind the ratio estimator might appear to be straightforward, there are in fact two very different ways of arriving at it, one obvious and one somewhat less obvious but no less important. We will examine the obvious one first.

It is indeed obvious that there is a close relationship between births and population. To begin with, most of the small geographical areas (there are a few exceptions such as military barracks and boarding schools) have approximately equal numbers of males and females. The age distribution is not quite so stable, but with a high probability different areas within the same country are likely to have more or less the same age distribution, so the proportion of females of child-bearing age to total population is also more or less constant. So, also with a reasonable measure of assurance, one might expect the ratio of births in a given year to total population to be more or less constant, which makes the ratio estimator an attractive choice.

We may have, therefore, a notion in our minds that the number in the population in the ith commune, Y_i, is proportional to the number of births there in an average year, X_i, plus a random error, U_i. If we write that idea down in mathematical form, we arrive at a set of equations similar to (1) above, though possibly with a more general variance structure than that implied by Eqns. (1c) and (1d), and that set would enable us to predict the value of Y_i given only the value of X_i together with an estimate of the ratio β. Laplace's estimate of β was a little over 28.35.

The kind of inference that we have just used is often described as "model-based," but because it is a prediction model and because we shall meet another kind of model very shortly, it is preferable to describe it as "prediction-based," and this is the term that will be used here.

1.4. The randomization approach to survey sampling inference

As already indicated, the other modern approach to survey sampling inference is more subtle, so it will take a little longer to describe. It is convenient to use a reasonably realistic scenario to do so.

The hypothetical country of Oz (which has a great deal more in common with Australia than with Frank L. Baum's mythical Land of Oz) has a population of 20 million people geographically distributed over 10,000 postcodes. These postcodes vary greatly among themselves in population, with much larger numbers of people in a typical urban than in a typical rural postcode.

Oz has a government agency named Centrifuge, which disburses welfare payments widely over the entire country. Its beneficiaries are in various categories such as Age Pensioners, Invalid Pensioners, and University Students. One group of its beneficiaries receives what are called Discretionary Benefits. These are paid to people who do not fall into any of the regular categories but are nevertheless judged to be in need of and/or deserving of financial support.

Centrifuge staff, being human, sometimes mistakenly make payments over and above what their beneficiaries are entitled to. In the Discretionary Benefits category, it is more difficult than usual to determine when such errors (known as overpayments) have been made, so when Centrifuge wanted to arrive at a figure for the amounts of Overpayments to Discretionary Beneficiaries, it decided to do so on a sample basis. Further, since it keeps its records in postcode order, it chose to select 1000 of these at random (one tenth of the total) and to spend considerable time and effort in ensuring that the Overpayments in these sample postcodes were accurately determined. (In what follows, the number of sample postcodes, in this case 1000, will be denoted by n and the number of postcodes in total, in this case 10,000, denoted by N.)

The original intention of the Centrifuge sample designers had been to use the same kind of ratio estimator as Laplace had used in 1802, namely

$$\hat{Y} = \frac{\sum_{i=1}^{N} \delta_i Y_i}{\sum_{i=1}^{N} \delta_i X_i} \sum_{i=1}^{N} X_i, \tag{3}$$

with Y_i being the amount of overpayments in the ith postcode and X_i the corresponding postcode population. In (3), δ_i is a binary (1/0) indicator of inclusion into the sample of size n: for any particular sample, all but n of the N elements of the population will have a value of $\delta = 0$ so that the sum of $\delta_i Y_i$ over $i = 1 \dots N$ yields the sum of just the n values of Y_i on those elements selected into the sample.

However, when this proposal came to the attention of a certain senior Centrifuge officer who had a good mathematical education, he queried the use of this ratio estimator on the grounds that the relationship between Overpayments (in this particular category) and Population in individual postcodes was so weak that the use of the model (1) to justify it was extremely precarious. He suggested that the population figures for the selected postcodes should be ignored and that the ratio estimator should be replaced by the simpler expansion estimator, which was

$$\hat{Y} = (N/n) \sum_{i=1}^{N} \delta_i Y_i. \tag{4}$$

When this suggestion was passed on to the survey designers, they saw that it was needed to be treated seriously, but they were still convinced that there was a sufficiently strong relationship between Overpayments and Population for the ratio estimator also to be a serious contender. Before long, one of them found a proof, given in several standard sampling textbooks, that without reliance on any prediction model such as Eqns. (1), the ratio estimator was more efficient than the expansion estimator provided (a) that the sample had been selected randomly from the parent population and (b) that the correlation between the Y_i and the X_i exceeded a certain value (the exact nature of which is irrelevant for the time being). The upshot was that when the sample data became available, that requirement was calculated to be met quite comfortably, and in consequence the ratio estimator was used after all.

1.5. A comparison of these two approaches

The basic lesson to be drawn from the above scenario is that there are two radically different sources of survey sampling inference. The first is prediction on the basis of a mathematical model, of which (1), or something similar to it, is the one most commonly postulated. The other is randomized sampling, which can provide a valid inference regardless of whether the prediction model is a useful one or not. Note that a model can be useful even when it is imperfect. The famous aphorism of G.E.P. Box, "All models are wrong, but some are useful." (Box, 1979), is particularly relevant here.

There are also several other lessons that can be drawn. To begin with, models such as that of Eqns. (1) have parameters. Equation (1a) has the parameter β, and Eq. (1c) has the parameter σ^2 that describes the extent of variability in the Y_i. By contrast, the randomization-based estimator (4) involves no estimation of any parameter. All the quantities on the right hand side of (4), namely N, n, and the sample Y_i, are known, if not without error, at least without the need for any separate estimation or inference.

In consequence, we may say that estimators based on prediction inference are parametric, whereas those based on randomization inference are nonparametric. Parametric estimators tend to be more accurate than nonparametric estimators when the model on which they are based is sufficiently close to the truth as to be useful, but they are also sensitive to the possibility of model breakdown. By contrast, nonparametric estimators tend to be less efficient than parametric ones, but (since there is no model to break down) they are essentially robust. If an estimator is supported by both parametric and nonparametric inference, it is likely to be both efficient and robust. When the correlation between the sample Y_i and the sample X_i is sufficiently large to meet the relevant condition, mentioned but not defined above in the Oz scenario, the estimator is also likely to be both efficient and robust, but when the correlation fails to meet that condition, another estimator has a better randomization-based support, so the ratio estimator is no longer robust, and the indications are that the expansion estimator, which does not rely upon the usefulness of the prediction model (1), would be preferable.

It could be argued, however, that the expansion estimator itself could be considered as based on the even simpler prediction model

$$Y_i = \alpha + U_i, \tag{5}$$

where the random terms U_i have zero means and zero correlations as before. In this case, the parameter to be estimated is α, and it is optimally estimated by the mean of the sample observations Y_i. However, the parametrization used here is so simple that the parametric estimator based upon it coincides with the nonparametric estimator provided by randomization inference. This coincidence appears to have occasioned some considerable confusion, especially, but not exclusively, in the early days of survey sampling.

Moreover, it is also possible to regard the randomization approach as implying its own quite different model. Suppose we had a sample in which some of the units had been selected with one chance in ten, others with one chance in two, and the remainder with certainty. (Units selected with certainty are often described as "completely enumerated.") We could then make a model of the population from which such a sample had been selected by including in it (a) the units that had been selected with one chance in ten, together with nine exact copies of each such unit, (b) the units that had been selected with one chance in two, together with a single exact copy of each such unit, and (c) the units that had been included with certainty, but in this instance without any copies. Such a model would be a "randomization model." Further, since it would be a nonparametric model, it would be intrinsically robust, even if better models could be built that did use parameters.

In summary, the distinction between parametric prediction inference and nonparametric randomization inference is quite a vital one, and it is important to bear it in mind as we consider below some of the remarkable vicissitudes that have beset the history of survey sampling from its earliest times and have still by no means come to a definitive end.

2. Historical approaches to survey sampling inference

2.1. *The development of randomization-based inference*

Although, as mentioned above, Laplace had made plans to use the ratio estimator as early as the mid-1780s, modern survey sampling is more usually reckoned as dating from the work of Anders Nicolai Kiaer, the first Director of the Norwegian Central Bureau of Statistics. By 1895, Kiaer, having already conducted sample surveys successfully in his own country for fifteen years or more, had found to his own satisfaction that it was not always necessary to enumerate an entire population to obtain useful information about it. He decided that it was time to convince his peers of this fact and attempted to do so first at the session of the International Statistical Institute (ISI) that was held in Berne that year. He argued there that what he called a "partial investigation," based on a subset of the population units, could indeed provide such information, provided only that the subset had been carefully chosen to reflect the whole of that population in miniature. He described this process as his "representative method," and he was able to gain some initial support for it, notably from his Scandinavian colleagues. Unfortunately, however, his idea of representation was too subjective and lacking in probabilistic rigor to make headway against the then universally held belief that only complete enumerations, "censuses," could provide any useful information (Lie, 2002; Wright, 2001).

It was nevertheless Kiaer's determined effort to overthrow that universally held belief that emboldened Lucien March, at the ISI's Berlin meeting in 1903, to suggest that randomization might provide an objective basis for such a partial investigation (Wright, 2001). This idea was further developed by Arthur Lyon Bowley, first in a theoretical paper (Bowley, 1906) and later by a practical demonstration of its feasibility in a pioneering survey conducted in Reading, England (Bowley, 1912).

By 1925, the ISI at its Rome meeting was sufficiently convinced (largely by the report of a study that it had itself commissioned) to adopt a resolution giving acceptance to the idea of sampling. However, it was left to the discretion of the investigators whether they should use randomized or purposive sampling. With the advantage of hindsight, we may conjecture that, however vague their awareness of the fact, they were intuiting that purposive sampling was under some circumstances capable of delivering accurate estimates, but that under other circumstances, the underpinning of randomization inference would be required.

In the following year, Bowley published a substantial monograph in which he presented what was then known concerning the purposive and randomizing approaches to sample selection and also made suggestions for further developments in both of them (Bowley, 1926). These included the notion of collecting similar units into groups called "strata," including the same proportion of units from each stratum in the sample, and an attempt to make purposive sampling more rigorous by taking into account the correlations between, on the one hand, the variables of interest for the survey and, on the other, any auxiliary variables that could be helpful in the estimation process.

2.2. *Neyman's establishment of a randomization orthodoxy*

A few years later, Corrado Gini and Luigi Galvani selected a purposive sample of 29 out of 214 districts (circondari) from the 1921 Italian Population Census (Gini and Galvani, 1929). Their sample was chosen in such a way as to reflect almost exactly the whole-of-Italy average values for seven variables chosen for their importance, but it was shown by Jerzy Neyman (1934) that it exhibited substantial differences from those averages for other important variables.

Neyman went on to attack this study with a three pronged argument. His criticisms may be summarized as follows:

(1) Because randomization had not been used, the investigators had not been able to invoke the Central Limit Theorem. Consequently, they had been unable to use

the normality of the estimates to construct the confidence intervals that Neyman himself had recently invented and which appeared in English for the first time in his 1934 paper.

(2) On the investigators' own admission, the difficulty of achieving their "purposive" requirement (that the sample match the population closely on seven variables) had caused them to limit their attention to the 214 districts rather than to the 8354 communes into which Italy had also been divided. In consequence, their 15% sample consisted of only 29 districts (instead of perhaps 1200 or 1300 communes). Neyman further showed that a considerably more accurate set of estimates could have been expected had the sample consisted of this larger number of smaller units. Regardless of whether the decision to use districts had required the use of purposive sampling, or whether the causation was the other way round, it was evident that purposive sampling and samples consisting of far too few units went hand in hand.

(3) The population model used by the investigators was demonstrably unrealistic and inappropriate. Models by their very nature were always liable to represent the actual situation inadequately. Randomization obviated the need for population modeling.[2] With randomization-based inference, the statistical properties of an estimator are reckoned with respect to the distribution of its estimates from all samples that might possibly be drawn using the design under consideration. The same estimator under different designs will admit to differing statistical properties. For example, an estimator that is unbiased under an equal probability design (see Section 3 of this chapter for an elucidation of various designs that are in common use) may well be biased under an unequal probability design.

In the event, the ideas that Neyman had presented in this paper, though relevant for their time and well presented, caught on only gradually over the course of the next decade. W. Edwards Deming heard Neyman in London in 1936 and soon arranged for him to lecture, and his approach to be taught, to U.S. government statisticians. A crucial event in its acceptance was the use in the 1940 U.S. Population and Housing Census of a one-in-twenty sample designed by Deming, along with Morris Hansen and others, to obtain answers to additional questions. Once accepted, however, Neyman's arguments swept all other considerations aside for at least two decades.

Those twenty odd years were a time of great progress. In the terms introduced by Kuhn (1996), finite population sampling had found a universally accepted "paradigm" (or "disciplinary matrix") in randomization-based inference, and an unusually fruitful period of normal science had ensued. Several influential sampling textbooks were published, including most importantly those by Hansen et al. (1953) and by Cochran (1953, 1963). Other advances included the use of self-weighting, multistage, unequal probability samples by Hansen and Hurwitz at the U.S. Bureau of the Census, Mahalanobis's invention of interpenetrating samples to simplify the estimation of variance for complex survey designs and to measure and control the incidence of nonsampling errors, and the beginnings of what later came to be described as "model-assisted survey sampling."

[2] The model of Eqns. (1) above had not been published at the time of Neyman's presentation. It is believed first to have appeared in Fairfield Smith (1938) in the context of a survey of agricultural crops. Another early example of its use is in Jessen (1942).

A lone challenge to this orthodoxy was voiced by Godambe (1955) with his proof of the nonexistence of any uniformly best randomization-based estimator of the population mean, but few others working in this excitingly innovative field seemed to be concerned by this result.

2.3. Model-assisted or model-based? The controversy over prediction inference

It therefore came as a considerable shock to the finite population sampling establishment when Royall (1970b) issued his highly readable call to arms for the reinstatement of purposive sampling and prediction-based inference. To read this paper was to read Neyman (1934) being stood on its head. The identical issues were being considered, but the opposite conclusions were being drawn.

By 1973, Royall had abandoned the most extreme of his recommendations. This was that the best sample to select would be the one that was optimal in terms of a model closely resembling Eqns. (1). (That sample would typically have consisted of the largest n units in the population, asking for trouble if the parameter β had not in fact been constant over the entire range of sizes of the population units.) In Royall and Herson (1973a and 1973b), the authors suggested instead that the sample should be chosen to be "balanced", in other words that the moments of the sample X_i should be as close as possible to the corresponding moments of the entire population. (This was very similar to the much earlier notion that samples should be chosen purposively to resemble the population in miniature, and the samples of Gini and Galvani (1929) had been chosen in much that same way!)

With that exception, Royall's original stand remained unshaken. The business of a sampling statistician was to make a model of the relevant population, design a sample to estimate its parameters, and make all inferences regarding that population in terms of those parameter estimates. The randomization-based concept of defining the variance of an estimator in terms of the variability of its estimates over all possible samples was to be discarded in favor of the prediction variance, which was sample-specific and based on averaging over all possible realizations of the chosen prediction model.

Sampling statisticians had at no stage been slow to take sides in this debate. Now the battle lines were drawn. The heat of the argument appears to have been exacerbated by language blocks; for instance, the words "expectation" and "variance" carried one set of connotations for randomization-based inference and quite another for prediction-based inference. Assertions made on one side would therefore have appeared as unintelligible nonsense by the other.

A major establishment counterattack was launched with Hansen et al. (1983). A small (and by most standards undetectable) divergence from Royall's model was shown nevertheless to be capable of distorting the sample inferences substantially. The obvious answer would surely have been "But this distortion would not have occurred if the sample had been drawn in a balanced fashion. Haven't you read Royall and Herson (1973a and b)?" Strangely, it does not seem to have been presented at the time.

Much later, a third position was also offered, the one held by the present authors, namely that since there were merits in both approaches, and that it was possible to combine them, the two should be used together. For the purposes of this Handbook volume, it is necessary to consider all three positions as dispassionately as possible. Much can be gained by asking the question as to whether Neyman (1934) or Royall