

P.V. Sukhatme's Contributions to Survey Sampling¹

Ranjan K. Som²

SUMMARY

Attempt is made to briefly encompass Pandurang V. Sukhatme's contributions to survey sampling under two broad headings: (1) theoretical and applied work in survey sampling; and (2) institution building specially for sample surveys in the field of agriculture and related subjects.

1. Theoretical and Applied Work in Survey Sampling

The earliest paper by P.V. Sukhatme (PVS) in survey sampling was "Contributions to the theory of representative method" in 1935, when he was working towards his Ph.D. Degree at London University. He was thus a pioneer in advocating the use of pilot studies for providing estimates of cost and variance functions, required for an efficient design for the survey proper. Prof. P.C. Mahalanobis had also been studying this problem in connection with the estimation of crops in Bengal, and later on the designs of large-scale sample surveys (Mahalanobis [13], [14], [15]). It is noted that PVS's paper followed Neyman's paper [19] setting forth the concept of stratified sampling, adumbrated in Tschuprow's then relatively little known paper [23]. PVS's guidelines for designing a sample survey apply equally well to stratified sampling.

One may discern the following strands among PVS's work on principles and methods of survey sampling: the use of pilot studies; multi-stage sampling; regression estimators in double sampling and study of non-sampling errors. In 1947, the United Nations Sub-Commission on Statistical Sampling put its imprimatur on the importance of pilot studies (to obtain preliminary estimates of costs and variances). PVS's work on double sampling was among the first set of works to be published on the topic; and his (G.R. Seth's) model used to study non-sampling errors in surveys, starting in 1952, were again contemporaneous with the pioneering work on the topic by Hansen and his colleagues at the U.S. Bureau of the Census (Hansen *et al* [6], [7]). The topic of non-sampling errors remained an abiding interest of PVS; witness his paper

- 1 Expanded version of an invited paper, presented at the Special Session on Survey Sampling in Memory of Professor P.V. Sukhatme, organized by the International Indian Statistics Association, on 10 August 1997, in Anaheim, California, U.S.A.
- 2 Consultant to the United Nations and the United Nations Population Funds in the fields of survey sampling and demography; 30 Waterside Plaza, Apt. 24A, New York, NY, 10010, U.S.A.

on some methodological aspects of sample surveys of agriculture in developing countries.

At the imperial (later, Indian) Council of Agricultural Research (ICAR), where he worked first as the Statistician and then as the Statistical Adviser from 1940 to 1952, PVS gathered around himself a group of statisticians to work in agricultural sample surveys. Among the reports that bear his name are: surveys on wheat in Punjab, North-West Frontier Province; surveys on paddy/rice in Central Provinces and Berar, Madras, Kolaba district in Bombay, and in the Bombay Province and surveys on the catch of sea fish. He also provided overviews of crop surveys in particular, and sample surveys in general.

As an interlude of the history of statistical development in India, it was noted that the optimum size of cut required to estimate the yield of crops became, in the mid-1940s to the early fifties, a subject of controversy between Dr. P.V. Sukhatme and Prof. P.C. Mahalanobis, as will be witnessed from Sukhatme's papers and Mahalanobis's papers [15], [16], [17]. Mahalanobis recommended the use of smaller sized cuts (in particular, circular cuts of radii 2', 3' and 4' providing estimates from the two concentric circles and the annular ring) to provide yield estimates; Sukhatme favoured larger sized cuts in view of the possibility of a greater border bias in smaller cuts (See Ghosh [4]). Deming [2], Sengupta [20] and Yates [25] have also discussed this topic.

The Indian National Sample Survey (NSS), that was started by the Government of India in 1950, to obtain economic, social, demographic, and agricultural characteristics of the country and States, had, for its crop estimation inquires, two sets of samples: the "*Central Sample*", that was under the direct technical and operational supervision of the NSS (the Central Government and the Indian Statistical Institute) and the "*State Samples*", under the supervision of the States Governments. During the first two decades of its operation, the NSS had used small cuts for the "*Central Sample*"; for the "*State Sample*", the State of West Bengal used small cuts but other States continued the tradition of using large sized cuts.

Now crop surveys are conducted by the States alone, without any matching "*Central Sample*", and the NSS plays only a supervisory role: the shape and size of cuts for various crops vary from State to State, most of the States use 5m × 5m square or 5m × 10m rectangular cuts; in the State of West Bengal, small cuts continue to be used, with concentric circles of radii 2', 4', and 5.625' providing three estimates of crop yields.

Another topic of debate was the technique of interpenetrating networks of sub-samples, introduced by Mahalanobis. Interested readers are referred to Mahalanobis [15], Sukhatme *et al* (1984), Ghosh [3], Koop [9], [10], [11], [12] and Som [21], [22].

In 1951-52, PVS was a Visiting Professor at the Department of Statistics and Statistical Laboratory at Iowa State University when he consolidated his work on theory and application of sample surveys (Kale, cited by Milton [8]). The result, his *Sampling Theory of Surveys with Applications*, was first published in 1952 (by the Iowa University Press) and has gone through three editions: the second, with his brother, the late B.V. Sukhatme, published in 1970, and the third, with Shashikala Sukhatme, and C. Asok joining in, came out in 1984. Systematizing principles and practices, this book has been a perennial favorite with students and researchers alike.

It should be of interest to look at the historical backdrop of the book. When the United Nations Sub-Commission on Statistical Sampling met in 1947, there were no books on sampling theory and methods, and students and survey practitioners had to search scientific journals, such as *Journal of American Statistical Association*, *Journal of the Royal Statistical Society*, and *Sankhya*, the *Indian Journal of Statistics* for papers on the topic. The very first session of Sub-Commission recommended that a manual on statistical sampling be prepared to assist in implementing the projected 1950 World Census of Agriculture and the 1950 World Census of Population. This recommendation resulted in a flowering of books by early pioneers of sampling —Yates [25], Deming [2], Sukhatme (1952), Cochran [1] and Hansen *et al* [6], [7].

In 1952, Sukhatme joined the Food & Agriculture Organization (FAO) of the United Nations, and retired in 1972 as the Director of the Statistics Division. The succeeding revisions of his book reflected his increasing world-wide experiences of survey sampling, particularly in agriculture, reflected also in his book, *Statistical Methods for Agricultural Research Workers*.

In recognition of his work in survey sampling, PVS was elected President of the International Association of Survey Statisticians in 1977.

2. Estimates of People Suffering from Malnutrition and Hunger

In addition to continuing research on theoretical and applied survey sampling at the FAO, PVS gave attention to the measurement of malnutrition and hunger in the world. Alarming estimates of the world's hungry population—such as 60 per cent of the world's population had been floating. But PVS had noted that a traditional Indian diet provided the necessary nutrients to enable many Indians to carry out their daily work efficiently without showing symptoms of chronic malnutrition or hunger (Kale cited by Milton [8]). PVS presented his finding to the U.K. Royal Statistical Society in 1961 and was awarded the prestigious Guy Medal of the Society.

After his retirement from FAO, PVS continued his research on nutrition and hunger at the Maharashtra Association for Cultivation of Science (MACS), Pune. He was also instrumental in establishing the Indira Community Kitchen

Programme, which provided to impoverished Indians engaged in hard manual labour a cheap, but without government subsidies, and balanced vegetarian diet - that may not have conformed to Western standards - (Kale, cited by Milton [8]). His analysis of the data led to the Sukhatme-Margen Hypothesis:

At low levels of calorie intake, energy is used with greater metabolic efficiency and that efficiency decreases as calorie intake increases over the homeostatic range.

The results of his analysis became the subject of a one-on-one lively debate between PVS and V.M. Dandekar, an eminent Statistician in India.

In this connection, the publication of the book *Diet, Disease and Development* in 1992 deserves mention. His 1984 paper showed how the lack of precise hypothesis and the failure to take interaction into consideration led to erroneous conclusions in studies of behavioral traits of man, including the relationship of IQ with social class. His concluding words are salutary, reflecting his deep humanism and social concern, and remain as valid as they were when he uttered them thirteen years ago:

“Even the concept of education as a process has not been clearly set forth by [the contending] schools. Unless this is done, we will not know how our education should be reformed to mould national character. Even the policy for reservation of jobs for scheduled castes ... in India depends on the outcome of this research. There is a real need to set up long term experiments to study this problem.”

3. Institution-building for Survey Sampling

3.1 The Indian (formerly Imperial) Council of Agricultural Research

After a brief stint during 1939-40 as Professor at the All India Institute of Hygiene and Public Health, Calcutta, PVS joined the Imperial (now Indian) Council of Agricultural Research (ICAR) in 1940. The talented statisticians that PVS had gathered around him at the ICAR included V.N. Amble, K. Kishen, R.S. Koshal, R.D. Narain, P. Narain, V.G. Panse, K.V.R. Sastry, G.R. Seth, and D. Singh, who jointly and severally developed sampling theory and practice in agriculture forestry, and fishery.

PVS was the prime mover in the founding in 1947 of the Indian Society of Agricultural Statistics, a scientific body with its headquarters in New Delhi. It started a statistical journal - *Journal of Indian Society of Agricultural Statistics* - where the results of theoretical and applied work by ICAR staff were often published. Professor Sukhatme was the first Secretary of the Society and its guiding spirit and he was able to have illustrious Dr. Rajendra Prasad, with his interest in agriculture, as the first President of the Society.

As a founder of the Society, P.V. Sukhatme guided it in pioneering and populating statistical methods in agriculture and related sciences.

ICAR had instituted training in statistics, including survey sampling. Most of the statisticians working on agriculture in the Central and State Government have gone through the training process, thus creating a huge impact on government statistics in India.

In recognition of the research and training it provided, converted into Indian Agricultural Statistics Research Institute (IASRI) thus becoming the apex organization for statistical research and training in agriculture in India.

3.2 The Department of Statistics, Bombay University

To impart training in statistics at post-graduate levels, leading to the Master's and doctoral degrees. The Universities of Calcutta, Kerala and Madras had established a Department of Statistics each respectively in 1941, 1943 and 1946. PVS took the lead in establishing the Department of Statistics at the University of Bombay in 1948: this prompted several other universities in Western and Central India, such as Poone (1950), Karnataka (1951), Delhi (1957), Ujjain (1960) and Rajasthan (1961), to open their own departments of statistics.

3.3 The Food and Agriculture Organization of the United Nations

At the Food and Agriculture Organization of the United Nations (FAO), PVS gathered talented statisticians from amongst the international community, such as D. Basu, El-Mahdi Said, H.M. Hussein, G.T. Jones, S.H. Khamis, C.H.M. Morojele, R.D. Narain, R.S. Sampford, W.S. Schulte, J.B. Simaika, P.C. Tang, S.S. Zarkovich, and R. Zasepa.

PVS was in a unique position to advise and assist governments of developing countries to use survey sampling techniques in agriculture, and he used the opportunity fully; FAO's technical assistance program covered most of the developing countries, from Afghanistan to Ethiopia to Zambia, through a network of regional commissions on agricultural statistics (and their equivalent bodies in Europe and Latin America) FAO advisers posted in regions and sub-regions, and FAO country expert-all working to ensure that census and survey programs and related international recommendations were adapted to meet local conditions and requirements.

In order to increase national capacities for survey sampling in agriculture, the FAO Statistics Division organised, all over the world, national, sub-regional, and regional training courses and published several research studies on census

and sampling, that are considered definitive. At the headquarters level, the FAO Statistics Advisory Committee was established, comprising leading world statisticians in the field, such as D.J. Finney, J. Neyman, V.G. Panse, C.F. Taeuber, and G. Theodore, to review advise and guide FAO's work in statistics.

PVS also gave special attention to further develop and expand FAO's work in the field of household food consumption surveys, design of experiments (stressing those in farmer's fields) and to prepare for the decennial FAO World Agricultural Census Programme with emphasis on the role of sample surveys as an integral part of the programme. Indeed, most of the agricultural censuses are based on sample surveys.

Under PVS's leadership, FAO played a significant role in coordinating activities in Statistics and Population of the United Nations and the U.N. agencies under the umbrella of the U.N. Administration Co-ordination Committee (ACC). Indeed, he chaired the 1967 session in Rome of the ACC Sub-Committee on Statistics and otherwise strengthened coordinating ties, including Family Expenditure Surveys with the Statistical Office of the U.N. and of other agencies; in the field of population, at the Sub-committee on Population, he was responsible for harmonizing FAO's agricultural and rural populations with the U.N.'s population projections and on population and food supplies.

3.4 The Maharashtra Association for Cultivation of Science (MACS), Pune

After retiring from FAO, PVS served as Regents Professor, University of California at Berkeley, and then joined the Maharashtra Association for Cultivation of Science (MACS) in Pune, where he founded the Department of Biometry and Nutrition, with a master's degree programme. He was also the Honorary Professor of Biometry at the Association.

4. P.V. Sukhatme's Legacy

The legacy that P.V. Sukhatme has left behind covers two areas. First national and international capacity for the scientific collection and analysis of agricultural statistics, owing to his pioneering leadership at ICAR for India, and at FAO for all over the world.

Second, his emphasis of the importance of studying not only sampling errors but, often most importantly, non-sampling errors, ranging from the fear of the "evil eye" to "border bias". Note his observations on the importance of evaluating and adjusting for non-sampling errors:

“To reduce sampling errors while allowing a bias several times as large to enter into the census [and survey-RKS] results is questionable from the viewpoint of the efficient use of resources,” and “In extreme cases, [response errors] may even vitiate the numerical results of the estimates to such an extent as to distort the picture of the agricultural situation altogether.”

These statements deserve to be enshrined by every survey- practitioner.

ACKNOWLEDGEMENTS

I am grateful to PVS's son, Prof. Vikas Sukhatme of Harvard University, and Dr. Salem H. Khamis, PVS's former colleague at FAO, for supplying me with material on PVS's life and work. For additional information, comments and suggestions, my thanks are due to: Prof. Subir Ghosh, University of California at Riverside, Prof. Vidyadhar Godambe, Department of Statistics, University of Waterloo, Dr. J.C. Koop, Consultant, Mr. Madhukar D. Palekar, retired Assistant Director, Statistical Office of the United Nations; and to Dr. P. Mukhopadhyay, Member-Secretary International Statistical Education Centre, Indian Statistical Institute, Calcutta, for the information on the present crop- cutting practices at the NSS and the States in India.

REFERENCES¹

- [1] Cochran, William G., 1953. *Sampling Techniques* (3rd ed., 1977). John Wiley and Sons, New York.
- [2] Deming, W. Edwards, 1950. *Some Theory of Sampling*. Reprint (1984), Dover, New York.
- [3] Ghosh, B., 1949. Interpenetrating (network of) samples, *BCSA*, 2, 108-119.
- [4] Ghosh, B., 1956. A model for perimeter errors, *BCSA*, 6, 189-192.
- [5] Gore, A.P., 1997. P.V. Sukhatme: a social statistician, *Eco. & Pol. Week*, Mumbai, February 8, 1997.
- [6] Hansen, Morris H., William N. Hurwitz and William G. Madow, 1953. *Sample Survey Methods and Theory*, I, *Methods and Applications*, II, *Theory*. John Wiley and Sons, New York. Reprint (1993).
- [7] Hansen, Morris H., William N. Hurwitz, Eli S. Marks and W. Parker Mauldin, 1951. Response errors in surveys, *J. Roy. Statist. Assoc.*, 46, 147-190.
- [8] Kale, B.K., 1997. Tribute to P.V. Sukhatme, excerpted by Milton, Barbara.
- [9] Koop, J.C., 1960. On the theoretical questions underlying the technique of replicated or interpenetrating samples. *Proceedings of the Social Statistics Section, American Statistical Association*, 196-205.

1 The papers by Prof. P.V. Sukhatme, referred to in this article are omitted from the 'References' to avoid duplication, as these are printed separately in this publication.

- [10] Koop, J.C., 1984. The technique of replicated or interpenetrating samples, In: *Handbook of Sampling, 6, Sampling*, edited by P.R. Krishnaiah and C.R. Rao, Elsevier, Amsterdam, 333-362.
- [11] Koop, J.C., 1993. Conditional inference about an existing universe from independent interpenetrating samples. *Commun. Statist. Theory and Methods*, **22** (4), 985-1006.
- [12] Koop, J.C., 1997. On conditional inference from sample survey data about an existing universe. *Sankhya (B)*, **59**.
- [13] Mahalanobis, P.C., 1938. Statistical Report on the Experimental Crop Census, (1937). Indian Central Jute Committee, Calcutta.
- [14] Mahalanobis, P.C., 1944. On large-scale sample surveys. *Philosophical Transactions of the Royal Society*, London, **231 (B)**, 329-351.
- [15] Mahalanobis, P.C., 1946a. Recent experience in statistical sampling in the Indian Statistical Institute. *JRSS,(A)* **108**, 326-378. Reprint (1958) with a Preface, *Sankhya*, **20**, 1-68, and 1961, Asia Publishing House, London, and Statistical Publishing Society, Calcutta.
- [16] Mahalanobis, P.C., 1946b. Use of small-size plots in sample surveys for crop yields. *Nature*, **158**, 798-799.
- [17] Mahalanobis, P.C. and Sengupta, J.M., 1951. On the size of sample cuts in crop-cutting experiments in India, 1939-50. *BISI*, **33** (2), 359-404.
- [18] Milton, Barbara, 1997. Obituaries: P.V. Sukhatme. *Amstat News*, **245**, 46.
- [19] Neyman, J., 1934. On two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. *JRSS*, **97**, 558-606.
- [20] Sengupta, J.M., 1964. On perimeter bias in sample cuts of small size. *Sankhya (B)*, **26** (1 & 2), 69-88.
- [21] Som, R.K., 1965. Use of interpenetrating samples in demographic studies. *Sankhya (B)*, **27** (3 & 4), 342-349.
- [22] Som, R.K., 1995. *Practical Sampling Techniques* (2nd ed.), Marcel Dekker, New York.
- [23] Tschuprow, A.A., 1923. On the mathematical expectation of the moments of frequency distributions in the case of correlated observations. *Matron*, **2**, 4.
- [24] United Nations, 1947. *Report of the Sub-Commission on Statistical Sampling to the Statistical Commission*. New York. Reprint (1948), *Sankhya* (8), 393-402.
- [25] Yates, Frank, 1949. *Sampling Methods for Censuses and Surveys*. Charles Griffin, London, 4th ed., (1981), Edwin Arnold, London and Oxford University Press, New York.