Journal of Official Statistics, Vol. 26, No. 2, 2010, pp. 215–231

Official Statistics in India: The Past and the Present

T.J. Rao[1]

In this article we shall trace the history and evolution of official statistics in India during the British period and immediately after independence. We shall also discuss briefly the teaching and training of statisticians in the country. In particular, the role played by P.C. Mahalanobis, C.R. Rao and P.V. Sukhatme among others, will be highlighted. Next, we shall look into the present scenario in the light of some of the recommendations, released in 2001, of the National Statistical Commission.

Key words: Statistical system in British India; Statistical system after Indian independence; Central Statistical Organisation; National Sample Survey; Indian Statistical Institute; Institute of Agricultural Research Statistics; P.C. Mahalanobis; P.V. Sukhatme; C.R. Rao; National Statistical Commission; Gap in theory and practice.

1. Introduction

Early origins of statistical data collection, compilation for various characteristics, and need for cross checking by an independent set of agents working under disguise are mentioned in Kautilya's great treatise in economics, *Arthasastra* (attributed to 321 – 296 B.C.). In addition to these official statistics, in particular, one finds in Chapter XXXV, p. 159 in *Arthasastra*: "*Spies under disguise of householders (Grihapatika, cultivators), who shall be deputed by the Collector-General for espionage, shall ascertain the validity of accounts (of Gopas, the village officers and Sthanikas, the district officers) regarding the fields, right of ownership and remission of taxes with regard to houses, and the caste and profession regarding families. . . . . .*" (Shamasastry 1929). One could make a guess that perhaps this prompted Mahalanobis to have an independent supervisory staff during the conduct of field operations of a survey.

During the great Moghul Emperor Akbar's period of rule (*Circa* 1590), we find that the details of official statistics from various departments of the Emperor including the land classification, crop yields, measurement systems, revenue, etc. are available in the masterpiece *Ain-i-Akbari* written by Abul Fazal. In the English translation, Jarrett (Vol. II, 1894, p. vii) notes that this wealth of information was "*a unique compilation of the systems*

[1] Adjunct Professor, C.R. Rao Advanced Institute of Mathematics, Statistics and Computer Science, University of Hyderabad Campus, India. Email: tjrao@hotmail.com

Acknowledgments: This article is based on talks given at the National Statistics Day Celebrations and 2nd Lecture Series on "Statistics: Past, Present and Future" organised by the C.R. Rao Advanced Institute of Mathematics, Statistics and Computer Science, University of Hyderabad and Indian Statistical Institute, Hyderabad during 29 June – 1 July, 2008. The author also wishes to thank the referees and an associate editor for their valuable comments, which improved the presentation.

of administration and control through the various departments of Government in a great empire, faithfully and minutely recorded in their smallest detail, with such an array of facts illustrative of its extent, resources, condition, population, industry, and wealth as the abundant material from supplied official sources could furnish."

2. The Past – Official Statistics During the Pre-Independence Days

The instability of the governments of the subsequent Moghul rulers during the 18th century led to the increased presence of the British who came to India as traders, businessmen, missionaries, plantation owners, etc. Statistical data collection and its maintenance suffered and the Britishers initially interested in trade finally established the East India Company in Eastern India, which turned out to be an organ of political power. Eventually, almost all of the undivided India came under the rule of British Empire (1757–1947). During the British Raj, in Eastern India, collection of revenue by intermediate tax collectors called Zamindars under a "permanently settled" system was introduced. The amount of revenue to be paid to the British treasury by the Zamindar was made permanent and was supposed to be paid by the sun set on a particular date. This "permanently-settled" system following the "sun-set law" suffered owing to the nonavailability of the village level functionaries called *patwaris*. There was no elaborate arrangement or an official statistical agency for collecting primary statistics, let alone the revenue.

Shaken by this, the Court of Directors of East India Company sent an urgent dispatch asking for carrying out a statistical survey of the Presidency of Fort William. In 1807, the Survey of Eastern India was conducted by the Governor-in-Council, Dr. Francis Buchanan. This survey covered an area of 60,000 sq. miles and of 15 million British subjects at a cost of 30,000 Pounds. Buchanan's Report (Buchanan 1807) was directed to a number of subjects such as topographical account of each district, conditions of the inhabitants, their religion and customs, the natural produce of the country, details on fisheries, mines and forests, vegetables grown, manure types, farm sizes, state of landed property, progress made in fine arts, commerce and a list of rare, useful and curious plants and seeds.

Thirty years later, Montgomery Martin was sent to India to assess the report containing all this detailed information of the region, collected with a "*critical attitude, keen scientific spirit and an experimental approach*" (cf. Mahalanobis 1950, p. 199). Martin (1838) published *The History, Antiquities and Statistics of Eastern India* in three volumes covering nine districts of Bengal. This work not only brought in England an awareness of the condition of the people in British India but was instrumental in recommending moderate land revenue rates, duties as per free trade, establishment of a banking system and municipalities. A few years later, Col. Sykes established a small department of statistics in 1847 in India House. This is perhaps the beginning of an official statistical system for (the then British) India. Baines (1918, p. 417), while tracing the development and progress of official statistics in India during the early British days, comments that "*Official Statistics, in the present sense of the word, are the offspring of British rule.*" He further adds: " *from the introduction of the British rule, attention has been paid to the record of fiscal transactions and of the sea-borne trade. Administrative statistics followed fitfully, it is true, but keeping pace with the development of the Province, and dating back,*

accordingly, to different periods in each local series of returns" (Baines 1918, p. 421). The Department of Statistics initiated by Sykes, started to collate data and release a series of statistical papers on India from 1853. One could also trace the first census published by a government officer named A. Shakespeare in 1848 which referred to the area and revenue of each *pargana* (district) in North West Provinces. Later Census Reports were published on 1 January 1855 and 10 January 1868. It is interesting to note that W.H. Carey gives a total count of Calcutta city as 179,917 in 1822 and a round figure of 400,000 in 1847!! The East India Company had taken a few censuses during 1769–1855 and the Crown during 1858–1869 but these were fragmentary and lacked uniformity. In 1862 a Statistical Committee was set up for preparing model statistical forms for compilation of a uniform system of Imperial Statistics.

The years 1867 to 1872 witnessed the first systematic attempt to enumerate the whole population of India by *actually counting heads*. But this was not synchronous and also lacked uniformity. The second census of 1881 was synchronous, more complete and modern and continued every 10 years, ever since. The next Census of India is due in 2011. Dynes (1899) notes that: "*In the Census of India, in Bengal, for example in 1881, steps were taken to cover all cases so effectively that everyone in the country on a given night was believed to be recorded, travelers on river boats or railways included, thus loosing, as the officials claimed, from a direct count only those absent on voyages by sea.*" Kingsley Davis (1951, p. 4) while correcting and adjusting the census figures, notes that "*the Indian censuses are remarkable not only for the information they reveal but for the special obstacles they have had to overcome*." For example, the 1931 reports appear in no less than 46 separately bound volumes of huge size (over a foot high!). During 1941, Census Commissioner Yeatts proposed to cut down the scope of census volumes and resorted to the "*humbler role of just collecting the statistical data, doing as little as possible of writing and leaving the figures to speak for themselves*" (Davis 1951, p. 5).

Going hand in hand with census reports are the Gazetteers of India. It is remarked that: "*No comparable area of the world has anything like the prodigious compilation of statistical data and demographic and historical material as a country which is almost a continent in the immensity and diversity of its character*" (Chaudhuri 1964). Hamilton's gazetteers of 1815, and its revision in 1828 and Thornton's of 1854 are some of the early ones. It was not until 1869 when the first Director General of Statistics in India, Dr. W.W. Hunter was appointed, that a plan was submitted to the Governor General in Council for a comprehensive and coordinated statistical survey. This resulted in the *Statistical Account of Bengal* (Hunter 1875) in 20 volumes. Hunter's scholarship is reflected in his other works, *Annals of Rural Bengal* in three vols. (Hunter 1868) and *A History of British India* (Hunter 1899) in two vols. This survey was followed in other regions and by 1881 the Imperial Gazetteer of India was released in nine vols. covering 15 British Indian provinces comprising 240 districts. By 1885–1887 the size increased to 14 vols. and by 1901 to 26 vols. and the process of releasing gazetteers continued.

It is said that "*if the census officials in India have had any defect at all, it is that they have been scholars and government men first and statisticians second*" (Davis 1951, p. 5). The Census Commissioners, Risley in 1901, Gait in 1911 and Hutton in 1931 were all interested in anthropology. A rich collection of data is available in *Statistical Abstracts for the several colonial and other possessions of the United Kingdom 1854–1868*, sessional

papers, cmd. 146 and the *Statistical Abstract relating to British India, 1861–1870* (London, H.M.S.O. 1870). For a later period, data on prices and cost of living is found in *Statistical Summary of the Social and Economic Trends in India* (Subramaniam 1945). Interesting analyses relating to data on various industries, demographic trends, levels of living of workers, estimates of national income among others are presented in *Economic History of India*: 1857–1956 (ed. Singh (1965)).

Before 1882, there was no systematic collection of agricultural data. The Secretary of State forwarded a set of tables compiled in India Office. Even though they lacked uniformity and were incomplete, they were published in *Agricultural Statistics of India*, in two volumes one for British India and the other for Indian States. The job of scrutinising and summarising was given to a Statistical Bureau established in 1895 under the Director General of Statistics who also coordinated trade, industrial and price statistics. Lord Curzon in 1905 bifurcated the department and a separate Directorate General of Commercial Intelligence and Statistics, (which completed 100 years in 2005) was established in Calcutta. Soon, *The Indian Trade Journal* appeared in 1906. Data on exports and imports, balance of trade and the role of banks during the mid 19th century to 1918 was reported in Shirras's book on *Indian Finance and Banking* (Shirras 1919).

With more and more data being collected, it was recommended by two committees, one headed by Visweswarayya in 1925 and another by Bowley and Robertson in 1934 that the government should set up an Inter Departmental Committee, which should look into the possibility of formation of a Central Statistical Office. Such an office should be chiefly for coordination and maintenance of all important statistics for the entire country.

3. Official Statistics Immediately After Independence

Immediately after independence, an urgent need was felt for a statistical structure towards a socio-economic development. By then Mahalanobis had already established himself as an expert in data collection through sample surveys especially in relation to crop acreages and yield, besides developing theoretical multivariate statistical tools. He founded the Indian Statistical Institute in 1931. His work described in the 1946 paper of the *Journal of the Royal Statistical Society* earned him an FRS. Thus it is not surprising that he was appointed as the Honorary Statistical Advisor to the Indian Government in 1949. A Central Statistical Unit was created under the charge of the Statistical Advisor in 1949. In May 1949, a permanent office for the Census and Vital Statistics was established under a Registrar General and ex-officio Census Commissioner. In 1951 the Central Statistical Organisation was formally established. The Central Statistical Organisation is now responsible for the coordination of statistical activities in the country, and for evolving and maintaining statistical standards. The most important activities of the organisation are National Income Accounting, Annual Survey of Industries together with follow-ups, as also the compilation of Indices of Industrial Production and Consumer Prices. It also publishes Human Development and Gender Statistics, participates in developing Trade, Energy, Construction and Environment Statistics, and provides training in official statistics.

Around the same time, the National Income Committee established in 1949 and chaired by Mahalanobis, together with the Standing Committee of Departmental Statisticians,

found large gaps in the statistical information. Not many attempts were made when it came to the estimation of the National Income (cf. Rao, V.K.R.V. 1940) of the country.

It was on 18 December 1949 that the first Prime Minister of Independent India, Pundit Jawaharlal Nehru, expressed the desire that a sample survey should be organised covering the whole country to collect essential information. By 25 December an abstract scheme for organising a National Sample Survey (NSS) had been prepared by Mahalanobis and was passed on to C.D. Deshmukh who was the Finance Minister at that time. This scheme was approved by the government in January 1950 and on 10 March the National Income Committee recommended the use of sampling methods to fill the gaps in information. Thus, in October 1950 the first round of data collection in the NSS took place. Rudra (1997) in his biography of Mahalanobis discusses the details of these developments.

One of the very first challenges in collecting official statistics immediately after Indian independence was to estimate the number of the unknown large number N of persons of a minority community who took refuge in the Red Fort in Delhi during the communal riots of 1947. It was impossible and risky either to take a census or to use a sample to estimate this number. An unconventional and ingenious method of obtaining this number was suggested by J.M. Sengupta, a close associate of Mahalanobis. He gathered data on quantities of rice, pulses and salt billed by the supplier and knowing the per person requirements of these items from nutrition surveys, then calculated the relevant ratios to obtain three parallel and valid estimates of N. It was found by cross-checking of data that the supplier had exaggerated the quantities of the expensive items rice and pluses, while Q, the quantity of salt, was correctly billed. Thus Q/q, where q is the per capita salt requirement, gave a good estimate of N.

Mahalanobis saw the need for a decentralised statistical system and encouraged the setting up of State Statistical Bureaus (now known as Directorates of Economics and Statistics) as well as District Level Statistical Offices. He involved C.R. Rao in the creation of the SSBs and District Offices and the training of the officials. These Directorates release estimates of area, production and yield in respect of principal crops. Yield estimates are obtained through crop-cutting experiments organised by them. The Directorates also function as the statistical authority and as the nodal agency for collection, compilation, analysis and dissemination of statistical information relating to all sections of the economy of the states.

While Mahalanobis played an important role in establishing the Central Statistical Organisation and the National Sample Survey for socio-economic issues, it was P.V. Sukhatme whose monumental work on Sampling Theory of Surveys, with applications especially in agriculture, at the Indian Council of Agricultural Research had a great effect on the release of official statistics in agriculture. The success of training programmes and consultancy at the Statistics branch of the Council eventually led to the establishment of the Institute for Agricultural Research Statistics.

P.V. Sukhatme graduated from Fergusson College, Pune, in 1932 and later studied at University College, London. His work, published as three original contributions in Statistics earned him a Ph.D. in 1936 and the innovative results on bipartitional functions which were published in *Philosophical Transactions of the Royal Society of London*, Series A in 1938 earned him a D.Sc. in 1939. During 1939–1940, after a brief stay at the All India Institute of Hygiene and Public Health, Calcutta, he joined the Indian Council of

Agricultural Research in 1940 as a statistician and subsequently acted as a Statistical Advisor of the Council till 1951. During his stay at the Council, along with V.G. Panse, he got involved in crop-cutting experiments for estimating area and yield of paddy, wheat, etc. While Mahalanobis was using smaller circular cuts (such as cuts of a radius of $4'0$) to be surveyed by ad hoc trained survey personnel, Panse and Sukhatme advocated the use of rectangular cuts of size $33' \times 16.5'$ in their experiments to be conducted by the administrative agency. The pros and cons of these two methods were much debated (Panse and Sukhatme 1948, 1951). This resulted in a major scientific dispute and Adhikari (1990) cites the different regions of India these scientists came from and the corresponding practices of collection of revenue and official statistics since early days in these two regions as the reason for their preferences. For a detailed account of this controversy and the joint compromise studies by the government, see Ghosh et al. (1999) and Joan Fisher Box (1978).

Sukhatme became the chief of the Food and Agriculture Organisation in Rome (1952–1971). After his retirement from the FAO, he got interested in World's Hunger and Protein Deficiency Models which later led to the Sukhatme-Margen Models (Sukhatme and Margen 1978). Sukhatme took an active part in planning the 1960 World Census of Agriculture (Sukhatme 1958) and participated in the seminars organised by the FAO. In his later days, he developed newer concepts in nutrition. He was responsible for a state-sponsored Community Kitchen Society project and also suggested improvements in the collection of official statistics relating to government-sponsored Integrated Child Development Services.

Like Mahalanobis, Sukhatme also stressed the importance of controlling nonsampling errors in order to obtain reliable official statistics. Besides training and emphasis on error-free collection of data, along with G.R. Seth, he proposed observable error models and estimated "observer effects" (Sukhatme and Seth 1952). These models were later applied in the analysis of longitudinal data relating to nutrition.

4. Teaching of Statistics and Training

It is Mahalanobis again who started the oldest Department of Statistics in 1941 at Calcutta University. Apart from teaching, faculty members were also involved in research and made fundamental contributions in multivariate analysis, construction of designs, sample surveys and inference. Mahalanobis thought that a post graduate programme could be much strengthened if the students were trained in statistics at the undergraduate level itself. Thus in 1944, an undergraduate Department of Statistics was established in the Presidency College, Calcutta, the syllabus of which became a model for all other later statistics departments in the country.

The first batch of students of the post graduate course of Calcutta University included C.R. Rao, "*who influenced the growth of the Indian Statistical Institute (ISI) more than any one else except Mahalanobis and, along with Mahalanobis, is the most famous statistician to come out of the Indian subcontinent*" (Ghosh et al. 1999, p. 23). After C.R. Rao returned from Cambridge in 1948, he became the Head and later the Director of the world famous Research and Training School of the Indian Statistical Institute and organised the research and training activities. Not only advanced level programmes in statistics were organised

for post graduate students, but at the same time statistical officers from various central and state governments were trained at the Research and Training School for improving their skills in dealing with official statistics relating to their jobs.

The Indian parliament passed the Indian Statistical Institute Act, 1959, which declared the Institute to be an "Institution of National Importance" and empowered the Institute to award degrees and diplomas in Statistics. The structure and syllabi of these courses, which started in 1960, were devised mostly by Mahalanobis and C.R. Rao along with J.B.S. Haldane, who worked at the Institute at that time. However, the training courses for statistical officers still continued as before.

Around the same time, the erstwhile Institute of Agricultural Research Statistics was organising a Professional Statistician's Certificate Course, which mainly focused on subjects relating to agricultural statistics, initially under the guidance of P.V. Sukhatme and V.G. Panse. This eventually grew to become a full-fledged Indian Agricultural Statistics Research Institute, which now has M.Sc. and Ph.D. programmes. For further details on the developments at this Institute see the Panse Memorial Lecture by G.R. Seth (Seth 1996).

5. The Role Played by Mahalanobis and His Colleagues in the Early Days

During the initial rounds of the NSS, under the general guidance of Mahalanobis, the sampling design section of which was housed in the Indian Statistical Institute, C.R. Rao, D.B. Lahiri, S. Raja Rao, and M.N. Murthy, among others, took an active part in the design aspects of the surveys. Early field work organised by S.S. Bose and J.M. Sengupta, among others for the crop-cutting experiments of Mahalanobis led to important applied work in the design of experiments by K.R. Nair on quasi-factorial experiments and other designs (Mahalanobis and Nair 1940), which improved official statistics on the yield of crops. D.B. Lahiri, whom Mahalanobis described "*as his conscience keeper in sampling theory,*" collaborated with Mahalanobis to make an assessment of the errors in censuses and surveys especially in the context of Indian official statistics (Mahalanobis and Lahiri 1961). Encouraged by Mahalanobis, Lahiri presented a paper on sampling with probability proportional to total size at the conference of the International Statistical Institute in 1951 (Lahiri 1951) as well as at the meeting of the UN Subcommission on Statistical Sampling around the same time. This method (also known as the Hájek-Lahiri-Midzuno-Sen method) is very much quoted in sample survey literature. M.N. Murthy along with his colleagues in the survey design section of the National Sample Survey devised techniques of obtaining unbiased ratio estimators and applied the circular systematic sampling method in surveys to improve the precision of estimates. While discussing the contributions of Mahalanobis, M.N. Murthy describes the multi-stage designs for crop yield employed by him, which sometimes involved four stages, namely villages, grids, plots and sample cuts. Further, Murthy (1964) explains how Mahalanobis treated the total error in sample surveys as consisting of sample to sample variation, fluctuations in measurements and observations, and more importantly deliberate mistakes by the investigators for which special cross-checks and controls were incorporated by him. N. Chakravarti helped Mahalanobis in the analysis of 60 years data on river floods in Orissa, the report based on which resulted in the construction of the Hirakud dam on the river Mahanadi in Orissa (Mahalanobis and Chakravarti 1931). This analysis is a good

reference material for releasing official statistics by the Central Water Commission of the Government of India and the River Research Institute. Important contributions in other applied areas such as agriculture, animal husbandry, fisheries, etc. were due to statisticians at the Institute of Agricultural Research Statistics – V.G. Panse, K. Kishen, V.N. Amble, G.R. Seth and P. Narain among others.

Around 1947, Stuart Rice, past President of the International Statistical Institute and member of the UN Statistical Commission, was strongly advocating an international programme for education in statistics, especially in those countries, which did not possess the present or potential resources of trained personnel. Mahalanobis, another member of the Commission, further suggested that such an organisation should be set up for Asian countries or for India and neighbourhood. The efforts of these two bore fruit with the setting up of the International Statistical Education Centre in 1950 in Calcutta. This Centre used to be jointly operated by the Indian and International Statistical Institutes under the auspices of UNESCO and the Government of India. At present, the Centre is operated by the Indian Statistical Institute under the auspices of the Government of India. Since its inception, it has provided training for more than 1,450 trainees from around 75 countries. Mahalanobis was the first chairman of the joint Board of Directors till his death in 1972. C.R. Rao has held this position as chairman of the Board of Directors since 1972.

With the growing need for development of official statistical systems, C.R. Rao chaired a UN Committee for the development of Statistics in South East Asia and recommended the establishment of an institute for this purpose. Thus in 1970, the Asian Statistical Institute was established in Tokyo, renamed in 1977 as the Statistical Institute for Asia and the Pacific (SIAP). While the ISEC centre at Calcutta conducts a regular course of ten months duration in Theoretical and Applied Statistics for strengthening the national statistical offices, the Tokyo institute SIAP organises modules of shorter duration.

During his visits abroad in the mid-forties, Mahalanobis was impressed by the quality control (QC) techniques that were being used, and he advocated the use of these in the still growing Indian industries. Later C.R. Rao was also associated with the QC movements in India. A start was made in 1946 when a special course on QC was organised by ISI, attended by twelve delegates. Encouraged by C. Tattersall of the Ordnance Testing Laboratory, Mahalanobis invited Walter Shewart to India in 1947. Under his guidance, ISI organised a one-week conference on "Standardization in Industrial Statistics" in Calcutta in February 1948 jointly with the Indian Standards Institution and this was attended by 190 delegates. The Indian Statistical Institute now has a full-fledged Statistical Quality Control and Operations Research Division with service units throughout the country to provide training, professional expertise and consultancy in all areas of Quality Management and Quality Systems related to the certification of ISO series and to conduct research. Orthogonal arrays developed by C.R. Rao and Taguchi methods have been successfully utilised in industrial experimentation.

Encouraged by Pundit Jawaharlal Nehru, Mahalanobis prepared the draft of the Second Five-Year Plan at the ISI. Many important studies concerning planning and national development were conducted by the Planning Unit of the Research and Training School of the Institute.

It is interesting to note that African Ministers for Planning and Economic Development adopted an Action Plan for Statistical Development in Africa in 1990, and they have even

decided to celebrate 18 November as Africa Statistics Day each year. Japan has been observing 18 October as National Statistics Day since 1973. Also the State Institute of Statistics in Turkey has proclaimed 9 May as National Statistics Day UN Statistics Division (UNSD) reports that around 102 countries and areas have been celebrating a Statistics Day, Week or Month to raise the profile of official statistics. A Working Group of the UN Statistical Commission on International Statistical Programmes and Coordination at its meeting in February 1997 deliberated on the need for holding a World Statistics Day. The UN Statistics Division's proposal for observing a World Statistics Day was overwhelmingly supported by 70 countries and international agencies. It is now decided that 20.10.2010 would be observed as the World Statistics Day focusing on the subthemes: Service to the nation and the world, professionalism and integrity. During the Berlin Session of the International Statistical Institute in 2003, the International Conference on "The Future of Statistical Theory, Practice and Education" held in 2004 and the subsequent conferences, the present author suggested the possibility of organising 29 June, the birthday of Prasanta Chandra Mahalanobis, as India Statistics Day each year. The Council of the Indian Statistical Institute discussed this in their meetings and requested the Government of India to consider the proposal. On 24 December 2006 while inaugurating the Platinum Jubilee celebrations of the Indian Statistical Institute, the Prime Minister Dr. Manmohan Singh pronounced 29 June as the National Statistics Day and the first Statistical Day was celebrated on 29 June 2007.

Mahalanobis was also influenced by Hubback's (1927) work on crop-cutting experiments done in 1923 and 1925 in Eastern India and he used random sample cuts in the forties for the estimation of areas under crop and crop yield. As early as 1934, Mahalanobis used the first area sample in the whole world for jute forecast. Mahalanobis's sample estimate of jute production in Bengal, obtained with a tenth of the cost and a fiftieth of the work force, compared to the Government's complete enumeration, made a convincing effect with respect to the accuracy and economy of sample surveys. During 1937–1945, he introduced several innovative techniques and preferred to call them "experiments in statistical sampling." He concentrated on the problems of organisation in large-scale sample surveys. Just as the large-scale production of a chemical is a matter of chemical engineering rather than pure chemistry, the organisation of large-scale sample surveys was equated to "statistical engineering" rather than pure theory of sampling by Mahalanobis. These experiments covered wide-ranging areas.

Several of these experiments that were taken up by Mahalanobis in the forties are very much relevant even in the present-day context:

In the Foreword to the paper entitled "Bengal Anthropometric Survey 1945: A Statistical Study" (Majumdar and Rao 1958), Mahalanobis defined a "group" as those belonging to the same caste, religion and tribe and living in the same district, which leads to a two-way classification. A further subdivision by subcaste, clan, etc. would result in very small sample sizes leading to "unreliable" estimates. This is exactly the situation that occurs in small domain (area) estimation, which is now an important technique being used in the collection and analysis of official statistics for small domains of many countries.

In the forties, Mahalanobis studied the demand for currency notes and coins as well as the average life of currency notes. A recent project of the Reserve Bank of India (RBI) on the very same topic was taken up by the ISI.

Economic and demographic factors leading to indebtedness were looked into by the professor and we now have the All India Rural Indebtedness and Landholdings Survey of RBI, the National Sample Survey Organisation and others.

Public Opinion (including tea drinking habits) and Radio Listeners Preference Survey during wartime were taken up by Mahalanobis and ISI. Opinion polls, market research surveys, exit polls, etc. are very popular these days.

The reference period for the collection of Monthly Consumer Expenditure Data – whether it should be a week or a month – was examined by Mahalanobis with S.B. Sen (Mahalanobis and Sen 1954) in the early NSS days. The debate continues even now and a pilot survey of an NSSO Expert Committee chaired by N. Bhattacharya (N.S.S. Expert Group on Non Sampling Errors 2003) was organised recently.

In 1926, a catastrophic flood occurred in the Brahmani river in Orissa and several low-lying areas got flooded. A committee of engineers concluded that the river bed had risen and hence the height of embankments should be raised by several feet. Mahalanobis found a significant correlation between the rainfall in the catchment area and height of the river flood in the delta area. His suggestions based on a report submitted in 1930 resulted in the construction of the multi-purpose project Hirakud dam, thus avoiding the raising of the embankments. This early work in operations research is perhaps very much applicable in the current crisis of frequent flooding of Visakhapatnam (Vizag) Airport in South India. Even though a new elevated runway was built, still the entire surrounding areas get flooded. It is not the rainfall in Vizag that is causing the floods, it is a problem arising at the Meghadrigedda and Kondagedda reservoirs and their operations during heavy rainfall at the catchment area. Just as Mahalanobis thought, perhaps there is no need to raise the river embankments all the way up to the point where the river meets the sea in Vizag as was suggested. Further research is needed on this aspect.

Thus some of the official statistics collected today have their roots in these early experiments of Mahalanobis.

The concept of "Total Sample Survey Design" was advocated by both Hansen at the U.S. Bureau of the Census and Mahalanobis in India (Rao 1999) independently. It is quite interesting to see the similarities in the approaches of these two great survey specialists (Rao 1999).

Mahalanobis did not approve any government administration regulated by the principle of authority in most developing countries. He thus insisted on built-in cross-checks such as inter-penetrating network of subsamples (IPNS). Deming, who has been using this technique in the U.S. as "replicated samples" since 1950, acknowledges Mahalanobis's work thus: "*The main feature of the IPNS is simplicity in the calculation of the standard error of an estimate. It also enables one to estimate rapidly the mathematical bias, if any, in the formula of estimation . . . . It helps to detect gross blunders in selection, recording and processing. It permits evaluation of variances between investigators, coders and other workers in the various statistical stages of processing*" (Deming 1964). This technique is now considered as the curtain-raiser for "resampling procedures" like bootstrap.

It is very interesting to note that while piloting the bill relating to the Indian Statistical Institute Act in the Indian Parliament on 14 December 1959, the Prime Minister Jawaharlal Nehru, praising the work done by Mahalanobis at the Institute, stated:

"*If we take two sample surveys and the samples are different – not in the same place but in the same wide area – and you get the same results, the results of the sample survey are more definite and more reliable. They may differ by one percent or half percent. Therefore, there is a necessity . . . .*"

For economy and accuracy while collecting data in many countries, Mahalanobis favoured multi-subject surveys, which are integrated. Seng (1951, p. 230) observed that "*in India, the conditions approximate more nearly . . . to the conditions of those countries . . . for which the experience of India will serve as a guidance and as an example worth imitating.*"

Though the Indian membership elected to the International Statistical Institute was a mere three in 1937 out of a total of 211 when Mahalanobis was elected in 1937, these being J.E. O'Conor representing British India (1886), G.F. Shirras (1924) and K.B. Madhava (1935), around the year 2000 there were 87 members from India out of a total membership of 2,000 plus. Both Mahalanobis and C.R. Rao were elected as presidents of the International Statistical Institute and served to improve the official statistical systems worldwide. In recognition of his pioneering work relating to statistics of developing nations, C.R. Rao was awarded the first P.C. Mahalanobis award in 2003 at the Berlin Session of the International Statistical Institute, for life-time achievement in statistics and the promotion of best statistical practice. For a biography of C.R. Rao, see Krishnankutty (1996).

6. The Present Scenario

Ever since the National Sample Survey was shifted from the Indian Statistical Institute (ISI) to become a part of the Government of India in 1971, the interaction between the academics at ISI and the officers in the government became less and less and the role of the ISI was just to nominate two of its faculty to the Governing Council of the National Sample Survey Organisation apart from the participation of ISI members on various technical advisory, steering and other committees of the government.

To discuss the present scenario in relation to the official statistical system in India, it would be best to refer to the Report of the National Statistical Commission (NSC 2001). In the year 2000, for the first time in India, the Government set up a commission to address appropriately the growing statistical needs of the society. This commission was chaired by Dr. C. Rangarajan. Most defects in the system were summarised in the report of the commission thus:

"Apart from the quality of data, there are other problems such as data gaps, duplication leading to conflicting statistics and inordinate delays in transmission and publication of data. In the field of Agricultural Statistics large data gaps exist with reference to the output of fruits and vegetables and other minor crops, and estimates of meat, meat products and fish. The results of the 16th Livestock Census, scheduled to be completed in 1997, are still not available for a number of States. The representativeness of the Index of Industrial Production has been considered as questionable due to *inter alia* inadequate information on the Small Scale Sector. One of the perennial issues relating to national income in our country has been the difference between the National Accounts and National Sample Survey estimates of consumption expenditure. In the area of External Sector Statistics also, the reconciliation of the data on exports and imports between the Directorate General

of Commercial Intelligence and Statistics and the Reserve Bank of India is essential. Further, the present system has not been able to provide adequate information on basic socio-economic indicators required for micro-level planning. For example, although the Civil Registration System was envisaged as a mechanism to provide annual estimates of the infant mortality rate, death and birth rate, etc. at the district level, it has failed and as a result, such estimates are not available at the decentralized level. No reliable information is available on many aspects in the Health Sector like problems of the aged, contribution of the private sector to health care, disease-specific expenditure on health, etc. The whole area of the Services Sector is undergoing far-reaching changes with the application of Information Technology. More of intangible goods are getting exchanged. However, huge data gaps exist with reference to such transactions."

The recommendation of the NSC regarding the creation of a permanent and statutory apex body – National Commission on Statistics – through an Act of Parliament, independent of the Government, in respect of policy-making, coordination and maintaining quality standards of core statistics has been implemented by a Government Order. The new apex body, still called the National Statistical Commission, was set up on 1 June 2005 and assumed charge on 12 July 2006. There is a new post called Chief Statistician of India, the holder of which also acts as the Secretary of the Ministry of Statistics and Programme Implementation. Consequent to the setting up of the new Commission, the erstwhile Governing Council of the National Sample Survey Organisation has been dissolved with effect from 30 August 2006 as all the functions of the Governing Council are assumed by the National Statistical Commission. However, initially, a Steering Committee has been constituted for National Sample Surveys.

According to the report, Directorates of Economics and Statistics in some states do not play a nodal role in the coordination of statistical activities within the state and lack survey sampling and data processing capabilities. A lot needs to be done with the huge amounts of data collected through state samples matching with the central samples. The Industrial Statistical Act of 1942 is very old and the present Collection of Statistics Act, 1953, is found to be weak.

Economic censuses (ECs) have been the backbone of several enterprise surveys of India. The four ECs conducted so far during 1977, 1980, 1990 and 1998 were not free of deficiencies. The operations of ECs and the designing of the follow-up enterprise surveys should be strengthened. In order to improve the utility and quality of data, new initiatives were introduced in the fifth EC 2005 such as recording the type of structures occupied by the establishments and making a directory of establishments with ten or more workers. Data transcription errors were to be minimised due to the use of ICR technology for the first time in data processing. Data on livestock surveys has been released and the 18th survey has already started.

Thus far, the National Sample Survey Organisation has released a little more than 530 reports in its 60 years of existence, each consisting of 200 pages on average. The reports have been released within one year after completion of data collection. These are available as hardware as well as software copies for the users. Unit level data is also supplied to researchers after a year of completion of the surveys. It is also reported that the reports are to be made available on the website of the Ministry of Statistics and Programme Implementation. However, it is very important that these reports should be carefully

checked for any inconsistencies and errors both in statistics and in style of presentation and should be corrected before release.

One of the fast-growing areas of recent origin is financial statistics. It is relevant here to point out that the Reserve Bank of India plays a very important role in the National Statistical System. Some of the recent surveys undertaken by the RBI are Survey of Small Scale Industries, Survey of Traders and Transport Operators, Survey of Non Profit Institutions Serving Households, *ad hoc* surveys on Remittance of Non Resident Indians, etc.

One of the recent major projects of the Government of India is the National Family Health Survey – 3, which was conducted from December 2005 to August 2006. Like the earlier surveys, this one provided estimators of indicators of population, health and nutrition by background characteristics at the national and state levels. Around 100,000 women and men were tested for HIV and more than 200,000 adults and young children for anaemia in this survey.

The District Level Household and Facility Survey – 3, also initiated by the Ministry of Health and Family Welfare, is another large-scale demographic and health survey carried out in India with a sample of about 700,000 households covering all districts of the country. Information on the National Rural Health Mission, a programme of the Government of India is also collected in this survey.

7. Looking Ahead

Mahalanobis in his lecture on "Statistics as a Key Technology" delivered at the 125th Anniversary Celebration of the American Statistical Association in 1964, made the following observation (Mahalanobis 1965, p. 43):

> "*Indian statisticians have high reputation in the world and yet the statistical system in India is weak and is lagging behind in the timely flow of information required for social and economic affairs. Why?*"

and then tries to diagnose this question. This was almost five decades ago. The situation now is quite different. The statistical system in India is no longer weak, but has become more complicated. With the advent of information technology in India, there is now a timely flow of the information required, as seen above. In the same lecture, Mahalanobis comments thus: "*It is not difficult to see what is wrong with official statistics in India. There is a gap between theory and practice*" (Mahalanobis 1965, p. 43). He quotes his own experience of interviewing an applicant who is a Ph.D. for a high post in sample surveys conducted by the Government of India and who had no occasion to use his results or knew or could think of any situation in which his results could be used in practice.

As the former President of the International Statistical Institute, C.R. Rao (1977) declared: "*It is said that India is a poor country. . . I would like to emphasize that we are poor only in the economic sense of the word and not in our resources of human talent*."

We have mentioned earlier that Mahalanobis's remarks heralded the problems of small domain estimation. Though there are a considerable number of papers published on these aspects, one does not find many practical applications of small domain estimation techniques in the context of Indian official statistics. While analysing the A-B-O blood

group data on Bengalees of the same Bengal Anthropometric Survey, C.R. Rao observed that the design amounted to stratified sampling and not simple random technique, thereby recognising the need for analysis of complex surveys. Not only were many theoretical papers published on this subject, but several practical applications including the development of computer software for analyses of complex official survey data were considered.

The importance of a sound official statistical system in any country is not something for a debate. The whole gamut of economic development and planning undoubtedly depends on that. It would then be interesting to see where one stands and it is suggested (Rao 2003) that a Statistical Development Index for each country along the lines of the Human Development Index could be constructed. Some components needed for such a construction could be statistical literacy in the country, the strength of the statistical system, the amount of dissemination of data, per capita statistical output (defined suitably), etc. Based on such a Statistical Development Index, India would have one of the top rankings, even though its Human Development Index is on the lower side. Also if a Statistical Poverty Line were to be constructed, one may still find several countries below the line, and the international bodies together should strive for the eradication of statistical poverty especially in the Third World.

The current trend is different. Most of the government statisticians now pass the Indian Statistical Service Examinations and are thoroughly trained before they are given a posting. It is true, however, that there is not much of an interaction between the academics and the government statisticians. Statisticians from government departments should be encouraged to visit universities and institutes to refresh their knowledge and learn new and useful methodologies, while teaching faculty should visit government offices to understand the problems of current interest and importance. It is interesting to note that Pitamber Pant, who acted as Secretary to Pundit Jawaharlal Nehru during the preindependence days, was sent by Nehru to the Indian Statistical Institute to learn Statistics in 1946. Pant later became the Head of the Perspective Planning Division of the ISI housed within the Planning Commission and guided the development of various models for long-term planning.

A law to enforce the collection of statistics, which was recently enacted by the Government of India under the "Collection of Statistics Act 2008," would be helpful in certain situations only. Awareness is most important and more attempts should be made to popularise the collection of statistics. All government departments, educational departments and others dealing with statistics should take part in celebrations such as National Statistics Day and World Statistics Day. There should be an official statistics component in the curricula of statistics courses in colleges and universities. The recently inaugurated National Academy of Statistical Administration near Delhi would be helpful in enhancing skill development and capacity building of statistical personnel in the country.

The National Informatics Centre of the Department of Information Technology provides a complete network for e-Governance of the federal and state governments and administrative regions such as districts and other government establishments. This department disseminates information on the e-Governance projects to the public, especially in the fields of education, health, employment, skill-building and entrepreneurship.

The e-Choupal model conceived by the ITC Agri Business Division in India is designed to address the problems faced by the Indian agricultural farmer who is burdened with weak infrastructure, profit-making intermediaries and fragmented pieces of land. Internet kiosks in villages are operated by *Sanchalaks*, a self-help group of farmers, from where a variety of official statistics for decision-making such as market prices, farm practices and risk factors are easily accessed. Thus e-Choupals do away with middle men and help the farmer in meeting the market demand with quality goods at low cost. It is reported that e-Choupal internet services today reach over four million farmers in and around 40,000 villages through 6,500 kiosks in ten Indian states. Certain other business houses have also started providing such advisory services on agricultural inputs and financing. This model could be of use in other areas of official statistics as well. The government has launched the National e-Governance Plan to support the growth of e-Governance within the country. It has recently been reported that 100,000 Common Service Centres across the country are being established at the *Gram Panchayat* (village administration) level with broadband internet. Also State Wide Area Network (SWAN) with minimum 2 Mbps bandwidth capacity per link will be connecting all government offices up to *Block* level (administrative zone above village level).

With a changing economy and rapidly increasing IT avenues, a modern outlook in the statistical system is indeed relevant, and one hopes that the Indian statistical system will be a leader in the international systems. Let us hope that by the time the 15th Conference of the Commonwealth Statisticians is organised in India, the country will have made further and fruitful advances in its attempts to modernise its statistical system.

8. References

Adhikari, B.P. (1990). Social Construction of the Statistical Estimation of the Crop Yield. Paper presented at the XII World Congress of Sociology. International Sociological Association, Madrid, Spain.

Baines, A. (1918). The History of Statistics; Their Development and Progress in Several Countries. In Memoirs to Commemorate the Seventy Fifth Anniversary of the American Statistical Association, collected and edited by John Koren, Book II, 417–426. N.Y: Burt Franklin.

Buchanan, F. (1807). Survey of Eastern India. Report submitted to the Court of Directors, London.

Chaudhuri, S.B. (1964). History of Gazetteers of India. New Delhi Publications Division.

Davis, K. (1951). The Population of India and Pakistan. Princeton: University Press.

Deming, W.E. (1964). On Some Contributions of Interpenetrating Network of Subsamples. In Contributions to Statistics, C.R. Rao (ed.). Calcutta: Statistical Publishing Society, 57–66.

Dynes, J.H. (1899). Development of the Plan for a Census of the World. Journal of the American Statistical Association, 6, 357–371.

Fisher Box, J. (1978). R.A. Fisher, The Life of a Scientist. New York: Wiley.

Ghosh, J.K., Maiti, P., Rao, T.J., and Sinha, B.K. (1999). Evolution of Statistics in India. International Statistical Review, 67, 13–34.

Hubback, J.A. (1927). Sampling for Rice Yields in Bihar and Orissa. Imperial Agricultural Research Institute. Bulletin, 166, Pusa. (Reprinted in Sankhyā (1946) 7, 281–294).

Hunter, W.W. (1868). The Annals of Rural Bengal. London: Smith, Elder and Co.

Hunter, W.W. (1875). A Statistical Account of Bengal. London: Trübner and Co.

Hunter, W.W. (1899). A History of British India. London: Longmans, Green and Co.

Jarrett, H.S. (1894). Translation of Ain-i-Akbari, Vol. II, p. vii. Calcutta: Asiatic Society of Bengal.

Krishnankutty, N. (1996). Putting Chance to Work. A Life in Statistics: A Biography of C.R. Rao. State College, PA: Dialogue.

Lahiri, D.B. (1951). A Method of Sample Selection Providing Unbiased Ratio Estimates. Bulletin of the International Statistical Institute, 33, 133–140.

Mahalanobis, P.C. (1950). Why Statistics? Presidential Address, Indian Science Congress, Thirty seventh Session, Poona. (Reprinted in Sankhyā, 10, 195–228).

Mahalanobis, P.C. (1965). Statistics as a Key Technology. The American Statistician, 19, 43–46.

Mahalanobis, P.C. and Chakravarti, N. (1931). Statistical Report on River Floods in Orissa during the Period 1868–1928. Submitted to Government of Bihar and Orissa.

Mahalanobis, P.C. and Lahiri, D.B. (1961). Analysis of Errors in Censuses and Surveys with Special Reference to Experience in India. Bulletin of the International Statistical Institute, 38, 401–433.

Mahalanobis, P.C. and Nair, K.R. (1940). A Simplified Method of Analysis of Quasi-factorial Experiments in Square Lattice with a Preliminary Note on Joint Analysis of Yield of Paddy and Straw. Indian Journal of Agricultural Science, 10, 663–684.

Mahalanobis, P.C. and Sen, S.B. (1954). On Some Aspects of the Indian National Sample Survey. Bulletin of the International Statistical Institute, 34, 3–14.

Majumdar, D.N. and Rao, C.R. (1958). Bengal Anthropometric Survey, 1945: A Statistical Study. Sankhyā, 19, 201–408.

Martin, M. (1838). History, Antiquities, Topography and Statistics of Eastern India. London: W.H. Allen and Co. (Also available: Cosmo Publications, Delhi, 1976).

Murthy, M.N. (1964). On Mahalanobis' Contributions to the Development of Sample Survey Theory and Methods. In Contributions to Statistics, C.R. Rao (ed.). Calcutta: Statistical Publishing Society, 283–316.

National Sample Survey Organisation Expert Group on Nonsampling Errors (2003). Suitability of Different Reference Periods for Measuring Household Consumption: Results of a Pilot Survey. Economic and Political Weekly, XXXVIII, 4, January 25.

National Statistical Commission (2001). Report Submitted to the Government of India.

Panse, V.G. and Sukhatme, P.V. (1948). Crop Surveys in India-I. Journal of Indian Society of Agricultural Statistics, 1, 34–58.

Panse, V.G. and Sukhatme, P.V. (1951). Crop Surveys in India-II. Journal of Indian Society of Agricultural Statistics, 3, 98–168.

Rao, C.R. (1977). Speech by President-Elect. Bulletin of the International Statistical Institute, 41, Book 1, 51–53.

Rao, T.J. (1999). Early Developments in Sample Surveys: A Tale of Two Personalities. Paper for the 52nd Session of the International Statistical Institute.

Rao, T.J. (2003). Statistics During Early Days in Developing Countries with Special Reference to India. Bulletin of the International Statistical Institute, Berlin.

Rao, V.K.R.V. (1940). Estimation of National Income in British India: 1931–1932. London: MacMillan.

Rudra, A. (1997). Prasanta Chandra Mahalanobis: A Biography. New Delhi: Oxford University Press.

Seng, Y.P. (1951). Historical Survey of the Development of Sampling Theories and Practice. Journal of the Royal Statistical Society, Series A, 114, 214–231.

Seth, G.R. (1996). Validity and Reliability of Statistical Procedures. Journal of Indian Society of Agricultural Statistics, 48, 274–283.

Shamasastry, R. (1929). Translation of Kautilya's Arthasastra, (3rd Edition). Mysore: Wesleyan Mission Press.

Shirras, G.F. (1919). Indian Finance and Banking. London: Macmillan and Co. Ltd.

Singh, V.B. (ed.) (1965). Economic History of India: 1857–1956. New Delhi: Allied Publishers.

Subramaniam, S. (1945). Statistical Summary of the Social and Economic Trends in India (1918–1939). Washington D.C: India Information Services.

Sukhatme, P.V. (1958). The 1960 World Census of Agriculture. Bulletin of International Statistical Institute, 36, 239–250.

Sukhatme, P.V. and Margen, S. (1978). Models of Protein Deficiency. The American Journal of Clinical Nutrition, 31, 1237–1256.

Sukhatme, P.V. and Seth, G.R. (1952). Non Sampling Errors in Surveys. Journal of Indian Society of Agricultural Statistics, 4, 5–41.

Received July 2008
Revised July 2009