Evolution of Statistics in India

J.K. Ghosh, P. Maiti, T.J. Rao, and B.K. Sinha

Indian Statistical Institute, Calcutta, India

Summary

This is a brief history of the evolution of official and academic Statistics in India which focuses mainly on the period 1930 to 1960 but traces its origins in antiquity and recent history. We also comment on how Statistics has continued to evolve since the 1960's. This is a history of both institutions and people, who built and shaped them, and of ideas.

Key words: Statistical System in India; Mahalanobis; Statistics as a key technology; Five Year Plans; Data Requirements; Central Statistical Organisation; National Sample Survey; Statistics in Indian research institutions and universities.

1 Introduction

While statistics have been collected and used in the Indian subcontinent from antiquity, major changes in collection and use took place during the British period (1757–1947) in Indian history. Some of this change was due to new imperial needs, but much of it occurred indirectly as a result of western education and a spirit of scientific curiosity and experimentation. Interest in rapid social, economic and technological development added a new dimension after India's independence in 1947. Half a century after that momentous event seems a good time to take stock of how Statistics has developed in India. The following account is meant to be a brief history rather than a current assessment. To us the most important period after independence is the decade 1950 to 1960 when so many things were happening at the same time. Our account begins in antiquity, focuses on the period 1930 to 1960 and ends with a brief sequel.

The architect of modern statistical methods in the Indian subcontinent was undoubtedly P.C. Mahalanobis, but he was helped by a galaxy of very distinguished scientists that included C.R. Rao, R.C. Bose, S.N. Roy, S.S. Bose, K.R. Nair, D.B. Lahiri and many others. There were also others like P.V. Sukhatme, and V.G. Panse who worked independently of Mahalanobis. Our history is a history of some of these persons as well as a history of institutions and interactions between persons and institutions.

In a concluding section, we try to assess what was unique about the growth of Statistics in India in the earlier part of the century and what may have been the historical reasons for what C.R. Rao has called a golden period for Statistics in India.

2 Historical Background

Early Origins

It is interesting and illuminating to note that statistical knowledge and probabilistic ideas were attributed to the kings and rulers mentioned in the great Indian epic, the Mahabharata, as is evident from the following quote of Ian Hacking from the History and Philosophy Science Seminar (quoted

by Godambe, 1976): "King Bhangasuri wanting to flaunt his skill in numbers, estimates the number of leaves, and the number of fruits, on two great branches of a spreading tree. There are, he avers, 2095 fruits. Nala counts all night and is duly amazed by morning. Bhangasuri accepts his due:

> "I, of dice possess the science, and in numbers thus am skilled."

That the concept of probability was recognized in the Indian-Jaina philosophy is clear from the writings of Bhadrabahu, who lived during the period 433–357 B.C., on syadvada or 'the assertion of possibilities' (syat = 'may be', vada = 'assertion'). Mahalanobis (1954) and Haldane (1957) refer to the actual text in Sanskrit of the dialectic of seven fold predication (Saptabhanginaya) and relate it to the concepts of probability theory with examples of 'tossing of a coin' and 'study of the physiology of the sense organs'.

The great treatise in Economics, the *Arthasastra* by Kautilya (normally attributed to 321–296 B.C.) during the Mauryan period had a detailed description of the system of data collection relating to the agricultural, population and economic censuses in villages and towns during the period . To illustrate, Chapter XXXV (Shamsastry, 1929, p.158) gives details such as:

> "It is the duty of *Gopa*, village accountant, to attend the accounts of five or ten villages, as ordered by the Collector-General . . . Also, having numbered the houses as tax paying or non-tax paying, he shall not only register the total number of inhabitants of all the four castes in each village, but also keep an account of the exact number of cultivators, cowherds, merchants, artisans, labourers, slaves and biped and quadruped animals, fixing at the same time the amount of gold, free labour, toll and fines that can be collected from it (each house)."

As observed by Mahalanobis (1950), not only were data collected in such a fine detail, but the need for cross-checking by an independent set of agents working incognito was mentioned in Chapter XXXV, p.159 thus: "Spies under the 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 . . . ."

It may be mentioned here that the Chinese pilgrim and traveller Hieuen Tsang's writings (dated late seventh to early eighth century) give a detailed description of the plan of cities, construction of houses, and an account of common products of India and data on the area of kingdoms and the distances between them.

Moghul Period

Let us take a leap forward to the Moghul period. An important masterpiece written by Abul Fazal during this period was *Ain-i-Akbari*. Abul Fazal belonged to the court of the great Moghul Emperor Akbar around 1590 A.D.. This had details of several government departments including the system of legalised measurements, land classification and crop yields by season among others. Abul Fazl was "regarded as a statistician, no details from the revenues of a province to the cost of a pine-apple, from the organisation of an army and the grades and duties of nobility to the shape of candlestick and the price of a curry-comb, are beyond his microscopic and patient investigation" (Jarrett, 1894). Revenue guides known as *Dastur-ul-amls*, maintained during Akbar's period, continued to be compiled even during the times of Shah Jahan and Aurangzeb. *Zawabit-i-Alamgiri* or the regulations of the Emperor Aurangzeb which was prepared in 1690 was a good chronicle of statistics. Other works which contained statistical information were due to Jag-Jivan Das, Rai Chatar-mal and Sujan Rai Bhandari (*Khulastu-i-Tawarikh* (1695–96)).

According to the system of land tenure and land revenue during the Moghul period, all land legally

belonged to the emperor and the cultivator was a tenant with full liberty of exploiting his piece of land. A proportion of produce, fixed from time to time, would have to be paid to the state as land revenue. However, with the decline and fall of the Moghul empire shortly after Aurangzeb, many of the officials called Jagirdars, Inamdars etc. appointed by the emperor became independent Nawabs, and Kings with no or nominal allegiance to a central authority.

3   Statistical System in British India

Early British Period

It was during the decline of the Moghul Empire that the British set foot in India as traders, plantation owners, businessmen and the like while the Indian political scenario consisted of a fragmented nature of numerous small and large independent or quasi-independent kingdoms.

British political power was first established by the East-India Company (EIC) in Eastern India. Eventually, all of undivided India except the nominally independent states and a few French and Portuguese settlements came under British rule (1757–1947).

In Eastern India the British introduced a 'permanently settled' system wherein the intermediate tax collectors, called Zamindars were made responsible for the payment of revenue of the large tracts under them to the British treasury. The amount of revenue was made 'permanent', *i.e.*, fixed for perpetuity, to be paid before the 'sunset' of a fixed date. Thus under the 'sunset-law', the system of village revenue officials, *Patwaris*, as functionaries of government ceased to exist. Therefore, in 'permanently settled' areas there was no elaborate official agency for collecting primary statistics. Hence, the need for acquisition of accounts and a detailed knowledge regarding the territories occupied was strongly felt by the East India Company (EIC).

A despatch from the Court of Directors of the EIC in 1807 read thus :

> "We are of the opinion that a statistical survey of the country, under the immediate author-
> ity of your Presidency, would be attended with much utility: We therefore recommend
> proper steps to be taken for carrying the same for execution."

In 1807, a survey of the provinces, subject to the Presidency of Bengal was commenced by the Governor-General in Council, Dr. Francis Buchanan covering an area of 60,000 square miles and about 15 million British subjects (Buchanan, 1807). Dr. Buchanan spent around £30,000 and submitted a report to London in 1816. This report contained a detailed information on topography of each district, the condition of the inhabitants along with their religions, customs, the natural produce of the country, fisheries, mines and quarries, the agricultural situation, the state of the landed property and tenures, the progress made by Indians in arts and the state of manufacturers, the operation of commerce and, in addition, an indication of rare, useful and curious plants and seeds. After a relatively long gap, in 1838 Mr. Montgomery Martin was sent to India to study the area surveyed in Buchanan's report. Impressed by the 'critical attitude, keen scientific spirit, and the experimental approach' of Dr. Buchanan, Martin (1838) published 'The History, Antiquities and Statistics of Eastern India' in 3 volumes consisting of 2400 pages covering 9 districts of Bengal. The main objective of Martin was to bring to light the fantastic work done by Buchanan and 'to arouse in some measure the people of England to some sense of feeling for the condition of the myriads of their fellow subjects . . . . . . in British India which is as much a part and parcel of the Empire as Scotland or Ireland'. The main recommendations that stemmed out were to fix a moderate land revenue rate, to levy duties on equivalent produce as per free trade, to encourage a sound and judicious banking system and finally to establish municipalities in principal cities.

A government officer named A. Shakespeare published in 1848 the first census relating to the area and revenue of each *pargana* (district) in North-West (N.W.) Provinces. A small department of statistics was started in the India House in 1847 by Col. Sykes. In 1853, the department released the

first series of statistical papers on India. Census reports of 1 January, 1855 and 10 January, 1868 were published. Earlier censuses of Calcutta taken in the year 1822 and in 1847 by W.H. Carey gave a total count of 179,917 and 4,00000 respectively! (see Chaudhuri (1964)).

The censuses taken during 1769–1855 by the EIC or those taken by the Crown during 1858–1869 were fragmentary, hardly systematic and lacked any uniformity. The first systematic attempt to ascertain the whole population of India by 'actually counting heads' was made between 1867 and 1872. It was not a synchronous census for the whole country, nor was it complete. The operation of a decennial census for the whole country started in 1881 and is continuing ever since. The report on the Census of British India taken in 1881 was published in three volumes.

Kingsley Davis (1951) remarks that 'the Indian censuses are remarkable not only for the information they reveal but for the special obstacles they had to overcome. ... they (the census officials in India) have enormously enriched our knowledge of India in nearly every branch of scholarship, from anthropology and sociology to geography and religion'.

Even though W. Hamilton published the first gazetter in 1815, he revised it and published East India Gazetter in 1828. However, Thornton's gazetter in four volumes published in 1854 was considered to be complete at that time.

Impressed by the trend in statistical activities, the Secretary of state ordered the Governor-General in Council to prepare a 'comprehensive and coordinated scheme of statistical survey' for each of the twelve great provinces of the then British India and Dr. W.W. Hunter was appointed as Director-General of Statistics in India in 1869 to carry out this work.

In 1870, Hunter gave a plan for an Imperial Gazetter of India. The local governments had planned differently for conducting this work with bigger budgets and manpower while several public bodies such as the Asiatic Society insisted on a systematic and coordinated effort so that the work was executed under a uniform plan. It was agreed upon to provide a data base collected by each local government as a common basis for comparison of statistics of the country and to suggest quick compilation methods from the data collected on a uniform plan. Thus the Statistical Account of Bengal (the present Bangladesh, West Bengal, Bihar and Orissa) was published in 20 volumes under Hunter's supervision.

For each district there were details on topographical data, ethnic divisions and creeds, agricultural situation, commerce, working of district administration and finally the sanitary and medical aspects and such meteorological data as could be procured. Statistical accounts for the provinces of Assam, N.W. Provinces, Punjab and others followed. Thus about 100 printed volumes aggregating to 36,000 pages covering 240 districts comprising 15 British Indian provinces were published and later condensed in the Imperial Gazetter of India which was released in 1881 in nine volumes.

S.B. Chaudhuri (1964) in his comprehensive work 'History of Gazetters of India' comments thus:

> "No comparable area of the world has anything like this 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."

The need for timely and accurate collection of agricultural data was felt by the Indian Famine Commission and agricultural departments were organised in various provinces which resulted in the publication of 'Agricultural Statistics of British India' in 1886. To scrutinize and summarize these data collected by the agricultural departments, a statistical Bureau was formed at the centre in 1895 to coordinate the agricultural, foreign trade, prices, wages and industrial statistics. The Director General (DG) of Statistics was in charge of this operation.

Later British Period

During the turn of the century in 1905, Lord Curzon abolished the post of DG of Statistics, reorganized the department by separating out the statistical data collection jobs and constituting the Directorate General of Commercial Intelligence and Statistics (DGCI & S). Its main functions were to collect commercial statistics to help trade and business, act as a liaison between Indian businessmen and their foreign counterparts and to publish journals and adhoc bulletins on trade statistics. The year 1906 saw the first issue of Indian Trade Journal. An important contribution to price statistics was a survey conducted in 1910 by Datta, Shirras & Gupta (1913). The book on Indian Finance and Banking by Shirras, who was the Director of Statistics with the government as well as a Fellow of the University of Calcutta contains very interesting data on exports and imports, balance of trade, growth of business, production of gold, silver, paper currency and details on banks for the period ranging in several cases from mid 1850's to 1918 (Shirras, 1919). The Economic Enquiry Committee set up in 1925 under the Chairmanship of Visweswarayya and more importantly the Bowley–Robertson Committee set up later in 1934, were mainly responsible for the government's decision to set up an Inter-Departmental Committee with the Economic Adviser to the Government of India as the chairman. The Inter-Departmental Committee recommended the formation of a Central Statistical Office for coordination, institution of a statistical cadre, establishment of State Bureaus at State Head Quarters and maintenance of important statistics for the entire country (Statistical System in India, CSO (1979)).

4 The Statistical System after Independence

As seen in the foregoing paragraphs, throughout the British period the statistical development was geared towards administration, trade, commerce and such other activities. It is only after the independence in 1947 that the country saw an urgent need for a statistical framework suitable for economic and social development. Mahalanobis was appointed as a Honorary Statistical Adviser to the Indian Cabinet in 1949 and a Central Statistical Unit was set up in the Cabinet Secretariat in 1949 under his technical guidance. A couple of years later the Central Statistical Organisation (CSO) was formed in 1951 to coordinate the statistical activities in independent India. The National Sample Survey (NSS) was created in 1950 as a multi-faceted fact-finding body. During 1961, the CSO and NSS were put under a full-fledged Department of Statistics.

Central Statistical Organisation

The Central Statistical Organisation (CSO) was set up mainly to coordinate the statistical work done in various ministries and other government agencies and to advise them, to maintain standards with regard to definitions, concepts and procedures, to provide consultancy, keep in touch with international statistical organisations, to prepare and publish an Annual Statistical Abstract and Monthly Statistical Abstract, to act as a liaison with United Nations Statistical Office and to disseminate annual statistics by graphs and charts as well as tables for public use.

The National Income Committee recommended in 1954 that the National Income Unit be transferred from the Ministry of Finance to the CSO and since then the estimation of national income has become an important activity. Similarly, a planning cell was organised at the CSO to look into the plan activities of the government. A population unit was added in 1956 mainly to examine the schemes and enquiries relating to population that are often referred to the CSO for technical advice from time to time and preparation of briefs and memoranda on censuses and vital statistics. The CSO also took part in training the central and state statistical officers to improve official statistics. In 1957, the Directorate of Industrial Statistics was transferred from the Ministry of Commerce and Industry. In order to have an integrated approach between the planning, statistical and survey organisations

the Department of Statistics, including the CSO and NSSO, has been transferred to the Ministry of Planning and Programme Implementation from February, 1973. B. Ramamurti, S. Subramanian, P.C. Mathew, and K.R. Nair headed the CSO in the early years.

National Sample Survey

The Standing Committee of the Departmental Statisticians as well as the National Income Committee (NIC), established in 1949, felt that there was an urgent need for improving the quality of statistical information. In 1950, the NIC recommended the use of sampling methods to fill the gaps in the estimation of national income.

The Indian Statistical Institute (ISI) which had acquired experience and expertise in large scale surveys since 1935 was approached by the Government of India to prepare the design with detailed plans and estimates for a comprehensive socio-economic national sample survey covering rural areas of India. The first round of data collection started in October 1950 and was completed by March 1951 with a sanctioned strength of 607 personnel. The Directorate of the National Sample Survey was transferred from the Ministry of Finance to the Cabinet Secretariat in 1957. As mentioned earlier, it has been in the Department of Statistics since 1961. While the field work was done by the Directorate, technical design including drawing up of questionnaires/schedules, instructions to field workers, details of scrutiny, data processing and tabulation were entrusted to the ISI under the general direction of Mahalanobis. In January 1971, the design and analysis wing was reorganised by shifting it from the Indian Statistical Institute to the Department of Statistics and forming the National Sample Survey Organisation which also included the Field Operations and Data Processing Divisions. The NSS is the largest multi-purpose socio-economic survey in the world. The activities were coordinated by four main divisions, viz., Survey Design and Research, Field Operations, Data Processing and Economic Analysis. Currently, besides socio-economic surveys, data is also collected through Annual Survey of Industries, Crop Yield and Area Estimation Survey, Urban Frame Survey and Prices Survey. J.M. Sengupta, D.B. Lahiri, S. Raja Rao, M.N. Murthy were associated with NSS since the early rounds.

According to D.B. Lahiri who is one of the chief architects of the NSS right from the beginning "the NSS chose a multi-purpose and multi-subject frame work, and the survey design was gradually so evolved as to permit study of the inter-connections between the various components of the socio-economic picture of the country and its constituent regions and states. The NSS has in the main been a population survey in a comprehensive sense, although there has been a sizeable effort on the estimation by an area survey of crop acreage and production to which Mahalanobis attached great importance because of chronic food shortage". Commenting on the complexity and scope of Mahalanobis's plans for the National Sample Surveys of India, Deming (1973) remarked: 'No country, developed, under-developed or over-developed, has such a wealth of information about its people as India . . . . '.

Other Statistical Divisions and Activities in the Government

Among the other important statistical wings in the Government of India, the office of the Director General of Commercial Intelligence and Statistics is one of the oldest establishments. It continues to be responsible for commercial intelligence and foreign trade statistics. The office of the Registrar General which was created in 1948 carries out its decennial Census Operations as well as Sample Registration System and publication of other demographic and vital statistics. During the 1951 and 1961 Censuses, several major changes were adopted either in collection or analysis of data under the leadership of the Registrars General, R.A. Gopalaswami and Ashok Mitra. For providing reliable vital rates to meet the needs of planning and policy decision, the Office of the Registrar General,

initiated the Sample Registration System to cover the rural and urban areas of the country since 1970. This scheme has been very successful.

The Labour Bureau set up in 1946 is responsible for collection and dissemination of labour statistics and publication of consumer price indices. Apart from these departments at the centre and host of others in various central ministries, the State Statistical Bureaus (SSB), which play the same role as the CSO at a state level, also collaborate with the NSSO in conducting multi-purpose surveys.

The need for forming a statistical cadre was recognized by the Inter-Departmental Committee on Official Statistics and in 1964 an Indian Statistical Service was organised to cater to the needs of the Central Ministries. Recently, the Department of Statistics, which covers new activities such as environmental statistics, service statistics, and gender statistics, has undertaken a programme for modernization and complete computerization of the existing data processing systems.

Perspective Planning Division

At the request of the Government of India, a draft of the second Five Year Plan was prepared at the Indian Statistical Institute in 1954 by Mahalanobis. Following a forward looking Harrod–Domar type of model, he modelled the net output of the economy as originating in two sectors, (Mahalanobis (1953, 55)), the investment goods producing sector and the consumer goods producing sector. His model is based on a growth curve for the economy given by

$$Y_t = Y_0[1 + \alpha_0\{(1 + \lambda_i\beta_i)^t - 1\}\{\lambda_i\beta_i + \lambda_c\beta_c\}/\lambda_i\beta_i]$$

where Y_t is the National Income in year t, α_0 is the initial rate of investment in the base year 0, λ_i and λ_c are respectively the shares of investment towards the investment goods industries sector and consumer goods industries sector while β_i and β_c are respectively the ratios of increment of income to investment for the two sectors. Over a longer period, a larger λ_i gives a higher rate of growth. In the Second Plan, λ_i was taken as around 0.3 giving a limiting investment rate of about 18 percent. The model thus led to a priority to the development of investment goods industries over consumer goods industries in order to have a high rate of growth of consumption in the long run. Later on, he prepared a four sector model in which the consumer sector is further divided into factory type, hand type inclusive of agriculture, and services of all sorts. The draft Second Five Year Plan was based on this model. It emphasized a rapid industrial development as the right strategy for economic growth.

As pointed out by C.R. Rao, "Mahalanobis never claimed that his model was a contribution to economic theory: he meant it as 'a conceptual framework which would be of help for practical purposes', and in revealing the broad characteristics of the system under consideration without getting lost in details". The conceptual framework proved to be useful for subsequent plans also.

In 1955, Mahalanobis was appointed as a member of the Planning Commission. In order to carry out further studies on planning, he envisaged the need of a Perspective Planning Division (PPD) which was established next door to the Planning Commission in Delhi. Pitambar Pant who was secretary to Jawaharlal Nehru during the pre-Independence days, was chosen to look after the PPD. Pant helped the Institute with respect to all negotiations with the government in addition to his responsibilities at the PPD. At the PPD, various models were developed for long-term planning under the guidance of Pant. A paper thus prepared entitled "Perspective of Development: 1961–76, Implications of planning for a minimum level of living" was circulated in 1962 and subsequently published in *Sankhya*, 1974. The depth and output of the Division and its role in the policies of Planning Commission during the early sixties could be considered as the 'high water mark of Pitambar's career' (Minhas *et al.*, 1974).

One of the interesting aspects of the preparation of the draft of the Second Five Year plan was the massive use of data collected by CSO and NSS. Records of this survive in the large number of working papers prepared during this period.

In "Studies Relating to Planning for National Development" working paper No. 1/P.U.1.1, released

in 1954, Mahalanobis writes as follows:

> "... Intensive studies have been started on the basis of the data collected by National
> Sample Survey to find out how the consumption of particular commodities or services
> actually changes with increasing levels of per capita expenditure ... "

In the same paper, Mahalanobis relates this to production as follows: "One industry would sell its products to various other industries. Also it would get its needs from other sources .... The whole industrial structure is closely interlocked and in order to conceive of a change in the level of production of one commodity, it is necessary to give consideration to the change in output of many other industries. When an approximate allocation of investment is ready, the anticipated consumer expenditure is known, and the requirement of final flows of consumer goods have been settled, it would be necessary to work out the total output of the different industries".

"This can be done with the help of inter industry relations (sometimes called input-output tables). Work is already in progress in 12 sectors (i.e. a 12 × 12 table) and arrangements are being made to prepare a 90 × 90 table."

One of the key tools in this analysis were to be input-output tables which are updated or constructed afresh and utilised for national accounts. Mahalanobis's remarks remain relevant:

"In order to consider the detailed breakdowns of production of commodities and the supply of services, the economic and technological relations between investment, income and employment in different industries would have to be used for which work on a small scale has already been started."

One of the early sources of input-output tables is a 12 × 12 table (Studies relating to Planning for National Development: working paper No. 1/P.U.1.1, November 19, 1954, Indian Statistical Institute).

Over the years, ISI has made many more detailed and sophisticated studies of consumption and income elasticity but the link with planning and policy is less clear.

5   P.C. Mahalanobis and the Indian Statistical Institute

The Early Period (1915–1931)

Mahalanobis, whose name has cropped up several times in this article, was born in a well-to-do progressive *Brahmo* family in 1893. *Brahmos* were an enlightened, reformist group who preached for monotheism and against castes and various superstitious rituals within Hinduism. Mahalanobis went to study in Cambridge in 1913 and in 1915 finished his Tripos in Natural Science with a first class. His first encounter with Statistics, which took place at this time, is described as follows by C.R. Rao (1973): 'At the time of Mahalanobis's departure to India from Cambridge, the first world war was on and there was a short delay in his journey. Mahalanobis utilized this time browsing in the King's College library. One morning, Macaulay, the tutor, drew his attention to some bound volumes of Biometrika ... Mahalanobis got so interested that he bought a complete set of Biometrika volumes ... he started reading the volumes on the boat during his journey and continued to study and work out exercises on his own during spare time after arrival in Calcutta.' He tried to look for problems where he could apply the new knowledge he was acquiring. In these pursuits, Acharya Brajendra Nath Seal had a great influence on Mahalanobis. He had been one of the first men to appreciate the significance of the new discipline. The first important work in Statistics in the modern sense to be undertaken in India was possibly the statistical analysis of examination results in Calcutta University. Seal, in 1917, as Chairman of the Committee for examination reforms in Calcutta University sought Mahalanobis's help in the above analysis. Another person who provided support to Mahalanobis in his choice of a new, untrodden path was the poet Tagore, who even wrote a poem for one of the early issues of *Sankhya*, the Indian Journal of Statistics (vol.2 (1934) p. 1).

During the session of the Indian Science Congress at Nagpur in 1920, Mahalanobis had a meeting

with Annandale, the Director of Zoological Survey of India who had collected data on Anglo-Indians in Calcutta and got interested in statistical analysis of Annandale's data. This resulted in the first paper (Mahalanobis, 1922) relating to the statistical analysis of Anglo-Indian stature. He continued to work on the anthropological data (Mahalanobis, 1925, 1930, 1931, 1936) and built up new methodologies for classifying or distinguishing populations characterized by such measurements. The famous Mahalonobis D^2 emerged in course of this work.

If Σ denotes the common dispersion matrix of the measurements, then the Mahalanobis measure of distance between two populations is given by

$$D^2 = (\mu_1 - \mu_2)'\Sigma^{-1}(\mu_1 - \mu_2)$$

where μ_i is the mean vector for the ith population, $i = 1, 2$ and Σ^{-1} is the inverse of Σ. Mahalanobis (1930) considers only the case of a diagonal Σ. In Mahalanobis (1936), he considers also the correlated case, i.e. general Σ and also introduced the familiar studentized version where the parameters μ and Σ are replaced by their estimates. He also introduced what amounts to graphical cluster analysis.

There is evidence (Pearson, 1928, quoted in Rudra, 1996) that Pearson had expressed reservations about the D^2-statistic and did not want to publish Mahalanobis's paper on the D^2-statistic which had been originally submitted to *Biometrika*. Mahalanobis was disappointed but did not give up (Mahalanobis, 1929, quoted in Rudra, 1996). He published his paper elsewhere (Mahalanobis, 1930). Mahalanobis's confidence in this work has been amply justified in subsequent theory and applications. One of the first major theoretical contributions of the Indian School was the proof that the studentized D^2-statistic has a non-central F distribution (Bose & Roy, 1938). The D^2-statistic remains a powerful and fundamental tool in multivariate analysis, classification problems and cluster analysis.

Some of the conclusions of Mahalanobis in his anthropological papers have also stood the test of time. He was right in claiming that the Bengali Brahmins resemble other Bengali castes far more closely than they resemble Brahmins elsewhere in India. However, in some other cases, later evidence points in a different direction. For example, as far as the Anglo-Indian Community is concerned, it is now believed that Mahalanobis had probably confined his study to a sample from upper stratum of the community and hence his conclusion of resemblance to upper caste Hindus is applicable to the upper class Anglo-Indians only.

With this background in analysis of anthropological data, it is not surprising that Mahalanobis (1933) turned his attention to Risley's (1891) famous data set which was collected on 5784 individuals belonging to 87 castes and tribes of Northern India summarized by 11 means and 8 indices. In a total of 20797 values, he found 142 serious discrepancies of which 133 were corrected by 'cross-examination of data' and 'internal consistency checks'. Also the importance of standardization in measurements was stressed by him (Mahalanobis, 1928).

During the early twenties, an officer of the Indian Civil Service, J.A. Hubback observed that the crop cutting system was quite inadequate and defective and conducted extensive crop cutting experiments on the paddy crop, which he called a 'random sampling method', by demarcating areas by a specially devised detachable triangular metal frame as distinct from the traditional tapes to demarcate the rectangular areas. This first experiment in 1923 related to the Godda Thana of Santal Parganas of Bihar State where 400 samples were harvested over an area of 100 square miles. Later, in 1925 he extended his method to 8 subdivisions each of about 1000 square miles in Santal Parganas District and also in the State of Orissa. Mahalanobis got interested in the use of random sample cuts for the estimation of areas under crop and crop yield. Hubback's work (Hubback, 1925) seems to have had influence on Fisher, Yates, Cochran and others of that time in Britain.

On the suggestion of Sir Gilbert Walker, the Director General of Observations, Mahalanobis (1923) looked at the correlations between upper air variables. He got appointed as a Meteorologist in

Calcutta, besides his usual duties as Professor of Physics at the Presidency College. Thus during the late twenties, Mahalanobis got involved in various directions of the growth of the new discipline of Statistics—the conceptual developments in multivariate anthropometric data analysis, the acceptance of sample surveys as a method of data collection as stressed by Kiaer (1895, 1897), Bowley (1906), Jessen (1926) among others, meteorological studies, crop cutting experiments to name a few. It is only natural that the 'Statistical Laboratory', which was being run as a 'workshop' at the Presidency College, Calcutta, should be given a status of an institution of research and higher learning.

Two young colleagues, Subhendu Sekhar Bose and Harish Chandra Sinha urged Mahalanobis to approach some important persons of the country to start a Statistical Society. On the fourteenth of December 1931, Professor Pramatha Nath Banerjea, Nikhil Ranjan Sen and P.C. Mahalanobis issued the following notice: (cf. *Sankhya*, 1, p.124, Annual Report)

> "A meeting will be held to consider steps to be taken towards the establishment of an Indian Statistical Institute on Thursday, the Seventeenth instant at 2.30 p.m. in the Board Room of M/s. Martin & Company, 12, Mission Row, Calcutta. Sir R.N. Mookerjee, K.C.I.E, K.C.V.O will preside."

From the minutes of the meeting, we note that 'it was unanimously resolved that an Indian Statistical Institute be started and that Sir R.N. Mookerjee be requested to accept the office of the President of the Institute . . . '.

The Indian Statistical Institute was founded as a society on 17 December, 1931. *Sankhya*, the Indian Journal of Statistics, was founded two years later.

The Second Period (1931–1950)

The second period is marked by the emergence of sample surveys, multivariate analysis and design of experiments as major statistical tools for practical work. They were also subjects for research at cutting edge. Another notable feature of the period was the introduction of undergraduate and postgraduate courses in Statistics. Last but not least, training programmes and practice of Shewart's Statistical Quality Control were introduced during this period.

Starting with exploratory surveys confined to a few square miles in Bengal in 1937, Mahalanobis was perhaps the first person to organise and carry out an objectively defined large scale survey covering the whole of Bengal (about 59000 square miles) in 1941. This scheme was designed to estimate the yield of jute crop and acreage under jute in Bengal. The survey in its earlier years was able to provide important information on the variability of the characteristics under study and costs of different survey operations. He recognised the need for assessing and controlling the non-sampling errors.

In his report to the Indian Central Jute Committee after a careful assessment of Mahalanobis's method H. Hotelling states:

> ". . . no technique of random sample has, so far as I can find, been developed in the United States or elsewhere, which can compare in accuracy or in economy with that described by Professor Mahalanobis . . . ."

Also, Fisher commented:

> ". . . The ISI has taken the lead in the original development of the technique of Sample Surveys, the most potential fact finding process available to the administration."

Mahalanobis's (1946) Sample Survey of Jute production in Bengal gave a figure of 7540 bales (1 bale = 400 lbs.) while the plot to plot enumeration by the Government which was ten times more costly and had a fifty fold manpower compared to the sampling method gave an underestimate of

6304 bales. The customs and trade figure which independently takes into account nearly all of the produce gave an answer of 7562 bales.

D.B. Lahiri collaborated with Mahalanobis on the analysis of errors in Censuses and Surveys in the Indian context (Mahalanobis & Lahiri, 1961). According to Lahiri (1973) the three notable contributions to sample survey techniques by Mahalanobis are "pilot surveys, concept of optimum survey design, and interpenetrating network of subsamples (IPNS)". All the three concepts are forerunners of important practical statistical contributions that emerged later—'pilot surveys' as a prelude to Wald's 'sequential analysis', 'optimum survey design' stressing the philosophy that all the resources provided for a survey should be used optimally as a precursor to 'operations research' and 'IPNS technique' as one of the curtain raisers for 'resampling procedures' like Bootstrap. Edward Deming (1964) acknowledged thus . . . ". . . for 14 years I have used only interpenetrating network of samples (IPNS), initiated by him (Mahalanobis), as everyone knows, about 1936 . . ." ". . . 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". Mahalanobis, (Mahalanobis, 1938) was aware of the probability proportional to size (pps) selection (Hansen & Hurwitz (1943)) even in 1937. He realised that, in agricultural surveys, it would be necessary to select plots using the cumulative totals of their areas, since these areas vary considerably. However, he assumed that excessive work load would make the selection impracticable and taking into account the high costs of travel between plots which are widely scattered, he recommended the use of 'grid sampling'. On a different level, he had considered the possibility of air surveys 'using specially sensitised films' for estimation of crop acreage in 1937 itself—a technique which has now become popular as 'Remote Sensing'.

Mahalanobis's work on D^2 and studentized D^2 led to very innovative use of matrix and n-dimensional geometric methods for derivation of the distribution of these statistics under a multi-variate Normal model. The identification of the studentized D^2 as a non-central F by R.C. Bose and S.N. Roy was the first major breakthrough in theoretical statistics in the thirties by the Indian school.

Mahalanobis's work with S.S. Bose on crop cutting experiments and yield estimates paved way for fundamental discoveries in construction of design of experiments by R.C. Bose using finite geometries and Galois fields. R.C. Bose derived new methods of construction for balanced incomplete block designs, orthogonal latin squares, confounded factorial designs and much else. For many years, India remained a leading contributor in this area. Others who made important contributions were K.R. Nair (Mahalanobis & Nair, 1940; Nair, 1992) and C.R. Rao. It was C.R. Rao who introduced the notion of orthogonal arrays (Rao, 1947) which in the hands of Taguchi had a profound effect on industrial experimentation.

At Mahalanobis's initiative, the first post graduate course in Statistics was introduced at Calcutta University in 1941. The first batch of students included C.R. Rao, who influenced the growth of the ISI more than anyone else except Mahalanobis and, along with Mahalanobis, is the most famous statistician to come out of the Indian subcontinent. This is how he (Rao, 1992) recalls his early years at the ISI:

> "I passed the M.A. degree examination with a first class, securing the first rank and a high percentage of marks. I was thus among the first five to receive the M.A. degree in Statistics from any Indian University. The Professor offered jobs to all of us in the ISI as technical apprentices on a salary of Rs.75 a month. I joined the ISI in December 1943." (At current exchange rates Rs. 75 amounts to about two U.S. dollars).

Prior to this, C.R. Rao had an M.A. degree in Mathematics from Andhra University in 1940. He received his Ph.D. from Cambridge in 1948 under the guidance of R.A. Fisher and Sc.D. from

the same University in 1965. When Rao returned from Cambridge in 1948 he took charge of the Research and Training School (RTS) which was one of the Divisions of the Institute with the sole responsibility of organising the research and training activities. D. Basu was among the first batch of scholars, who joined in September, 1950. Gradually with fresh admissions of trainees every year, the RTS expanded its activities further. C.R. Rao held the post of Professor and Head of the Division of the Theoretical Research and Training in the ISI from 1949 to 1963. In 1963 he became the Director of the RTS. After Mahalanobis's death in 1972, Rao became the Secretary and Director of the ISI, the designations which Mahalanobis had.

During the mid-forties Mahalanobis foresaw the need for introducing Quality Control (QC) in Indian industries and later C.R. Rao also had been associated with the QC movement in India. In 1945–46 a special course on QC was organised by the ISI which was attended by 12 persons. There was encouragement from a few men like C. Tattersall of the Ordnance Testing Laboratory who fully realized the importance of using QC in industry. But government departments were apathetic. Influenced by the pioneering work done by Walter Shewart in Statistical Quality Control, Mahalanobis invited him to India. Shewart arrived in Calcutta on December 22, 1947 and took the lead in organizing a one-week conference on 'Standardization in Industrial Statistics' in Calcutta from 8th to 14th February 1948 under the auspices of ISI and Indian Standards Institution. This was attended by 190 persons. All these efforts finally culminated in starting the first SQC Unit at Bombay in 1953 followed by two units at Bangalore and Calcutta in 1954. The main objective of these units initially was promotional—to visit industries and act as consultants. This service was later on extended to many principal industrial cities in the country. Today the Division, known as the SQC and OR Division, is engaged in both teaching and research besides consultation and promotional activities.

After devastating floods in the Brahmani river in the state of Orissa in 1926, an expert committee of engineers attributed this to the rising river bed and recommended a corresponding rise in the height of embankments. When the problem was referred to Mahalanobis, he (Mahalanobis, 1931, Mahalanobis & Chakravarti, 1931) studied the data on rainfall in the catchment areas of the river during the period 1868–1928 and related this to the level of the rivers. Contrary to the engineers' suggestion, he recommended construction of dams in the upper reaches of the river to stop the excessive rainfall from flooding the plains. Based on his calculations for a multipurpose scheme of flood control, hydroelectric power generation and irrigation facilities, the Hirakud dam was constructed in 1957. The Chief Minister of Orissa wrote a letter thanking Mahalanobis for his efforts. This work may be regarded as one of the earliest case studies in Systems Analysis and Operations Research, subjects which flourished after the second world war.

Fifties and Early Sixties

The substantial contributions of the Institute to theoretical and applied work, its training and promotional activities culminated in recognition by the Government of India. The parliament passed the Indian Statistical Institute Act, 1959 which declared the Institute as an "Institution of National Importance" and empowered it to award degrees and diplomas in Statistics. The Prime Minister of India, Pandit Jawaharlal Nehru who piloted this bill in Parliament, made a speech that remains relevant as one of the most inspiring defences of science and academic freedom ever made by the head of a Government:

> ... "Now we want science to grow , and I think it is quite essential that we should accept this broad approach to this question that scientific work should have a certain latitude. Therefore, we have decided that in this particular matter, this should continue to be an autonomous organization . . . ".

Soon after, in June 1960, the Institute introduced the Bachelor of Statistics (B.Stat.), Master of Statistics (M.Stat.) and Ph.D. degree courses. In view of its current expertise in the related areas of Quantitative Economics, Mathematics and Computer Science, the Act has been amended by the Parliament in September 1995 permitting the Institute to give degrees in these related disciplines as well.

Fisher's view that "teaching, instruction or training in Statistics, at whatever level, is bound to be, on the one side with fact finding projects in the traditional statistical fields of demography and economics, and on the other side with opportunities to gain first hand familiarity with at least some field in natural sciences . . . " was also shared by Mahalanobis. The syllabus for the B.Stat. degree thus includes an acquaintance with biological, physical and geological sciences. With Professor J.B.S. Haldane as a regular staff member, the Biometric unit expanded further. This unit had been set up earlier under the leadership of Masuyama before his departure in August, 1954. The Institute now has a Biological Sciences Division consisting of 37 scientific workers.

As early as in 1950, the Computing Machines and Electronics Laboratory (CMEL) was started in the ISI. In March, 1956, HEC-2M was installed while a Russian Computer URAL was gifted in 1959. Later on the Institute possessed IBM 1401 and Honeywell systems. A joint project of the Indian Statistical Institute and the Jadavpur University, Calcutta resulted in the first Indian-made solid state general purpose digital electronic computer ISIJU-1 in April 1966. Further improvements had to be abandoned on the advice of the Government. However, the Institute remains a leading centre of research in Image Processing, Pattern Recognition and various other aspects of theoretical and applied Computer Science.

Mahalanobis saw the need for training statistical officers from the Middle East, South and South East Asia, the Far East and from the Commonwealth Countries of Africa. Under the auspices of UNESCO and the Government of India, the International Statistical Education Centre (ISEC) was opened in 1950 and is jointly operated by the International Statistical Institute and the ISI, Calcutta. Since its inception the centre has provided training to 1239 trainees from over 50 countries.

After Mahalanobis's death, C.R. Rao continued as Secretary and Director of the ISI from 1972 to 1976. In 1976, the Institute got a distinguished probabilist as its new Director—Gopinath Kallianpur. Kallianpur was a Professor at the ISI in the fifties, but had left permanently to work in the U.S.. After thirteen years at the University of Minnesota, Kallianpur returned to the ISI as the Director in 1976, a post he held until 1979.

The ISI also had a new Constitution in 1976. The new Constitution puts more stress on teaching and research and less on societal activities. The post of Secretary was abolished, but the Director was to function in future also as the Secretary of ISI. The ISI was to remain both as an Institution of National Importance governed by the Act of Parliament and a Society governed by the Registration of Societies Act. Currently, the Institute has its headquarters in Calcutta and two other centres at Delhi and Bangalore. It also has a network of service units of the SQC and OR Division at Vadodhara, Mumbai, Thiruvananthapuram, Pune, Coimbatore, Chennai, Hyderabad, Calcutta, Delhi and Bangalore.

Arguably the golden period of the ISI was the fifties (cf. Rao, 1973). In addition to Mahalanobis and Rao, the faculty included R.R. Bahadur, D. Basu, G. Kallianpur, D.B. Lahiri, M. Mukherjee, R. Mukherjee and many other distinguished luminaries. Bahadur was a professor at the ISI from 1956–1961, when he returned to the University of Chicago. D. Basu left Dhaka (now in Bangladesh) after independence to join the ISI as a student and later became a Professor as well as the first Dean of Studies. Between them Rao, Bahadur, Basu and Kallianpur and a new group of brilliant students including K.R. Parthasarathy, R. Ranga Rao, V.S. Varadarajan, S.R.S. Varadhan made fundamental contributions to probability and classical inference during this period which were as important as the earlier contributions of Bose and Roy to design of experiments and multivariate analysis. Among the other earlier students of the Institute who achieved international reputation are G.P. Patil, T.N.

Srinivasan, R.G. Laha, J. Roy, Sujit Kumar Mitra, D.K. Roy Choudhury, I.M. Chakraborty. No treatment of such topics as linear models, estimation, maximum likelihood estimates, complete sufficient statistics, conditioning and ancillarity, probabilities on locally compact commutative groups, weak convergence and Edgeworth expansions, can be complete without an appropriate coverage of work done at the ISI during those years. The Cramér–Rao lower bound, Rao–Blackwell Theorem and Basu's theorem on independence of an ancillary and a complete sufficient statistic have been part of any undergraduate or graduate course in Theoretical Statistics. The Cramér–Rao lower bound has sophisticated recent applications to bounding rates of convergence of Bayes risk and density estimates. It was a sign of intellectual vitality of those times that Lahiri was both a distinguished number theorist and a sampling expert *par excellence*. M. Mukherjee was India's leading expert in national income and R. Mukherjee one of India's foremost quantitative anthropologist and sociologist.

During the fifties Mahalanobis himself had turned to planning but found time to introduce a new statistical tool called Fractile Graphical Analysis. He also worked on demographic problems. The Second Five Year Plan drafted by Mahalanobis with the help of Indian and Foreign collaborators was to remain a model for Indian planners for many years. While planning in its old somewhat rigid form is no longer in fashion, it cannot be denied that the policies advocated by Mahalanobis helped in India's rapid post independence industrialisation. Most experts agree that this is an essential step for economic development.

It was during the fifties that the ISI attracted a host of famous visitors. Among them was Norbert Wiener. Exposure to Wiener's prediction theory, generalised harmonic analysis and chaos expansion changed the directions of Kallianpur's research which culminated in the now-famous Kallianpur–Striebel function space version of Bayes formula and foundations of the theory of optimal filtering in the context of stochastic differential equations.

At a suggestion from Wiener, Masani, a distinguished young mathematician of Bombay (now Mumbai), came to ISI Calcutta during the October holidays of 1955 and started working on several problems in factorization and in multivariate prediction. Masani's collaboration with Wiener was a major event in his career.

Kolmogorov visited the Institute in April 1962 and was a great inspiration to the probabilists. A frequent visitor to the ISI in the fifties, the Indian combinatorial mathematician, S.S. Srikhande was a student of R.C. Bose at Chapel Hill during 1947 to 1950. Along with the other "Euler Spoilers", Bose and Parker, he settled in the negative the famous Euler Conjecture on Orthogonal Latin Squares.

6   Indian Agricultural Statistics Research Institute (IASRI)

On the recommendation of the Royal Commission of Agriculture, a small Statistics section was set up in the Indian Council of Agricultural Research (ICAR) in 1929. P.V. Sukhatme joined the section in 1940. After completing school education in Pune, P.V. Sukhatme graduated in 1932 from Fergusson College of the same city with Mathematics as the principal subject and Physics as the subsidiary. During 1933–36, he studied at the University College of London and was awarded a Ph.D. in 1936 and a D.Sc. in 1939 for his work on bipartitional functions. Before joining the Statistics section in ICAR, he was a Professor at the All India Institute of Hygiene and Public Health at Calcutta during 1939–40. Towards the end of 1943, the enquiry committee set up by the Government of India to look into the causes of the devastating Bengal famine observed that one of the main factors responsible for the famine was the defective statistics of crop production available at that time. The Statistics section, under the guidance of Sukhatme, began research in the methods of collection of yield statistics of crops by developing survey techniques of yield estimation under random sampling. Official forecasts were also being released from the results of sample surveys by 1949. The work done by the section was recognized by the United Nations and the FAO and a special training programme was organised for the statistical officers from the South-East Asian governments

for 14 weeks in 1949 on appropriate census and sampling techniques as applied to population and agriculture. Sukhatme left ICAR to take over as Chief of Statistics Branch, FAO in Rome in 1951.

V.G. Panse, Director of the Institute of Plant Industry, Indore, who was closely associated with the work of the Statistics section of the ICAR, replaced him soon after. Panse was collaborating with Hutchinson at the Institute of Plant Industry on the use of quantitative techniques in Agronomy. They adapted the randomized block and split-plot designs to the plant breeding material at Indore and developed a 'replicated progeny' (Panse & Hutchinson, 1935, 1937). Panse (1940) demonstrated how the genetic component of observed variability could be estimated by taking the regression of progeny means on parental value and explained the importance of selecting plots based on their deviations from plant means rather than basing on their own values. Panse also introduced appropriate genetic models which brought out the effects of the number of segregating genes, the magnitude of their action, the modification due to dominance and the environmental influence on progress due to selection. Using Panse's methods, a statistical analysis of data collected over a ten-year period in a goat breeding project at Etah in U.P. was done at the ICAR. It was observed that improvement in milk yield was not so much due to a genetic improvement of the stock through selection as due to extraneous factors. Consequently, this led the animal breeders to realize the need and importance of statistical methods in planning animal breeding experiments and analyzing them. It is the success of this statistical appraisal which led to the expansion of ICAR Statistics section to become a full-fledged Institute of Agricultural Research Statistics. Panse moved to Delhi in 1951 as Statistical Adviser to ICAR.

During 1941, Panse was approached by the Indian Central Cotton Committee (ICCC) for objectively estimating the yield per acre of cotton production in place of the subjective method being followed in the country. Panse insisted that any sampling method must fit into the existing administrative structure and take cognizance of the fact that the departmental staff and the farmers were already familiar with crop cutting procedures. Mahalanobis, as pointed out earlier, having conducted several crop-cutting experiments preferred plots of small sizes—three or four concentric circular plots with different radii. However, ICCC adopted Panse's method. The merits and demerits of the two methods of Panse and Mahalanobis were extensively debated (Panse & Sukhatme, 1948, 1951). Panse also carried out a number of sample surveys to check the accuracy of area statistics and observed that the estimates by the *patwari* agency in the surveyed areas were quite satisfactory, even though it involved extra burden to the *patwaris* in addition to their normal work.

Panse was one of the founders of the Indian Society of Agricultural Statistics. It was established on 3 January 1947 with a view to promoting the study of and research in statistical theory in the widest sense and its application to Agriculture, Animal Husbandry, Agricultural Economics and allied areas. The Journal of the Indian Society of Agricultural Statistics released its fiftieth volume in 1997.

The IARS, renamed as the Indian Agricultural Statistics Research Institute (IASRI), has contributed significantly over the years to the fields of Experimental Designs, Sampling Methods, Statistical Genetics, Biostatistics, Forecasting Techniques, Statistical Ecology etc. under the guidance of Panse and Sukhatme. During the recent years, Prem Narain, the noted geneticist and biometrician became the Director of IASRI.

A Major Scientific Dispute

We conclude this section with an account of a major scientific dispute between ISI and ICAR on the best method of crop-cutting experiments and agricultural surveys, which is of interest from the point of view of Sociology of Science (cf. Adhikari, 1990).

After considerable experimentation with cuts of different shapes and sizes, Mahalanobis recommended the use of circular cuts of radius 4′ for yield surveys and the ISI had been using the same

circular cuts in the National Sample Survey and other surveys. As against this, the ICAR had been using the rectangular cuts of size $33' \times 16.5'$ in the crop yield surveys conducted through the state agency.

Keen interest in resolving this technical controversy on this issue was shown by Mahalanobis and he suggested joint studies both by the ICAR and the ISI (Mahalanobis, 1946):

> "I may mention, however, that for some considerable time, I have been pressing on ICAR authorities the need for carrying out crop cutting work by both ISI and ICAR methods in the same region with a view to studying the relative efficiencies of the two systems."

Adhikari indicates that these studies did not reveal significant differences between the two methods. However, the controversy continued.

Both the ISI under the leadership of Mahalanobis and the ICAR under the leadership of Panse differed significantly about the investigating agencies in conducting the field work also.

In the scheme of Panse, it was emphasized that any sampling method must fit into the existing administrative structure. Mahalanobis, on the other hand, advocated that the field work should be done by well trained investigators recruited for the purpose of the survey.

Adhikari (1990) relates this to the fact that Panse and Mahalanobis came from different parts of India with very different systems of collecting revenue—one going back to the Moghul period and the other introduced by the British.

> "The Moghul Emperor appointed officials called Jagirdars, Inamdars over large tracts responsible for the collection of land revenue. At the level of the village, there was the village accountant (now called a *patwari*), also a state employee who would actually determine the land revenue of each cultivator and collect it on behalf of the officials.
> ..."

> "At the end of the 18th century, the British government introduced a new system in a fairly large part of India, particularly Eastern India, called the permanent settlement. In this system the intermediate tax collectors, called the Zamindars, were made responsible for the payment of revenue of the large tracts under them to the British treasury. The amount of revenue was made permanent"... "... The system of *patwaries* as functionaries of Government ceased to exist..."

What appeared to be a scientific controversy was rooted in the social background of the scientists and, in a way, both were right and both were wrong.

Sastry (1977) refers to joint studies of the Ministry of Agriculture, CSO and ISI conducted in 1960–61 and the studies by a Technical Committee set up by the Planning Commission conducted on 4 crops during 1963–66. No marked differences in the yield rates with the two types of cuts were observed.

7   Statistics in some Indian Universities

The Department of Statistics of Calcutta University is the oldest in the country and one of the pioneering Departments in the world which introduced Statistics as a separate and full-fledged discipline and not as a part of Mathematics or Economics. It was started in 1941 at the initiation of Mahalanobis in collaboration with the Indian Statistical Institute, which was at that time functioning from the premises of the Presidency College, Calcutta. The post-graduate Master's degree programme in Statistics of Calcutta University was the first of its kind introduced in Asia. Initially, Mahalanobis was the honorary Head of the Department of Statistics which consisted of one Lecturer, one Assistant Lecturer and some part-time lecturers. As regards resources, it had a few hand-operated calculating

machines and a library consisting of one single book shelf. The first batch of Master's degree holders, which came out in 1943, included among others C.R. Rao and H.K. Nandi.

In addition to offering a post-graduate course, the department carried on research activities right from the start. Members of the faculty included, apart from Mahalanobis, R.C. Bose and S.N. Roy, who between them made fundamental contributions to the application of classical multi-dimensional geometry and finite geometrics to multivariate analysis and construction of designs. After Mahalanobis, R.C. Bose became the Head of the Department in 1945. When Bose left for the U.S.A. in 1949, S.N. Roy acted as Head till 1950 when he also left for the States. Thereafter the Headship devolved on P.K. Bose. Two other active members of the faculty at that time were H.K. Nandi and M.N. Ghosh who made fundamental contributions in design of experiments, inference, decision theory and asymptotics.

Three years after the post graduate Department of Statistics of Calcutta University was started, an undergraduate Department of Statistics was established in 1944 in the Presidency College, Calcutta. The department which offered Honours level courses in Statistics also initially had the support from the ISI faculty in the form of part-time teachers. Mahalanobis himself would lecture on some general topics concerning this new discipline and motivate them. K.B. Madhava of ISI was the honorary Head and around 1946, A. Bhattacharyya and B.N. Ghosh joined the Department, the former serving as the departmental Head. They framed and followed a syllabus and a curriculum which became a model for several other undergraduate Statistics Departments in the country that followed suit. Bhattacharyya is well known for his contributions to estimation theory, measures of divergence and characterization theory. The characterization of the bivariate normal has led to interesting recent developments in multivariate analysis (cf. Arnold, 1994).

The Post Graduate Department of Statistics was started in 1948 in Bombay University. M.C. Chakravarti was appointed as the Head of the Department. Under his guidance, the department grew up to be one of the prominent teaching and research centres in Statistics in the country. Chakravarti also founded the Indian Statistical Association and became the editor of the Journal of Indian Statistical Association. Chakravarti's work in Design of Experiments is well known and several students of his have occupied important positions in various Indian and foreign universities and industries.

The Department of Statistics at the University of Pune was started in 1953 with 11 students enrolling in the first batch under the leadership of V.S. Huzurbazar, who is the first Bayesian in India. He had earlier obtained a Ph.D. Degree from Cambridge working under the guidance of Harold Jeffreys, the famous astrophysicist and Bayesian. The Department of Pune University gradually expanded and made notable contributions to Probability Theory, Stochastic Processes, Inference and other areas. It is currently one of the most active departments of Statistics in India. The University Grants Commission has selected the Department for Special Assistance for Statistics under the programme for Centres of Advanced Studies.

Other major Universities which have played a leading role in the area of Statistics are Universities of Madras, Mysore, Kerala, Patna, Guwahati, Andhra, Lucknow. All these Departments of Statistics were formed roughly between early forties through early fifties more or less in the order mentioned. Together the Indian Universities have produced some of the world's most well-known statisticians.

8 A Sequel: The Period After 1960

Most of the earlier discussion centered on the period before the sixties. In this section, we take a quick look at the sequel.

One of the reasons for rapid growth of Statistics in India was the close interaction between the ISI and various technical wings of the Government of India. It did not last very long after 1960. Nehru died in 1964, Mahalanobis in 1972, ending an era of close cooperation that began in 1947.

There were other historical reasons for fewer interactions. The Planning Commission of the Government of India was strengthening its in house research. Also, the Planning Unit of ISI Delhi, which was stationed in the same building as the Planning Commission moved to a new campus, far from the Planning Commission. Finally, differences in perception about what would be an optimal path of development led to parting of ways.

Something similar happened to weaken relations with the NSS. During the final years of Mahalanobis's life, ISI's management of the Design Division of the NSS had come in for criticism. Reports of surveys were far behind schedule. A year before Mahalanobis died this section was taken out of the ISI and made a part of the NSS Organisation. In a sense this was a natural administrative step leading to the unification of the field operations and theoretical wings of the NSS. But it did not bring about any improvement in publication of reports. The relation with the NSS was never the same again even though ISI is represented on the Governing Council of NSS.

Globally too, specially in the United States, theoretical and applied statistics had been drifting. This had its effect on Indian statisticians many of whom were trained in the U.S. or had interactions through visiting assignments.

There were many achievements on the theoretical side to compensate for this isolation. Indeed there was an exponential growth in fundamental theoretical work by Indian statisticians in India and abroad. Some of the notable contributions by the post-C.R. Rao generations include development of new tests in multivariate non-parametric inference by S.K. Chatterjee, M.L. Puri and P.K. Sen, a critical evaluation of classical survey sampling from different points of view by V.P. Godambe, D. Basu and J.N.K. Rao and introduction of a new class of designs called search designs by J.N. Srivastava. Moreover, R.N. Bhattacharya, J. Sethuraman, M. Ghosh, J.K. Ghosh and many others contributed to various aspects of asymptotics and inference ranging from rigorously derived Edgeworth expansions and higher order asymptotics to sequential analysis, reliability and life testing. A paradigm shift involving Bayesian Analysis had arrived in India.

The scenario on the applications side showed less activity. It was here that the effects of breach between academia and government, the prime user of Statistics in India, is most visible. But there have been outstanding exceptions as well as modest but sustained growth in certain areas.

We would place in the first category Sukhatme's (1965) new hypotheses about nutrition which suggest that an individual can adjust to varying intakes of calories and that the usually stipulated norms may be higher than necessary. A bivariate analysis of available data on protein and calorie intake for Asian countries, particularly for India, revealed (Sukhatme, 1965) that when diet is adequate in energy, the protein intake is usually satisfactory. It was shown that protein deficiency is the indirect result of inadequate energy in the diet. His scientific beliefs as well as philanthropy led to the establishment of 'Indira Community Kitchen' in Pune. Based on a survey carried out by the NSS, Minhas analyzed the data which seems to lend empirical support to this hypothesis but results of similar enquiries later have been ambiguous.

Other substantial contributions have been in the assessment of the extent of poverty and its decline after independence. Contributors include Nikhilesh Bhattacharya, S.D. Tendulkar and others. (See for example, Pal, Chakravarty & Bhattacharya (1986), Tendulkar (1989), Minhas, Jain & Tendulkar (1991)). The monograph by Bhattacharya, Coondoo, Maiti & Mukherjee (1991) deals with time trends in poverty and inequality in rural India using NSS budget data from 1952–53 to 1983. In this work an econometric model is fitted to explain the observed variation in the incidence of this poverty.

There have also been extensive studies of official statistics and data collected by NSS. An innovative study by Minhas (1988) shows how estimates obtained from these two sources can be reconciled. Among other important contributions we would include a definitive study of one of India's largest dams by Minhas *et al.* (1972).

We now leave aside these outstanding but mostly individual achievements and turn to areas that saw a modest but sustained growth. These include many improvements in the preparation of national

accounts by the CSO, and applications of Taguchi's methods to industry, mainly by the Division of Statistical Quality Control and Operations Research in the ISI. Even more important has been the development of Medical Statistics. India now has an Institute for Research in Medical Statistics and several active Departments of Biostatistics. Though no new methodology has emerged, India has gained experience and expertise for conducting clinical studies for new treatments in the environment of a developing country. The ongoing controversy of the failure of the polio vaccination programme (Wyatt, 1996) highlights the dangers of WHO programmes that ignore local conditions and local expertise.

India has also had a strong population studies programme, conducted by the International Institute for Population Sciences in Mumbai and various Population Research Centres set up in certain Institutions and Universities.

A very large scale National Family Health Survey (NFHS) covering 24 States and National Capital Territory consisting of 88562 sample households was launched by the Ministry of Health and Family Welfare, New Delhi in 1992–93. This provides a memorable demographic snapshot of the world's largest democracy coming to terms with problems of marriage, family planning, old age and growing need of health care, infant and child mortality among other things (see IIPS, 1995).

From the late seventies onward there have been many discussions in India on how this fruitful interaction between theoreticians and practitioners can be restored and what could be possible frameworks for such dialogues. While no clear framework or concensus has emerged many individuals started doing such joint research in the eighties. Here, too, a global change in the perception of priorities has helped.

A significant new development has been the setting up of two centres in the ISI—the Policy Planning and Research Unit (PPRU) at ISI Delhi and Survey Research and Data Analysis Centre (SURDAC) at Calcutta to revive a close interaction between academia and government. The SURDAC is expected to be a modest version of the Survey Research Center of the University of Michigan. Similar joint consultancy projects with industry, health professionals, epidemiologists, ecologists, environmentalists and computer scientists have become very popular in the ISI, and other academic institutions.

The nineties have seen the impact of Information Technology on academia, government, business and industry. A significant innovation has been the use of Palm Top Computers (PTC) in the North Indian state of Haryana for collection of socio-economic survey data besides the traditional data collection by investigators of the FOD of the NSSO. It is envisaged that PTC's would be used in some other states as well. Necessary modifications in software are being developed by the National Informatics Centre. District Head Quarters are being linked up for electronic communication. The Directorate General of Commercial Intelligence and Statistics at Calcutta is planning to transmit data electronically to the Central Government at Delhi. The reports of NSS are now available to the users on floppies. Even though the use of computers in teaching Statistics is already gaining momentum in various Universities and Institutions, the "Multimedia Approach" is still in an infant stage.

The apex national committee for Statistics, National Advisory Board on Statistics (NABS), was set up in September, 1982 to provide technical guidance for policy issues concerning development of Statistics. This has been reconstituted six times, most recently in 1998 with a membership of 32. So far its influence on the Statistical System in India is not properly felt. However, it is beginning to take an active step in encouraging new initiatives. The Indian Official Statistical System itself is receiving a great deal of critical attention both from scholars and from the media. We would regard this too as a positive sign of the relevance of Statistics and the need to change with times. One of the important new priorities is to provide short term estimates of the national income and industrial and agricultural growth. The CSO's short term estimates have not been as reliable as its annual estimates.

In any case a right framework is now in place for interaction between academia, industry and government as well as interdisciplinary research involving Statistics and other sciences. Only the

future can tell how we make use of these new opportunities.

9 Conclusion

India had a long historical tradition of collection and use of various kinds of statistics. The system was strengthened during the British period. Nonetheless, the development in Statistics that took place between, say, 1930 and 1960 is quite remarkable. We do not know of any other developing country where this happened. Nor was such growth visible in any other discipline in India during the same period. What were the possible reasons? One can only hazard a guess.

We believe there were several important ingredients for success.

In Statistics, unlike other disciplines, India was not a late starter. Indeed, much of the development even in the United States came later. Only Britain had started earlier. This helped creation of an Indian school of Statistics with its own mix of theory and applications.

The blending of theory and applications, and interaction between institutions, academia and the Government was a source of live problems, excitement and funding. In hardly any other discipline in India does one find greater opportunities for research or the excitement of an emerging discipline with significant applications backed by a new theory. This attracted the best Indian minds from among mathematicians, physicists, economists, sociologists and anthropologists.

The single most important reason was probably the appearance of a right man at the right time. By one of those extra-ordinary but recurring coincidences in history, Mahalanobis switched interests from Physics to Statistics. He was able to see the future for Statistics as one of the key technologies of our time, and took up its study and applications with all the passion of an inventor and entrepreneur.

The rest is history.

Acknowledgements

References

Adhikari, B.P (1990). Social construction of the statistical estimation of crop yield. Paper presented at the *XII World Congress of Sociology of the International Sociological Association*, Madrid, Spain.

Arnold, B.C. (1994). Bhattacharyya's normal conditional distribution. *Essays on Prob. and Stat.*; Festschrift in honour of Prof. Anil Kr. Bhattacharyya, 1–13, Dept. of Statistics, Presidency College, Calcutta.

Bhattacharya, N., Coondoo, D., Maiti, Pradip & Mukherjee, R. (1991). *Poverty, Inequality, and Prices in Rural India*, Sage Publications, India.

Bose, R.C. & Roy, S.N. (1938). The distribution of Studentized D^2-Statistic. *Sankhya*, **4**, 19–38.

Bowley, A.L. (1906). Address to the Economic and Statistics Section of the British Association for the Advancement of Science, York, 1906. *J. Roy Statist. Soc.* **69**, 540–558.

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

Carey, W.H. (1822). *The Good Old Days of the Honorable John Company*, (New ed. Edited by A.N. Mookerjee, 1964, Calcutta Quins Book Co.).

Central Statistical Organisation (1979). *Statistical System in India*, Dept. of Statistics, Ministry of Planning, New Delhi.

Chaudhuri, S.B. (1964). *History of the Gazetters of India*. New Delhi Pub. Division.

Datta, K.L., Shirras, G.F. & Gupta, B.K. (1913). *Survey on Price Statistics: 1910*. Report submitted to the Government of India.

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

Deming, W.E. (1964). On some contributions of interpenetrating net-work of samples. *Contributions to Statistics* (ed. C.R. Rao). 57–66. Stat. Pub. Society, Calcutta.

Deming, W.E. (1973). P.C. Mahalanobis—A Profile. *Commemoration Brochure issued on the occasion of the 80th Birth Anniversary of P.C. Mahalanobis* by the ISI, Calcutta.

Godambe, V.P. (1976). A historical perspective of the recent developments in the theory of sampling from actual populations, *Jour. Ind. Soc. Agr. Stat.*, **28**, 1–12.

Haldane, J.B.S. (1957). The Syadvada system of predication. *Sankhya*, **18**, 195–200.

Hansen, M.H. & Hurwitz, W.N. (1943). On the theory of sampling from finite populations. *Ann. Math. Statist.* **14**, 333–362.

Hubback, J.A. (1927). Sampling for rice yields in Bihar and Orissa, *Imp. Agri. Res. Inst. Bull*; Pusa, 166 (reprinted in *Sankhya* (1946) **7**, 282–294).

Huchinson, J.B. & Panse, V.G. (1935). Studies in the technique of field experiments I,II,III,IV. *Ind. Jour. Agr. Sc.* **5**, 523–538, 545–553, 554–558, 671–692.

Huchinson, J.B. & Panse, V.G. (1937). Studies in plant breeding technique. *Ind. Jour. Agri. Sc.*, **7**, 531–564.

Hunter, W.W. (1875) . *A Statistical Account of Bengal.* London: Trübner & Co., (also reprinted in India by D.H. Publishing House, Delhi, 1973).

International Institute for Population Sciences (IIPS, 1995). *National Family Health Survey: India, 1992–93.*

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

Jessen, A. (1926). Report on the representative method in statistics. *Bull. Int. Stat. Inst.*, **22**, 359–437.

Kiaer, A. (1895). Observations et expériences concernant des dénombrements représentatifs. *Bull. Int. Stat. Inst.* **9**, 176–183.

Kiaer, A. (1897). *The Representative Method of Statistical Surveys.* (Norwegian version—also available 1976 English translation, Central Bureau of Statistics of Norway, Oslo).

Lahiri, D.B. (1973). Prasanta Chandra Mahalanobis and Large scale sample surveys. *Sankhya*, **35**, (suppl.) 27–44.

Mahalanobis, P.C. (1922). Anthropological observations on the Anglo-Indians of Calcutta. Part I, analysis of male stature. *Rec. Ind. Museum*, **23**, 1–96.

Mahalanobis, P.C. (1925). Analysis of racemixture in Bengal, Presidential Address. Anthropological Section, Indian Science Congress. *Jour. and Proc. Asiat. Soc. Bengal*, **23**, 301–333.

Mahalanobis, P.C. (1928). On the need for standardization in measurements on the living. *Biometrika*, **20A**, 1–31.

Mahalanobis, P.C. (1929). *Personal Communication to R.A. Fisher*, dated 21 November, 1929.

Mahalanobis, P.C. (1930). On tests and measures of group divergence, Part I, Theoretical formulae. *Jour. and Proc. Asiat. Soc. Bengal*, **26**, 541–588.

Mahalanobis, P.C. (1931). Anthropological observations on the Anglo-Indians of Calcutta, Part II, Analysis of Anglo-Indian headlength. *Rec. Ind. Museum*, **23**, 97–149.

Mahalanobis, P.C. (1931). *Statistical study of the level of the rivers of Orissa and the rainfall in the catchment areas during the period 1868–1928.* Submitted to the Govt. of Bihar and Orissa.

Mahalanobis, P.C. (1933). A revision of Risley's anthropometric data relating to tribes and castes of Bengal. *Sankhya*, **1**, 76–105.

Mahalanobis, P.C. (1936). On the generalized distance in statistics. *Proc. Nat. Inst. Sci., India*, **2**, 49–55

Mahalanobis, P.C. (1938). A statistical report on the experimental crop census of 1937. *Indian Central Jute Committee*, 13–83.

Mahalanobis, P.C. (1944). On large-scale sample surveys. *Phil. Trans. Roy. Soc., London*, Ser. B, **231**, 329–451.

Mahalanobis, P.C. (1946). Sample surveys of crop yields in India. *Sankhya*, **7**, 269–280.

Mahalanobis, P.C. (1950). Why Statistics? Presidential Address. Indian Science Congress, Thirtyseventh Session, Poona, (reprinted in *Sankhya*, **10**, 195–228).

Mahalanobis, P.C. (1954). The foundations of Statistics. *Dialectica*, **8**, 95–111; (reprinted in *Sankhya*, **18**, 183–194).

Mahalanobis, P.C. (1955). The Approach of Operational Research to Planning in India. *Sankhya*, **16**, 3–130.

Mahalanobis, P.C. & Chakravarti, N. (1931). *Statistical report on river floods in Orissa during the period 1868–1928.* (In 20 chapters with 47 tables and 27 maps). Submitted to Govt. of Bihar and Orissa.

Mahalanobis, P.C. & Lahiri, D.B. (1961). Analysis of errors in censuses and surveys with special reference to experience in India. *Bull. Int. Stat. Inst.* **38** (2), 401–433 (reprinted in *Sankhya*, Ser. A, **23**, 325–358).

Mahalanobis, P.C. & Nair, K.R. (1922). 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. (Statistical notes for agricultural workers, No.25). *Ind. Jour. Agri. Sci.*, **10**, 663–685.

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

Minhas, B.S. (1988). Validation of large-scale sample survey data—Case of NSS estimates of household consumption expenditure. *Sankhya*, Ser. B., **50**, Pt. 3, 1–63.

Minhas, B.S. , Jain, L.R. & Tendulkar, S.D. (1991). Declining incidence of poverty in the 1980's—Evidence versus artefacts. *Econ. and Pol. Weekly*, July 6–13, 1991.

Minhas, B.S., Parikh, K.S., Srinivasan, T.N., with Marglin, S.A. & Weisskopf, T.E. (1972). Scheduling the operations of the Bhakra system, *Sankhya* Ser. B., **34**, 105–204.

Minhas, B.S., Rudra, A. & Srinivasan, T.N. (1974). 'Pitambar Pant', *Sankhya*, Ser. C., **36**, 1–8.

Nair, K.R. (1992). In Statistics by Design. In *Glimpses of India's Statistical Heritage.* (Eds. Ghosh, J.K., Mitra, S.K. & Parathasarathy, K.R.) 103–150, Wiley Eastern.

Pal, P., Chakravarty, S.R. & Bhattacharya, N. (1986). Poverty in rural India: A decomposition analysis. *Ind. Econ. Rev.*, **21**, 2, 149–173.

Panse, V.G. (1940a). A statistical study of quantitative inheritance. *Ann. Eug.*, **10**, 76–105.

Panse, V.G. (1940b). The inheritance of quantitative characters and plant breeding. *J. Genetics* **XL**, 283–302.

Panse, V.G. & Sukhatme, P.V. (1948). Crop Surveys in India–I. *Jour. Ind. Soc. Agr. Stat.*, **1**, 34–58.

Panse, V.G. & Sukhatme, P.V. (1951). Crop Surveys in India–II. *Jour. Ind. Soc. Agr. Stat.*, **3**, 98–168.

Pearson, K. (1928). *Personal Communication to Mahalanobis* dated 22 May, 1928.

Rao, C.R. (1947). Factorial experiments derivable from combinatorial arrangements of arrays. *J. R. Stat. Soc.*, B, **9**, 128–140.

Rao, C.R. (1973). Mahalanobis Era in Statistics. *Sankhya*, **35** (suppl.), 12–26.

Rao, C.R. (1992). Statistics as a last resort. In *Glimpses of India's Statistical Heritage*. (Eds. Ghosh, J.K., Mitra, S.K. & Parthasarathy, K.R.), 151–213, Wiley Eastern.

Risley, H.H. (1891). *Tribes and Castes of Bengal*. Government Printing, Calcutta.

Rudra, A. (1996). *Prasanta Chandra Mahalanobis: A Biography*. New Delhi: Oxford University Press.

Sastry, N.S. (1977). Estimation of agricultural production by sample survey for estimating crop acreages and crop yields: The Indian experience. *Bull. Int. Stat. Inst.*, **47**, 3, 425–439.

Shakespeare, A. (1848). *Census of N. W. Provinces*. Govt. of N. W. Provinces, India

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

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

Sukhatme, P.V. (1965). *Feeding India's Growing Millions*. Bombay: Asia Pub. House.

Tendulkar, S.D. (1989). Economic inequalities and poverty in India : An Interpretative Overview. In P.R. Brahmananda and V.R. Panchamukhi (eds.): *The Development Process of the Indian Economy*. Bombay: Himalaya Pub. House.

Wyatt, C.S. (1996). Poliomyelitis in Indian children who have received oral polio vaccine: Vaccine failure or low potency? *Current Science*, **70**, 6, 437–440.

Résumé

Ceci est une courte histoire de l'évolution de la Statistique officielle et académique en Inde. Nous retraçons l'histoire à partir des origines de l'antiquité jusqu'à l'histoire plus récente mais portons plus d'attention sur la période des années 1930 à 1960. Ceci et l'histoire d'institutions, des gens qui les bâtirent et les formèrent et d'idées.