thoughtful introduction. Through our joint collaboration, a set of articles on relativity was published in the form of a booklet by the Calcutta University. It had been in circulation in different countries for quite some time. It is perhaps not available now.

Saha went away to Allahabad and I to Dacca. Prasanta retired from Government service after working for a long time as a professor and later made adequate arrangement for the teaching of statistics after coming over to the University. Afterwards, he served his country with devotion for about twenty years. He had the firm conviction that, through the measure of statistics, we would be able to understand the realities of all problems and with it, would be able to think as to what the country should do. His talent was acknowledged in many countries. He had visited most of the countries of the world and his deep knowledge was recognized by world bodies. Till the last day of his life, Prasanta's only thought was statistics.

Scientific contributions of Professor P. C. Mahalanobis

Dwijesh Dutta Majumder

Background and perspective

A person's contribution in any sphere of life and society can better be considered in the context of the personality and the environment that led to his or her contributions. This is very much applicable to the case of the Professor, as P. C. Mahalanobis was popularly known to us. In this respect see the brief biography of P. C. Mahalanobis by his cousin and a colleague, Late Anikendra Mahalanobis[1], his scientific contribution by Prof. C. R. Rao[2] and a brief biography by the present author[3].

In order to understand the versatility of P. C. Mahalanobis not only in respect of his contributions to anthropology, meteorology, flood control, statistics, large-scale sample survey, planning, agriculture, national income, information processing technology, operations research, quality control, educational testing methods, demography, sociology, ancient reptiles of India, speech research, linguistics, genetics, psychometry, manpower planning, haematology, social change and economic development, unemployment and underemployment, organization of science and technology, but also his contributions to social, cultural and intellectual movements, one has to visualize the background and historical perspective of the then Bengal *vis-à-vis* India. Because of the limitations of space, I am not attempting to provide the exhaustive list of bibliography which is 271 scientific papers and five books as is given in refs. 1, 2, along with ref. 4 which contains appreciations and reminiscences of Mahalanobis by prominent scientists such as Simon Kuznets, Leon H. Keyserling, Edward H. Colbert, Alex Comfort, H. L. Shapiro, W. Edwards Demming, Academician I. Stefanov, A. T. A. Learmouth, Ashok Rudra, K. P. S. Menon, C. R. Rao, D. B. Lahiri, M. Mukherjee, Catherine A. Galbraith, and others.

P. C. Mahalanobis, M. N. Saha, S. N. Bose, N. R. Sen, J. C. Ghosh and their predecessors P. C. Ray and J. C. Bose are all products of renaissance which is also known as nineteenth century's new awakening in religion, art, culture, literature and science that spread throughout the country. His family background, his close personal acquaintance with Brojendra Nath Seal, Rabindra Nath Tagore in early life and C. D. Deshmukh and Jawaharlal Nehru afterwards was very important in not only making such massive scientific contribution but also in building up Indian Statistical Institute for advanced research and training, in starting the Indian Journal of Statistics — *Sankhyâ*, in establishing Statistical Publishing Society and a large Electronic Computer Research and Development Centre.

Mahalanobis graduated with honours in physics from Presidency College, Calcutta in 1912 and passed tripos examination from Cambridge University part I in mathematics in 1914 and part II in natural science (physics) in 1915 with a first class first in order of merit. In Cambride, he came in close contact with eminent scholars like Srinivasa Ramanujan, G. H. Hardy, J. M. Keynes, Lowes Dechinson, Bertrand Russel and many others. His study at Cambridge had great impact on his life and works. He was awarded a senior scholarship to work at Cavendish Laboratory under the guidance of C. T. R. Wilson and J. J. Thompson, came to Calcutta to spend the vacation, joined the physics department of Presidency College as a teaching staff, got involved in many social activities and did not return. In 1948 he retired from Presidency College when he was its principal, and was made professor emeritus of the college. In the meantime he laid the foundation of statistics in India and of the ISI, which brought him and the country world-wide recognition, some of which we shall review in the next few sections.

Mahalanobisian view of statistics

If we look at the developmental history of mathematical systems or structures we see that they are in general suggested by situations which, while they are different, have some basic features in common so that their emergence is essentially the result of a process of unification and abstraction. A mathematical system, thus, lays bare the structurally essential relations between

otherwise distinct entities, so that the system remains valid for each of those otherwise different situations which provided motivations and inspirations for the same, leading to a better and fuller understanding of the motivating situations. Mahalanobis indicated the conceptual framework of his view of statistics in relation to physics, mathematics and philosophy in some of his papers[5-8]. According to him though analytical statistics and probability are retrospectively linked to physics and its foundations laid by mathematicians like Fermet, Pascal, the Bernoullis, Laplace, Gauss, Poisson, Cauchy Tchebycheff and others, statistics cannot be considered as a branch of pure mathematics. He cited Bertrand Russel's description of mathematics as statements of the form:

If P, then Q

With the condition that it is immaterial what P and Q are. Whereas statistical statement can only be made in the form: If P, then Q will follow 9 times out of 10 or 99 " " " 100, etc.

If the predictions do not fail to an extent then the statistical reasoning is not valid. Some of these distinctive aspects of pure mathematics and statistics are related with deductive and inductive logic and are traceable in the history of physics. The spectacular developments of classical mechanics in 17th and 18th centuries gave bias towards materialistic determinism, though the complete solution even for three bodies not to speak of many bodies problem could be worked out. Statistical concepts became imperative in connection with 2nd law of thermodynamics, in mid-nineteenth century and Gauss–Laplacian theory of errors provided the foundation of Maxwell–Boltzmann distribution of velocities in kinetic theory of gases, by the end of 19th century statistical mechanics became integrated with science. With the emergence of quantum theory and the principle of relativity in the opening years of 20th century, the materialistic determinism was replaced by Heisenberg's principle of uncertainty in the form:

$$\Delta p\, \Delta q \geq h \qquad (1)$$

in which p and q are generalized momentum and space co-ordinate and h is Planck's constant. This inequality relation had a profound impact on physics and is of great significance in statistics. It asserts finiteness and discreteness of physical changes and also denies the possibility of either Δp or Δq assuring infinitesimal value. Another significant thing is its analogy with correlational expression in a limiting form:

$$E(\Delta x, \Delta y) = \rho, \qquad (2)$$

where Δx *and* Δy are a pair of statistical deviations, E is the expectation and ρ is the correlation coefficient. If the relation is valid for a pair one can drop E and if ρ is minimum then we may introduce the sign of inequality and write

$$\Delta x\, \Delta y \geq \rho. \qquad (3)$$

From the relations (1), (2) and (3) Mahalanobis is inclined to think of an irreducible 'correlation' with a minimum of h that shows: (a) explicit denial of independence of Δp and Δq, (b) bounded structure of the world of physics, and (c) mechanistic determinism has no place and foundation of physical science is statistical in nature. Mahalanobis also traced certain interesting resemblances of the ancient Jaina philosophy to the probabilistic and statistical view of reality. He found that the Jaina logic of *Syādvāda*, which was current over two thousand years ago, used to keep a margin of uncertainty for all predictions similar to the modern theory of statistical estimation and inference.

After asserting this philosophical foundation Mahalanobis with an inimitable frankness developed his models for sample surveys along with its theoretical and experimental tools, methodologies, implementations, sharpening of the tools, minimization of the cost, error analysis, presentation of results and their interpretation. He wanted a multi-disciplinary approach consisting of researchers from mathematics, statistics, physics and engineering for developing general principles of the design of sample of survey.

Contribution to statistics

Rao[2] in his memoirs of The Professor elaborated Mahalanobis's contributions to statistics in the following subsections. (a) The Indian Statistical Institute, (b) Mahalanobis distance, (c) Meteorological research, (d) Early examples of operations research (which he actually meant was research on flood control), (e) Educational tests, (f) Errors in field experimentation, (g) Large-scale sample surveys, (h) Perspective planning, and (i) Fractile graphical analysis. I intend to add to this list: (j) Indian Journal of Statistics — *Sankhyā* series A, B & C, (k) Mechanized data processing and computational mathematics, and (l) Statistical quality control and operations research for Indian industries.

The growth story of ISI, fascinating though, is beyond the scope of the present paper, but I intend to start with *Sankhyā*, because establishing and running an internationally reputed journal on a specialized subject from India in those days was a remarkable contribution to Indian science. It also reflected his own view of statistics.

Sankhya: The Indian journal of statistics

Mahalanobis felt the need for starting a journal for stimulating research and advanced studies in statistics since the establishment of ISI in 1931 and its first issue was brought out in June 1933 with himself as its Founder–Editor. He elaborated the reason for naming *Sankhya* in the editorial of its first issue citing from *Atharva Veda*, *Amar Kosha*, and *Bhagvad Gita*, in its dualistic meaning of 'determinate knowledge' and 'number'. The fundamental aim of statistics being to give determinate and adequate knowledge of reality with the help of number and numerical analysis, the ancient Indian word Sankhya embodies the idea adequately.

Mahalanobis distance (anthropometric research)

The first opportunity to use statistical methods came to Mahalanobis in connection with the analysis of anthropometric measurements data taken on Anglo-Indians in Calcutta which was his first scientific paper[9] followed by about fifteen more papers[1,2] out of which the most important ones in my view are (10) (11) (12). The work led to the formulation of the well-known D^2-statistic known in the literature as Mahalanobis distance, and one of the most widely

used mathematical tools to investigate problems of cluster analysis, taxonomical classification, pattern recognition and many other fields of data analysis for its mathematical simplicity and elegance.

If we express the general multivariate normal density as:

$$p(x) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(x-\mu)' \Sigma^{-1} (x-\mu) \right], \quad (4)$$

where x is a d-component column vector, μ is the d-component mean vector, Σ is the $d \times d$ covariance matrix, $(x-\mu)'$ is the transpose of $(x-\mu)$, Σ^{-1} is the inverse of Σ and $|\Sigma|$ is the determinant of Σ. The quantity

$$D^2 = (x-\mu)' \Sigma^{-1} (x-\mu) \quad (5)$$

can be called the squared Mahalanobis distance from X to μ.

$$\text{Formally, } \mu = E[x] \quad (6)$$

$$\text{and } \Sigma = E[(x-\mu)(x-\mu)'] \quad (7)$$

$$\mu_i = E[x_i] \quad (8)$$

$$\sigma_{ij} = E[(x_i - \mu_i)(x_j - \mu_j)] \quad (9)$$

Σ is always symmetric and positive semidefinite.

The multivariate normal density is completely specified by $d + d(d+1)/2$ parameters, the elements of the mean vector μ and the independent elements of the covariance matrix Σ. Samples drawn from a normal population tend to fall in a single cluster. The centre of the cluster is determined by the covariance matrix. It follows from equation (4) that the loci of points of constant density are hyperellipsoids for which D^2 is constant. Thus the contours of constant Mahalanobis distance to μ and volume of these hyperellipsoids measures the scatter of the samples about the mean. Mahalanobis also lay down an axiom for the cluster validity[12] called dimensional convergence of D^2. Mahalanobis's extensive theoretical and experimental work with anthropometric data generated many problems for research in multivariate analysis and his students and colleagues at ISI made valuable contributions in this field.

Meteorological research

While functioning as a reviewer of Mahalanobis's papers on anthropometric survey, Sir Gilbert Walker, the then director general of observatories, got attracted by his work and sent to him some meteorological problems, that not only resulted in a number of pioneering publications[2,13] but also in his appointment as meteorologist in Calcutta in charge of the eastern region, which he carried out from 1922 to 1926 in addition to his duties as professor of physics at the Presidency College.

Flood control — North Bengal and Orissa

The next most important work in which Mahalanobis got involved concerned flood control[4,14]. Government referred the expert committee report on the disastrous North-Bengal flood of 1922 for his opinion. He contradicted the report on the basis of a statistical study of the previous 50 years of flood and rainfall data which he carried out and submitted his entirely different recommendations which were implemented and found effective. Similarly, he contradicted the expert committee report of the 1926 Orissa flood on the basis of a statistical study of the previous 60 years of flood and rainfall data and recommended constructing drains at upper reaches for flood control and generation of electricity which formed the foundation of the Hirakud Project after independence.

Educational tests and psychometry

Mahalanobis developed methodologies for construction, performance and data analysis of different aspect of educational tests and published a series of six papers[2]. He established a Psychometric Research and Service unit at ISI which caters to the need of public sector organizations for recruitment of skilled personnel whenever required.

Error analysis in agricultural field experimentation

In 1925 Mahalanobis initiated a study on the probable error in field experiments[15] which turned out to be a pioneering work, as the same method was independently used by J. Neyman. Mahalanobis, also in complete ignorance, was in line with the Fisherian design of experiments introduced a few years earlier. This paper[15] turned out to be responsible for one of the most deep-rooted and lasting collaborative friendship between Fisher and Mahalanobis in respect of foundations and methodological aspects of large-scale sample surveys. In this area of research Mahalanobis and his colleague Late Subhendu Sekhar Bose contributed about thirty papers, all published in *Indian Journal of Agriculture* and *Sankhya*[2].

Large-scale sample surveys

Rao[2] characterized the Professor as a physicist by training, a statistician by instinct, and a planner by conviction. This statement is amply corroborated in Mahalanobis's contribution on sample surveys. He also expressed the same desire for such collaborative work by scientists of these three fields and also of engineering[5].

Mahalanobis along with a band of dedicated 'workers' (as Mahalanobis used to call everybody including himself as 'worker'), planned and executed numerous surveys in the period 1935–46. There are about fifty papers published[1,2] in this period in this field of study. Some were on consumer expenditure, tea-drinking habits, public opinion and public preference, acreage under a crop, crop yields, incidence of plant diseases, rupee census, and the rest on related methodologies. Mahalanobis's classic papers in this field were published — one in 1944 in *Philosophical Transactions*[6] and the other in 1946 in *J. Royal Statistical Society*[16]. In these papers he dealt with fundamental problems of randomness, identification of levels of randomness and basic theory of sample surveys and estimation procedure along with the operational aspects of planning, executing, processing and presentation of data for dissemination in science and society.

Mahalanobis developed the concepts of 'pilot surveys' and 'optimum survey design' that was forerunner of Abraham Wald's sequential analysis. In the context of the acreage surveys, he conjectured the variance function of the form $V_x = cpqx^{\theta}$, where x is the area of

the sampling unit, p the proportion of area under a particular crop, $q=(1-p)$ and c, and g parameters to be estimated. The formula was found to fit reasonably well with experimental data. In order to give the sample survey the rigour of a scientific experiment, to avoid personal bias in taking measurements and instrumental errors, to infuse built-in cross checks in the validation of results, he introduced the concepts of repetition and interpenetrating network of samples (i.p.n.s). This was the most controversial of his methods—as a physicist he was convinced of such practices—but the survey statisticians considered it unnecessary and expensive. But some scientists abroad, including veteran American statistician Demming supported the idea and practiced it. Lastly as chairman of UN Subcommission (1947–51), he introduced the use of sample survey methods in developing countries for the collection of socio-economic and demographic data and laid down specifications for conducting large-scale sample surveys.

Approach to planning for national development

One can trace Mahalanobis's involvement in national planning in 1940s. When Subhas Bose was Congress president he appointed the Congress Planning Committee with J. L. Nehru as its chairman and M. Visweswaraya, M. N. Saha, K. T. Shah, A. D. Shroft, Ambulal Sarabhai, H. V. Kamath, etc. as members. Nehru asked Mahalanobis to write a statistical supplement to the committee's report. Though Mahalanobis could not have systematic or complete acquaintance with economics as a subject, he made several meaningful and effective suggestions on regional planning in his papers on flood control[4,14]. There are also enough evidences of his acquaintances with economics in the large number of papers and reports on sample surveys reviewed in the previous section. He became concerned with macro-economic problems of India since 1949 when he became chairman of Indian National Income Committee.

Mahalanobis introduced a simple concept of capital-output ratio[5] in 1950 and recommended a Harrod–Domar type of model[17] in 1952. His idea of two-sector and four-sector model was introduced in two papers[18,19], where the net output of economy was conceived as originating in two sectors, one producing all investment goods and the other all consumer goods. In four-sector model, the consumer goods sector was made three types: (i) factory type, (ii) manual type including agriculture, and (iii) services sector. Solving the related simultaneous equations and estimating the investment and income coefficients, Mahalanobis determined the investments to be made in different sectors during the Second Five-Year Plan of India.

Mahalanobis in his mathematical model considered national economy as a 'system' and a 'conceptual framework which would be of help for all practical purposes' and in 'revealing the broad characteristics of the system under consideration without getting lost in details'[19]. He wanted to use the model only as 'scaffolding to be dismantled once the building was erected'. He was free from economic dogma. His views are amply explained in two books and a large number of papers[1,2] in relation to education, scientific and technical manpower, labour, unemployment and demographic problems of India.

Modern computational tools and mathematical methods

Mahalanobis was acquainted with advancements in mechanization of calculations and statistical data processing taking place abroad and from the very beginning tried to acquire best available contemporary tools, be it desk calculators, calculating and tabulating machines or digital computers. His idea was not only to provide modern computing aids to scientists but also to provide a first hand knowledge of state-of-the-art technology so that ultimately the country could be self-sufficient in the information processing technology. From the many personal discussions that I had with Professor it can be said that some of his ideas were ahead of the technology/methods available at the time that created problems for his and others' lives.

As a result of Ist ISI Review Committee Report (1967) the computer developmental activities had to be shrunk. He advised us to concentrate on sophisticated applications research using statistical methods and to develop newer statistical tools. This resulted in new areas of research on speech/image processing, statistical pattern recognition, artificial intelligence, etc. at ISI.

During the last years of his life he was personally engaged in developing a new computational tool for statistical data analysis which he called Fractile Graphical Analysis. This was a semi-non-parametric method for comparison of two samples, first developed for comparison of socio-economic conditions of a group of people at different points of time or of two groups of people at two different places[20]. The method was also applied in fields like psychometry, biometry and demography and was extended to higher dimensional data[21,22].

Last years, days and hours

Some of the writings of his last years[23,24] indicate the shape of India and the world he was dreaming of. He was mentally alert and physically strong throughout his life. From early 1972, he started feeling unwell, and on 24 May 1972 he was admitted to a nursing home in Calcutta for an abdominal operation on 7 June 1972. He stood the operation well. He discussed with some of his colleagues about his new statistical tool fractile graphical analysis in the nursing home but told his wife Rani Mahalanobis that he may not be alive on his next birthday namely 29 June 1972. On 28 June while he was discussing some administrative problems of ISI, he felt uncomfortable and within an hour he expired. So Mahalanobis was active and contributing till the last hours of his life.

1. Mahalanobis, A., *Prasanta Chandra Mahalanobis*, National Book Trust, New Delhi, India, 1983, (Saka 1905).
2. Rao, C. R., *Prasanta Chandra Mahalanobis, Biographical Memories of the Fellows*, Royal Society of London, 1973, vol. 19.
3. Dutta Majumder, D., *Prasata Chandra Mahalanobis—A Scientist with a Social Commitment*, P. C. Mahalanobis Birth Centenary Volume, ISI Club, Calcutta, 1993.
4. *SAMVADADHAM, P. C. Mahalanobis Memorial Volume*, House Journal of the Indian Statistical Institute, Calcutta, December 1974, vol. 10, nos. 1–4.
5. Mahalanobis, P. C., Sample Survey, Presidential Address, Section of Mathematics and Statistics, Indian Science

Congress, Baroda 1942. Published in *Why Statistics and Other Essays* (ed. Bose, P. K.), Statistical Publishing Society, Calcutta, 1986.
6. Mahalanobis, P. C., *Philos. Trans. R. Soc. London*, 1944, **B231**, 329–451.
7. Mahalanobis, P. C., Why Statistics, Address of the General President, Indian Science Congress, Poona, 1950.
8. Mahalanobis, P. C., *Dialectica*, **8**, 95–111; *Sankhya*, 1957, **18**, 189–194.
9. Mahalanobis, P. C., *Rec. Indian Mus.*, 1922, **23**, 1–96.
10. Mahalanobis, P. C., *J. Proc. As. at. Soc. Beng.* (*n.s.*), 1930, **26**, 541–588.
11. Mahalanobis, P. C., *Rec. Indian Mus.*, 1931, **23**, 97–149.
12. Mahalanobis, P. C., *Proc. Natl. Inst. Sci. India*, 1936, **2**, 49–55.
13. Mahalanobis, P. C., *Nature*, 1923, **112**, 323–324.
14. Mahalanobis, P. C., *Sankhya*, 1940, **5**, 1–20.
15. Mahalanobis, P. C., *Agric. J. India*, **20**, 96–116.
16. Mahalanobis, P. C., *J. R. Statist. Soc.*, **A109**, 325–378.
17. Mahalanobis, P. C., National Income, Investment and National Development, Lecture delivered at the National Institute of Sciences of India, New Delhi, 1952.
18. Mahalanobis, P. C., *Sankhya*, 1953, **12**, 307–312.
19. Mahalanobis, P. C., *Sankhya*, **16**, 3–130.
20. Mahalanobis, P. C., *Econometrica*, 1960, **28**, 325–351.
21. Mahalanobis, P. C., Proc. Int. Cont. Quality Control, Tokyo, ICQC, 1969, 515–518.
22. Mahalanobis, P. C., Int. Cont. Computer Science, Cairo University, December, 1969.
23. Mahalanobis, P. C., *Sankhya*, **B31**, 435–458.
24. Mahalanobis, P. C., Social Change, Science, and Economic Growth. One Asia Assembly on New Directions for Asia. Press Foundations of Asia, Manila, Philippines, April 1970.

Dwijesh Dutta Majumder is in the National Centre for Knowledge Based Computing, Indian Statistical Institute, Calcutta 700 035, India

Edited and published by Prof S. Ramaseshan, Current Science Association, Bangalore 560 080.
Typeset by Book Makers (I) Pvt. Ltd., Bangalore 560 009. Printed at Printek Printers, Bangalore 560 079.