South East Asian J. of Math. & Math. Sci. **ISSN : 0972-7752**
Vol.12, No.2 2016, pp. 49–58

PRESENT STATUS OF STATISTICAL EDUCATION IN INDIA

D.S. Hooda

Honorary Professor (Mathematics),
GJ University of S & T, Hisar and
Adviser(R) to ABV Hindi University, Bhopal
E-mail: ds_hooda@rediffmail.com

Dedicated to Prof. A.M. Mathai on his 80^{th} ***birth anniversary***

Abstract: In the present paper a brief account of Indian Statistical System and contributions of Indians to development of statistics are presented. Statistical education at school and undergraduate levels is discussed in detail. Educational and research activities in statistics at premier institutions and universities are highlighted. In the end some suggestions for future planning in statistical education are made for discussion.

Key words and Phrases: Agricultural statistics; survey techniques; design of experiment; analysis packages; and data mining.

2010 Mathematics Subject Classification: 62.

1. Introduction

Statistics as a subject is just a century old science. During last few decades it has penetrated into almost all sciences like agriculture, biology, business, social, engineering, medical sciences, etc. Its wide and varied applications have lead to the growth of many branches, such as Industrial Statistics, Biometrics, Biostatistics and Agricultural Statistics. These branches have emerged as distinct entities or subjects with a bulk of statistical techniques specific to their application areas.

Agricultural Statistics comprises the area of statistical science that deals directly with the problems of field experimentation and interpretation of results in agricultural sciences. It was R.A. Fisher whose dedication and continued efforts put the statistical science on a strong mathematical foundation. The theory of experimental designs and other statistical techniques developed by Fisher constitute the backbone of the Agricultural Statistics.

In advanced countries with established scientific tradition there is continuing concern with validity of data and validity of conclusions. In underdeveloped countries the principle of authority is still dominant; the question of validity scarcely arises. Statistics, therefore, remains a matter of formal or administrative sanctions; anything having the official stamp must be accepted as authoritative.

In a village economy the information available within the village or in neighboring villages is enough for all practical purposes. The question of comparability or compilability does not arise. But the scenario has changed after globalization and industrialization, now the need of statistics of increasing coverage and accuracy becomes more and more insistent. India has already reached this stage; however, the statistical system in other developing countries is unable to meet the need.

The future of statistics needs to be discussed in view of recent development in information technology such as data mining, data communication and information processing networks and requirement of end users. The current statistical methodology based on probabilistic models developed for the analysis of small data sets appears to be inadequate and requires some more methods to be put forward in the name of data mining for such purposes. In the present communication we give a brief account of Indian Statistical System and Statistical Education at various levels in India. In the end we have made some suggestions and recommendations for future planning and these are open for discussion.

2. Indian Statistical System

Collection and use of statistics for administrative purposes in India has a long history spread over many centuries. The *Arthasastra* and the *Ain Akbari* mention the practice of numerical data collection for purposes of statecraft in ancient and medieval India. The Mughals had a system of collection and compilation of crop statistics to help them in land revenue collection. Later on, Britishers created their own data generating system to serve their specific ends. During the British period, consolidation efforts were made for the collection of socio-economic data. Their system was restricted to a few specific fields like trade and commerce, selected industrial products, population, some basic crop statistics and livestock.

The Indian Statistical System is one of the largest institutional framework having a vast wealth of information, not all of which gets due attention of decision-makers. It is one of the oldest and has a large network supported by competent personnel and with a adequate facilities for data management. The system claims a wide coverage of information items as well as of area and people. When we talk about a national statistical system, we can reasonably ask whether such a system possesses the four characteristics of a system viz. content, structure, communication and control? Such a system should be able to:

- Identify its long-term and short-term objectives as well as a strategy to achieve these objectives.
- Ensure a structure in terms of job differentiation by recognizing different types of information to be collected and presented on the different aspects of information processing for which the responsibility and authority should be clearly vested, possibly in different groups of people.
- Establish clear lines of communication within itself to take care of inter- group or collective responsibilities as well as with the environment.
- Review its contents, structure and communication regarding their effectiveness as well as efficiency and modify or control one or more of these as and when found necessary.

It is not difficult to see what is wrong with official statistics in India. There is gap between theory and practice. There is gap between the means and the end in the absence of any clearly perceived purpose. There is lack of appreciation of the need of cross-examining the data, although it is the first responsibility of a statistician. Unfortunately, collection and scrutiny of primary information or processing and handing of data were usually not liked by government officials and statistics was, therefore, not considered a popular subject for acquiring skill and experience in pre-independence days.

Just after independence in 1947, the system of data collection followed by the Britishers was found inadequate to meet the necessity of a strong database covering a variety of social and economical aspects. The existing system even did not provide the basic data required for estimation of national income, which is essential for assessing performance and progress of the economy.

The immediate task, therefore, was to set up a statistical system capable of filling the large gaps in the data essential for formulating economic plans. A very important step in this direction was the creation of the Directorate of National Sample Survey in 1950. Its aim was to collect essential statistics related to the socio-economic conditions and agricultural production in India. Due to the sustained efforts of academicians and official statisticians, the Indian Statistical System has attained the height of the most comprehensive statistical system of a developing country.

To-day India has a decentralized statistical system between centre and the states. The demarcation has been done partly on the functional basis and partly on the regional basis. Statistics of items like foreign trade, banking and currency and census are wholly allocated to the centre and that of like agriculture and education

are assigned to the States. There is also a common category of items known as concurrent list which where both the centre and the States can operate simultaneously to meet their respective requirements. The department of Statistics (DOS) of the Central Government is the apex body in the official statistics of India. . The Indian statistical system pertains to the collection, compilation and dissemination of data relating to socio-economic, agricultural and industrial statistics in India.

Ministry of Agriculture, the Central Statistical Organization (CSO) and the National Sample Survey Organization (NSSO) are some of the important agencies at the national level involved in collection, compilation and dissemination of data.

The CSO is mainly responsible for coordination of statistical activities as well as evolving and maintaining statistical standards. The NSSO has been a leading sample survey organization since its establishment in 1950 and continues to conduct major multi-subject surveys that provide valuable data required by the policy makers. The NSSO conducts large-scale surveys at the national level and collects and disseminates information on different areas. The NSSO, under the scheme of improvement of crop statistics, also provides technical guidance to the states in respect of the crop estimation surveys and performs sample checks to assess the quality of primary work done by the state agencies in area and crops estimation surveys.

In addition to the Department of Planning and Statistics and the Ministry of Agricultural some other Government Departments, institutions and autonomous bodies and non-government organizations are actively involved in the creation of large database. Many of them compile and publish information on agricultural Statistics and livestock population collected through periodic censuses.

3. Contribution of Indians to Statistics

Statistics is one of the subjects in whose development Indian researchers and practitioners have played a significant role. Many of them have shifted to the advanced countries like USA and enriched the science of statistics and have indeed left an indelible mark in the statistical literature. Consequently, the current status of statistics in India is not up to the mark it had to be. In fact 'Brain-Drain' is the main reason for its deterioration as most of the eminent statisticians have shifted to the advanced countries. This had caused a detrimental effect on statistical education and research in India.

Next, we present the list of well-known Indian statisticians who made outstanding contributions to the development of statistics as a separate subject or discipline:

Name	Landmarks in Statistics
D Basu;	Bayesian Statistics
V.S. Huzurbazar;	Bayesian inference properties of the exponential family of dis-

tributions and the uniqueness of the likelihood equation.
D.B. Lahiri; Lahiris Method of sample selection with PPS
P.C. Mahalanobis; Concept of optimum survey design, Pilot survey, D2 Interpenetrating network of samples
C.R. Rao; Cramer-Rao Lower bound, Rao Blackwell theorem , First and second order efficiency , g-inverse, growth curves, MINUE, Unified theory of linear models...
P.R. Masani; Wierner-Masani theory of Multivariate Stationary Stochastic Processes.
P.V. Sukhatme; A pioneer in Agricultural Statistics and sampling through his excursion to the discipline of sociology, nutrition and genetics.

Mahalanobis Contribution to Statistics

Professor P.C. Mahalanobis put statistical practice on a scientific basis in India. He made pioneering contributions to the design of largescale sample surveys and to the use of interpenetrating sub-samples for handling measurement errors. Professor Mahalanobis demonstrated that statistics is an essential tool for rational decision making based in fallible information in all activities related to sciences, industry, trade and administration. He called statistics as a key technology; Mahalanobis(1965). In words of Professor P.C. Mahalanobis:

...statistics must have a purpose. It had its origin in the counting of men or of cattle or in the measurement of land, food grains etc., for purpose of management and administration from immemorial time. The very word statistics shows the connection with statecraft. With emergence of the concept of probability, statistical has been characterized by a dual motivation, one of which is utilitarian and concerned with economic gains and the other scientific or logical and concerned with the question of validity of data or conclusions.

Professor P.C. Mahalanobis is mainly responsible for the introduction of statistics in India and establishment of institutes like ISI on the India soil. On theoretical side, Professor P.C. Mahalanobis in 1930,and 1936 gave D2 Statistics (1930,36) often known in statistical literature as Mahalanobnis distance and has been widely used in classification and discriminant analysis.

Professor Mahalanobis also laid down an important axiom for validity of cluster analysis (1936) called dimensional convergence of D^2. It states that of D^2p and D^2_∞ are the Mahalanobis distances between two populations based on p-characters and all possible characters respectively, then under certain conditions $D^2p \to \infty$ as $p \to \infty$

The larger scale sample survey technique as practiced today also has its in the pioneering work of Professor P.C. Mahalanobis in 1940's and 1950's. He observed

that, there is a great need for sample surveys in collecting information especially in developing countries, where, official statistical system are poor and data are treated as an integral part of the administration system.

The methodology of large-scale sample surveys was developed during 1937-1944 in connection with the surveys planned and executed by the ISI Calcutta. The surveys were spread in many dimensions like consumer expenditure, tea drinking habits, public opinions and public preference, crop areas and crop yield and Incidence of plant diseases. The basic results on large scale surveys were published in 1944 and also presented at a meeting of the Royal Statistical Society (1961).

4. School and Under-Graduate Statistical Education

In India a controversy has cropped up in tertiary statistical education between the curricula and approaches of those actively associated with statistical profession. There are misguided beliefs that statistical understanding is not necessary to teach statistics. Profession emphasis on the real problems of statistics; real data and interdisciplinary interaction; focus on tertiary teaching; greater awareness of linking with students past, present and future are salient features of statistical education in India.

The ideas of problem based and collaborative teaching packages (ref. Densie, 1998) have influenced statistical teaching but there seams to be general acknowledgement that an approach of structure and real problems in parallel (ref. Sowey, 1991, 1998) helps students in developing their concepts and practical confidence. Computer technology has expanded horizons for both learners and teachers accessibility to techniques, demonstrations and visualization, enhancing both skill and conceptual development

Realizing the importance of statistics, some elementary topics were introduced at school level in the subject of mathematics. The topics were included at 10+1 and 10+2 level are: Definition of statistics, Measures of central tendency, Measure of dispersion, Probability, Correlation and Regression.

At under graduate level statistics is taught as one paper in B.A. pass and honour courses in mathematics in most of the universities, while some universities have introduced statistics as a separate subject and the contents included in the curriculum are: Descriptive Statistics, Correlation and Regression, Probability Theory, Random Sampling, Statistical Estimation, Tests of Hypothesis and Analysis of Variance.

To meet the challenges of teaching statistics to undergraduate students we suggest the following:

- Relevant teaching material be prepared and handed over to the students in

the beginning of the course.

- Text books may be made available and these should be written keeping in view the objectives of the courses and background of the students.

- The teacher should make efforts to create interest among the students for statistics learning by correlating its teaching with real life problems.

- Teaching should be made experiments oriented and enjoyable illustrating by live examples and using audio- visual aids.

The above mentioned objectives can be fulfilled if prior to teaching a particular statistical technique, the teacher introduces a situation for the need of particular technique. This will help the students to understand the need for the particular technique and relevance of statistics to their own discipline. Audio-visual aids like overhead projector, slide projector and LCD's can help in better presentation and condensing information from bulk and thus provide more information in short time.

5. Higher Education and Research in Statistics

Statistics is taught in most of the Indian universities at the post-graduate level. Universities in some states offer both UG as well as PG degrees in statistics. ISI Calcutta with its branches in Delhi and Banglore and Indian Agricultural Statistics Institute at New Delhi are engaged in research and teaching both and so is IASRI. In order to impart training at post-graduate level, leading to M.Sc. and Ph.D. degrees, Calcutta University established the departments of statistics in 1943. The several other universities also established separate departments of statistics and started postgraduate program in statistics. The universities which are well known nationally and internationally in statistical education are Calculta, Delhi, Panjab, Mathurai, Bombay, Madras, Rajsthan, etc.

In addition to these universities Indian Statistical Institute at Calculta, Delhi and Banglore are imparting statistical education at postgraduate level. These institutions and universities mentioned above deserve the major share of credit for the past and current developments of statistical education in India. There is an International Centre located at Calculta where eminent statisticians from every corner of the world are invited to deliver seminar and talks. In Indian Statistical Institutes degree courses of statistics are taught to cover a wide range of subjects somewhat analogous to courses in medicine and engineering. Pure mathematics and theory of probability have important place.

Statistical theory and different branches of applied statistics and economic planning to suit special needs of India form a large part of the teaching program. In

these three groups of subjects mathematics, statistics and economics include great deal of content of knowledge, besides theory and techniques. Facilities are provided to students to become familiar with a number of scientific subjects. Here the emphasis is also given to practical courses like statistical analysis and interpretation of data.

Most of the universities in India award M.A./ M.Sc. degree in statistics without thesis, while in some universities a thesis or dissertation is partially requirement for postgraduate degree in statistics. In ISI and IASRI the postgraduate students in statistics are required to specialize in one of the fields viz. economics, public health, agriculture, business management, marketing and computer applications. Experts in their own settings teach them two papers on the chosen subject. A course ending project work in the real life setting is a part of curriculum at these universities.

In Agricultural Statistics, the most significant contribution of Indians is towards the methods of estimation of crops yield through crop-cutting experiments. In addition to the statistical wing of the Indian council of agricultural research and the Indian Statistical Institute, States Agricultural Universities have also contributed significantly to the development of experimental designs for judging the optimum size and shape of plots through crop cutting experiments. The techniques developed by the Indian Statisticians in areas of Sample Survey, Design of Experiments, Statistical Genetics, Crop Forecasting, Categorical Data analysis and Linear Models have also been adopted by other Asian and African countries.

Adequate sampling techniques have been developed for the estimation of area and crop yields. Methodologies have been developed for the estimation of fruits and vegetables productions that are perishable in nature and require multiple harvesting. Suitable sampling techniques have also been developed for the estimation of fish catch from both marines as well as in land resources that require sampling over time and space.

Advanced research has been pursued in both theoretical and practical aspects of design of experiments. In theory, important fields such as asymmetrical factorial designs, PBIB designs, mating designs and response surface designs including Rota table designs have been investigated extensively during the last four decades. Besides construction, other aspects like optimality and robustness of designs against loss of observations and presence of outliers have also been investigated.

On applied front, aspects of planning, designing, analysis and interpretation of data on agricultural crops have been covered to a large extent. In recent years much research has been done in statistical genetics, in particular, applications of genetic statistics to plant and animal breeding have received attention of many

researchers. Crop forecasting methodologies have been developed for obtaining advanced estimates of crop production on the basis of biometrical characters and weather parameters. In particular, for crops like rice, wheat, cotton, jute, sugarcane etc. forecasting models have been developed using various plant characters measured at different stages of crop growth.

6. Suggestions and Recommendations for Future Planning

a. The time has come to introduce educational programs appropriate for statistics as a fully developed new technology that calls for the utilization of a wide range of scientific knowledge to help in solving scientific or practical problems. As Fisher had pointed out, " a profession statistician, as a technologist, must talk the language of both theoretician and practitioner". The education of a statistician, like that of other technologists, must have a broad base.

b. The challenges in statistics education are never trivial, but offer opportunities inherent to the richness of statistics as a discipline and a servant. University structure should try to encourage competition rather than cooperation among faculties and thus not to reward good service teaching; the juxtaposition of a general decline in quantitative preparation versus increased quantitative needs of business and industry.

c. The awareness in the statistical profession and importance of statistics education is the need of the hour. The commitment of teachers to their students and good practice of statistics will lay excellent foundation for the future. However, the challenge to increase awareness and acknowledgement across the country needs cooperation with mathematical sciences and other disciplines. The future of statistics needs to be discussed in view of recent development in information technology such as data missing, data communication and information processing networks and requirement of end users. The current statistical methodology based on probabilistic models developed for the analysis of small data sets appears to be inadequate and require some methods to be put forward in the name of data mining for such purposes.

d. A statistician must have rigorous training on the analysis packages such as SPSS, SAS etc. During the training periods he may be asked to handle practical problems, case studies or small research projects of applied nature. Training should also include the preparation of layout for field experiments and their actual implementation in the field.

e. Developments in Statistics are beneficial for both the government departments and private industry. Therefore, research scholars and faculty from the universities and research institutions involved in statistical research must be provided with fellowship and research grants from government departments and private industries to the tune of advanced countries.

References

[1] Dansei, B (1998) Using Collaborative Learning Packages to Teach Introductory Statistics at Post- Dansie Secondary Level. *Pro. 5th International Conference on Teaching Statistics,* 125-130.

[2] Mahalanobis, P.C.(1965) Statistics as a Key Technology. *The American Statistician,* 43-46.

[3] Sowey, E.R. (1991) Teaching Econometrics as Though Coherence Matters. *Pro. of Third International Conference on Teaching Statistics,* 321-331.

[4] Sowey, E.R. (1998) Statistical Vistas: Perspectives on Purpose and Structure. *Journal of Statistics Education,* vol. **6(2)**.