PRESENT STATUS OF STATISTICAL EDUCATION IN INDIA

D. S. Hooda and B.K. Hooda
CCS Haryana Agricultural University
India

In the present paper statistical education in India at school, college and university levels is discussed in detail. A brief account of statistical education and research activities at some premier institutions is presented. Steps taken by Government of India to popularize statistics are highlighted. Ways and means for further improvement in statistical education at national and international levels are also suggested.

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, 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 past, Indian Statisticians do have enjoyed a high status at the international level.

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 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. 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 improvement of status of statistical education.

The Indian Statistical System is one of the largest institutional framework, coming with a vast wealth of information, not all of which gets due attention of decision-makers. It is one of the oldest has a large network supported by competent personnel and with an 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:

 (a) Identify its long-term and short term objectives as well as a strategy to achieve these objectives.
 (b) Ensure a structure in terms of job-different ion by recognizing different types of information to be collected and presented the different aspects of information processing for which the responsibility and authority should be clearly vested, possibly in different groups of people.

(c) Establish clear lines of communication within itself to take care of inter-group or collective responsibilities as well as with the environment.

(d) 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.

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. Since then, due to the sustained efforts of academicians and official statisticians the Indian Statistical System have 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. It comprises the CSO and NSSO. The Department is also responsible for all policy matters of the Indian Statistical Institute. 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, the Department of and Statistics and the Ministry of Agricultural some other Agencies like the 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. CONTRIBUTIONS 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 following list of well-known Indian statisticians and their outstanding contributions to the development f 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 distributions and the uniqueness of the likelihood equation.
D.B. Lahiri	Lahiri's Method of sample selection with PPS
P.C. Mahalanobis	Concept of optimum survey design, Pilot survey, D^2 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

It was in the 20th century (in 1939) that Professor P.C. Mahalanobis put statistical practice on a scientific basis in India. He made pioneering contributions to the design of large–scale 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, foodgrains etc., for purpose of management and administration from time immemorial. The very word statistics shows the connexion 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 D^2 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^2_p and D^2_∞ are the Mahalanobis distances between two populations based on p-characters and all possible characters respectively, then under certain conditions $D^2_p \longrightarrow D^2_\infty$ as $p \longrightarrow \infty$.

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 dimension 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. STATISTICAL EDUCATION AT SCHOOL AND UNDER-GRADUATE LEVEL IN INDIA

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 levels are Definition of statistics, Measures of central tendency, Measure of dispersion, Probability, Correlation and Regression. At under graduate level statistics are 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;

(i)   Relevant teaching material be prepared and handed over to the students in the beginning of the course.

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

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

(iv)   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.   INDIAN UNIVERSITIES AND INSTITUTIONS IMPARTING STATISTICAL EDUCATION

Statistics is taught in most of the Indian universities at the postgraduate 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.

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. They are taught two papers on the chosen subject by experts in their own settings. A course ending project work in the real life setting is a part of curriculum at these universities.

6. STATUS OF AGRICULTURAL STATISTICS IN INDIA

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 ICAR and the Indian Statistical Institute, SAU's 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.

In areas of 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.

7. SUGGESTIONS AND RECOMMENDATIONS

The statistical system in the India is comparatively weak and is lagging behind in the timely flow of information required for social and economic affairs. This is mainly due to the existing gaps between and practice and between the means and ends. The validity of the data published by a government office is not questioned. On the hand, statisticians working in universities and research institutes, often, scientists from other disciplines and undertake methodological investigations utilizing the data generated them. They have to work with people from other professional backgrounds to solve practical problems. For instance, a statistician who works with the agricultural scientists must have an understanding of biological phenomena and should develop the ability to coordinate with other people. We therefore, purpose that all statisticians must be trained in diversified areas.

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.

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.