Sankhyā : The Indian Journal of Statistics
1979, Volume 41, Series B, Pts. 3 and 4, pp. 129-137.

PERSPECTIVES IN STATISTICS*

By C. RADHAKRISHNA RAO

SUMMARY. The address refers to the expanding frontiers of knowledge in statistics and the current controversies, and outlines the role that the International Statistical Institute can play in directing future research to make statistics a socially meaningful and viable science. Suggestions are made about the training of statisticians with a proper blend of theoretical knowledge and skill in applications, development of statistical courses for specialists in other disciplines, the role of government statisticians, and the use of computers in statistical research. Some examples are given to highlight the difficulties involved in defining the efficiency of an estimator and the possible dangers in the uncritical use of methods developed by academic statisticians in practical work.

We, the members of the International Statistical Institute and its Sections are greatly honoured by the invitation extended to us by the government of Phillipines to hold the 42nd biennial session in Manila. We specially esteem the honour bestowed on the Institute by His Excellency President Ferdinand E. Marcos and Honourable Imelda Romualdez Marcos graciously consenting to be patrons of this session. We deeply appreciate their messages welcoming us and offering their best wishes for the success of our conference.

It is my special privilege to thank the members of the Honorary, Advisory and Organizing Committees, on behalf of the delegates of this session, for the cordial welcome given to us. The hospitable and historic city of Manila has many attractions, and the Organizing Committee has announced a number of social programs and excursions. I may remind the delegates that we also have a stimulating scientific program and a variety of business meetings which may be profitable to attend.

Dr. Tito Mijares, the Chairman of the Organizing Committee, has done a remarkable job in making the arrangements for the meeting. Having been in his position, not long ago while organizing the Delhi session, I am deeply aware of the responsibilities, and strain and stresses involved in such an undertaking. I have my sympathies for Dr. Mijares, and I am sure you would like me to express our appreciation for all he has done to make our visit to Manila fruitful and enjoyable.

---

*Presidential Address, 42-nd session of The International Statistical Institute, Manila, Phillipines, 4–14 December, 1979.

Key words :

Training of statisticians, mean square error, mean absolute error, James-Stein estimator.

I would like to thank the International Statistical Institute for holding two consecutive sessions in developing countries. We are all aware of the primary need in these countries to strengthen the national statistical systems and improve the data base for socio-economic planning. The presence of a large number of distinguished statisticians at this meeting and the deliberations on a wide spectrum of statistical theory and applications will no doubt stimulate the statistical development in this region.

Statistical science is relatively young compared to other disciplines, but it is a rapidly growing one. There is no field of human endeavour, whether it is regulating our daily lives, managing the affairs of an enterprise, administering a state, planning for the socio-economic development of a country, diagnosis of diseases, designing military operations, forecasting catastrophes or probing the secrets of nature, which does not come under the purview of statistics.

It is inevitable that when demands arise for statistical analysis from such widely different fields, some diversification of interests takes place leading to the creation of different branches of statistics and eventual categorization of statisticians by narrow areas of specialization. The process of diversification is a natural one and, indeed, necessary for the growth of any discipline. Perhaps, in the case of statistics, the diversity was in its very roots; it emerged as a unitary discipline in the first quarter of this century by the fusion of parallel lines of developments in actuarial science, probability applied to games of chance, and theory of errors in astronomical measurements. But as soon as the underlying principle of statistics was perceived as inductive inference by which a coherent body of natural knowledge can be built-up by experiment and observation, it began to branch off rapidly in many directions. There is some concern expressed in the statistical profession that diversification had led to the creation of several subdisciplines within statistics, with little communication between them.

What is the future of statistics in view of the current trend towards possible fragmentation and specialization ? This is the question that is engaging the serious attention of the ISI. What should be done to preserve the means of communication for exchange of ideas between statisticians specializing in different areas of theory and application ? How do we abstract and codify the tools developed in one field for possible use in another ? How do we train statisticians to enable them to do specialized jobs and to equip them with a broadbased knowledge to meet new situations and challenges in the future ?

The ISI has always shown leadership by periodically assessing the needs of the statistical profession and taking necessary steps for the development of statistics. When it was founded about 95 years ago, the main aim of the ISI was in achieving "uniformity in methods of compiling and abstracting statistical returns and in inviting the attention of the governments to the use of statistics in solving problems". At the 25th session of the ISI in Washington, considering the emergence of statistics as a scientific discipline, constitutional amendments were made laying greater emphasis on "promoting the use of statistical methods, the training of statisticians, the encouraging of statistical research and the growth of interest in the advancement of statistics and in the furthering of international statistical relations". Following these changes, the ISI has established up to now separate Sections, with open membership, in four specialized areas : regional and urban statistics (IARU), mathematical probability and statistics (BSMP), survey sampling (IASS) and computing (IASC). This has enabled the ISI to bring under its wings a major part of statistical activities and to provide a common forum for discussion for all categories of statisticians.

From the passive role of reviewing the developments in statistics and arranging conferences, the ISI has launched a major project, the World Fertility Survey (WFS), with direct involvement in the collection of primary data, their tabulation and interpretation. During the last few years, the technicians employed in the WFS have made valuable contributions to methods of data collection, production of computer programs for tabulation and analysis of data and to the development of new statistical methodology for the interpretation of fertility data. The ISI is in the process of launching similar large scale projects and setting up an organization for promoting major research activities on survey methodology and analysis of socio-economic data.

Peter Jacob Bjerve, in his presidential address at the Vienna session, expressed some concern about "a growing gap between theoretical and practical statisticians" and about "problems relating to the increasing specialization in various subject matter areas in statistics". To study these problems and to make recommendations on "bridging the gap", the ISI immediately appointed a Committee on the Integration of Statistics. At the Delhi session of the ISI, a new Committee on Future Directions was formed to examine the broader question of how ISI can play a more effective role in the development of statistics. The reports of these committees have been circulated to all the members, and sufficient time has been alloted for their discussion at this session.

Each one of you has a chance to shape the future of the ISI. It is indeed gratifying to note that the membership has been taking an active interest in the affairs of the Institute and the officers of the Bureau and the permanent office have been prompt in implementing the recommendations of the general body.

For about 40 years, I have been associated with the research and training activities of the Indian Statistical Institute where, under the leadership of the founder P. C. Mahalanobis and advice of R. A. Fisher, attempts have been made to develop statistical training programs and research in an inter-disciplinary way providing a proper interface with other subjects. Based on my experience I would like to make a few remarks on Perspectives in Statistics and indicate the role that ISI can play in the development of statistics as a truly viable and socially meaningful science.

Statistical research has not reached a self generating stage where the statistical problems under current discussion constitute the main source for providing new statistical problems. Statistical research can flourish only in an atmosphere where the existing methodology has been found useful in solving problems in other disciplines and there is demand for the development of further methodology to meet new situations. If this premises is accepted, we have to critically review the present practice of teaching statistics through a set of courses in abstract statistical theory, and also explore avenues for promoting motivated research so that the available talent can be effectively used for solving problems of relevance to the real world.

It is generally agreed and pious resolutions have been passed by several committees to the effect that all statistical courses must have practical orientation. But very few attempts have been made to evolve integrated courses with the desirable blend of theory and applications. Perhaps the task is a difficult one and may involve considerable discussion and some experimentation. I hope the Education Committee of the ISI will give some priority to this project.

While the discussion of practical problems in teaching statistics is essential, it must not be done at the expense of weakening the theoretical and mathematical contents of the courses. No doubt, a statistician should be equipped with the practical skills to enable him to choose the appropriate methodology for analysis of given data; at the same time, he should have the theoretical knowledge to formulate a new statistical problem for study if the existing methodology is inadequate in any given practical situation. The gap

between theoretical statisticians and practitioners may be partly bridged if we are able to produce statisticians with the proper knowledge of theory and flair for applications and give them opportunities to improve their skills during the course of their work.

The bulk of statistical activity in any country will be in the socio-economic sector and will be the responsibility of the government. What role should the statistical officers and government statisticians play in this activity ? Should they merely collect data for possible use by others or take a more active part by also undertaking such analyses of data useful in decision making ? Opinions still differ on these issues, at least in the developing countries. The reasons are partly historical. The role of the Director of Statistics was specified as that of only a "compiling officer" expected to produce data which are "timely, accurate and without gaps". His relation with statistics should be merely "arithmetical", that is, he should not comment on them.

The distinction between suppliers and users of data is an unfortunate one. "The gap between official statisticians and academic statisticians" referred to as the lack of communication between two groups is perhaps the result of different roles they have been playing, one group producing ostensibly useful data and the other producing perhaps inapplicable theory. The situation can be remedied by first recognizing that the arithmetical work of a statistical officer in compiling data constituted only the preparatory stage, and he can render valuable service to the community by assisting the policy makers with proper interpretation of data. To argue that a statistician has no place at the policy making table and that he is merely a "purveyor of facts" is to ignore the whole body of specialized techniques developed in statistics for making optimum decisions on the basis of available evidence and continuously monitoring the results for feedback and control. It is also inconceivable how a statistical office can effect improvements in data and fill the gaps if it is not directly involved in data analysis and problem solving.

It is not suggested that the government statistical offices should be turned into research departments for producing research papers. The desired purpose can be served in a different way. First, a statistical office should develop the necessary expertise to undertake research on methods of data collection and improving the quality of data, to compile meaningful social and economic indicators, and to undertake data analysis to help the policy makers in taking optimal decisions. Second, a statistical office should identify and encourage outside experts to undertake certain types of research useful to the government for long-range planning. The possibility of inviting experts from

universities and research institutes to work in government statistical offices and to send government statisticians to teach in the universities on practical aspects of statistics should be explored. Third, there should be a continuous training program for government statisticians to bring them abreast with the ever expanding statistical techniques and applications.

A government needs good statisticians as well as good statistics. One is not a substitute for the other. The administrators must realize that public policy making is no longer a gamble with an unpredictable chance of success or a hit and miss affair and that it is now within the realm of statistical techniques, whereby optimal decisions can be taken on the basis of properly collected data.

How can the suggestions outlined above for giving a more dynamic role to government statisticians and statistical offices be implemented ? Much depends on the initiative of the Directors of Statistical Offices and how responsible the government bureaucracy is on these issues.

Judging from the current trends of research reflected in published papers and the increase in the number of meetings held to discuss specialized areas of statistics, further diversification of statistics is likely to take place. We may be witnessing in the near future not only a wider variety in the application of statistics but greater polarisation of ideas between different schools of statisticians on what constitutes statistical methodology. We should welcome such developments as conducive to the growth of our subject and its usefulness to society. The problems of identifying the varied fields of application of statistics, recognizing divergent lines of thought and bringing all categories of statisticians to a common platform for discussions will be the responsibility of an international organization like the ISI. Creation of sections, as the ISI has done in recent years is perhaps necessary, but care must be taken to see that it does not lead to sectionalization of the ISI. This may be avoided by developing integrated programs for the biennial sessions with greater concentration on topics of common interest, whose discussion may bring about the ultimate synthesis of different lines of developments in statistics. Discussions of special and more technical topics could be left to intersessional meetings of the Sections of ISI.

We have identified the gap in the making of a statistician (*first type gap*, the need to produce a statistician with a proper blend of theoretical knowledge and practical skills) and the gap between statisticians (*second type gap*, the need to develop communication between specialists in different branches of statistics for exchange of ideas). There is a *third type gap*, interdisciplinary

in nature, which we must not neglect, between statisticians and subject matter specialists. Statistics helps in solving problems, but the problems belong to basic fields of research like Biology, Physics, Economics, Sociology and so on. The success of research projects, therefore, depends on effective collaboration between statisticians and subject matter specialists. Ill designed experiments, imprecisely defined measurements, unwarranted editing of data by investigators, lack of checks and cross checks in the observations, deviations from planned experimental procedures can lead to misinterpretation of data. A scientist should have some appreciation of the statistical approach, and a statistician some knowledge of the subject matter of research to facilitate a discussion between the two on what the relevant issues and possible approaches are, and design experiments to generate data having the desired information. The inter-disciplinary gap is a much wider problem which touches on the education and training of both the statisticians and scientists and I hope the ISI will initiate some discussion on this subject.

I have already made a reference to unmotivated research in statistics with no relevance to immediate practical problems or to general statistical thought. In some cases results are obtained for illposed problems without sufficient understanding of the practical situations. I may refer to the confusion caused by treating the problem of estimation in a decision theoretic set up choosing a particular loss function like the mean squared error (called the Gaussian loss function to indicate the authority behind it). Let us consider the simple problem of estimating σ^2, the variance of a normal population, from a sample of n observations. If S^2 is the corrected sum of squares, then

$$E\left[\frac{S^2}{n+1} - \sigma^2\right]^2 < E\left[\frac{S^2}{n-1} - \sigma^2\right]^2 \text{ for all } \sigma$$

so that the unbiased estimator, $S^2/(n-1)$, is "inadmissible" under the Gaussian loss function. But it is also true that

$$P\left\{\left|\frac{S^2}{n+1} - \sigma^2\right| < \left|\frac{S^2}{n-1} - \sigma^2\right|\right\} < 0.5 \text{ for all } \sigma$$

so that the biased estimator $S^2/(n+1)$ is "inadmissible" if we consider the probability of greater closeness to the true value as the criterion. With the mean deviation (Laplace loss function) as the criterion, both $S^2/(n-1)$ and $S^2/(n+1)$ become inadmissible since there exists a constant c different from $n+1$ or $n-1$ for which $E|(S^2/c)-\sigma^2|$ attains a minimum. Faced with such

a situation it is not clear what justification there is in emphasizing, as is done in many text books, that $S^2/(n+1)$ is a good estimator of σ^2, except of course in the "unlikely situation" where the Gaussian loss function strictly holds. Perhaps different estimates are useful in different situations and a judicial choice depends on a proper understanding of a given practical problem.

The situation becomes more complex when several parameters are estimated using a compound quadratic loss function as criterion as in James-Stein or ridge regression procedures. Dramatic improvements are claimed over the usual unbiased estimators, which are declared as "inadmissible". The results are no doubt theoretically extremely interesting and have opened up new lines of research. However, their use is strictly governed by the limited appropriateness of the compound quadratic loss function in practical problems. Recently, in a problem where the estimates of a given characteristic (say y) have to be found for a number of geographical units periodically over time, the possibility of making the James-Stein modification of the unbiased estimates for the units obtained at each time period was considered. The estimates were meant for use on each occasion for sharing given resources among the units in proportion to the y values. It is known that the James-Stein procedure introduces an upward bias in the estimation of smaller parameters and a downward bias for larger parameters. If such biased estimates are used there is a possibility of units with smaller y values being favoured in the long run at the expense of units with larger y values. This may be desirable but defeats the purpose for which estimates are made. I have given one illustration where caution is needed in "transfer of methodology between academic and government statisticians".

May I take this opportunity to say a few words about the role of computers and their impact on research in statistics. Computers can do 80 million calculations a second (about 6.9 trillion a day). The enormous speed appears to be both a boon and a hindrance to statistical research. It enabled us to undertake the complex computations needed to apply some of the multivariate techniques developed in the past and could not be exploited due to lack of computational devices. On the other hand, it has also encouraged uncritical use of statistical methods through the commercially available computer package programs. It is also beginning to change the character of theoretical research. "Extraction of all the available information from given data", the slogan given by R. A. Fisher, is no longer the guiding principle in statistical analysis. It is thought that what is lacking in sophistication of methodology can be made up by acquiring more data and processing by computers using less efficient procedures. Those of us who have experience of working in applied

areas are aware that data rarely meet the requirements for specified statistical analyses. There are always special features in individual cases which have to be taken into account in the analysis of data and interpretation of results. All these require considerable skill and experience on the part of a statistician. The use of an existing computer program should not be the end of statistical analysis of given data. It can be a good beginning for getting an insight into the problem to make further probes. I am reminded of an aphorism mentioned in a lecture by Dr. David Cox : There are no routine statistical questions, there are only questionable statistical routines. I hope the IASC, our newly formed Section, will give some thought to these problems and provide guidance in the proper use of computers and development of interactive research.

I have mentioned only some of the broad issues which are likely to confront the statistical profession in the near future and have repercussions on the progress of statistics. The responsibility of shaping the future of statistics and preparing the profession to meet new challenges lies with the ISI and its membership. If the enthusiasm of the members displayed in recent years continues, I can predict a bright future for statistics with a high confidence coefficient.

ACKNOWLEDGEMENT

This address is reprinted from the Proceedings of the 42nd Session of the International Statistical Institute with the permission of the International Statistical Institute.

Paper received : December, 1979.