CRISIS OF STATISTICS PEDAGOGY IN INDIA

Madhu S Paranjape
Kirti College, University of Mumbai (Bombay), India
msparanjape@yahoo.com

The daunting task for a Statistics teacher in India at tertiary level, to remain relevant in a world driven by technology and increasing demand for data based understanding of reality is two- fold. First is the challenge of motivating students to specialise in Statistics. This is especially so in an environment where the curriculum remains trapped in extensive mathematical derivations with very little scope for students to handle real data-based problems. Secondly, the teacher must develop her expertise as a Statistics professional through research projects and tool building based on primary data. This needs urgent attention in an era when the application of statistical methodology is rapidly penetrating other disciplines and the statistical expertise is outstripping the boundaries of this discipline. This paper looks at the experience and efforts of the author, on both fronts, in her thirty years of teaching and research in Statistics at the undergraduate level.

INTRODUCTION

The period immediately after India's independence and ushering in of the Five Year Plans may be described as the golden era of Indian statistics. Professor P.C. Mahalanobis, who is regarded as a pioneer in both theoretical and professional statistics, was appointed as the first Statistical Adviser to the Cabinet, Government of India in January 1949. He was the architect of the statistical system of independent India. Professor P. V. Sukhatme, as Statistical Adviser to the Ministry of Agriculture, was responsible for the development of Agricultural Statistics. There was a sizeable growth in the number of institutions teaching statistics, most notable among them being the Indian Statistical Institute at Kolkatta. The Indian statistical system and contributions of Indians to statistics is extensively covered in Hooda and Hooda (2002).

However, the last three decades have seen a decline which is best captured by Professor A.P.Gore, former Head of Statistics department, University of Pune in his letter as President of Indian Society for Probability and Statistics (Gore, 2007). He has said "*The days when statisticians played a leading role in the affairs of our society seem to be over..... All in all, the one common thread that runs through the Sukhatme story appears to be that society at large is far more interested in general issues and difficulties faced by the public than in any technical aspect of statistics. It responds warmly if there is any attempt to analyze societal issues and to devise solutions for them. In contrast, the ambience in a typical statistics course or department is far removed from any kind of reality, let alone a current issue.*"

In this background, in the University of Mumbai (Bombay), it is twenty years since some of us in the faculty teaching statistics at the undergraduate level started to debate on the one hand, the methodology of teaching Statistics; and on the other hand the content of our courses. The growing concern was fuelled by the facts that learning and application of several descriptive statistical methods were available at the click of a computer mouse at first in MS Excel etc in the early eighties and later on in user friendly statistical packages by the beginning of nineties. Secondly, these methods were in extensive use by students of other disciplines like Psychology and Geography who handle real data and context. The students of Statistics too used data, but it was from structured problems and ironically they never got the feel of real data.

PRESENT SETUP

After the Higher Secondary level (Grades 11-12), the students in our University, spend three years in an undergraduate course. A student, opting for three electives at the first year, majors in one of them at the third year. The concepts of frequency tables and diagrams/graphs are introduced at the Secondary school levels that get repeated at the Higher Secondary level. The same are revisited by students opting for statistics as an elective at the first year of the undergraduate course. The syllabus framed by University is common for all the constituent (affiliated) colleges. University's Syllabus committees for various courses design/ modify the syllabi that have to be adopted by all colleges. *Thus there is no scope for individual colleges/ departments/ teachers to incorporate changes in the content and structure of courses.*

As the students move from the basics to specialisation, they grapple with rigorous derivations in Probability distributions, Regression analysis, Sampling theory, Design of Experiments, Statistical Inference etc. The laboratory hours test their skills in using stated methodology for solving numerical examples. In case of topics in Applied Statistics such as Reliability, Biostatistics–Genetics, Epidemiology etc too, the students need to master extensive mathematical derivations. Although students learn sampling theory in detail, including principles of constructing Questionnaires and Schedules, they rarely get the opportunity to design a survey and analyse primary data. *Even more disturbing is the fact that almost all Statistics teachers share the lack of experience of conducting even a single sample survey or carrying out analysis of real data in their entire teaching career.*

This handicap appears most stark when researchers from other disciplines approach a Statistics teacher with some problems in their data analysis or when persons from small scale industry need some statistical help in testing differences between techniques/machines/products. Very often researchers in Humanities or certain Social work groups undertaking socio-economic studies of specific communities require assistance with data analysis. For a teacher of Statistics, even an informal association with such studies offers challenges to sharpen one's understanding of the context and to suggest appropriate methods of data analysis and statistical inference. Unfortunately, due to rigid compartmentalization of subjects and disciplines such interactions between academia are rare. This author had opportunity to associate with social scientists conducting real data based studies on two occasions: socio-economic status of women from a minority community and survey of households of the rural poor. Grappling with appropriate statistical tools for their data analysis that necessitated interplay of qualitative and quantitative features was an enriching experience.

TACKLING PROBLEMS IN A DEVELOPING COUNTRY

A developing country like India faces many challenges in the field of higher (tertiary) education especially in the State funded institutions where the majority of the students enroll. The resource crunch in education has prevented hands-on use of computers by students in various disciplines including Statistics, except those doing a course on Computer Systems/ Computer Science/ Information Technology. In case of Statistics courses, the laboratory hours have to make do with scientific calculators, draw charts manually, etc. In the absence of adequate computing facilities, introduction of a statistical package at the undergraduate level becomes pointless.

In order to confront head-on the daunting two-fold task of remaining relevant in a society that is increasingly demanding data-based understanding of reality, and of motivating students to specialize in Statistics, attempts were made through discussions in the association of Statistics teachers and in the syllabus committees to introduce component of hands-on –data experience and training for the students. However, such attempts were always put in cold storage due to expected rise in financial commitment of the State/ Institutions as well as due to a perception among teachers –arising though out of their own lack of training–that such activity would be unmanageable and difficult to assess for performance appraisal of students.

This situation was the compelling factor for this author's decision in the early nineties, to introduce students to hands-on survey techniques, through simple projects. Over the years other colleagues also got involved and our department has built up a rich collection of primary data-based projects with varied themes. Some of these are:

- Utilisation of college library
- Survey of patients admitted to Public hospitals
- Television (T.V.) Channels–popularity of programmes among college students
- T.V. Viewing habits among students within the college
- Opinion poll on Politics, Sports, Internet and General Issues
- Preferences for soft drinks- popular brands
- Extent of mobile usage
- Utilisation of pocket money
- Distribution of travel time and mode of transport to college
- Shopping patterns at Malls.

The students are informed that though these projects are not part of the curriculum, participating in them would enhance their overall understanding of the subject. In each project based on primary data, mainly involving students at First/Second year level, the class builds a consensus on the theme, followed by a discussion on the questions, sampling frame and sample units. Using a single common Questionnaire, each student selects a random sample of 5-6 sample units; thereby the class gets a random sample of about 150-200 sample units. Simple analysis is carried out using a spreadsheet such as MS Excel. The report is generally prepared by the guiding teacher and is printed in the annual magazine of the college. Our experience shows that getting involved in a project of this kind, gives the students an exposure to handling context based real data and makes learning of various techniques more meaningful. Illustration of the project on "Extent of mobile usage" carried out by the First year students, in the academic year 2008-09 is given below.

ILLUSTRATION

During classroom discussion, a Questionnaire was constructed to elicit information on various characteristics of usage of mobile (cell) phones from a sample of students on the college campus like: whether the student uses a mobile phone and if so the brand of the handset, whether the student's family has a telephone, the billing options, monthly charges, monthly income of the family, number of family members possessing mobile phones and total monthly expenses of the family on mobile usage.

The class collected data on 100 students from different classes and faculties. Three groups were formed and each group was assigned the task of classifying and tabulating information on specific questions. During a laboratory session, the class generated simple frequency tables and 2x2 cross tables; created graphs using MS Excel; calculated Yule's measure of association and tested significance of association using Chi-square test.

RESEARCH EXPERIENCE

While training the students, it is imperative that the teacher too develops her expertise in survey and data analysis techniques. This requires handling research projects based on contextual primary data and building relevant statistical tools. An Index is an important tool to measure variations in socio-economic data and makes it an important input in policy making. References in media (newspapers) to UNESCO's Human Development Indices (HDI) for judging comparative development of countries, is a case in point.

At their post graduate level, students of various social science disciplines like Psychology, Geography learn to use Index as a tool of comparative analysis. However, building an Index as a tool to measure variation is never taught to students specialising in Statistics. This handicap continues to hamper them in their profession as teachers of Statistics. As a result, when required, they find themselves ill-equipped to offer a critique of the statistical tools in use in a particular context.

This researcher has over a period of time developed Questionnaires, devised research designs and conducted sample surveys to capture employment patterns of the educated, in particular graduates. Moreover, an Employability Index (EI) has been developed, based on HDI methodology, to compare inter and intra- group variations in employability (Paranjape, 2007). In this study present income, job-mobility (the number of jobs that a sample unit may have switched) and unemployment period (period for which a sample unit remained unemployed from the time one started looking for work) are considered as characteristics of employability and therefore are inputs in the construction of the EI.

The EI is a weighted average of three indices that collectively distinguish between the employability of different groups of individuals. These indicators are Index of Unemployment Period (U'), Index of Income / Emoluments (E') and Index of Job–Mobility (M'). The general formula for each of these indicators is given as

$$X' = \left(X_i - X_{min}\right)/\left(X_{max} - X_{min}\right);$$

where, X_i is the value of variable X for i^{th} sample unit and, X_{max} and X_{min} refer to the highest and lowest values of X. Then the Employability Index EI = $\sqrt{\left(E' + M'\right)/2U'}$.

Additionally, rates of return for graduates have been estimated for comparative cost-benefit analysis of graduates, separately by region, gender, faculty and socio-economic background. These rates are utilised to critically review the application of Rate of Return approach advocated by the World Bank in matters of policy decisions in the field of higher education in developing countries.

At present, this researcher has undertaken a micro level study based on a sample survey of households to examine the sharp divergence between perception and reality with reference to human development and substantive employment in a district of Maharashtra.

DISCUSSION AND CONCLUSION

It would be unfair not to report here the encouraging instances of more systematic efforts made in this regard by some teachers in Pune University, most notably by Professor A.P.Gore. Here the component of a Project is introduced in the syllabus of Statistics. Moreover, Project competitions are organized for students in their final year of graduation or those pursuing a Masters' course in statistics. Among the most inspiring contribution of Dr. Gore and his colleagues to the teaching of statistics is the compilation of over 100 real life data sets together with brief descriptions and suggestions for analysis (Gore et. al., 2006). In some universities, teachers conduct statistics quiz competitions, essay competitions etc. A few years back, our department organized a "job fair" for interaction between statistics students from neighbouring colleges and several organizations offering jobs, mainly those in Market Research.

However these efforts are few and far between. While they do help to motivate a small group of students to take up statistics for further study, they have not impacted the overall pedagogical practices in statistics education in the universities. In sharp contrast with this scenario is the extensive use of quantitative techniques like Operations Research in the courses offered by Management schools that have mushroomed in India during the last decade. Secondly, in various knowledge-based industries, careers are emerging that require statistical applications but graduates of statistics are not equipped to take advantage of these opportunities. This is mainly due to the incongruence between handling of real data by non-statisticians and over emphasis by statisticians on extensive mathematical -derivations based courses.

The lack of available expertise has resulted in institutes outside the university system, offering highly specialized courses in Clinical Trials, Biostatistics, and Applications of GIS etc at very high fees; which are beyond the reach of most students.

The teaching of statistics in India is at a crossroads. The rising expectations among students in order to become competitive in a knowledge driven job market are at loggerheads with the ossified syllabi in most institutions. A ray of hope as noted above lies in the path charted out by number of statistics teachers working independently in various universities. This optimism is supported by the National Statistical Commission (NSC) set up by the Government in July, 2006 on the recommendations of a Commission constituted to review the statistical system and the entire gamut of official statistics in the country. The NSC is to serve as a nodal and empowered body for all core statistical activities of the country, evolve, monitor and enforce statistical priorities and standards and to ensure statistical co-ordination among the different agencies involved. The current challenges offer opportunities for rejuvenation of teaching and research in statistics.

REFERENCES

Gore, A. P. (2007). Why statistics has lost its central role in societal affairs in India and What can we do about it? http://ces.iisc.ernet.in/hpg/nvjoshi/isps/apg-isps-june2007-roleofstat.doc.

Gore, A. P., Paranjpe, S. A., & Kulkarni, M. B. (2006). *100 Data Sets for Statistics Education*. University of Pune. http://stats.unipune.ernet.in.

Hooda, D. S., & Hooda, B. K. (2002). Present Status of Statistical Education in India. IASE. http://www.stat.auckland.ac.nz/~iase/publications.php?show=1.

Paranjape, M. S. (2007). Determinants of Employability of Graduates: A Case Study of University of Mumbai. *The Indian Journal of Labour Economics, 50*(2).

Paranjape, M. S (2008). Rate of Return to Higher Education in India: A Case Study of Graduates of University of Mumbai. *Journal of Educational Planning and Administration, XXII*(4).