CONVOCATION ADDRESS

by Jerzy Neyman,

University of California, Berkeley, U.S.A.

PROBLEMS OF YOUNG PEOPLE PREPARING FOR RESEARCH WORK IN STATISTICS

Chairman Haksar, Professor Rao, Professor Mukherjee, Mrs. Rani Mahalanobis, Ladies, Gentleman and, last but not least, the Young Graduates :

On this solemn occasion when you, the Young Graduates, turn a leaf of the history of your progress in scholarship, it is a pleasure to offer you my hearty congratulations and my wish of More Power To You. Also I wish to offer you a few items of advice, based on my own experience.

As you continue in your careers, you will have many problems. First, you will have to work for your living and meet such duties as will be imposed on you in whatever jobs you will have. I am sure that you will perform to the best of your ability. In addition, however, you, as young scholars, will have some particular research problems of your supreme personal interest on which to work in your spare time. The identity of the "problem of supreme personal interest" is a very subjective matter. However, a little circumspection and the exercise of some will power might be helpful. I will tell you something about my own experience.

The next thing to think about in the present stage of your development is the acquisition of tools for your research in the future. As I shall illustrate on some examples, the tools in statistical research, whether purely theoretical or, so-called, applied are a variety of mathematical disciplines, not only those that are commonly taught right now, but also some of those that are currently being developed by our brothers, pure mathematicians. Quite frequently, opinions are expressed that the latter disciplines now in the making are too esoteric to be used by statisticians. As I will illustrate on some examples, these opinions are shortsighted. The mathematical disciplines that now appear esoteric, with their new jargon, may well be of common use at the time you are ready to become leaders in statistics. If you do not acquire these disciplines now, the future of your scholarly maturity may well be reduced to performing routine functions. Yes, the latter are useful and must be performed. But, young as you are, why not reach for leadership ? I wish I could translate to you adequately a relevant line of a Polish poet. Approximately, it is as follows :

"The Young—Fly above the ordinary levels !"

All I just said are generalities. Let me now proceed to some specifics.

Selection of some specific "own problems"

Do this with some circumspection. Fifty-one years ago, I was somewhat in your present position. I earned a Ph.D. degree at the University of Warsaw, Poland. My thesis consisted of a paper on "my own" problem and of a few other papers that originated from the job I had. Just as was the case with Professor R. C. Bose. I was a pure mathematician. While Professor Bose might be described as a number theorist, I was passionately interested in set theory, a different discipline, but equally purest of pure mathematics. In the paper I produced on my "own problem" [1], I proved that point sets and intervals in one dimension have a certain remarkable property. Also I proved that sets and intervals in more than one dimension do not necessarily possess the property in question. Now, isn't this remarkable ? At that time I was fascinated by this result and felt convinced that everyone ... well, perhaps with some exceptions, will share my enthusiasm. Actual developments proved different.

As far as I know, of all the millions of people on this planet only two exhibited some interest in my findings. One was Professor W. Sierpinski, my chief Ph.D. examiner, who seemed to have liked my theorems. The other gentleman with similar thoughts was a Dr. Moore, professor in one of the universities in the south of the United States. Some years after my Ph.D. while I was in England studying mathematical statistics, possibly in 1925, I received a brief letter from Professor Sierpinski advising me to have a glance at three recent issues of the Bulletin of the *American Mathematical Society*. I did look them up and found three articles by Professor Moore. In the first Professor Moore contended that my theorems are false. In the second he said that these theorems are correct and produced new proofs. The third article indicated that my theorems can be improved. As I said, to the best of my belief, no one else ever noticed my results.

What is the conclusion ? This seems to be that (a) my theorems do have something that might interest people with an appropriate mentality (not a very common one !), and (b) that the problem itself was, so to speak a fringe problem, not in any way influencing large scale developments either in set theory or in any other scholarly domain. On the other hand, the parts of my thesis on problems I picked up at my job as a statistician at the Agricultural Research Institute in Bydgoszcz, Poland, proved to be different. Here, in due course, I managed to pick up subjects that did have some consequences on the development in the mathematical statistics in which I became very deeply involved emotionally. And the moral ?

The moral for you is to look around carefully when selecting your "own problem." Read the relevant literature critically, listen to other people and, if need be, be prepared to alter your initial preoccupations.

Applied statistical work : routine and innovative

When, in connection with some research in science, be it astronomy, biology, meteorology, etc., I use some *technique* established earlier, and I emphasize the word "technique," I perform routine work (to my regret, with very frequent mistakes in arithmetic, etc.). This kind of work is useful and, in fact, necessary. But it is not very inspiring.

The routine work just described must be clearly distinguished from innovative statistical work in science. To illustrate my point I will describe to you, very briefly, two outstanding applied statisticians of our epoch. One of them is David Kendall of Cambridge University, England, and the other Herbert Robbins of Columbia University, New York. Many people are likely to disagree with me and call the two gentlemen theoreticians, having little to do with applied statistics. It's rather too bad, but I persist. I rather think that the disagreement is due to semantics. What my opponents call applied statistics I call routine statistical work.

I am impressed by the innovative applied statistical work of Kendall because (a) it is so broad, and (b) it is so successful. Suffice it to say that it extends from the study of the chance mechanism governing the proliferation of living cells (perhaps bacteria) to the mechanism of interpenetration of cultures, tastes and styles of successive generations of people who lived in the very distant past. I have in mind the problems of archeology, to determine that this burial place is more ancient than the other, etc.

With a degree of imagination and some skill, it is not very difficult to construct what we call a mathematical model of a natural phenomenon. Thus, when thinking of proliferation of living cells it is not unusual to compare it to the so-called birth-and-death stochastic process. Not infrequently, I do so myself ("routine work") and I do not mean to blame the others. However, I have a great respect for Kendall for his asking the question, and for investigating it, whether and to what extent the actual proliferation of cells conforms with the then customary theory. Undoubtedly in cooperation with a biologist, Kendall studied the matter, and got results. The time intervals between births of cells and their subsequent division into two daughter cells behaved in a manner incompatible with the birth-and-death process.

Too bad ! However, if not a birth-and-death process, then what might the actual process be ? Kendall worked on the question and, in due course, arrived at a novel stochastic model which fits the observations better [2]. In due course, this model of Kendall, or its possible improvements, will generate a statistical "technique" to be used routinely. Thus far, however, Kendall's results are only rarely familiar to classical applied statisticians.

The process of questioning and then testing an accepted theory, and then of producing a more satisfactory one is what I call "innovative" work in science. In order to be able to do such work, the applied statistician must have a mastery of the relevant mathematical discipline, a mastery comparable to the rich mathematical tool box of David Kendall. Also, the applied statistician must have scholarly initiative and talent.

Now about Herbert Robbins. His mathematical tool box is very rich, like Kendall's, but very different in character. Here again many colleagues in the profession will insist Robbins is a theoretician, far away from applications. I disagree. To me Robbins is a red-blooded applied statistician with a mathematical equipment much better than that of his contemporaries. Of the many brilliant achievements of Robbins, I will mention two. One is that, in about 1955, he managed to invent a method, now described as "empirical Bayes' procedure", to incorporate into the empirical statistical studies something that is vaguely described as "earlier experience." In so doing, he solved a problem which baffled statisticians for something like two centuries and continues to baffle right now. Robbins' feat is magnificent and I like to talk and to write about it [3], but this would be much too long for the present address.

The other feat of Robbins, also strictly "applied," is something very new. This is a design of sequential clinical trials intended to diminish the frequency of giving a patient an inferior treatment. The work, still in progress, involves very delicate mathematics, but can hardly be treated other than distinctly applied statistical research of excellent style.

My hope is that in the future some of you will develop to do innovative applied statistical work comparable to that of David Kendall and Herbert Robbins. What to do about it now ? One answer is : acquire the necessary mathematical tools. Unfortunately, tools alone will not do the trick. In addition to tools you have to have talent. But some people do have talent, why not you ? See that your talent is not wasted !

How to acquire a suitable mathematical tool box

A specific answer to this question is impossible. The point is that it is impossible to predict now which of the continually proliferating and developing novel mathematical disciplines will be particularly relevant to statistical problems not now formulated, of which you may become aware, say, in 10 or 20 years from now. One thing that is sure is that these will be disciplines that are not covered in the mathematical courses customarily offered in the universities. Your difficult job is multiple. On the one hand, you have to consume and to put into your blood stream and bones all the classical mathematical material taught in your school. Next, or in parallel, you have to familiarize yourself with novel ideas, not yet in text books, but discussed in journal articles and in monographs. Some conversations with productive and broad-minded scholars are likely to be very useful. The only specific suggestion that I can make with assurance is for you to do what I myself do when in need. I go for a chat with Professor L. M. LeCam, one time my student and now my respected colleague and friend. But LeCam is far away, in Berkeley, and all you can do here is to read his writings (a substantial book is forthcoming !). While we are on this subject, I suggest that whenever you see books by such as W. Feller, T. Harris and M. Loeve, from the West and as B. Gnedenko, A. Kolmogorov and Yu. Linnik, from the East, to name a few, read them avidly and do your best to own them.

However instructive is the reading of good books, personal conversations are needed with someone here at the ISI. My stay in Calcutta has been too brief to be sure, but some conversations I had with Dr. A. Maitra suggest that his opinions might be helpful.

How lucky you are !

In addition to familiarity with appropriate novel mathematical disciplines, your success as productive statisticians depends very much on contacts with important scientific research. Ordinarily, in the universities that I know, including the University of California, such fruitful contacts are not readily established. True, they do occur from time to time, usually on the initiative of a substantive scholar and I benefited from them. However, occurrences of this kind are infrequent and when they do materialize, all I can do is either engage in some cooperation or say sorry, I am not interested. I cannot choose among several possibilities. Contrary to this, here at the ISI, through the wise direction of Professor Mukherjee and through the broad activities of Dr. Rao, all of you are exposed to a great variety of research projects going on immediately under your eyes, occasionally with your own active participation and generally with the participation of your immediate colleagues.

This arrangement within the ISI, being partly a school with a post-graduate programme, and partly a multisubject research institution, offers you opportunities for contacts with applied research that are not paralleled in any other institution I know of.

So, be alert, appreciative of the institution you are in, study "your own problems" and interesting novel mathematical disciplines, and enjoy life !

Good luck and more power to you !

References

1. Jerzy Neyman, "Sur un theorame metrique." First published in ***Fundamenta Mathematicae***, Vol. V. (1923). Republished in *A Selection of Early Statistical Papers of J. Neyman*. Cambridge University Press and University of California Press, 1967.
2. David G. Kendall, "On the role of the variable generation time in the development of a stochastic growth process." *Biometrica*, Vol. 38 (1948), pp. 316-330.
3. J. Neyman, "Two break-throughs in the theory of statistical decision making", *Rev. Intern. Statistical Institute*, Vol. 30 (1962), pp. 11-27.