THE PLACE OF THE NATURAL SCIENCES IN A STATISTICAL COURSE[1]

Extracts from the lecture delivered on the occasion of the inauguration of new degree courses, 16 August, 1960

J. B. S. HALDANE

This Institute is today starting a course in statistics leading up to a degree. There is nothing unique in that. What is unique is that, besides statistics and mathematics, we are going to teach the natural sciences, and some of the social sciences to statisticians. I am in charge of the teaching of the biological sciences, and hope to teach some chemistry. So I am vitally concerned.

Now why should a statistician learn anything about science? There are two reasons. First, some knowledge of natural science is just as much part of general education as the knowledge of history in the narrow sense. An Indian who has never heard of Asoka or Akbar is uneducated. But the history of India goes back far beyond Mohenjodaro. The Deccan trap is the record of one of the most terrible series of disasters which we can imagine, in which a huge area was repeatedly buried under red-hot molten lava, and every living creature on it destroyed. The great plain of the Ganges valley and the Punjab was formed comparatively recently. An arm of the sea south of the rising Himalayas was filled up with silt from them. Probably the area round Lucknow was once very like that round Calcutta before men had made canals and bunds, and a little earlier was like the Sunderbans today. To realise this, however dimly, is an important part of a liberal education. We cannot adjust ourselves to the great changes taking place around us unless we know a good deal, not merely of what science has achieved, but what it is likely to achieve. This is equally true if we are traditionalists who regard change with abhorrence. We cannot resist it unless we understand its causes.

But a statistician has other reasons to study science, and our course will be designed primarily to meet the needs of statisticians. What are these needs? Some of our students will find posts under the Central and State governments, some in the private sector. A few will take posts in the Institute, and others will go to universities. It is my earnest hope that not all of these will teach statistics, but that some will become physicists, geologists, biologists, and so on. We must, however, think primarily of the needs of those who are going to work in government offices or for firms, since they are the majority. Both of these groups may have to deal with numerical data arrived at by scientific methods.

Now Professor Mahalanobis' desire is that our scientific course should be integrated. My emphasis is on its being quantitative. We are not quarrelling. The shape of a tree is, at present, a matter of what may be called pure botany. The moment we ask about its size, why, for example, it can grow to 20 metres, but not to 50, we are involved in fairly elementary physics, which are

[1] From the Souvenir published by Indian Statistical Institute Alumni Association on the occasion of celebration of the Golden Jubilee of Indian Statistical Institute.

not very different from the physics involved in the design of tall buildings. The moment biology or geology becomes quantitative it necessarily becomes integrated with physics, chemistry, statistics, or all three. One can read a textbook of zoology without learning why no birds are as heavy as a cow or as light as a bee. I hope nobody will go through our course without doing so. I hope that we shall instal into our pupils a feeling for the meaning of numbers applied to anything whatever. I want our students to say 'nonsense' if they are told that the area of forests in Madhya Pradesh is 8 lakhs of square kilometers, that a lakh of people died of cholera in India in 1959, that the density of lead is 45 times that of water, that the temperature in a blast furnace is 400° Centigrade, that a certain sample of pig iron contained 6% silicon, and so on. This does not mean that they have to memorize the correct figures, but to get a feeling for numbers. If such a feeling were widespread in Delhi, I doubt if, to take one example almost at random, the total area under sal in Part 'C' states in 1954-55 would have been given as 22,504 square miles when the total of the columns concerned was in fact 2,504. This is no doubt a misprint. But I find it a little hard to believe that between 1953 and 1954 the area of the Punjab under sal forests fell by 15%. However, both these statements are found in 'Indian Forest Statistics', issued by the Ministry of Agriculture. Such publications certainly make the task of a biologist who takes planning seriously rather difficult.

A feeling for numbers can only be acquired by practice. What should the practice be? As an example I want our students to make a census of all the trees in the compounds of numbers 202 and 203, and even, with the Director's permission, of 204. They will come up against real difficulties. Is this a tree or a bush? Is this one banana plant or a dozen? This is no harder than deciding what is a factory or a household. I estimate that there are rather under 100 betel-nut palms in the compound of No. 203. I may be wrong, I haven't counted them. But I want our boys and girls to get the feel of what a hundred trees look like, and constantly to be asked 'How many?', 'How much ?', 'How often?', 'How powerful?' and so on.

A statistician in a government office may have a deal with figures on a vast variety of topics. Whenever he gets a set of figures he ought to know at least roughly what they are about. My colleagues are busily planting a garden with the more important Indian crop plants. Our students, when they leave us, may have to deal with figures about jowar and bajra. I dare say a number of Bengalis have never seen either, and Punjabis have never seen jute or tapioca. They will; and they will learn at least a little about them. I doubt if many of our students know the exact meaning of such words as anticline, bauxite, a reversed fault, and monazite. But if they have to deal with mineral resources they will need to.

A statistician concerned with quality control in a factory requires a much more specialized knowledge; but at least we hope that any one of our graduates will have made some physical and chemical measurements not too unlike those whose results he has to analyze. If he has not, he may either take the analyst's results as something not to be questioned, or regard them as so much nonsense which he is bound to tabulate. The first thing to ask in such cases is how accurate the results are. The analyst may say that the mean difference between duplicate measurements is under one per cent. I got an answer of this sort a few months ago. It turned out that the worker in question had merely measured some photographs twice. He had not taken the photographs twice. So the error measured was perhaps only a small part of the total. Unless our statisticians have made measurements of this kind themselves, they will not be able to improve the performance of

the experts. Of course they will live much quieter lives if they do not ask awkward questions, until their firm goes bankrupt.

Now the inevitable criticism of such a course as we are proposing is that it will be hopelessly superficial. It is true that our graduates will be unable to pass a B.Sc. examination in zoology, botany, or physiology in any Indian university, or its equivalent in any British one. It is also true that by the third practical exercise in the course which I have designed, our students will be doing something which is not part of any biological course in India, and is a part of very few elsewhere; a systematic study of variation in plants and animals. Later, I hope, they will do a good deal more quantitative biology of various kinds. It is most unlikely that a gold medalist of an Indian university in any of the biological sciences would secure a 50% mark in our biological examination.

The reason for this is two-fold. In the first place, in Indian universities at the moment, students who choose a biological course must give up the study of mathematics, not to mention statistics, at an early stage. This means that graduates in the biological sciences are automatically debarred from most of the types of research which would be of value in developing our agriculture and husbandry. Secondly, while in England the biological teaching lags about fifteen years behind research, the corresponding lag in India is about fortyfive years. The growing points of European biology today are no longer the study of form with a microscope or otherwise, but the quantitative studies of animal and plant physiology, animal and plant populations, and animal behaviour. These do not require much higher mathematics. One can use the ordinary methods of statistics without knowing point set theory, just as one can use $\sqrt{2}$ or π without having mastered Cantor's theory of irrationals. But they do require a mathematical outlook.

Biology is a science in which innumerable methods can be applied. In the U.S.A. there is a tendency to use as expensive methods as possible; for example electron microscopy, measurements of nuclear magnetic resonance, and so on. These are needed to solve certain problems. But there are equally important problems requiring no more apparatus than a foot rule, a cheap balance, and a stop-watch; and, as I shall show later, some of these are more important for India. In our biological course we shall ask for very little apparatus; and if we use advanced physical methods, they will, I hope, be the methods of electronics, in which my colleagues on the ground floor are experts. We shall also, I hope, do better than university teachers of zoology (though not of botany) in familiarizing our students with the common animal species around them, many of which are of great economic importance, either by eating crop plants, or by eating the eaters of crop plants.

I have strong views on the teaching of chemistry, which, I believe, should be quantitative from a very early stage. But I do not wish to impose my views on my colleagues. As for other subjects, I have sometimes acted without consulting my superiors. Some of you may have noticed a graduated post in the tank of No. 203. I hope this will give us a graph of the annual variation of the tank level; though unfortunately the most interesting part of the annual change, the rapid rise of level during the first month of the monsoon, will occur during vacation. I consider it of the utmost value that our students should make quantitative records of such simple phenomena as changes in water level, which are, of course, of immense importance in India.

During the three years that I have been a member of the staff of the Institute I have been building up a small biological research department. In doing so I have had five principles in mind. First, the work done should require statistical methods at least one stage, and the more statistics the better. Second, it should not overlap in excellent work done by my other colleagues. It should not even appear to do so. Thirdly, it should require little imported apparatus, and should, therefore, serve as a model which could be copied by biologists all over India. Fourthly, it should, so far as possible, have practical applications, or look like having them. One of the founders of the Royal Society three centuries ago wrote that it 'valueth no knowledge, save as it hath a tendency to use', and I have kept this principle before me. Fifthly, it should be intelligible to students without a highly specialised knowledge of biology.

... ...

I give an example of co-operation with a biologist elsewhere. A scientist of the Saha Institute for Nuclear Physics in Calcutta, has been photographing bactoriophages with an electron microscope, and measuring the photographs. His paper in 'The Nucleus' aroused my attention, because I happen to have a taste for numbers, and 10 Angstrosms and 100 Angstroms mean something nearly as different to me as 10 metres and 100 metres. He was kind enough to give me further details. I will remind you that a bacteriophage is an organism which is or is not living according to the criteria one adopts, which parasitizes bacteria, and is much too small to see with ordinary light. We find that the average length of the so-called head of his strain is 964 Angstroms, but the standard deviation is only 22-3 based on a sample of 400. These results are startling for three different reasons. They are the smallest biometrical figures even obtained. A human skull is about two million times as long as a bacteriophage head. Secondly, the coefficient of variation, 2.3%, is not significantly above the lowest value known to me for any matrical character in a living organism, and is far lower than most. Thirdly, 22.3 Angstroms is only about 14 times the diameter of a carbon atom. And a good deal of the observed variation was certainly due to errors. It seemed possible that the bacteriophage was standardised not as an organism, but as a molecule.

The scientist concerned knows little of statistics, and had not seen the full interest of his discoveries. I have therefore arranged for him to collaborate with one of our research scholars, in analysing his results as thoroughly as possible, and perhaps tracking down some of the sources of error. I hope that we shall turn out at least one student a year who can help any Indian biologist with any problem requiring statistical treatment. Many of the so-called statisticians employed by agricultural research stations have not the needed qualifications.

I consider that it should be an important part of the education of our students to follow the research being done around them. It is only so that they will see how a need for statistical treatment arises in any particular case, and how the most appropriate methods are or are not chosen. These are only samples of the work which our students will see being undertaken. Some of it is just as interesting statistically, but I have mentioned research of general interest. I hope that my physical and chemical colleagues will also carry out research which will be equally, or even more, stimulating. I know that my colleagues on the economic side are doing so.

Not only will our students spend some time at Delhi, Giridih, and so on, seeing aspects of applied statistics which cannot be studied here, but in Calcutta at least an afternoon a week is to be devoted to visits. I think the most important of these will be visits to factories and other establishments where statistical quality control is practised. But I shall certainly take parties to the Botanical Garden, which is world famous, and the Zoo, which deserves to be; and I am trying to make arrangements for visits to hospitals. We must certainly go to the Indian Museum for palaeontology, and I venture to hope we may also go there for more humanistic studies.

Teaching is a kind of activity, requiring a different kind of organisation from either the collection and interpretation of statistics, or the ordinary kinds of scientific research. It is quite essential that this should be understood. Let us take such a simple matter as punctuality. If the same punctuality is expected from research workers as from computers, many kinds of research are impossible. An experiment may last for 18 hours; if so, two such experiments a week, with some preparatory work and calculations in the intervals, are all that a man can do. If he or she must sign books between certain hours, as my colleagues have to do, such work cannot be done. A computer should be fairly punctual, but no harm is done if he arrives ten minutes late; and if he finds a mistake which takes him some time to track down, he had better for. be late than inaccurate. But in a teaching course both teachers and students must be punctual within a minute or so, not only in starting work, so long as the demand for lecture rooms is high, in ceasing it. This is not the case at the case present.

... ...

... ...

To sum up, the Institute is embarked on a great adventure. If it succeeds, we shall set an example, not only to India, but to the world. But it will require effort and even self-sacrifice. I for example, see no prospect of doing serious research, even of a theoretical kind, in the next two years. Unfortunately, some workers in the Institute appear to be making deliberate efforts to prevent this course involving them in any increased work or even any change of their habits, and are therefore, in effect trying to wreck it. They may succeed. If so they may have very serious adverse effects on the intellectual and economic development of India. There are others, both in the RTS organization, and outside it, who are making every effort to get our scheme to work, even though this annoys some others. But we are already behind schedules in regard to various facilities for teaching. For myself, I can only promise to do my utmost to fulfil my duties to the students whom I welcome today.

The late Professor J.B.S. Haldane first visited the Indian Statistical Institute in 1951. Thereafter he visited it four times and after his fourth visit in 1957-58 he stayed on till 1961 as a worker of the Institute. The curriculum of the B.Stat. degree course started in 1960 was mainly based on his idea of a unified science and statistics course. Prof. Haldane died in Bhuvaneswar in 1964.