STATISTICAL TOOLS IN RESOURCE APPRAISAL AND UTILIZATION*

by P.C. Mahalanobis

I am glad to have the opportunity to speak at this Conference. As time is short, I shall only try to put across briefly a few leading ideas.

I have come here to sell statistics to resource engineers and economists. I hope you will find me an honest broker. Two grounds on which I plead for statistics, I may mention at the outset. First, we are bold enough to claim that the statistical method supplies the real foundation of inductive logic or inductive inference in all scientific reasoning. Secondly, at a more concrete level, the aim of the statistical method is to secure the maximum economy of cost in the appraisal, utilization and conservation of resources.

First, the abstract aspect. In deductive logic, one starts with certain premises from which the conclusions inevitably follow. The counterpart in modern thought is pure mathematics which Bertrand Russel, the mathematician and logician, defined as the class of all propositions of the form "P implies Q". That is, if P, then Q must follow. There is complete formal certainty of the deductive conclusions; but we must not enquire what P and Q are, or whether they are real in any mundane sense at all. This is pure mathematics.

Classical physics was based on a deterministic logic. An invariable sequence of cause and effect is in many ways similar to the formal-deductive process in pure mathematics. The classical method is to isolate one factor at a time and write down the equations of motion of a single particle. Prediction becomes, in principle, absolutely certain. Following, step by step, by mathematical reasoning what happens to the particle, it is possible to predict the future. Newton in his great treatise on the principles of natural philosophy adopted this mathematical form. The construction of a rational model of the whole universe was, no doubt, a great triumph.

However, when one has to deal with not one, or two, but three, or more, particles, the equations of motion become unmanageable, and the next steop in the history of the physico-chemical sciences was the study of not one single but an aggregate of particles or molecules. This led to the kinetic theory of gases and thermodynamics. The object now was not to ascertain with complete certainty the motion of each single particle separately, but to investigate the collective properties of a large number of particles taken together. Again there were great triumphs of theoretical conclusions corroborated by experiments. A parallel development led to the emergence of the concept of entropy; and the principle of increase of entropy was recognized as nothing but the increasing degree of randomness of the physical universe. The growing complexity within the field of the physico-chemical sciences thus led to the gradual replacement of the deterministic-mathematical model by the probabilistic-statistical view.

A further shift to a statistical model occurred with the development of statistical and quantum mechanics culminating in the enunciation of the uncertainty principle by Heisenberg which denied the possibility of absolutely certain knowledge of both the position and the motion of a particle at any given instant. This deliberately sets a limit on the deterministic chain by openly admitting the existence of residual uncertainty. In the physical sciences, the whole foundation thus changes and becomes fundamentally of a probabilistic or statistical nature.

At this stage, I ought to mention the concept of the random sample which supplies the theoretical foundation of modern statistics. The tossing of a coin supplies a simple example. It is quite impossible to predict anything about a single toss of a single coin, that is, whether it will turn up head or tail. However, if the coin is tossed a large number of times, it is practically (but never absolutely) certain that heads and tails would turn up approximately in equal numbers. The expectation of such roughly equal frequency of heads and tails essentially refers to the whole set of throws, that is, the prediction belongs not to any single individual throw but to a group or assemblage of throws taken together. Secondly, the relative proportion of heads and tails will inevitably fluctuate from one throw to another so that it is never possible, even in principle, to make any absolutely certain prediction. And yet, although prediction is never absolutely certain, it is possible to estimate the limits of uncertainty.

The position is similar in the case of physical observations and measurements. However careful the observer may be, even the simplest measurements, for example, of the length of a rod, when repeated, have been always found to vary. The average of a number of repeated measurements usually becomes more and more steady as the number of measurements is increased. The deviation from the average, i.e., the "error" of each individual measurement is sometimes positive and sometimes negative and behaves like heads and tails in the tossing of a coin. Fluctuations or variations are thus essential features of all measurements. This is true not only in the physical but also in the biolo ical and social sciences.

This is the crucial point. All knowledge in physical, biological and social sciences is ultimately based on measurements and observations. Every set of measurements is characterized by variation. In fact each set of measurements constitutes only one out of many possible similar sets. Also, the totality of all possible sets constitutes the "population" or "universe" under study. Each set of measurements is thus essentially a sample of the "universe". In order to reach valid conclusions about the population or the universe, it is necessary that the sample should be representative of the universe. In statistical language, the condition for such representativeness is supplied by the fundamental concept of randomness.

The aim of statistical theory is thus to reach general conclusions about the population or the universe on the basis of the sample. And, as all scientific knowledge is based on "samples" of observations and measurements, statistical theory supplies the only valid method of making inductive inferences.

So much for the statistical method in the abstract. I may now make a rapid survey of the scope of the statistical method in the concrete. We may try to construct a mental map, and start by representing pure mathematics by a geometrical point at the centre. A small circle round the centre can represent classical physics in which the form is mathematical but knowledge is, in fact, based on measurements and subject to errors of observation, and hence amenable to statistical treatment.

A second larger circle may represent the wider field of kinetic theory of gases, thermodynamics, and statistical mechanics. In this region of physico-chemical sciences, the factors of variation are often amenable to a large degree of control and the classical method of isolating and studying one single factor at a time is usually available.

We next come to the field of biological (and social) sciences which can be represented by a third larger circle. In 1900, Karl Pearson coined the word "biometry" for the methods appropriate to this region. It is no longer possible to isolate and study one single factor at a time. For example, in agricultural experiments to compare the yield of different varieties of a crop, there are innumerable factors which cannot possibly be isolated such as the fluctuation of soil fertility from one plot to another, root competition, the influence of weather conditions, etc. An altogether new approach became necessary, and was made by R.A. Fisher, just about a quarter of a century ago, in the development of the design of experiments and the analysis of variance. He explained quite clearly.

"No aphorism is more frequently repeated in connexion with field trials than that we must ask Nature few questions, or, ideally, one question at a time. The writer is convinced that this view is wholly mistaken. Nature, he suggests, will best respond to a logical and carefully thought out questionnaire; indeed, if we ask her a single question, she will often refuse to answer a single question, she will often refuse to answer until some other topic has been discussed."(1)[1]

Fluctuations and variations are known to be large and complex. The object now is not to eliminate them, but to allow them scope to come into play in a balanced fashion so that valid inferences may be

[1]Numbers within parentheses refer to items in the bibliography

drawn from the experimental observations (2). Variations are themselves of great importance, and can be used as yardsticks. In fact, once we succeed in measuring the chance variations, it is possible to reach valid conclusions with statistical and probabilistic rigour (but not with mathematical certainty).

Experience has shown the great advantages of the statistical design of experiments in economy of effort and time, and in adaptability not only in agricultural field trials but in technological experiments and resource appraisal of all kinds. For example, suppose we wish to study different methods of treating wood for its better preservation. There are many factors of variation, temperature, humidity, pressure, the use of different chemical agents, etc. The strategy of the statistical method here is to design the experiment in such a way that we may measure the effect of the different factors operating at the same time. Large developments have occurred during the last two decades, and biometrics are finding increasingly fruitful applications in biology and genetics; agriculture and forestry; education and psychology; and the medical and social sciences.

Variations always occur in the process of manufacture. You can not turn out two screws or two rods absolutely indentical in size or other specifications. The problem is to maintain the quality (or output) of the manufactured articles within permissible limits of variation. My friend, Dr. W.A. Shewhart, who has spoken on the subjec tackled this problem in the Bell Telephone System. So long as the fluctuations remain stable, production is under statistical control. This was used by Shewhart to develop the method of control charts (3) When fluctuations go beyond the control limits, there is clear indication of change in the conditions of production which is usually of great diagnostic value (4). "Statistical Quality Control" and allied methods created a kind of revolution in the efficiency of production in the USA during the war. In fact, I have been assured, and the evidence is there, that the large-scale war effort of this country could not have been possible without the help of such statistical methods.

The industrial system does not consist entirely of machines. The human factor is also there, and is a source of large variations. Here also logic indicates, and experience has shown, that statistical methods can be used most effectively in operational research and investigations dealing with the physical, psychological, and social conditions of work.

I may now come to another region, that of sampling surveys, which we may demarcate by drawing a fourth circle in our diagram. Factors of variation are now even more complex, and are not subject to experimental control. This is the field where traditionally the method of the exhaustive census has been used for a long time. In recent years, during the last 10 or 12 years, there have been important developments in the use of the sampling method.

The driving motive in the first instance was economy. The United States of America can afford to conduct an agricultural census which would aim at enumerating every single farm. But in a country like India this is simply unthinkable. In India, it was the overwhelming need of economy which led to the use of sampling methods. In crop production, for example, the experience in India has shown beyond dispute that one can get results with a margin of error not exceeding 2 per cent at a cost of one-tenth or even one-twentieth of the cost of an attempted complete enumeration. Dr. Harrison has referred to the efforts now being made by the FAO to organize a world-wide agricultural census in 1950 or 1951. In many countries, the only possible line of advance will be by the use of the sampling method.

Speed is another important factor. In a sample survey, both the field work and the tabulation of the data can be finished very quickly. In a complete count, the results are usually available when they have already become out of date or when the interest or the possibility of practical applications is mostly gone.

The advantage is not merely in lower costs or greater speed, but what is most remarkable, a sample survey when properly conducted has been usually found to be more accurate than an attempted complete count. The reason is very simple. In a sample survey, only two or three hundred workers are often quite enough against perhaps fifteen or twenty thousand required for a complete count. It is naturally possible to have better trained and more experienced investigators, better inspection, and all-round better quality of work in sample surveys so that we can get much better results. This has been the experience in India, in the USA., and also in the USSR, as far as I could gather from my colleagues on the UN Statistical Commission.

The sampling method has another great advantage. In a sample survey, it is always possible to calculate valid estimates of the margin of uncertainty. With a proper design, it is also possible to estimate (and hence often eliminate) fluctuations arising from investigator bias and other causes. In fact, the survey can almost always be arranged in a way to supply a direct measure of its over-all reliability. For example, consider an area, possibly a whole country, over which a sample survey is being conducted to ascertain, say, crop acreage. It is possible to throw a set of random points on the map and make a physical inspection of the crops at these points. Suppose these points are marked in green on the map, and suppose just one per cent of the land is investigated in this way. We then get a result based on what we may call the green sample. We can at the same time, throw at random another set of points which we may mark in red on the map and which cover just another one per cent of the land. The green and the red points are like two inter-penetrating clouds, and from each we get an independent estimate. The two estimates are not identical, but if they are in good agreement, we may reasonably conclude that we have got something which is objective.

From each sample, it is also possible to estimate a valid measure of the margin of error, and compare the difference between the two sample results with the margins of error of the difference. Such a comparison would show immediately whether or not the sampling was done under conditions of statistical control. In principle, the sampling method thus has within itself the possibility of ascertaining its own margin of error and hence assessing its own reliability. In a complete count, on the other hand, if any items are missed, or other mistakes are made, we can never know; that is, the margin of uncertainty is completely unknown.

Again, in a sample survey it is not only possible to ascertain the level of precision, but it is also possible to prepare the design of the survey in such a way that the desired precision can be attained at a minimum cost (5).

I have briefly indicated the advantages of statistical sampling. I should also mention that the three methods, the design of experime statistical quality control, and sampling surveys are inter-related and supplementary, and used in combination, can lead to great economy and efficiency in the survey of resources, not merely in the sense of cartography, but in the appraisal of the quality of the material, and also how one can control the utilization and hence the conservation of available resources. Statistical sampling (together with statistical experimentation and statistical control) is in fact finding increasing use in such diverse fields as agriculture, demography, commerce and industry, and economic and social studies of all kinds (6).

I may now turn to still another field which we may demarcate by drawing a fifth circle in our diagram. This is the area of free observations in which factors of variation are neither amenable to control, nor to experimentation, and are not even subject to sample surveys. The only thing to do is to undertake a patient collection of observations, and a painstaking investigation of possible statistical connexions between different factors of variation. Although experimentation is not possible, when significant relationships are discovered, it is possible to make predictions and compare such predictions with subsequent observations. This is the field in which statistical correlations and the analysis of time series have been used with great success, for example, in economic and business statistics, the study of weather and river records, or mortality data.

I may give one or two examples. Consider the control of floods. In India, in the province of Orissa, there was a catastrophic flood in the Brahmani river in 1926. The question arose whether the bed of the river had changed. A committee of expert engineers, after a careful study, reached the conclusion that the bed of the river had been raised by about 4 feet. In order to give the same protection from floods, the engineers recommended that flood embankments should be raised by at least 4 feet at a cost of many millions of rupees. This was the conclusion reached by ordinary methods of appraisal; but a fundamental difficulty was that you cannot directly measure the level of the river bed as you cannot have any bench marks.

An entirely different approach was, however, possible with the help of statistical methods. Detailed analysis revealed a close correlation between the rainfall in the catchment and the height of the river. The statistical evidence showed that heavy rainfall in the upper reaches of the river had caused the very high flood (which appeared to the engineers to be of a catastrophic kind). Purely on a statistical basis, the advice could be given with confidence that there was nothing wrong with the river, and it was not necessary to spend millions of rupees to raise the height of the embankments. I gave that advice in 1930. If a mistake had been made, catastrophic floods would have occurred and swept away large portions of the country during the last 20 years (7). This was a practical demonstration that statistics could save a lot of money; and in India, saving 40 or 50 million dollars was not a negligible affair.

I may give another example of a more abstract nature. The motion of the moon gives rise to tides in the sea and the rivers. The motion of the moon must, of course, also cause a tidal effect on the earth's atmosphere. That is due to the motion of the moon, the pressure of the atmosphere would undergo a cyclic change, and also the temperature (due to adiabatic heating up). The effect is naturally very small. Actual calculations show, for example, that the temperatu would rise and fall with an amplitude at certain places of .008^{0}C, that is, less than one-hundredth of a degree, due to the lunar tide. The amplitude of the cyclical change of barometric pressure in certain places would be less than one-thousandth part of an inch. Variations in temperature and pressure from hour to hour or from day to day are enormously larger in comparison. Yet if one has a long series of records, these gross variations, although enormously large, would cancel out and the effect of the lunar tide should be capable of being ascertained by purely statistical methods. This is exactly what Professor Sydney Chapman, F.R.S., of the University of Oxford and his associate workers have done. They have shown by the detailed statistical analysis of a long series of records extending over 50 or 60 years that the gross variations cancel out, and the observed effects of the lunar tide on the earth's atmosphere agree satisfactorily with the calculated values (8). This is a striking example of the great power of the method of correlation analysis.

In this brief review, my aim has been to indicate the wide scope for the use of statistical methods in science and technology. At the abstract level, the aim of statistical theory is to reach general conclusions (about the "universe") from a knowledge of the particular (i.e. the "sample"). At the concrete level, the aim is to collect relevant information (or extract such information from available data) with a view to choose, on a probability basis, the "best" out of two or more possible programmes of action. The statistical method thus supplies, both intheory and in practice, the true logic for making decisions on a scientific and objective basis.

At one end, in classical physics, it is often possible to isolate and study one factor at a time, and the statistical method consists primarily in the adjustment of observations with the help of the classical theory of errors. With more complex systems, the degree of control decreases, and the statistics of assemblages becomes more important as in the kinetic theory of gases, thermodynamics, and statistical mechanics. In the field of biometry, factors of variation become still more important and are inextricably mixed up and cannot possibly be studied in isolation. The strategy changes, and the aim becomes to study more than one factor at the same time with the help of appropriate designs of experiments, analysis of variance, and statistical correlation. A further extension is the use of control charts, and the methods of statistical quality control. Beyond this lies another wide field in which the most fruitful line of advance is by the use of sampling surveys (9). The degree of control is small, but predictions on the basis of statistical sampling are still possible and capable of being corroborated or refuted by subsequent observations. Finally, there is the field of free observations where conclusions can be drawn (and predictions, subject to verification, can be made), only on the basis of detailed statistical correlation and analysis.

Over the whole field, the statistical method supplies the basis of uncertain inference (as distinguished from the absolutely certain deductions of formal logic and mathematics). The margin of uncertainty in statistical inference, in principle, is ascertainable; and its magnitude depends on the degree of control to which the factors of variation are amenable and on the available quantity of information. The use of statistical method can be, therefore, commended with confidence to all persons interested in resource appraisal, utilization and conservation.

Bibliography

(1) R. A. Fisher, "The Arrangement of Field Experiments", Journal of the Ministry of Agriculture, (United Kingdom), vol.33, 1926, page 511.

(2) R. A. Fisher, The Design of Experiments, 1935, 4th edition, 1947, Oliver and Boyd, Edinburgh.

(3) W. A. Shewhart, The Economic Control of the Quality of Manufactured Products, 1931, Macmillan & Co.

(4) W. A. Shewhart, Statistical Method from the Viewpoint of Quality Control, 1939, Graduate School of the United States Department of Agriculture, Washington.

(5) P. C. Mahalanobis, "On Large Scale Sample Surveys", Phil. Trans., Royal Society, London, vol.B231, 1944, pages 329-451.

(6) P. C. Mahalanobis, "Recent Experiments in Statistical Sampling in the Indian Statistical Institute", Jour., Roy. Statistical Society, London, vol.109, 1946, pages 325-378.

(7) P. C. Mahalanobis, Report on Rainfall and River Floods in Orissa. Submitted in 1930 and printed by the Government of Bihar and Orissa.

(8) S. Chapman and K.K. Tschu "The lunar atmosphere tide at twenty-seven stations widely distributed over the globe", Proc. Royal Society, London, vol. A.195, 1948

(9) F. Yates, Sampling Methods for Censuses and Surveys, 1949, Charles Griffin & Co. Ltd., London

---

*Proceedings of the United Nations Scientific Conference on the Conservation and Utilization of Resources 17 August - 6 September 1949, Lake Success, New York, Volume 1, Plenary Meetings, pp. 196-200.