INFORMATION SCIENCE AND ITS CONNECTION WITH STATISTICS

TOSIO KITAGAWA
KYUSHU UNIVERSITY

1. Introduction

The purpose of this paper is to give an exposition of the general scope of information science and to make clear its connection with statistics. It is expected that this will be helpful in finding directions of active development of the latter. For this purpose we must develop several preliminary considerations in order to secure the understanding of the readers on the reasons why the author wishes to present the topics on such an occasion. By information science we mean a newly organized branch of science which has at least two characteristic aspects. The first aspect is shown by the fact that it is presently drafting a form of blueprint of its own future development. It is indeed a science planned in its scope and thus is sharply distinct from the natural growth which has been experienced in a majority of areas of pure sciences. The second characteristic aspect comes from its construction in which constituents range over a vast ensemble of individual sciences, and it is apparent that information science is an amalgamated science whose constituent branches have their respective scientific principles. In short, information science is a planned, consolidated, and integrated science having several different branches, and because of these two characteristic aspects, its methodology ought to be explained in some detail.

We shall explain our points of view on these two characteristic aspects, both from our experience and from somewhat more logical considerations. In giving our explanation we have to refer to various topics which will range over physics, biology, electronics, and so on, none of which seems at first sight to be an appropriate topic of the *Symposium on Mathematical Statistics and Probability*. Moreover, we do not intend to give a report on established results of some scientific area, but are merely trying to give our ideas on how to organize a new field of science. In spite of these two unusual circumstances, we intend to refer to a topic which has a definite connection with the development of statistics in the coming days.

2. The road leading to information science

The author has been working in the area of successive processes of statistical inference and control [37]–[45] for the fifteen years since 1950. The image of information science which we shall expound in the following sections is indeed, for the author, one of the natural directions of extension of the research areas which can be imagined from such developments.

We have explained in our previous papers [41], [45] the connection between our statistical theory and cybernetics in the sense of Wiener [78]. The former is concerned with successive processes of statistical inference and control, whereas the latter is concerned with communication and control in men, animals, and machines. In spite of the differences between the scopes of these two topics, their logical aspects have a deeper connection than might be imagined at first appearance. The intimate connection comes from the following circumstances.

(1) The theory of communication and control discussed in cybernetics is based on a mathematical apparatus which is nothing but that of stochastic processes. Indeed the science of communication so far developed in cybernetics by Wiener is essentially statistical science ([78], p. 10).

(2) The theory of successive processes of inference and control is concerned with an automatically controlled sequence of statistical procedures, as we have explained in detail in one of our recent papers [45]. Indeed, the theory is a cybernetical formulation of statistical inference and control whose logical aspects were explained in another paper of the author [41].

These two circumstances will be sufficient to explain the connection between statistical theory and cybernetics. Now it is also important to refer to the differences between them. The two crucial differences which have a certain connection with the birth of information science are the following.

(a) Cybernetics is a prototype of integrated science, whereas the theory of successive processes of statistical inference and control belongs to the area of statistics, although its fields of application cover various areas of science and technology.

(b) The theory of successive processes of statistical inference and control deals with learning from experience in a somewhat broader sense than that developed in the first 1948 edition of *Cybernetics.*

The road which has led us to a certain formulation of information science was suggested by our struggle to reach a higher level at which these two differences (a) and (b) can be "aufgehoben."

The crucial problems confronting us in this connection are the following:

(i) whether or not a methodological science can be developed in an efficient way without some systematic plan of connections with substantive sciences;

(ii) whether or not either cybernetics or the theory of successive processes of statistical inference and control can efficiently contain the theory of learning processes and self-organization in their respective frames.

Keeping these problems in mind, we shall enter in the following section into the more fundamental discussion regarding the topics of three categories of sciences. Indeed, in the area of statistics a long sequence of disputes on whether this science is methodological or substantive has never been settled. It is one of the byproducts of this paper that we can and shall be in a position to give a discussion, from a broad point of view, on this problem which has stood for three centuries.

3. Three categories of sciences

The following current classification of sciences into three categories has been adopted very frequently, and it seems to us to be at least convenient to start with the following description.

Category 1. *Object Science.* The sciences of this category have their respective assigned objects of research but no assigned methods of research. Their scientific information as a whole cannot be expected to be organized in a deductive form by one methodological procedure. On the other hand, they are always concerned with the same object. In this sense scientific information on objective sciences can be mostly descriptive.

Category 2. *Elementary Science.* These sciences have their respective assigned objects of research, and moreover, they have a certain unified methodology of attack on their assigned objects of research. Their scientific information admits an intention of research workers and scientists to organize them, at least partly, in a deductive form. It is one of their objectives to search for principles of how to organize, as far as possible, at each stage of its development, their information in a deductive form.

Category 3. *Methodological Science.* This type of science has no assigned object of research, but it has an organized information pattern which can be applied to certain sciences belonging to the first and the second categories. Its job may sometimes be to abstract a common feature of various methodologies developed in the first and the second categories, but its main function is to develop a unified methodology on its own systematic foundation with an intention to establish scientific achievements of its own with consequences applicable to other categories of sciences.

Cybernetics in the original sense of Wiener ([78], 1st ed.) is manifestly intended to be a science of the third category, and it has been understood as such by a majority of people. Statistics is understood by some statisticians, including most of the mathematical statisticians, to be a science of the third category, whereas other people working in substantive sciences, including economics, still insist that the science of statistics cannot be understood merely from its methodological aspects and that statistics cannot exist in itself apart from coexistence with sciences of the first category. Mathematics can be recognized to be one of the purest and most extreme cases of the third category. Logic is manifestly the methodology to be valid in every science. However, it so happens that one can dispute whether or not logic should be considered a science. Some may insist that logic is not an individual science, but rather a super entity without which no theoretical scientific reasoning can be secure. This viewpoint is indeed intimately connected with the definition of science in which experience must be accumulated through observation and experimentation. It is our purpose in this paper to collect these fundamental methodologies within broadly defined science families and then to compare them with each other in order to make clear the distinctions among them. This is in itself the principle of collection and differentiation which can be considered a fundamental method of logic. By such a method, we shall be able to make a systematic approach to analyzing learning processes. Indeed, our spectrum contains logic, mathematics, statistics, and cybernetics belonging to the third category of science in its broad sense.

4. Methodological aspects of statistics

In a previous paper of the author [41], logical aspects of successive processes of statistical inference and control were discussed in some detail. The three main logical aspects are concerned with (i) objectivity, (ii) subjectivity, and (iii) practices. Each of these consists of two fundamental principles, and hence the total logical aspects are concerned with six fundamental principles:

(1) Objectivity
  (a) principle of probabilistic scheme,
  (b) principle of deterministic scheme,
(2) Subjectivity
  (c) principle of strategy,
  (d) principle of valuation,
(3) Practices
  (e) principle of efficiency,
  (f) principle of successiveness.

The background upon which such a system of six principles was proposed is manifestly influenced by the particular studies of the author on successive processes of statistical inference and control, and cannot be claimed to cover the whole statistical activity of human beings and human society. Nevertheless, it is the constant pursuit of the author to adhere to the characteristic aspects of successive processes of statistical inference and control and then to endeavor a generalization of various ideas established in specific restricted areas to more general aspects of different areas in statistical research and statistical activity. In this sense data analysis was discussed in one of the papers of the author [41]. According to the author [41], data analysis does not belong to either descriptive statistics or statistical inference theory.

The decision function approach has its validity in some restricted area, but it cannot cover all the aspects of successive learning processes unless it should become equipped with modification procedures for cost functions, and the methods of pattern recognition under which it is formulated. In discussing data analysis in statistics the author [41] placed an emphasis on its learning procedures, and it is the consequence of these discussions that some fundamental aspects of learning processes in data analysis have several common features with those under which successive processes of statistical inference and control are discussed.

In 1963 a paper of the author [45] discusses the fundamental aspects of automatically controlled sequences of statistical procedures and their implications for statistical programming to be used in automatic digital computers and for statistics in its general sense. This series of papers [41], [44], and [45] since 1960 is entirely concerned with an elaboration and a generalization of the ideas developed in the series published in the Fifties. Therefore, we are always conscious of the fact that our logical considerations have been prepared to give a rearrangement of what we have obtained in our specific approaches and to give a bird's-eye view of the whole domain so as to reach a better understanding of what area we have covered and how we have obtained our results, in an expectation that further developments can be searched for leading to a still better understanding of what we have done. Having reached the standpoint of our 1963 paper, we are now very conscious of our distance from the classical theory of R. A. Fisher [24], [25], and from that of Neyman-Pearson [56], [57]. So much has been left undeveloped in an area which we pointed out to be worthy of cultivation. It is, however, the present opinion of the author that the undeveloped area of statistical science can be more realistically scheduled for development only after establishing an understanding of the science of information which should be deeper than anything we have hitherto achieved.

5. Methodological aspects of cybernetics (1)

A more direct and deeper understanding of the science of information and control may be expected to be more easily obtained in the realm of cybernetics than in that of statistics. It is true that the logical principles enunciated in the previous section were derived from our studies and reconsiderations of specific approaches on successive processes of statistical inference and control, but we believed that they were identical to the logical principles to be expected in the formulation of cybernetics, as we have explicitly emphasized in section 4 of our paper [41]. In short, our approach may be termed the "cybernetical formulation of statistical procedures." With this affinity and similarity in mind, it is still true that cybernetics has at least one more aspect which statistics neither covers nor is responsible for. This is the fact that cybernetics attempts to be a theory of communication and control in real entities, that is, in men, animals, and machines. That is to say, cybernetics has the responsibility of explaining and discussing the real phenomena in real existence. Therefore, it is quite natural to expect to find there some techniques, some ideas, and some achievements which we do not have in statistics but which can be carried over to the formulation of statistics, particularly in reference to learning procedures. It has turned out, however, that statistics has not benefitted much from cybernetics, at least in its early stage. This has left some feeling of disappointment toward a science on which high hopes had once been placed for help, or at least guidance in developing understanding of learning processes in statistics.

The reason why cybernetics has not shown achievement commensurate with the expectations of the Fifties derives from its methodology.

Fruitfulness of a science belonging to the third category can be expected only to the degree that it cooperates with the other sciences belonging to the first and second categories.

In the search for effective progress in the area of statistics we have now reached a notion of information science from which, we believe, we can obtain aid and guidance. At the same time, statistics is one of the indispensable constituent elements in a consolidation of various scientific methods in information science. Statistics has its reason for existence, but it can flourish and further develop by maintaining the intimate connection with information science which we are now going to explain.

6. Structural aspects and constituents of information science

Information science aims at being a unified scientific approach integrating various phenomena connected with information; that is, production of information, transmission of information, transformation of information, storage of information, deduction of information, pattern recognition due to information, retrieval of information, various operational uses of information, and so on. These information phenomena can be found in biological existence, social lives of human beings, as well as in machines manufactured by human beings. A science of information which treats information phenomena dealing with common aspects irrespective of particular details of real existences is naturally one of the sciences belonging to the third category in the sense of section 3. By this assertion we deny the idea that any science of the third category is an abstract theory which can be treated by symbols and notations such as in mathematics and in some of mathematical statistics. On the contrary, we are now proposing a science of information which has a certain structure, consisting of the following five indispensable research branches.

Branch 1. Physical information phenomena.
Branch 2. Theoretical formulation of information phenomena.
Branch 3. Information system analysis.
Branch 4. Information phenomena in biological existence.
Branch 5. Artificial realization of information phenomena.

Branch 1. This branch contains the following research divisions.
1.1 Metal and magnetic elements.
1.2 Semiconductor elements.
1.3 Cryogenic elements.
1.4 Optical information elements.
1.5 Fluid information elements.
1.6 Elastic wave information elements.
1.7 Chemical information elements.
1.8 Dielectric information elements.

Branch 2.
2.1 Logic.

2.2 Theory of recognition.
2.3 Theory of linguistics.
2.4 Theory of information networks.
2.5 Self-organizing systems.
2.6 Learning theory.
2.7 Theory of information transmission.
2.8 Information theory.
2.9 Mathematical programming.
2.10 Statistical theory.

Branch 3.
3.1 Information processing systems.
3.2 Computation systems.
3.3 Information control.
3.4 System analysis.
3.5 Operational analysis.
3.6 Man-machine systems.

Branch 4.
4.1 Neurophysiology.
4.2 Integration of the central nervous system.
4.3 Sensory information processing.
4.4 Transmission of genetic information.
4.5 Exhibition of genetic information.
4.6 Adjustment of genetic information.

Branch 5.
5.1 Information circuits.
5.2 Information transmission apparata (transducers).
5.3 Information transformation apparata.
5.4 Recognition apparata.
5.5 Language apparata.
5.6 Thinking apparata.
5.7 Adaptation apparata.
5.8 Education apparata.

Branch 1 is devoted to research on physical materials and phenomena associated with these materials, which have or will have some connection with information phenomena. Results and achievements obtained in Branch 1 have been and will be the basis of developments in Branch 5. Realization of information phenomena by virtue of cooperation between Branches 1 and 5 have been and will be a stimulus to a new theory of information in Branch 2, and will provide powerful experimental apparata to research in the area of Branch 4. An introduction of new information apparata will involve new problems regarding information systems to be discussed in Branch 3, and sometimes even revolutionary effects will be experienced as we have already seen in the case of electronic computers.

Branch 1 is responsible for the creation of new materials as carriers of information phenomena and as the basis for information apparata through which biological information phenomena can be investigated and by means of which new mathematical models can be introduced.

Branch 1 will contain various researches on various matters and on physical and chemical phenomena, including solid matter, fluid, elastic waves, optical, and chemical phenomena.

The role of Branch 2 is to prepare various theoretical models through which information phenomena can be described and analyzed. Its research domain will cover theoretical investigations on control, learning, self-organization, statistics, and programming, and hence it should also contain fundamental research on recognition, foundations of language and of mathematics. It is also to be noticed that the theory to be developed in this branch is responsible for giving theoretical foundation of automation, naturally with the cooperation of the other branches of information science.

Remarkable advances have been recently experienced in this area, and revolutionary effects are now anticipated in epidemiology, logic, linguistics, and psychology. More direct influences have already been observed in mathematical theories regarding statistics, computation, and control.

In view of recent remarkable progress, it may be almost impossible to overestimate the tremendous and far-reaching effects of information science in the coming ten years on social and biological sciences. On the other hand, it is frankly admitted that the social and biological sciences that are currently somewhat less developed are so because of the lack of development of information science in these areas, in spite of the evident facts that information plays the crucial role in their formulations. These are important problems which will be facing us in the near future.

The objects of research in Branch 3 are various information systems with information apparata which are to be invented and investigated in Branch 5. Human societies cannot continue to keep their various activities in sound condition without adequate systems of information. For instance, we can mention here production control, inventory control, economic planning, and so on. Among these information systems there can be found a set of fundamental information systems such that any information system is a composition of some of these fundamental ones. In the present state of our social and academic activities we can mention the following set of fundamental information systems: (i) statistical data processing, (ii) computation systems, (iii) documentation systems, and (iv) information control systems. Branch 3 also contains research on systems analysis, operations research, and man-machine systems as its essential constituent research techniques. Researches and problems in this area of Branch 3 will have direct and strong influence upon mathematics and statistics. It will also prepare various powerful techniques for the development of economic science and management science.

Branch 4 is responsible for investigations regarding information phenomena encountered in the biological sciences. There are various biological phenomena to be investigated in the light of information science, that is to say, through uses of theoretical information models and with the aid of information science techniques. On the other hand, results established regarding biological information phenomena will give rise to many problems to be discussed in Branch 2, and hence strong stimuli for development of theories in information science can be expected to come from Branch 4. Without deep investigation of information systems in the biological realm some characteristic aspects of biological phenomena cannot be made wholly clear, however far their material and energy aspects may have been investigated. In this sense Branch 4 will constitute one of the fundamental divisions upon which essential progress in biological science will depend.

Branch 5 is devoted to the development of artificial information apparata. At the present stage the central problems in this area are concerned with artificial intelligence, including various functions such as pattern recognition, translation of languages, information retrieval, adaptive control, learning procedures, and even thinking, in a certain sense. In such investigations Branch 5 will depend on the investigations of all other branches, and conversely, the achievements obtained in Branch 5 will provide powerful tools for investigations in other branches.

In short, each of the five branches constitutes an indispensable research department in information science, and it will give scientific or technical aid to every other branch directly. Hence, all these five branches are connected through mutual aid and cooperation. With any lack of such aid or cooperation, a science of information cannot be expected to realize its potential of sound and efficient development. With this idea in mind, let us make a series of surveys of various well-established sciences which are closely connected with information science. We shall do this in the next few sections.

7. Imbedding in and correlation with information science

Let us now turn to examining the actual character of various sciences which are, each in its own way, closely connected with the notion of information science described in the previous section. We shall then find two different situations: (i) imbedding and (ii) nonimbedding and correlation.

(1) There are many examples of sciences which are imbedded in the framework of information science. By an imbedding we mean here that the constituents of the science belong to one of the five branches enumerated in the previous section.

(2) There are again many examples of sciences, each of which has constituents belonging to one of the five branches but also has an area that is not covered by any branch of information science. These are the cases of nonimbedding and correlation.

We shall see that cybernetics and bionics can be recognized as sciences imbedded in the framework of the five branches, and so are examples of situation

(1). On the other hand, there are many other examples illustrating situation (2). Let us mention here a few examples: behavioral sciences, mathematical sciences, medical electronics, human engineering, and so on. The following few sections are devoted to analyzing these situations in some detail, with the idea in mind that such analysis can be useful in clarifying the role of information science in motivating scientific research of various kinds. These analyses will involve an objective observation on the present state of each individual scientific activity, and then there will follow naturally the task of classification of sciences. But this is not all that is to be done. We are rather seeking a strategy for promoting scientific research. We are not only a geographer who wants to prepare a map for everyone's use, but we are rather a mountain climber aiming at reaching the top of a mountain or a navigator going on the ocean with the intention of searching out a new continent. In this light there is an essential element that comes into our discussion. This is the subjective attitude of research workers regarding methodologies and motivations in their particular research areas. Each individual research worker in a team of research workers, and even a group of research workers itself, has a certain standpoint according to which a particular emphasis will be placed upon its activities. An objective description of research activities by a systematic analysis of mutual relationships has given us the structure of information science in terms of the five branches listed above. This much corresponds to the making of a map. But each research worker, and each group of research workers, is rather a traveler having his own intentions. Thus there is something more than just the map involved. This real picture will involve us in deeper considerations in discussing the two situations mentioned above.

8. Methodological aspects of cybernetics (2)

Cybernetics was introduced by N. Wiener in 1948, as a science of communication and control in man and machine. We understand cybernetics to be one of the prototypes of information science. In the earliest development of cybernetics its scientific emphasis was placed on the study of the principal aspects of communication and control by means of mathematical models based on stochastic processes and their transformation by linear (and later by nonlinear) filters. In the terminology of information science as we have set it forth above, cybernetics' main task was the establishment of Branch 2 to give an integrated picture of communication and control and their applications to Branch 4 and other branches of science and technology, including sociology and linguistics, computers, and automatic machinery.

In the later development of cybernetics there were many features different from those observed in early stages. For instance, deeper considerations on the foundations of cybernetics appeared in the French literature, while economic system analysis was stressed by the Russian school. Some of these investigations may be said to belong to either Branch 2 or Branch 3. Also it should be observed that Wiener and his collaborators were, from the earliest stages in the history of cybernetics, keenly interested in artificial intelligence machines; such investigations belong to Branch 5 of our information science. In the latest period of development various new features have emerged, as described in several publications of Wiener [79]–[83]. Thus, as regards the theoretical features of cybernetics, their scope has become large enough to cover the greater part of the constituent categories of Branch 2 detailed above and bio-cybernetics in particular [82] and [83] has shown a remarkable growth into many research areas.

In spite of these recent developments, the earlier characteristic aspects of cybernetics seem to be still most prominent in its present activity, for the reason that no systematic connection between Branch 1 and the others has been explored. This is why cybernetics is to be distinguished from bionics, and also why cybernetics cannot be identified with the whole range of information science.

9. Methodological aspects of bionics

Bionics—the name was coined by J. Steel in 1960—is one of the most interesting fields of scientific technology. Its main purpose is to realize simulations of biological systems by means of artificial apparata, and hence to invent a new field of engineering (cf. [14], [53]). There are many biological functions which engineers can probably bring to bear to solve their own problems. Particularly in the field of biological information phenomena we have many examples: sensory organs, pattern recognition, learning processes of animals, self-organizing mechanisms, and so on.

Now let us consider the five branches of information science. In its first appearance bionics contains only the two Branches 4 and 5, and its main job is to cultivate Branch 5 in the light of knowledge obtained in Branch 4. However, a more careful observation convinces us immediately of the fact that Branches 1, 2, and 3 are also indispensable in promoting its main job. Branch 4 is concerned with biological phenomena, but information obtained there is, in the first place, described in terms of physical and chemical concepts in order to secure the possibility of realizing corresponding information phenomena through artificial apparata. For this reason bionics must depend heavily on the development of Branch 1. Second, the main interest in bionics should be focused upon information phenomena in biological existence, and therefore the phenomena of chief concern in bionics should be described in the terminology of Branch 1 and explained in the light of various theories in Branch 1. Branch 3 is also indispensable in bionics. Indeed, without system analysis of information systems there can be no artificial intelligence that can be recognized as corresponding to biological information organs.

We have emphasized that the results coming out of Branch 4 will benefit Branch 5, through the cooperation of Branches 1, 2, and 3. However, the cooperation among these branches is by no means to be one-sided. For example, research in Branch 2 will gain from results in Branch 4, which can reveal many biological phenomena that will admit new directions for research and inspire new theory and fresh model building.

In this way we see that bionics has at least five branches in common with information science, although not all of the constituent fields of each branch are contained in bionics.

According to these observations, bionics, in terms of its objective structure, may be said to be imbedded in information science. However, we imagine that some researchers in bionics may not be completely satisfied with this identification. In addition to any such classification work there is need of deep understanding of the subjective attitude of scientific research workers. The subjective attitude common to most research workers in bionics may be said to be characterized by a conviction regarding the fruitfulness of results to be expected in the research direction "from biology to engineering," as a guiding principle. In this connection we want to make here the following three points which seem to be important.

(1) An emphasis on some subjective guiding principle in establishing a science and/or a technology can be expected to be fruitful in exploring a new domain of research, and sometimes it proves to be a powerful working hypothesis.

(2) At the same time it is necessary to have an objective picture of just where a researcher is working and of the complex of paths along which he can progress from there.

(3) In the actual course of development of research the principles of practice described in section 2 are also of considerable importance.

In short, there are three principal aspects of importance to be considered in establishing and developing a new field of science: (1) the subjective guiding principle, (2) the objective map, and (3) the practical procedure.

An original initiator starts with (1), and his successors tend to be overly concerned with (3). After many years of struggle, with many successes and many failures, a science will have attained its reason for existence. Then there will appear an author who will supply (2) in a systematic fashion. At this stage there arises the danger of losing the strong impulse of fostering further developments without an appeal to some guiding principle. In a period of science and technology which can open a new era for mankind, we should be conscious of the laws governing birth, growth, and death of a science, and we should have a "medical plan" which is effective throughout the entire life of a science. Bionics is a new field of science and technology, and it should be worthwhile to analyze its methodology from the standpoint of information science; this we have tried to do.

10. Methodological aspects of the mathematical sciences

The terminology of mathematics began, a few years ago, to show up increasingly in various fields of the natural and social sciences, thus giving evidence of extensive new achievements in the application of mathematical methods (cf.

[85]). There had been a long continuation of rather unfruitful and tedious disputes as to whether and how one could distinguish between pure and applied mathematics. During approximately a quarter of the century, from 1925 to 1960, remarkable progress has been made in areas where mathematical methods play an essential part. Econometrics in economics, biometrics in biology, and psychometrics in psychology have the common feature of emphasizing the role and effectiveness of mathematical and statistical methods in these substantive sciences. Econometrics and psychometrics established themselves and their respective domains in the realm of science during these years, whereas biometrics which had its birth in the first quarter of the century, and which also had experienced a somewhat slow if steady growth in that earlier period.

Moreover, various new topics, such as the design of experiments, decision procedures, amount of information, learning processes, computer of logic, systems analysis, and programming of operations, have become objects of mathematical studies. These extensions of the scope of mathematical study were remarkable in two ways. In the first place, they had not been introduced through any internal growth mechanism of the traditional fields of mathematics. In the past history of mathematics we have a number of examples showing the internal growth mechanism in traditional mathematics. For example, the birth of non-Euclidean geometry was stimulated by deep studies upon the foundation of Euclidean geometry and was realized as a consequence of the establishment of a new system of axioms. The growth and development of so-called abstract algebra in the German school in the early Twenties were likewise not influenced by any other science but came about through the accumulation of research in the area of algebra itself. Now, a sharp contrast can be seen between such spontaneous developments due to the internal mechanism of mathematics in the first quarter of the century and new developments which have been much influenced by contacts with other fields of scientific and technological research. These latter had stimulated new mathematical investigations and have led us to create a set of new fields of mathematics.

Indeed, these all belong to the topics under Branches 2 and 3 in our formulation of information science. We believe that this correspondence between new fields of mathematical science and the topics in Branches 2 and 3 of information science has a profound implication for the mathematician in regard to direction of research. First of all, this coincidence shows that new fields in mathematical science are concerned with model formulations and with the study of models of information phenomena in our broad sense. Second, this coincidence involves the question of whether possibly the ideas of the mathematical sciences can be used as tools in exploring further in present and future areas of purely mathematical study.

The distinction between mathematics and mathematical sciences can be of use to mathematicians, because the notion of mathematical sciences will make them conscious of the origins of their mathematical problems and of the need for contact with the substantive sciences, which provide them with fruitful problems. It is also of use to people working in the sciences and in technology, because it will remind them of how much they owe to mathematical methodology for their studies. Nevertheless, there does exist the fundamental question of whether the notion of mathematical sciences is in itself sufficient to motivate new fields of mathematical study.

The essentially new topics which the present mathematical sciences have to explore all belong to Branch 2 of information science, and their vigorous development can be anticipated from the cooperation of the various branches of information science, through a network of interrelations connecting the different scientific and technological lines of research. Without this background of systematic cooperation, we cannot expect to have a satisfactory presentation of realistic and essential problems for study in the mathematical sciences.

11. Methodological aspects of the behavioral sciences

Behavior may be understood as an action or a sequence or system or pattern of actions, which are induced by a decision, or by a sequence or system or pattern of decisions. Since such decisions are more or less based upon three uses of information—(a) storage of information, (b) pattern recognition, and (c) operational use of information—it is not surprising that the behavioral sciences have an intimate connection with information science. However, it is important to look into the logical aspects of the behavioral sciences in order to find out exactly how they and information science are mutually correlated and, moreover, in order to be able to find a sound procedure for establishing a mode of cooperation that may be beneficial to both of them.

The behavioral sciences were born in the realm of the social sciences, but their methodologies are quite different from those that are traditional in the social sciences. The behavioral sciences deal mainly with human behavior as studied in psychology, economics, sociology, and politics. More specifically, the behavioral sciences are mainly concerned with the psychological, economic, social, and political behavior of individual human beings and of certain groups and organizations. It seems to us, however, that a theory in behavioral sciences can be properly developed only after setting down the principles of objectivity, subjectivity, and practice. For instance, economic behavior is human behavior with respect to economic affairs in economic systems, as we see it in our usual experience. But a theory of economic behavior can exist and can have scientific value only after three logical aspects of objectivity, subjectivity, and practice have been duly established, according to which there can be constructed an adequate model for discussing specific behavior. For example, the game theoretic approach to economic behavior by J. von Neumann and O. Morgenstern [72] is a case in which certain types of economic behavior within certain specified economic situations are described according to their formulations under what we have called three logical aspects: objective world, subjective attitude, and practice. It should be remarked at the same time that the economic man specified by the game-theory model due to von Neumann and Morgenstern is merely one possibility; he by no means exhausts all possibilities.

With regard to politics, sociology, and psychology, the situations are more or less the same as those of economics as just exemplified. For instance, in politics, political behavior can be the subject of the behavioral sciences only after we have answered the following questions:

(1) How do we define the objective world of politics? What are its principal model schemes?

(2) How do we define objectively the subjective principles of political man?

(3) How do we define objectively the principles of practice of political men?

Any theory in the behavioral sciences should be developed by preparing and working with a model that answers to all three logical aspects. We neither think that such a model has already been set in any area of the behavioral sciences, nor imagine that such models will be easily obtained within the next few years. All we can admit at the present moment is that such models have been proposed and adopted tentatively in some areas of the behavioral sciences with some successes in giving somewhat reasonable explanation and predictions of particular phenomena.

In addition to such general observations on the present states of the behavioral sciences, the more fundamental problems concerning the following issues should receive our attention:

(a) Under what conditions can model building be useful as a theoretical tool in exploring a given area of our research in the behavioral sciences?

(b) What are the features which characterize a particular branch of the behavioral sciences?

(c) What procedures can be introduced by which to improve a model in the behavioral sciences in the light of accumulating information obtained by experience and/or by experiments (if it be possible)?

(d) Under what conditions are experiments possible in a branch of the behavioral sciences?

As far as we are aware, these four fundamental problems have not been duly discussed by behavioral scientists; it has rather been characteristic of the attitudes of these scientists during these twenty years that tremendous adventures have been pursued, but without any deep considerations or reflections upon their logical and scientific basis. We do not think the present situation can continue forever, but we do think that at least these four fundamental problems should have been carefully discussed and should have been resolved, leading us to a unified understanding of the logical and scientific basis of the behavioral sciences before their development became accelerated.

Having laid out explicitly these fundamental problems in behavioral sciences, let us concentrate our discussion on problem (b) which must be presented in this section. Behavior may, in the simplest case, be a response to a stimulus, but in general it should be recognized as being closely connected with information concerning our objective world, our pattern of subjective attitudes, and our strategic principles of practice. Learning processes, adaptation, and self-organization are the proper topics of Branch 2 of information science, and a deep theory in the behavioral sciences can scarcely be expected to develop unless such topics have been adequately deliberated beforehand. Information systems such as data processing systems, documentation systems, computer systems, and statistical systems are also among the fundamental factors specifying the objective world in which economic, social, political, and psychological behavior takes place. In this sense Branch 3 of information science is one of the basic foundation stones upon which any branch of the behavioral sciences is to be built.

On the other hand, the behavioral sciences deal with human behavior which is neither entirely nor uniquely determined by information on the objective world. As we have said, it also depends on the subjective attitudes of the human being faced with certain objective worlds in which game-theoretic approaches, for instance, may have some validity.

In short, whereas the behavioral sciences require some branches of information science as indispensable constituents, they also require other principles which are not included in information science. But it is extremely important to remark here that information science should be developed much further before any appreciable amount of encouragement can be given for the progress of the behavioral sciences.

This remark is based on three reasons. The first is that any deep scientific achievement in Branches 2 and 3 can only be secured by mutual cooperation with the other three branches of information science, that is, with the whole realm of information science. The second reason is that a theory in the behavioral sciences cannot be very substantial without some specification of the objective world in which the pertinent human behavior takes place and such specification of the objective world can be more or less described by its information systems, which may be sometimes quite primitive, and rather loose, and sometimes quite advanced and well organized. The third reason is that human behavior should be analyzed with reference to the individual's storage of information and to his pattern of recognition, and these can only be suitably discussed in Branch 2 of information science.

On the other hand, we will not be fair if we do not point out also the possibility of indebtedness of information science to behavioral sciences. This possibility derives basically from the very fundamental fact that information, in our understanding, cannot exist without connection with control in its broad sense. Now, it is true that there does exist a long distance between the two notions, control and behavior, at the present stage of our understanding. But it should be observed also that it is the remarkable trend of the last quarter of this century that the distance between these two notions has been becoming shorter and shorter. The recent topics in the discussion of control problems are adaptive control, control in learning processes, and automation; and these will be recognized as taking into consideration the human behavior of learning and adaptation. On the other hand, all the branches of behavioral sciences seem to have had the remarkable common tendency to strive for model building, each within its particular theoretical framework; this should have some connection with Branches 2 and 3 of the information science. The shorter their distance becomes, the clearer are their common features. But this is not all that can be expected. Having found a clearer picture of the features common to these two notions, it will be easier to discover also the distinctions between them and also to seek for a scientific tool to add to the theory of control in order to better analyze human behavior in psychology, economics, sociology, and politics.

In closing this section it should be remarked that the expression "behavioral science" can and should be coined after, and because of, the establishment of common scientific principles which would be valid in every one of various branches of what is currently called "the behavioral sciences."

12. Library and documentation science

Before entering into the main topics of the present paper a few remarks should be given here regarding library and documentation science which, we assert to be a subdivision of information science as we have defined it above. We do not believe that very lengthy arguments are needed to establish the assertion we have just made. On the other hand, it is extremely important to remark that library and documentation science is not to be investigated independently of information science. Holding to the view that library and documentation science is an important subdivision of information science, we can look forward to keeping this science in constant and close touch with every new development in the branches of information science. For instance, information apparata, such as translation machines, learning machines, and teaching machines, can be quite helpful in solving technical problems encountered in library and documentation science and technology. The systems analysis of library activity is one of the most important topics to be investigated within the general framework of information systems; indeed, here is a very promising field as yet undeveloped research topics where operations research can be successfully applied. Fundamental research in the mathematical theory of linguistics will find immediate application in the field of library and documentation science. At the same time, the science will provide us with many interesting problems in practical information retrieval and mechanical translations, which are to be discussed and solved in Branches 2, 3, and 5 of information science. We may remark that library and documentation activities are now so prevalent in human activity that they are suitable subjects for statistical investigation and surveys which are easily designed and performed. Indeed this division of information science is a promising area for the application systems analysis and information apparata with the purpose of gathering large amounts of data to be used for further improvement, not only of libraries but also of information service centers in general.

13. Characteristic aspects of some of the recent developments in statistics

It is the purpose of this section and the following ones to pin-point some of the characteristic aspects of recent developments in statistics and to show the connections between statistics and information science which we have already mentioned. It is not our intention to cover the whole of recent developments in statistics; we are going to look only into some of the principal ones, which we believe to be important as indicating exceptional new directions of research in statistics, and show their close connection with developments in the various branches of information science. It is also to be remarked that we shall restrict ourselves to observations on the most recent trends in statistics, of the last five years. Thus we shall talk neither about the statistical decision function approach [74] nor about the successive process approach of statistical inference and control ([38]–[40]), both of which had been developed essentially in the ten years up to 1960. We shall talk about some of the very recent remarkable advances that have already been realized and some of the new frontiers of research that have been proposed, confining ourselves, moreover, to topics which have features distinguishing them from developments that took place during the Fifties. That is to say, we are going to point out some evolutionary directions which may be said to have been already programmed for the future development of statistics. It is true that (i) some of these evolutionary strains have been there during the whole process of the development in statistics, or that, in the case of others, (ii) their origins can be traced to some point in the history of statistics, or, for still others, that (iii) they can be understood most clearly from the standpoint of successive processes of statistical inference and control. Nevertheless, they should be recognized as evolutionary directions of statistics rather than outgrowths of statistics.

Let us just briefly note what we mean by an evolutionary direction in statistics in contrast to an outgrowth of statistics. By an instance of evolution in statistics we mean (i) the fact that some recent advances in statistics came about by treating problems which seemed to be neither authentic problems of statistics nor traditional cases for the application of statistics, and (ii) the recognition that such an advance has involved the reformation of the framework of informative patterns which have been traditionally associated with statistics and by which statistical information has been gathered and processed.

Now, what are characteristic aspects of recent advances in statistics which are to be recognized as evolutionary? Before answering this question, let us first give here a list of some topics whose descriptions may help us toward finding an answer.

Our list is claimed neither to be the unique solution to the present problem nor to be the best one; it is merely the product of our own experience in statistics and of our evaluation and interpretation of the work of some contemporary researchers in various fields employing statistical approaches. The list reads as follows:

(1) statistical programming,
(2) data analysis,
(3) statistical control processes,
(4) total quality control system,
(5) operations research and integrated product quality control.

Each of the following five sections is devoted to discussing one of these five topics. There then follows discussion on the formulation of statistics which can suitably cover the characteristic new developments enunciated by the list. It will be shown why and to what extent modern statistics will have deep connections with the information science in the days to come.

14. Statistical programming and automatically controlled sequences of statistical procedures

The principles of statistical analysis using large electronic computers have become one of the most important topics in statistics, and several contributions have been worked out by Terry [70], Yates [86], [87], Cooper [17], and M. G. Kendall and P. Wegner [36]. Among others, Terry [70] explained the role of the statistical programmer in the following two sentences:

(i) "The statistical programmer does not know a priori the exact analytic path that his data must follow."

(ii) "The statistician may very well prefer to let the data speak for itself and suggest the appropriate transformation to exclude from consideration on measurement deemed discordant, or to replace such measurements by derived measurements."

In realizing these principles and the role of the statistical programmer, Terry [70] suggested the broad aspects of adequate statistical programming:

(iii) "Now, with the advent of the disc file, which has the effect of increasing the storage capability of the computer to the order of two million measurements or more, we believe that it will be possible to store in this ancillary device many different statistical strategies, computational techniques, and statistical decision rules as well as large blocks of data."

(iv) "Then, by writing a program of the analytical strategy to be employed, we could permit the data to call in the appropriate analytical techniques and rules, and thus produce a much more effective final analysis."

Let us now pick up two particular topics in statistical programming in order to observe the characteristic aspects of statistical programming. The first topic is concerned with screening and validation procedures which are particularly important in the logic of statistical approaches. The second topic is a review of comprehensive programming systems developed recently by several statisticians.

Regarding the first topic, many experts on census and large scale sample surveys have been keenly aware of different types of errors occurring in the case of large-scale sample surveys.

Deming [20] gave a detailed listing and description of the different types of errors which should be taken into consideration, both in designing and analyzing sample surveys. Mahalanobis [49] gave the classification of different kinds of errors into three types, and "revealed the great importance of controlling and eliminating as far as possible the mistakes which occurred at the stage of the field survey." The interpenetrating sample procedure was introduced by him as one way of doing this.

Yates [87] referred to the general problem of preliminary editing of data before analysis and enunciated the uses of electronic computers in the following sentence. "Once appropriately instructed a computer will perform any required tests on each item of data as it is read in, and can draw attention to anomalies, reject suspicious items, or even in some cases make the appropriate correction." Several papers and memoranda have been written by various authors with particular reference to screening and validation problems. We can mention (i) preliminary assessment, by Reed [62], (ii) autostat, by Douglas and Mitchell [21], (iii) treatment of spotty data, by Tukey [71], and (iv) analysis of residuals by, Anscombe-Tukey [2]. Cooper [17] has published the first set of rigorous procedures for validating and controlling the presentation of data to a computer.

Several comprehensive programming systems have been prepared by various statisticians and various institutions.

Yates [87] pointed out several important aspects of the use of computers in research, saying that: "In research statistics the analysis must in fact proceed step by step, the exact nature of the next step being determined after examination of the results of the previous step. This presents considerable problems, since the results at each step must be stored (clearly on magnetic tape, if available), and indexed in such a manner that the required item can be specified as data for the program performing the next step."

The MUSP prepared by statisticians at Harvard University is said to consist of a set of 19 subprograms which can be called in by a program in a special-purpose language specially designed for MUSP. The sequential operation of subroutines is directed by a control program called MUSP Control Program which "accepts as input the symbolic specification of the problem to be solved in terms of a sequence of subroutine names and parameter values, checks the specification for obvious errors such as missing parameters, translates the specification into a machine-oriented representation, and then executes the resulting set of the specifications interpretively" (M. G. Kendall and Wegner [36]).

We have quoted several important sentences of various statisticians in connection with statistical programming which has become an important task for statisticians during the past five years. It is our opinion that the problems of statistical programming presented here will lead us to fundamental reconsiderations of statistical analysis and statistical information, as we shall make clear in section 18, from a standpoint involving the introduction of learning processes in statistics.

15. Data analysis and automatically controlled sequences of statistical procedures

One of the most valuable contributions to the foundations of statistics is the thorough consideration of the various aspects of data analysis given by Tukey [71]. Data analysis has always been considered to be a collection of fragmentary techniques, and its importance has not been properly recognized in the history of statistics. Some scholars have habitually classified statistics into two divisions, namely, (i) descriptive statistics and (ii) statistical inference theory. This classification is currently adopted by a majority of statisticians. In the first place, speaking from the theoretical point of view, the domain of application of descriptive statistics has been understood to be sharply distinguished from that of statistical inference because the latter divisions are exclusively concerned with random samples from a hypothetical population, while the former does not rely upon the notions of population and sample. In the second place, the classification has had real significance on practical grounds because each of two divisions has had its individual domain of application distinct from the other.

It is one of the remarkable trends to be observed in statistical activity today that both of these reasons, theoretical and practical, are losing their force to maintain the classification of statistics into descriptive statistics and inference theory. For instance, there are indications that the tasks of official statistics are becoming more and more analytic, and that the gap between the two divisions of statistics is now becoming much narrower than it has heretofore been. Illustrative examples in technological research problems in engineering industries as well as examples in large scale sample surveys and in designed experiments on biological phenomena can easily be given to indicate the diminishing distance between these two presumed divisions in statistics. Indeed, on one hand, we are now faced with the need for handling masses of data, while on the other hand, we are equipped with the feasibility of analytical studies in virtue of the use of high-speed electronic computers having rich memories.

We stand in need of developing a methodology for handling large masses of data and for conducting analytical studies of such data. We believe this circumstance is the real basis for data analysis becoming so serious a problem in recent years.

It will be worthwhile to discuss the following assertions by Tukey [71] in this connection:

(1) "We need to give up the vain hope that data analysis can be founded upon a logico-deductive system like Euclidean plane geometry." Data analysis is intrinsically an empirical science (section 46, p. 63).

(2) "Some may feel that if data analysis cannot be a logico-deductive system, it inevitably falls to the state of a crass technology. With them I cannot agree" (section 46, p. 63). "There will be also the hallmarks of stimulating science: intellectual adventure, demanding calls upon insights, and a need to find out" how things really are "by investigation and the confirmation of insights with experience" (section 46, p. 63).

(3) "I look forward to the automation of as many standardizable statistical procedures as possible" (section 17, p. 22). He mentioned three arguments in favor of his views.

16. Statistical control processes

Statistical control processes are understood to be control processes based upon probabilistic formulations and/or control processes which require statistical information during their course of evolution.

In his formulation of cybernetics N. Wiener [78] emphasized both of these characteristics of statistical control processes. In the first place, the theory of stochastic processes was developed from the study of Brownian motion in the effort to build a mathematical apparatus for the design of controls. Second, by the use of stochastic processes Wiener aimed to define the essential aspects of the semi-exact science that became the domain of cybernetics.

These two aspects of cybernetics are present in every statistical control process, and are the features that distinguish such a process from a deterministic one. Most deterministic approaches to control processes could be made more realistic by adding stochastic elements in place of some of the deterministic elements, the latter being usually simplified or idealized pictures of the former. However, this task of introducing stochastic elements is not necessarily as easily done as one might imagine. In fact, the actual situation is that deeper analyses of mechanisms have been performed in the realm of deterministic control processes because of the possibility of appealing to well-developed techniques of mathematical analysis, for example, differential equations and the calculus of variations, and more recently the method of dynamic programming. The works by Pontryagin [61] and his school, and those by Kalman and others, are far more precise than those obtained in the area of statistical control processes. A systematic approach to control processes has been developed by Bellman [7] through his original method of appealing to dynamic programming. Now let us mention here some of the formulations of statistical control processes:

(a) stochastic approximation by Robbins [64] and his successors,

(b) evolutionary operations programs and adaptive process controls by Box [16] and his school,

(c) successive processes of statistical control by the present author [38] and [40].

Other examples can be cited but these will suffice to enable us to compare the characteristic features of statistical approaches to control processes with those of deterministic approaches. Now what are the essential differences between these two? Our answer to this question is this. In statistical approaches (a) the principle of objectivity, (b) the principle of subjectivity and, (c) the principle of practice can be and are clearly formulated in full awareness of our information pattern, the accuracy and precision of our information, the objective of control, the cost of control, the sequential or successive approaches of control procedures.

In short, statistical control processes are discussed in the realm of information science, whereas most deterministic formulations, if not all, are mainly or exclusively concerned with idealized situations which assume, mostly if not always, persons who have complete information on the subject matter and whose controls are completely exact or completely under command. Due to the fact that statistical approaches to control processes have their basis in consideration of information patterns, amounts of information, accuracy and precision of controls, and strategic principles of control procedures, it is obvious that statistical control processes can take from the branches of information science any results and techniques that can be beneficial to their theories. In particular, the recent advances in learning processes and adaptive controls are now challenging us to generalize the present scheme of statistical control processes to the point that these results, obtained mainly in deterministic schemes, can be combined with probabilistic formulations.

17. Statistical quality control (SQC) and total quality control (TQC)

Statistical quality control was introduced by W. A. Shewhart [67] in the first quarter of this century, and has made a great contribution to the development of statistical methods as well as to industrial engineering, where it has been successfully applied. The Shewhart principles of statistical quality control were rather precisely described in his monograph [67], and gave us the basis upon which statistical techniques could be invented and applied. During the past forty years, in which extensive experience with quality control has accumulated and much success has been enjoyed throughout the world and in all industries, there have appeared some new or revised viewpoints regarding the principles underlying quality control. It seems worthwhile to observe, among other things, how and why the original idea of statistical quality control has been transformed into the new idea of total quality control. The object of this section is to show the characteristic aspects of these two different ideas on quality control, particularly with reference to the notions and the systems of information control attached to each of them. It is our opinion that by doing this we can get a deeper understanding of how intimately quality control problems are connected with information science.

In the original Shewhart principles of statistical quality control, great emphasis was placed on the consequences of the applications of statistical methods to quality control. It seems to be rather exact to say that the essence of the Shewhart principles of quality control is the application of statistical methods. The applicability of statistical methods to engineering production was essentially based on the approximate admissibility of introducing into actual production circumstances a probability field in terms of which to interpret the fluctuation of product quality during production. This situation could be assumed to exist when specification and standardization were attained at a certain level of control of the effects on products. It seems fair enough to say that without statistical methods no systematic or scientific approach to quality control would have been introduced.

However, the accumulation of experience in various industries has occasioned reflection on the meaning and the implication of each word in the expression "statistical quality control." In the first place, it was argued that quality should be interpreted broadly enough to cover various stages of planning, production, inspection, marketing, and management. Instead of interpreting quality as simply the quality of the manufactured product, there were strong arguments prevailing in a wide range of industries in favor of generalized interpretation of the quality which would comprise, along with "quality" in the narrower sense, the completion date for production, cost, and afterservices.

The second point of issue was concerned with the meaning of "control." Control chart procedures were one of the fundamental tools in statistical quality control, and they worked quite well in various fields of industrial engineering at certain levels of control of production, before the stage when automation was introduced. Now the introduction of automatic control in some of these fields has produced a change in the fundamental aspects of production and sometimes also a restriction on the use of statistical quality control in that interruption of production by human action was required for controlling the process. At the same time, the combined development of large scale electronic computers and of theories and apparatus for control led to the introduction of new and more advanced features of control, including adaptive control, optimizing control, control with learning, and so on. It is now almost self-evident that quality control would not be effective in controlling quality, either in the narrow or in the broad sense, unless the meanings and the procedures of control were broad enough to cover these new notions of control. In fact, the more automation proceeds in an industry, the more urgently needed it is to generalize the notion of control beyond its original formulation in the SQC of Shewhart.

The third point of issue was concerned with the statistical methodology of statistical quality control. In view of the generalizations of the two fundamental notions, quality and control, it has become a matter of course to encounter needs for broader methodologies than those which sufficed under the narrow sense definitions. Thus, having the completion date as one quality of production in the generalized sense, we naturally require various techniques that may be available for production and inventory control, such as those belonging to the areas of linear programming, dynamic programming, queueing theory, scheduling and flow techniques, and so on.

In a word, after the elapse of only three years following the introduction of Shewhart's principles of statistical quality control (SQC) in Japanese industries, we experienced the need for a generalization of his principles with respect to each of S, Q, and C. We may conveniently use the symbols S^*, Q^*, C^*, to indicate the generalizations of the corresponding original notions, S, Q, and C.

Let us summarize here what we have explained, in the following abbreviated form:

$$(1)\qquad \begin{aligned} SQC &= \text{Statistical Control of Quality} \\ &= SCQ \rightarrow SCQ^* \rightarrow SC^*Q^* \rightarrow S^*C^*Q^* \\ &= M^*C^*Q^*, \\ &= M^*Q^*C^* \end{aligned}$$

where we have put, for the moment,

$$(2)\qquad Q^* = \begin{Bmatrix} Q\text{: Quality in the strict sense} \\ C\text{: Cost of production} \\ D\text{: Date of production} \\ S\text{: Afterservices and care} \end{Bmatrix},$$

$$(3)\qquad C^* = \begin{Bmatrix} \text{Adjustment control} \\ \text{Control by analysis} \\ \text{Adaptive control} \\ \text{Control by learning} \end{Bmatrix},$$

$$(4)\qquad \begin{aligned} S^* &= \begin{Bmatrix} \text{Statistical method} \\ \text{Operations research} \\ \text{Management Science techniques} \end{Bmatrix}, \\ &= M^*. \end{aligned}$$

The transformation of SQC into $M^*Q^*C^*$ was necessary in most of industrial engineering, and it was realized in a majority of actual cases by involving the cooperation of all persons in a given company. In the consequence, quality control has been concerned with every job in the company. Quality control must be performed in every job in the company by every person in the company. This understanding of the quality control movement implied the transformation of SQC into $M^*Q^*C^*$, and it was indeed the practical basis on which some remarkable benefits were brought about. Nevertheless, this understanding could not have been sufficient as a theoretical basis for the further development of quality control, unless it had been supplemented with stricter observations of real situations.

In this connection the idea of total quality control is worthy of discussion. The terminology "total quality control" has been used with somewhat different meanings by different authors in the field of quality control. In order to avoid misunderstanding, we shall set down certain characteristic aspects of total quality control according to our understanding of the term, which we believe is close to that of Juran [34] and Feigenbaum [22], [23], but which is not identical with either.

The characteristic aspects of total quality control are embodied in the following two properties:

(1) a total valuation system is introduced in terms of which every activity company can be evaluated;

(2) a total information system is introduced which makes information available for controlling every activity in the company.

These two aspects are characteristic of a total quality control system, they may be only approximately fulfilled in any actual company. By having a common unit of evaluation, the merit or demerit of each activity of the company can be evaluated, and the manager can give an exact indication for the job of any particular person. By having a total information system, the company can secure its self-regulation.

Here again we have come to the realm of information science, which can provide us with the effective organization for advancing total quality control.

18. Operations research and integrated control of product quality

Operations research [1] is concerned with operations that are applied to man-machine systems in which controls are required to attain the desired objectives. For instance, production lines in a plant, or railway systems, entail operations in this sense. Department stores are concerned with various kinds of selling operations; the wood and pulp industry has to do with operations that include transportation and tree planting. The development of a city implies a large assembly of operations, including the building of roads, houses, and electrical network, a water system, and a transportation system for its inhabitants.

Operations research is understood to be the scientific approach to operations. An operation research team endeavors to develop suitable theories for predicting the responses to operations indicated by a manager and performed by an organization; and also to introduce measuring procedures so that actual performances and states of the system can be measured, and to set up information systems through which such information can be transmitted to managers.

In short, operations research gives to a manager both a theoretical tool and a tool of observation by means of which the manager can determine suitable management objectives and can then put into effect operations that will be effective in attaining these objectives.

Now the essential role of an operations research team consists in its working as one constituent element in a closed cycle of a feedback system of information. An operations research team can and must be ready to improve proposed theoretical tools in the light of data accumulated, and it must take constant responsibility for checking whether the system is working well and for informing the manager of any significant change in order that the manager may make the necessary decisions to improve the situation and bring to bear suitable operations on the system before it turns out to cause any serious loss.

This understanding of the role of the operations research team will suffice to make clear the following two observations.

The first is concerned with the relationship between operations research and information science. Operations research is not to be viewed in terms of individual final solutions to restricted types of operational problems, but rather as being imbedded in the framework of an information system through which constant checking and improvement can and should be performed. In consequence, the following assertions can be deduced.

(1) Operations research should be organized and carried out in such a way that it is working routinely.

(2) Operations research should be recognized as a learning process in which experience can be accumulated and improvement of operations realized as a result of these experiences.

(3) Operations research should be systematized in such a way that total optimalization as well as suboptimation should be brought about or at least searched for in the handling of the various problems of management in its broadest sense.

The second observation is this. In terms of what we have said in our first observation, it is now quite clear that the notion of integrated control of product quality as advocated by various quality control engineers in Europe corresponds exactly to the role of operations research as we have described it. The remarkable differences between total quality control (TQC) and integrated control of product quality (ICPQ) arise from the following three factors.

(1) The notion of product is broader in ICPQ than in TQC. In ICPQ the building of a library, a hospital, a railway, or a harbor is considered as the creation of a product, or, more properly, a superproduct, to which quality control can and should be applied.

(2) The objective of control must be broad enough to cover social welfare and is not to be confined to individual manufacturing companies.

(3) The methodologies of controlling products are not confined to statistical methods, but should rather be concerned with operations in the sense of operations research as just explained.

In summarizing these two observations, we can say that operations research can work as a routine job in management and production, just as SQC has been working in industry, leading to the notion of ICPQ, and that this is possible by imbedding operations research in the information science scheme associated with management and production. If this assertion is granted, then the duties and the legal status of operations research workers in industry can be formulated and agreed on. These deliberations show how and why the information science approach can serve in essential ways in the present age of industrialization.

19. Automatically controlled sequences of statistical procedures

The purpose of the present section is to show to what extent one of the recent new developments in statistics, namely, automatically controlled sequences of statistical procedures, abbreviated by ACSSP, is connected with the various developments explained in the last few sections, and then to give a preparation for the introduction of the following section in which we shall summarize the relationship between statistics and information science.

(1) One of the characteristic features of recent developments can be seen in the fact that the notion of automatically controlled sequences of statistical procedures (ACSSP) is realistic and totally applicable in statistical activity. In our two papers [44] and [45] we have explained the needs and the usefulness

of the notion of ACSSP. Automatically controlled sequences of statistical procedures are defined with reference to computer programs. They have turned out to be applicable in virtue of the development of computers, which are now indispensable information apparata in the sense of information science. Statistical programming is nothing but the setting up and programming of ACSSP into a computer in the terminology of that computer's particular program language. Indeed, programming implies automatically controlled sequences of procedures. In characterizing programming as statistical programming, we need deep and thorough understanding of the implications of statistical methodology.

The concepts of statistical programming emphasized by Terry [70], Yates [87], and others, are extremely important for a clear distinction between mathematical and statistical programming.

In a previous paper [45] we have also given our own idea on how to define a statistical procedure. This runs as follows.

(1) A statistical procedure is a procedure that deals with a set of data under the following two specifications:

(a) composite regularities are described in terms of probability fields, as specified in subsection 6.1 of [45];

(b) quantitative evidence is obtained through the application of two principles for acquisition of data, namely, randomization and transformation, as explained in subsection 6.2 of [45].

Evidence obtained from statistical procedures is a collection of inductive determinations based on a set of rejection rules for statistical hypotheses, in the sense explained in subsection 6.3.

The third terminology of automatic data analysis, that of Tukey [71], is essentially of the same content as ACSSP.

Indeed, automatic data analysis is performed according to a family of ACSSP's which is programmed into a computer. The usefulness of automatic data analysis is explained very clearly by Tukey [71], and his assertion holds true for the usefulness of statistical programming and the ACSSP family.

In short, automation of statistical analysis is one of the characteristic features in various developments of statistics.

(2) The second characteristic feature of recent developments in statistics lies in the introduction of various formulations of learning processes into the domain of statistics.

We have so far explained three uses of information obtained from data with reference to a tentatively specified model. The tentativeness of specified models is rather common in actual statistical analysis. And, in fact, it has been an essential job of statisticians to decide what course of analysis can and should be chosen in view of the data. The difference between classical statistical analysis and more recent statistical analysis does not lie in the recognition of need for learning by experience. Both of them have been based on the understanding of this need. The difference between the two comes in the matter of the learning processes being or not being formulated in terms of programs in computers.

In recent work in statistics, some learning processes have been formulated to the point that ACSSP can be used for their analysis. For example, statistical procedures of TE type—that is, estimations after preliminary test(s) of significance—are one of the simplest kinds of learning processes that are being formulated in the framework of the ACSSP. In general, an ACSSP pertains to a successive approach in which a feedback principle is adopted in order to adjust the course of procedures in view of future data. One can quote many examples showing the recognition of processes as learning processes and their introduction into statistical inference and control. These examples are given in a series of papers of the author dealing with successive processes of statistical inference and control.

(3) The third characteristic feature of recent developments in statistics is the emphasis of a fundamental fact which has not been duly recognized by some schools of mathematical statisticians. This is the fact that statistics as a whole is an empirical science and hence is not a logico-deductive system at each stage of its development. This is by no means contradictory to the assertion that a theory of statistics should be a logico-deductive system. Indeed, any one of the ACSSP approaches whose importance in recent developments in statistics has been recognized provides us with a logico-deductive system that can be used for determining operating characteristics. Nevertheless, we must always reserve the freedom to enlarge and revise any ACSSP family in view of data we have gathered and of the experience we have accumulated in applying ACSSP families in various domains of statistical problems. Statistics has no limited areas of application—the development of science, technology, and industry, as well as that of the social sciences, has repeatedly introduced new problems which had not previously been treated at earlier stages in the growth of statistics.

For example, quality control had not been a topic of statistics until the first ten years of this century. But we should duly recognize that statistical quality control has had much influence on the development of statistical method.

In view of the various influences on the development of statistics that have come from outside, it should be the attitude of statisticians to keep open the doors of statistics in order that such influences may always be accepted, but at the same time to organize the structure of statistics as systematically as possible at each stage of its development. What we have established should be programmed in a computer as ACSSP, and we should always be searching for procedures that may be useful in particular types of statistical problems. It is contradictory to our fundamental understanding of statistics as an empirical science to restrict the problems of statistics to within a certain prescribed range. In this connection Tukey [71] asserts the importance of statistics in cultivating new fields of data analysis by pointing out "the needs for collecting the results of actual experiences with specific data-analytic techniques," and by recommending "free use of ad hoc and informal procedures in seeking indications" (Tukey [71], section 45, p. 62).

(4) The fourth characteristic aspect of the recent developments we referred

to in the previous sections is the intimate connection between statistical inference and statistical control. In the formulation of statistical decision functions, decisions and actions are closely connected, mainly for the reason that decisions are recognized as operational uses of information. Due to our position on the uses of information in statistics, we do not think that operational uses of information should be the sole purpose for which we reduce our data; two other uses, storage of information and pattern recognition, should also be taken into consideration. With this general understanding of the three uses of information in statistics, it may be remarked that ACSSP makes particularly easy the connection of inference with control. Although in both of the papers [44] and [45] we have referred exclusively to automatically controlled sequences of statistical inference, it can be readily seen that statistical procedures can be enunciated both with inferences and with control. This can be seen by observing that controls can be understood as transformations of one objective world into another, usually by changing a set of parameter values into another set or by introducing some more drastic change of the objective world to another one which is still a member of the class of objective worlds within our concern.

20. Statistics and information science

The object of this section is to summarize all the foregoing considerations and observations in order to show the relationship between certain recent developments in statistics and those in information science, and also to point out some underdeveloped areas in statistics which it would seem to be worthwhile to cultivate in view of the relationship between these two sciences.

For this purpose we shall prepare ourselves with three different points of view. The first point of view is that of the three logical principles of statistics enunciated in the preceding section: (1) objectivity, (2) subjectivity, and (3) practice.

The second point of view arises from the three uses of information: (a) storage of information, (b) pattern recognition, and (c) operational uses of information.

Finally, the third point of view refers to the five branches of information science: (i) information elements, (ii) information theory, (iii) information systems, (iv) biological information, and (v) information apparata.

The first of these three points of view will be found to give a primary classification, while the two other standpoints combine to give subsidiary categories under the headings assigned by the primary classification.

[1] *Objectivity.* This category is concerned with the question of how to formulate models of the objective world which is to be dealt with in statistics. Broadly speaking, there are two fundamental schemes for formulating models of the objective world, namely, chaos and cosmos. These two principal features can be observed in the theoretical schemes of information science by examining the subject matter and scope of all constituent research fields in Branch 1 of information science. At the same time it should be added that one of the characteristic features of certain advances in statistics during the last five years is the

increasing need for elaboration and/or generalization of cosmos formulations. Indeed, it is in the area of information science that deeper consideration concerning logic, pattern recognition, information networks, and systems have appeared, raising the notion of cosmos into greater prominence than ever before. These new developments in information science will have great influence on the general formulation of statistics. In particular, they are deeply concerned with pattern recognition in statistical approaches, a subject that has not been fully discussed in the traditional area of statistics.

Here a close cooperation between the two sciences, statistics and information science, can be usefully initiated. In fact, information science owes its chaos formulation of the objective world to the field of statistics. On the other hand, statistics can and must learn much from information science, where the notion of cosmos has been studied in various connections as explained just above. In particular, the statistician can and must maintain intimate contact with the methods and achievements in the fields of communication theory, control theory, mathematical linguistics, and the logic of computers, all of which belong to the subject matter of Branch 1 of information science.

[2] *Subjectivity.* This category is concerned with the principles upon which is based our choice subjective attitude. In statistics our subjective attitudes and performances should be coordinated and integrated under two principles: (i) the principle of strategy and (ii) the principle of valuation. We have explained the implications of these two principles, so that they are seen to be valid not only in decision function approaches in statistics but also in the formulation of controls. It is to be remarked here that no theoretical formulation of information science can work well without reference to these two principles of subjectivity.

This category is indeed indispensable in a formulation of information science. A piece of information cannot have meaning without reference to some subjective existence, which conducts itself according to a certain set of subjectivity principles. In this sense the two principles valid over a wide range of statistical theories are common to theoretical formulations under Branch 1 of information science. As far as these principles can be formulated objectively, the category of subjectivity can be applied to the formulation of information science. In this sense information science can expect to gain from the formulation of statistics. Here, in speaking of statistics we are not confining ourselves to just decision function approaches. As we have explained in our paper [41], the Fisherian spirit of scientific inference ([24], [25]) can be understood within the framework we have explained in the previous section. In order to establish a theory in statistics we must present as many automatically controlled sequences of statistical procedures as possible, which can be dealt with from the standpoint of the two principles of subjectivity. In this connection emphasis should be placed upon every aspect of the three uses of information (cf. [43]), including (a) storage of information, (b) pattern recognition, and (c) operational uses of information. In virtue of such a systematic review of statistical principles, we shall be conscious of the need for introducing and developing the two principles, strategy and valuation, in the case of each of the two information uses (a) and (b). Indeed, the two principles have been applied mostly, if not exclusively, to the information use (c), as we observe it in Wald's theory of statistical decision functions [74].

Although we do not intend to give any extended discussion of the application of these two principles to the first two categories of uses of information use, it should be pointed out that strategy and valuation regarding storage and pattern recognition are crucial subjects in both information science and statistics. The problems suggest an as yet undeveloped area of statistics, in which some new systematic approaches should be tried. This assertion holds also for Branch 2 of information science.

[3] *Practice.* Once the subjective aspects of our strategies and valuations have been settled upon for our objective world, various principles of practices will enter to play the role of providing us with practical guides in dealing with real circumstances. In this connection two principles seem to be essential: (i) the principle of efficiency and (ii) the principle of successiveness, or the principle of feedback. As we have explained in our previous paper [41], the characteristic features of statistical analysis can be explained in terms of these two principles. Now, as we have already explained, in view of the recent developments in information science, the need is urgent for a generalization of the principle of efficiency as applied in various statistical problems, including the three uses of information. Thus, the operating characteristic approach of ACSSP has to refer not only to operational uses of information but also to storage and pattern recognition. Now, how can we define an efficiency of storage of information? How can we define an efficiency of pattern recognition? These have not been duly discussed in statistics. But we are now coming to the stage where we cannot help but be conscious of the present lack of any adequate generalized notion of efficiency which can be applied to such generalized problems in statistics. We are now only in a position to be able to suggest the importance of the problems before us. We should not forget that we are faced with these problems. In this connection we have already discussed the purposes and the implications of systems analysis applied to ACSSP.

An ACSSP can be considered as a system that is decomposable into a set of component subsystems, each of which is also an ACSSP, while it can also be considered as a component subsystem of a more complex system which is also an ACSSP (cf. [45]).

Systems analysis is concerned with several types of analysis, such as (i) component analysis, (ii) composition analysis, (iii) stability analysis, (iv) flexibility analysis and (v) reliability analysis. We have explained the implication of each of these five types of analysis in our previous paper [45]. Equipped with all these methods of analysis applicable to an ACSSP, we have a rather broad view of the various possible functions of an ACSSP, not only with respect to operational uses of information, but also partly with respect to storage and pattern recognition, as, for example, (i) component analysis and (ii) composition analysis.

There has been a great deal of progess in the domain of information science bearing on the principle of successiveness. It is concerned with learning processes and self-organizing systems. On the other hand, one of the mainstreams of development in statistics can be evaluated as development along the lines of learning processes.

Data in themselves are indeed a bulk accumulation of information, and any statistical understanding obtained from data should have a certain integration pattern of statistical recognition.

An integration pattern of statistical recognition can be defined with reference to an information summary scheme and an information evaluation scheme. An information summary scheme is a recognition pattern according to which information obtained from data is stored and arranged. An information evaluation scheme is a set of utility functions associated with an information summary scheme, according to which the utility of information stored and arranged in the information summary scheme is measured. In our previous paper we gave a detailed analysis of the Wald decision function approach [74] in statistics with reference to these two schemes and discussed the reasons why some parts of statistics are not completely satisfactory with respect to this formulation. In order to achieve more freedom and a higher degree of flexibility, whereby both uses of information, storage and pattern recognition, can be admitted in the schemes of statistics, we have suggested in our previous paper [45] the formulation as follows.

(i) An ACSSP system can be decomposed into a set of subsystems each of which has an integration pattern admitting a risk function.

(ii) A coordination principle is set according to which all these subsystems are coordinated into a single whole. (This principle does not necessarily refer to the corresponding category in the utilitarian viewpoint, which is common in information evaluation schemes of subsystems.)

(iii) A coordination is not necessarily geared to any integration pattern having a risk function evaluation scheme, and hence, the whole may be decomposed.

(iv) A coordination is to be flexible enough to learn from experience so as to be able to introduce a tentative integration pattern over the whole as data accumulates. Such a tentative integration is always subject to replacement by another more suitable one as more information is gathered.

In short, a decomposable coordination of subsystems has been suggested in our previous paper [45]. We believe this will serve as a framework covering the three aspects of information use, and at the same time will admit learning processes, which are essential in the development of statistics.

A few words may be in order to clarify the relationship between statistics and the three particular branches of information science, information elements, biological information, and information apparata. We have already shown the intimate connections between statistics and the other two branches of information science, information theory and information systems.

Briefly, statistics has no direct connection with these three branches, but it has an important indirect connection with each of them. An understanding of this is, we believe, indispensable to the development of statistics in the coming years, because it provides us with a systematic way of maintaining contact with the achievements in information science.

21. Summary and comments

This paper consists of twenty-one sections, which can be broadly divided into three parts. In the first part of the first five sections, we explain the impetus and the reasons for the need to establish an information science. In the second part, consisting of the middle seven sections, from section 6 and section 12, a description of information science is given, and its connections with some neighboring sciences, such as cybernetics, bionics, the mathematical sciences, the behavioral sciences, library and documentation sciences, are illustrated in order to give a better understanding of the scope and role of information science. In the third part, consisting of the last nine sections, some of the characteristic aspects of recent statistical problems are described, which show up implications of information science for the development of statistics. We can summarize our points of view in the following assertions.

(1) There is an urgent need to develop an information science which will satisfy the following specific requirements:

(i) it is a planned area of consolidated (integrated) scientific approaches which must start with the five indispensable branches set forth in section 6;

(ii) its theoretical formulation has a scope which is broad enough to cover the three principal logical aspects: (a) objectivity, (b) subjectivity, (c) practices, as explained in section 20.

(2) Statisticians and mathematicians should have intimate contact with such an information science for the following two reasons:

(i) statisticians and mathematicians should be aware of the crucial problems in information science to whose solution they can contribute,

(ii) statisticians and mathematicians should be aware of essential achievements in the many branches of information science; these will in turn serve them in their selections of problems and thereby further their research.

(3) In particular, some of the crucial problems for formulations in mathematics and statistics are now coming to be common topics of discussion among scientists and engineers working in information science, and statisticians and mathematicians will have much to gain from new developments in information science.

In the conclusion of the paper the author takes the occasion to give several remarks on the contemporary works and views of other authors who deal with what we call information science.

REMARK 1. The scope and role of information science as we describe it may be said to be almost identical with those which the Russian scientists call "kibernetiki" (cybernetics). From the publications [46], [48], and [75] I have the impression that "kibernetiki" has larger scope and role than the cybernetics of Wiener ([78] through [83]), as in the case with our information science. The reason we adopted the terminology of "information science" rather than "cybernetics" is that we think it better to retain the latter term in the original meaning given it by the inventor and his collaborators. The substantial difference that may exist between kibernetiki and our information science is that we propose to construct our science in accordance with the requirements just explained. This position of ours may be explained as intending to keep our research within the area of science.

REMARK 2. In the later stage of the development of his cybernetics, Wiener emphasized that an important part of cybernetics in the future would lie in the field of organization (cf. [80], p. 12). He also referred to a problem of economic planning with the assertion that such a problem is of a statistical character, and therefore of an information character (cf. [80], p. 12). In view of these affirmations, it might be thought that cybernetics in the sense of Wiener has been growing to the point where there is no substantial difference of coverage between it and our information science. However, it is to be noted that he did not seem inclined to include certain econometric and psychometric research within the area of cybernetics. This is one of the crucial differences between our information science and cybernetics. Our position is to refer to three logical aspects, which include subjectivity and practice as well as objectivity. It is our view that no scientific theory of organization can be built without reference to a set of logical aspects including subjectivity and practice. The author is in complete agreement with the view of Ashby [3] when he suggests that the recognition of the limitation implied by the law of requisite variety may prove useful by ensuring that our scientific strategies in the case of a complex system shall be new strategies genuinely adapted to the special peculiarities of that system. He then suggested that the strategies appropriate to complex systems might be those already becoming well known under the title "operational research." However, it seems to us that the logical aspects of operations research can be reduced to those which, as we have already explained, are more fundamental in formulating the methodology of science.

REMARK 3. In recent years several symposia ([14], [15], [19], [51], [53], [68], [78], [82]) have been held in the areas of artificial intelligence, perception, bionics, and biocybernetics, and many works have been published concerning pattern recognition, learning processes, and self-organization; among these quite a few are executed from the standpoint of statistical approach (cf. [8], [14], [15], [18], [27], [28], [33], [35], [51] through [58], [59], [63], [68], [73], [76] through [78], [81] through [84]). In particular, decision function approaches have been applied to pattern recognition and learning processes by several authors (cf. [27], [28], [66]). The full range of statistical approaches is not covered by the Wald formulation of statistical decision functions, as we have explained in section 20, as well as in [37], (3), [47], [43], [44], and [45]. There are few works which are clearly statistical approaches to given problems but which cannot claim to come under the decision function approach (cf. [35], [58]). Furthermore, there do exist several recent statistical works searching for different possibilities (cf. [33], [59], [76], [77], [84]), and these seem to us to be more or less conscious of the problems which we have explained in section 20. There are also many stimulating pieces of research on mathematical models of learning and on neurophysiological models ([51], [52], [53] through [55], [58], [65]). We look forward to mutual assistance and cooperation among these areas with the consequence that there may be a more definitive realization of our ideas suggested in section 20.

REMARK 4. In section 18, we emphasized that any ultimate formulation of control theory should be given within the framework of information science. In spite of all the quite excellent work on control theory, we have to agree with A. T. Fuller ([29], p. 291) when he points out that at present there exists no theory of control which allows simultaneously for the noisiness, variability, and partial nonlinearity and the jelly-like behavior of plants. It must also be recognized that it is not possible to measure all of the components of a state vector and that its dynamic equation may not be known by the controller. In this connection the recent work of Bellman and his colleagues ([9], [11], [12], and [15]) seem to be of importance; it is devoted to the identification problem and the determination of unmeasurable state variables on the basis of the observations of a process.

REMARK 5. Our standpoint given in section 18 does seem to us to be almost in agreement with the views of Beer [4], [5], and [6].

REMARK 6. The terminology of information science(s) has sometimes been used in a sense different from ours. Tayler [69] gave a clear description of what he means by information science and it is broad enough to cover what we call library and documentation science. His definition is quite suitable for bringing together various areas which are essentially in need of librarians and documentators. From our explanations in section 12 it can be seen that all the areas suggested by Tayler are imbedded within our framework.

REMARK 7. M. C. Goodall [31] points out that the theory of cognitive systems deals with inductive inference from a new point of view, that of constructibility; that constructibility brings to light a basic paradox which is analogous to the situation that developed in logic with the discovery of Russel's [9] paradox. Bellman and Kalaba [8] have suggested the construction of a hierarchy of processes whereby to introduce precision into the concepts of "intelligence" and "thinking." Their standpoint can be understood in terms of ours because we have arrived at our information science by starting from the actual problems of statistics and step-by-step generalizing the notions which have been actually needed in this research. Although we have been most heavily concerned here with the abstract notions, one of the characteristic properties of our research strategy is to always have a basis which is adequately realistic and whose analytical methods promise ever new development.

This paper is dedicated to the celebration of the seventieth birthday of Professor Jerzy Neyman with congratulations on his scientific achievements, which are monumental in the history of statistics, and on his successful efforts in fostering international cooperation among statisticians and mathematicians.

REFERENCES

[1] R. L. Ackoff, *Progress in Operations Research*, Vol. 1, Publications in Operations Research, New York, Wiley, 1961.

[2] J. J. Anscombe and J. W. Tukey, "The criterium of transformation," unpublished manuscript, 1954.

[3] W. R. Ashby, "Requisite variety and its implications for the control of complex systems," *Cybernetica*, Vol. 1 (1958), pp. 83–99.

[4] S. Beer, "Operational research and cybernetics," *Proceedings of the 1st International Congress on Cybernetics*, Namur, Paris (1958), pp. 29–45.

[5] ———, "A technical consideration of the cybernetic analogue for planning and programming," *Proceedings of the 1st International Congress on Cybernetics*, Namur, Paris (1958), pp. 515–534.

[6] ———, "The impact of cybernetics on the concept of industrial organization," *Proceedings of the 1st International Congress on Cybernetics*, Namur, Paris (1958), pp. 535–554.

[7] R. Bellman, *Adaptive Control Processes, A Guided Tour*, Princeton, Princeton University Press, 1961.

[8] R. Bellman and R. Kalaba, "Dynamic programming, intelligent machines, and self-organizing systems," *RAND Memo* 3175-PR (1962).

[9] R. Bellman, H. Kagiwada, and E. Kalaba, "Quasilinearization, system identification, and prediction," *RAND Memo* 3812-PR (1963).

[10] R. Bellman, R. Kalaba, and R. Sridhar, "Adaptive control via quasilinearization and differential approximation," *RAND Memo* 3928-PR (1965).

[11] R. Bellman, B. Gluss, and R. Roth, "On the identification of systems and the unscrambling of data: some problems suggested by neuro-physiology," *RAND Memo* 4266-PR (1964).

[12] ———, "Segmental differential approximation and the 'black box' problem," *RAND Memo*, 4269-PR (1964).

[13] R. Bellman and B. Gluss, "Adaptive segmental differential approximation," *RAND Memo* 4314-PR (1964).

[14] E. E. Bernard and K. R. Kare, *Biological Prototypes and Synthetic Systems*, Vol. 1, Second Bionics Symposium, sponsored by Cornell University and the General Electric Company, Advanced Electronics Center, New York, Plenum Press, 1962.

[15] H. Billing, "Lernende Automaten," Bericht über die Fachtagung der Nachrichtentechnischen gesellschaft im DVE (NTG), Fachausschluss 6, *Informationsverarbeitung*, München, 1961.

[16] G. E. P. Box, "Evolutionary operation: a method for increasing industrial productivity," *Appl. Statist.*, Vol. 6 (1957), pp. 3–23.

[17] B. E. COOPER, "Designing the data presentation of statistical program for the experimentalist," *The 34th Session of the International Statistical Institute at Ottawa*, Vol. 40 (1963), pp. 567–587.
[18] P. W. COOPER, "The hyperplane in pattern recognition," *Cybernetica*, Vol. 5 (1962), pp. 215–238.
[19] L. COUFFIGNAL, "Le symposium de Zurich et les concepts de base de la cybernetique," *Cybernetica*, Vol. 1 (1958), pp. 15–31.
[20] W. E. DEMING, *Some Theory of Sampling*, New York, Wiley, 1950.
[21] A. S. DOUGLAS and A. J. MITCHELL, "AUTOSTAT: A language for statistical data processing," *Comput. J.*, Vol. 3 (1960), pp. 61–66.
[22] A. V. FEIGENBAUM, *Quality Control, Principles, Practices, and Administration*, New York, McGraw-Hill, 1961.
[23] ———, *Total Quality Control—Engineering and Management*, New York, McGraw-Hill, 1961.
[24] R. A. FISHER, "Statistical method and scientific induction," *J. Roy. Statist. Soc. Ser. B*, Vol. 17 (1955), pp. 69–78.
[25] ———, *Statistical Methods and Scientific Inferences*, London, Oliver and Boyd, 1959 (2d rev. ed.).
[26] J. D. N. DE FREMERY, personal communication.
[27] K. S. FU, "A sequential decision model for optimum networks," *Biological Prototypes and Synthetic Systems*, Vol. 1, edited by E. E. Bernard and M. R. Kare, New York, Plenum Press, 1962.
[28] ———, "A statistical approach to the design of intelligent machines—pattern recognition and learning," *Cybernetica*, Vol. 5 (1962), pp. 88–102.
[29] A. T. FULLER, "Directions of research in control," *Automatics*, Vol. 1 (1963), pp. 289–296.
[30] G. FÜRST, "Changing tasks in official statistics," *The 34th Session of the International Statistical Institute at Ottawa*, Vol. 15 (1963), pp. 196–206.
[31] F. H. GEORGE, *The Brain as a Computer*, New York, Pergamon Press, 1962.
[32] M. C. GOODALL, "Induction and logical types," *Biological Prototypes and Synthetic Systems*, Vol. 1, edited by E. E. Bernard and M. R. Kare, New York, Plenum Press, 1962.
[33] K. JANKOWSKI, "The dynamic investigation of stochastic sequences," *Cybernetica*, Vol. 5 (1962), pp. 155–197.
[34] J. M. JURAN, L. A. SEDER, and L. M. GRYNA, *Quality Control Handbook*, New York, McGraw-Hill, 1961 (2d ed.).
[35] L. KANAL, "Evaluation of a class of pattern-recognition networks," *Biological Prototypes and Synthetic Systems*, edited by E. E. Bernard and M. R. Kare, New York, Plenum Press, 1962.
[36] M. G. KENDALL and P. WEGNER, "An introduction to statistical programming," *The Review of the International Statistical Institute at Ottawa*, Vol. 31 (1963), pp. 305–360.
[37] T. KITAGAWA, "Successive process of statistical inferences," *Mem. Fac. Sci. Kyushu Univ. Ser. A*, Vol. 5 (1950), pp. 139–180; Vol. 6 (1951), pp. 54–95; Vol. 6 (1952), pp. 131–155; Vol. 7 (1953), pp. 81–105; Vol. 8 (1953), pp. 1–29; *Bull. Math. Statist.*, Vol. 5 (1951), pp. 35–50.
[38] ———, "Successive process of statistical controls," *Mem. Fac. Sci. Kyushu Univ. Ser. A*, Vol. 7 (1952), pp. 13–26; Vol. 13 (1959), pp. 1–16; Vol. 14 (1960), pp. 1–33.
[39] ———, "Successive process of statistical inferences applied to linear regression analysis and its specializations to response surface analysis," *Bull. Math. Statist.*, Vol. 8 (1959), pp. 80–114.
[40] ———, "A mathematical formulation for evolutionary operation programs," *Mem. Fac. Sci. Kyushu Univ. Ser. A*, Vol. 15 (1961), pp. 21–71.
[41] ———, "The logical aspects of successive processes of statistical inferences and controls," *Bull. Inst. Internat. Statist.*, Vol. 38 (1961), pp. 151–164.

[42] ———, "Estimation after preliminary test of significance," *Univ. California Publ. Statist.*, Vol. 3 (1963), pp. 147–186.

[43] ———, "The relativistic logic of mutual specification in statistics," *Mem. Fac. Sci. Kyushu Univ. Ser. A*, Vol. 17 (1963), pp. 76–105.

[44] ———, "Automatically controlled sequence of statistical procedure in data analysis," *Mem. Fac. Sci. Kyushu Univ. Ser. A*, Vol. 17 (1963), pp. 106–129.

[45] ———, "Automatically controlled sequence of statistical procedures," *Bernoulli, Bayes, Laplace, Anniversary Volume*, Berlin, Springer-Verlag (1965), pp. 146–178.

[46] L. J. Krieger, "Soviet philososphy, science and cybernetics," *RAND Memo* 3619-PR (1963).

[47] R. Levien and M. E. Maron, "Cybernetics and its development in the Soviet Union," *RAND Memo* 4156-PR (1964).

[48] A. A. Lyapunov, "General problems of cybernetics," *Problems of Cybernetics, I*, New York, Pergamon Press, 1960, pp. 1–21.

[49] P. C. Mahalanobis, "On large-scale sample surveys," *Philos. Trans. Roy. Soc. London Ser. B*, Vol. 231 (1944), pp. 329–451.

[50] ———, "Some aspects of the design of sampling surveys," *Sankhyā*, Vol. 12 (1952), pp. 1–7.

[51] M. Minsky and O. G. Selfridge, "Learning in random nets," *Information Theory*, Fourth London Symposium, edited by Colin Cherry, London, Butterworth, 1961.

[52] M. E. Maron, "Mechanisms underlying predicative behavior for an intelligent machine," *RAND Memo* 3011-PR (1962).

[53] G. A. Muses, "Aspects of the theory of artificial intelligence," *Proceedings of the First International Symposium on Biosimulation, Lacarno 1960*, New York, Plenum Press, 1962.

[54] A. Newell, "Learning, generality and problem-solving," *RAND Memo* 3285-PR (1962).

[55] ———, "Some problems of basic organization in problem-solving problems," *RAND Memo* 3283-PR (1962).

[56] J. Neyman, "On the problem of the most efficient tests of statistical hypotheses," *Philos. Trans. Roy. Soc. Ser. A*, Vol. 23 (1933), pp. 289–337.

[57] J. Neyman and E. S. Pearson, "The testing of statistical hypotheses in relation to probabilities a priori," *Proc. Cambridge Philos. Soc.*, Vol. 29 (1938), pp. 492–510.

[58] S. Papert, "Some mathematical models of learning," *Information Theory*, Fourth London Symposium, edited by Colin Cherry, London, Butterworth, 1961.

[59] G. Pask and H. von Foerster, "A predicative model for self-organizing systems," *Cybernetica*, Vol. 3 (1960), pp. 258–300; Part II, *Cybernetica*, Vol. 3 (1961), pp. 20–55.

[60] K. Pearson, *The Grammar of Science*, Oxford University Press, 1896.

[61] L. S. Pontryagin, *The Mathematical Theory of Optimal Processes*, New York, Prentice Hall, 1962. (Translated from the Russian.)

[62] S. J. Reed, "Screening rules," M.S. thesis, Rutgers University, 1959.

[63] K. H. Reid, "Design studies of conditional probability computers," *Biological Prototypes and Synthetic Systems*, Vol. 1, edited by E. E. Bernard and M. R. Kare, New York, Plenum Press, 1962.

[64] H. Robbins and S. Monroe, "A stochastic approximation method," *Ann. Math. Statist.*, Vol. 22 (1951), pp. 400–407.

[65] L. Rosenblatt, *Principles of Neuro-Dynamics; Perception and the Theory of Brain Mechanisms*, Washington, 1962.

[66] G. S. Sebestyen, *Decision-Making Processes in Pattern Recognition*, New York, Macmillan, 1962.

[67] W. A. Shewhart, "Statistical method from the viewpoint of quality control," Washington, D. C. (1939).

[68] K. Steinbuch and S. W. Wagner, "Neuere Ergebnisse der Kybernetik," Bericht über die Tagung Karlsruhe 1963 der Deutschen Arbeitsgemschaft Kybernetik, Munchen, 1964.

[69] R. S. Tayler, "The information sciences," *Library Journal*, Vol. 88 (1965), pp. 4161–4163.

[70] M. E. Terry, "The principles of statistical analysis using large electronic computers," *The 34th Session of the International Statistical Institute at Ottawa*, Vol. 40 (1963), pp. 547–552.

[71] J. W. Tukey, "The future of data analysis," *Ann. Math. Statist.*, Vol. 33 (1962), pp. 1–68.

[72] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton, Princeton University Press, 1953.

[73] R. D. Turner, "First-order experimental concept formation," *Biological Prototypes and Synthetic Systems*, Vol. 1, edited by E. E. Bernard and M. R. Kare, New York, Plenum Press, 1962.

[74] A. Wald, *Statistical Decision Functions*, New York, Wiley, 1950.

[75] W. H. Ware and W. B. Holland, "Soviet cybernetics technology: I Soviet cybernetics, 1959–1962," *RAND Memo* 3675-PR (1963).

[76] S. Watanabe, "Information theoretical analysis of multivariate correlation," *IBM J. Res. Develop.*, Vol. 4 (1960), pp. 66–82.

[77] ———, "A note on the formulation of concept and of association by information-theoretical correlation analysis," *Information and Control*, Vol. 4 (1961), pp. 291–296.

[78] N. Wiener, *Cybernetics or Control and Communication in the Animal and the Machine*, New York, Wiley, 1948 (1st ed.); New York, M.I.T. Press, 1961 (2d ed.).

[79] ———, "The mathematics of self-organizing system," *Recent Development in Information and Decision Processes*, edited by R. E. Machol and P. Gray, New York, Macmillan, 1962, pp. 1–21.

[80] ———, "My connection with cybernetics, its origins and its future," *Cybernetics*, Vol. 1 (1958), pp. 1–14.

[81] ———, *God and Golem, Inc., A Comment on Certain Points where Cybernetics Impinges on Religion*, New York, M.I.T. Press, 1964.

[82] N. Wiener and J. P. Schade, "Nerve, brain and memory models," *Progress in Brain Research*, Vol. 2, Amsterdam/London/New York, Elsevier, 1963.

[83] ———, *Progress in Biocybernetics*, Vol. 1, Amsterdam/London/New York, Elsevier, 1964.

[84] R. A. Wiesen and E. H. Shuford, "Bayes strategies as adaptive behavior," *Biological Prototypes and Synthetic Systems*, Vol. 1, edited by E. E. Bernard and M. R. Kare, New York, Plenum Press, 1962.

[85] S. S. Wilks, "New fields and organizations in the mathematical sciences," *The Math. Teacher*, Vol. 51 (1958), pp. 2–9.

[86] F. Yates and H. R. Simpson, "A general program for the analysis of surveys," *Comput. J.*, Vol. 3 (1960), pp. 136–140.

[87] L. Yates, "Computer in research-promise and performance," The Presidential Address, *Comput. J.*, Vol. 4 (1961), pp. 273–279.