[Reprinted from the Journal of the AMERICAN STATISTICAL ASSOCIATION, September, 1931]

STATISTICAL METHOD FROM AN ENGINEERING VIEWPOINT [1]

BY W. A. SHEWHART, BELL TELEPHONE LABORATORIES, INC.

The possibilities of scientific development must be viewed differently today than they were a few years ago. No longer do we believe that we can do *exactly* what we want to do by making use of *exact* laws of nature. At the beginning of the period of great industrial development, scientific principles were assumed to have an exactness which today we do not believe that they have. Since progress in engineering depends upon the application of physical laws in the use of physical properties, and since both physical laws and physical properties are statistical in character, there is great need for the development of a statistical method that will meet the demands of the engineer in the solution of his problems.[2] Engineering problems place demands upon statistical method that are not met by the discussions of the subject in the literature.

I have been forced to listen to some rather drastic criticisms of statistical method as customarily discussed insofar as such discussion fails to cover the kinds of problem an engineer must consider. Let us look at a few typical criticisms of this character before we consider the extension of the statistical method to cover these criticisms. As I have considered elsewhere [3] the criticism of some current concepts of the theory of random sampling, I shall not touch here upon this important subject.

One of the engineer's first criticisms is that there appears to be wide divergence in the opinions of different authors as to just what is meant by statistical method. Furthermore, he points out that the descriptions of the method are often vague. Let me give you a case in point. If we turn to one of the classics on statistical method, we find the following statement: "By statistical methods we mean methods specially

[1] Revision of a paper delivered at the Ninety-Second Annual Meeting of the American Statistical Association at Cleveland, Ohio, December 29, 1930.

[2] With this end in view, a Round Table Conference on the Analysis and Interpretation of Engineering Data was called by the American Society of Mechanical Engineers and the American Society for Testing Materials on December 5, 1929. This meeting was held in New York City and was attended by representatives from many fields of engineering. At this meeting it was resolved that a committee for the development of statistical applications in engineering and manufacturing be formed, consisting of one member from each of the following four societies: the American Society for Testing Materials, the American Society of Mechanical Engineers, the American Mathematical Society, and the American Statistical Association. It was further resolved that this Committee recommend to the several societies a program for continued activity.

In accord with this resolution, appointments were made by the four societies concerned. As the representative of the American Statistical Association on this Committee and as an engineer, I was asked to present the problem involved in applying statistical method from an engineering viewpoint.

[3] "Random Sampling," *The American Mathematical Monthly*, Vol. XXXVIII, pp. 245–270, May 1931.

adapted to the elucidation of quantitative data affected by a multiplicity of causes." I have been asked many times what the author really means by "elucidation of quantitative data affected by a multiplicity of causes." An engineer goes to this book to find out what the statistical method will do for him, and he fails to find a satisfactory answer in the above complex phraseology.

If the same engineer turns to a certain well-known treatise on the theory of probability, he finds the following statement: "By the science of statistics we understand the recording and subsequent quantitative analysis of observed mass phenomena." What does this mean? If in disappointment the engineer turns to a certain French treatise, he will find the statement that the statistical method consists of four parts: (1) registration of facts, (2) classification of results and construction of tables, (3) presentation of results, (4) critical study of results.

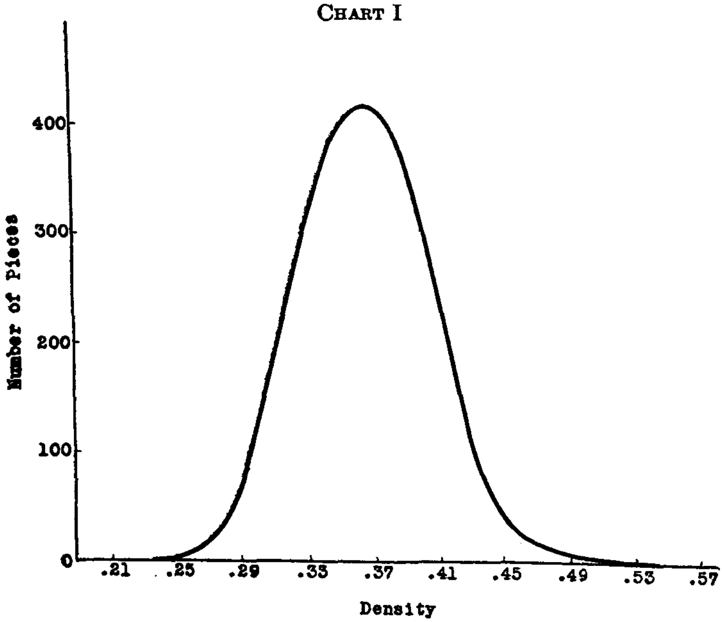
Finally if he looks at one of the best recent books on modern statistical method in the English language, he finds the statement: "Statistics may be regarded as (1) the study of populations, (2) the study of variation, (3) the study of the methods of reduction of data." To say the least, this description of statistical method does not appear on the face of it to be the same as that given in the French treatise.

If an engineer turns to another recent book on mathematical statistics, he will find the following statement: "By statistical methods, we shall mean methods of summarizing, comparing, and analyzing quantitative data of a certain type. The exact characteristics of such data will become clearer as we proceed; we state here merely that the data of the social sciences are normally of this type, and that the data of the exact and natural sciences have some of the characteristics of this type, insofar as errors of measurement or extraneous influences interfere with exact observation."

This book tells the engineer, as it were, that the only things that are statistical with which he has to deal are errors; and yet that engineer knows that the very reason why he is in search of the assistance of statistical method is that every physical property and every physical law is statistical in nature. For example, the density of Sitka spruce as determined by the measurements on 2,105 pieces made by the Forest Products Laboratory [1] is shown in Chart I. This distribution of density is not a distribution of errors: *it is the distribution function of a physical quantity*, errors of measurement being negligible in this case. It is typical of the large variation that we find in all physical properties even when considered from a macroscopic viewpoint.

[1] Technical Bulletin No. 158 of United States Department of Agriculture, *Comparative Strength Properties of Woods Grown in the United States*, by L. J. Markwardt.

CHART I



The author of one of the best texts on mathematical statistics has summarized what appears to me to be the content of classic statistical method. We may state his two objectives in statistical phraseology as: (1) the presentation of data, (2) the determination of the universe from the sample, or the theory of estimation.

Suppose an engineer feels satisfied with these two objectives and sets out to read all there is in the literature bearing upon these two phases of statistical method. He will likely be impressed with the fact that there is an indefinitely large number of ways of presenting data and an indefinitely large number of ways of going from the sample to the universe. After an excursion of this character, suppose that this engineer comes back to his job, and suppose that that job is the development of a new material, one of the important physical properties of which is its tensile strength. Suppose that his assistant, in accord with previously given instructions, has made five pieces of this new material which have been tested for tensile strength with the results expressed in pounds per square inch shown as follows:

29314
34860
38020
25810
30120

How can this engineer profit by all of the high-brow statistical theory which he has digested on the presentation of data and the theory of estimation? Carried away as he is likely to be with his new interest in life, he may do some fancy tricks with the data and after a lengthy discourse about modern distribution theory, make some prediction as to what he may expect to get in the future in the production of this kind of material. If, however, as is usually the case, this engineer statistician has one or more colleagues experienced in an engineering way in the production of this kind of material, he will likely be asked some embarrassing questions.

One of his colleagues may say, for example: "Yes, possibly this high-brow theory of estimation is all right if you are sure that the data with which you start are *good.*" That is apt to be a bad blow for the engineer statistician. Good? What does the engineer mean by good data? It is likely that our engineer statistician has not found any discussion of criteria for good data in his survey of the field. At first he may be inclined to take an argumentative stand and to say that statistical method applies to all data. At least I have heard such arguments presented under conditions of lapse of memory.

Suppose, however, that the statistician's colleague is a physicist. He is likely to make some remark like this: "Suppose you ask a Freshman to go into the laboratory and measure some physical quantity such as the charge on an electron, and he comes back with, let us say, ten measurements. Do you mean to tell me that the statistical method of presenting and interpreting these data is exactly the same as it is for presenting the original series of measurements made by Millikan?" Such a remark makes one realize that data may be good, bad, or indifferent, even though one has not met in his reading any description of the characteristics of good data.

Such a remark forces the engineer statistician to see that something is overlooked when we say that statistical method is divided into the two phases: presentation and estimation. Roughly speaking, this part of the theory applies if the data are good, but no account is taken of the fact that data in a great many instances are not good. *If statistical method is to prove of any great value, it must help the engineer (or anyone else for that matter) to get good data, as well as help him to present and to predict after he has obtained good data.*

In general, the goal of the engineer is to do a thing the way he wants to do it within limits that are economical. The things with which he has to work are physical properties of materials and physical laws, both of which are statistical. In anything that he tries to do, there is an element of chance. The engineer like the physicist or chemist tries to eliminate chance insofar as it is feasible for him to do so. He realizes, however, that after he has gone as far as he can in eliminating chance, he will still be a considerable distance from his goal. Chance we must have with us always, but the engineer's concept of chance is somewhat different from the chance that he sees discussed in books on statistics.

To an engineer chance merely stands for the *unknown;* a chance cause is an unknown cause. The engineer's problem is to eliminate all assignable or findable causes of variation in what he is trying to do. Let us therefore approach the subject of statistical method from this viewpoint.

We start with the three postulates:

Postulate 1: All chance systems of causes are not alike in the sense that they enable us to predict the future in terms of the past.

Postulate 2: Systems of chance causes do exist in nature such that we can predict the future in terms of the past even though the causes be unknown. Such a system of chance causes is termed constant.

Postulate 3: It is physically possible to find and eliminate chance causes of variation not belonging to a constant system.

The justification of these three physical postulates in the field of engineering has been considered elsewhere.[1] They give a basis for defining what we mean by good data and for developing criteria which have been found of great value in obtaining good data.

The engineer is interested in finding and eliminating assignable causes of variation in physical phenomena. However, after he has gone as far as he can in doing this, he is interested in predicting what will happen in the future. This he can do upon the basis of Postulates 2 and 3. The resultant data are of value in that they make prediction possible and are therefore worth presenting. On the other hand, an engineer appreciates full well that if the data have not come from a constant system of chance causes, prediction within limits is not possible on a rational basis, and hence they are of little fundamental value. The engineer therefore describes the statistical method in something like the following form:

1. Development of criteria for determining when variations in a phenomenon must be left to chance.
2. Estimation or prediction when variations in a phenomenon must be left to chance.
3. Presentation of data from which prediction is possible.

Let us illustrate the application of statistical method from this viewpoint.

[1] "Economic Quality Control of Manufactured Product," by W. A. Shewhart, *Bell System Technical Journal*, Vol. IX, April, 1930.

As a result of industrial research to develop a high insulation material, 204 pieces were obtained which gave the values of resistance in megohms presented in Table I.

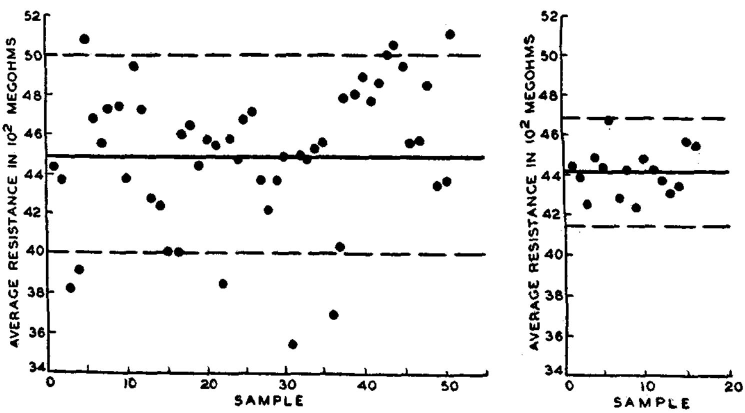
TABLE I

5045	4635	4700	4650	4640	3940	4570	4560	4450	4500	5075	4500
4350	5100	4600	4170	4335	3700	4570	3075	4450	4770	4925	4850
4350	5450	4110	4255	5000	3650	4855	2965	4850	5150	5075	4930
3975	4635	4410	4170	4615	4445	4160	4080	4450	4850	4925	4700
4290	4720	4180	4375	4215	4000	4325	4080	3635	4700	5250	4890
4430	4810	4790	4175	4275	4845	4125	4425	3635	5000	4915	4625
4485	4565	4790	4550	4275	5000	4100	4300	3635	5000	5600	4425
4285	4410	4340	4450	5000	4560	4340	4430	3900	5000	5075	4135
3980	4065	4895	2855	4615	4700	4575	4840	4340	4700	4450	4190
3925	4565	5750	2920	4735	4310	3875	4840	4340	4500	4215	4080
3645	4190	4740	4375	4215	4210	4050	4310	3665	4840	4325	3690
3760	4725	5000	4375	4700	5000	4050	4185	3775	5075	4665	5050
3300	4640	4895	4355	4700	4575	4685	4570	5000	5000	4615	4625
3685	4640	4255	4090	4700	4700	4685	4700	4850	4770	4615	5150
3463	4895	4170	5000	4700	4430	4430	4440	4775	4570	4500	5250
5200	4790	3850	4335	4095	4850	4300	4850	4500	4925	4765	5000
5100	4845	4445	5000	4095	4850	4690	4125	4770	4775	4500	5000

Reading from top to bottom beginning at the left column and continuing throughout the table gives the order in which the pieces of material were made. The question is: Has the engineer gone as far as he can go in eliminating unknown causes of variation? If he has, it should then be possible to use these data as a basis for predicting what he may expect to get in the future, and the tabulation of the results in satisfactory form should serve as standards for this particular quality. If, however, causes of variation not belonging to a constant system were present in this case, prediction would not be possible upon a rational basis.

Dividing these data into samples of four, for reasons which we cannot go into here, and applying a criterion which has recently been developed for determining when variation must be left to chance, we get the results shown graphically at the left of Chart II. The black dots

CHART II



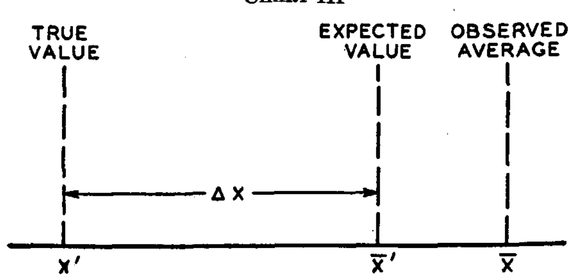
represent the successive averages of four. If one of these falls outside the dotted limits, it is taken as an indication of the existence of trouble in the form of an assignable cause or one that does not belong to a constant system. The fact that several of the points fall outside the limits in this case indicates trouble. Further research found and eliminated certain assignable causes of variation after which another test of like nature gave the results shown in the control chart at the right of this Chart. This is indicative of what extensive investigations of this character in our own work have revealed. In other words, our experience indicates that we may make use of recent distribution theory and the fundamental law of large numbers to establish criteria which materially assist in detecting the presence of assignable causes of variation when they exist, thus assisting us to get good data.

The successful operation of these criteria involves the choice of statistics to be used, the choice of ways of using these statistics, and the choice of limits. The basis for these three selections is, in the last analysis, of an experimental nature, and needless to say, cannot be considered here.

Still more important is the fact that we must make use of intuition and judgment in dividing the data into rational subgroups. We cannot do more, in this connection, than illustrate what we mean by a simple case familiar to all.

Suppose, for example, we wish to measure some physical quantity whose true but unknown magnitude is X'. Such a quantity, for example, might be the charge on an electron. Suppose we make a series of n measurements, the average of which is $\bar{X}$. If there is no constant error of measurement and if the errors of measurement are controlled by a constant system of causes with an expected value of zero, the law of large numbers tells us that the observed average $\bar{X}$ will approach the true value X' as a statistical limit as the number n of observations approaches infinity. If, however, assignable causes of error are present, this limit does not hold. In a similar way, if a constant error ΔX is present, the limit approached in the statistical sense is $X'+\Delta X$ and not X', Chart III. To eliminate the constant

Chart III



error requires the use of intuition and judgment in discovering the error; it cannot be eliminated statistically.

The kind of intuitive guidance required in eliminating the constant error is required in the division of data into rational subgroups, where for the purpose of our discussion *a rational subgroup is taken to be the objective set of data coming from a constant system of chance causes.*

In this way we arrive at a rational basis for obtaining good data, involving the application of criteria for determining when variations must be left to chance. After the good data have been obtained, the engineer may apply existing theory of presentation and estimation with full assurance that the use of such theory helps him to do something that he could not do as well otherwise.

I have tried to show why it is that without the use of recently developed criteria for determining when a phenomenon must be left to chance, we are liable to make applications of existing theory of presentation and estimation in a way that will lead to serious fallacies.

In the light of this study it appears that the statistical method is threefold in character and that it is capable of fitting into the engineer's program of doing what he wants to do within limits that are economical by helping him to weed out assignable causes of variation; to predict after assignable causes of variation have been eliminated; and to present the resultant data in a way to contain the essential information. These recent developments appear to open up important avenues of progress in the extension of the theory of statistics as a tool for all who are trying to make use of physical properties and physical laws, and should prove of even greater economic importance than the many applications of statistical theory that have already been made in industry.