APPLICATION OF STATISTICAL METHOD IN MASS PRODUCTION

by

W. A. Shewhart

Paper prepared for presentation at the Conference of Engineering and Industrial Statistics to be held at Massachusetts Institute of Technology, Sept. 8-9, 1938.

INTRODUCTION

Sailing Directions

When Professor Freeman gave me sailing directions for the preparation of this paper, he told me that most of the audience would be composed of three groups: budget directors, executive vice presidents, and directors of research. He also emphasized the fact that few in the audience would be statisticians.

The fact that you are here I take as evidence that you have perhaps heard or read some of the outpourings of that enthusiastic and rapidly growing class known as ~~of~~ mathematical statisticians and you have come in the hope of finding out what, if anything, statistical method can do for you. I assume that you at the moment are pretty much in the state of mind of one of my managerial friends who greeted me on a recent occasion as follows: "Say Shewhart, I want the low-down on all this shouting of statisticians like yourself about the value of your wares. What's more, I want the answer in plain everyday English and not in that foreign lingo that you fellows like to spout. Of course, if you can convince me that what you have to sell is half as good as you think it is, I'll gladly take time to learn your lingo because I realize that every science has its special vocabulary. However, I haven't time to go on any wild goose chase and I am not going to put time on statistics until I am convinced that fellows like yourself are not seeking a pot of gold at the end of a statistician's rainbow.

"As a scientist, I think I know something about scientific method. As a director of research, that method is my stock in trade. Now how does this thing you fellows call statistical method differ from scientific method? If it is a part of scientific method, then what part? Answer me these questions and tell me in general terms what I could do with statistics that I cannot do without and I'll buy you a good lunch."

Well, I had more than an hour to talk to this friend, and I got my lunch. Today in somewhat less than an hour, I shall try to sketch in outline my answer to this doubting Thomas friend of mine before he became a convert to the ways of the statistician.

<u>Choosing a Subject of Common Interest to All of Us</u>

I think that most of you are like my friend in not being interested in the discussion of methodology ~~abstract principles~~ unless this is ~~these are~~ presented against a background of practical problems like those <u>you</u> meet in your daily work. Hence let us start by painting such a background. Everyone of us is called upon to interpret today's experiences in terms of tomorrow's; to infer what is going to happen tomorrow in the light of what has happened in the past. Our success in our job - no matter what that job is - depends a lot upon our being able to make valid predictions about future experience in terms of past experience.

One fly in the ointment is that we are surrounded everywhere we go with little demons of CHANCE. Even in the field of <u>exact</u> science, these little demons make it impossible for us to get exactly the same results when we do our very best to repeat even the most elite measurements of physical science. What we would give to be able to predict the whims of these little fellows! What a still more glorious world this would be if we could get rid of them altogether! So far as what I shall say has to do with the method of getting rid of the demons of chance and of predicting their whims when we cannot get rid of them, we shall be on somewhat common ground. Further than this, the subject matter used in illustrating the principles is drawn from the field of quality control of manufactured products and that is a field in which all of us are either directly or indirectly interested.

We live in an age of applied science. The food we eat, the clothes we wear, the houses we live in, the cars we drive, in fact almost everything we buy and use is either directly or indirectly the result of applied science. Whether we are producers or consumers of goods made possible by ~~the results of~~ applied science, we are affected by anything that affects the control of these goods ~~our physical environment~~ through the use of scientific principles.

Now, one of the greatest contributions of this machine age is mass production through the manufacture of interchangeable parts. The introduction of interchangeable manufacture made possible improved quality and cheaper costs of the goods we use and it has given us many new products that we would not otherwise have had. This was accomplished without any help from statistical method.* However, by using statistical method, the advantages accruing from mass production can be materially increased.

In what way, you ask, does statistical method perform this service? Broadly speaking, ~~the answer is two ways:~~ a) through its use, certain desired ends can be attained in the control of quality with less cost than these ends could be attained otherwise; b) through its use, some desired ends can be attained that could not possibly be attained otherwise. In one case, it is a matter of dollars and cents as for example, when we reduce the cost of inspection and the number of rejections. In the other case, it is a matter of doing something with statistical method that we want to do but that we cannot do without it. For example, in building a structure such as a bridge or machine, we may be limited by some property such as tensile strength. Because of the variability of any

---

* For an interesting summary of the advantages secured by mass production before the days of statistics, see ~~an article by~~ J. W. Roe, Mechanical Engineering, Oct. 1937, pp. 755-758.

Interchangeable Manufacture

such quality, we must often allow quite a wide tolerance range. By reducing the variability, we may make more efficient use of material and effectively extend the field of possible structures.

Three Propositions

In particular, I shall try to establish the following three propositions:

A. Statistical Method + Mass Production = New Tool of Research.

B. With this new tool of research, some useful things can be done that cannot be done otherwise. In particular, we can track down and get rid of some of the little demons of chance that could not otherwise be budged from their lair, thus making it possible to make the most efficient use of raw and fabricated materials and to attain the highest possible quality assurance.

C. The use of statistical method in mass production provides a scientific method for minimizing the cost of producing things to satisfy human wants.

PROPOSITION 1:
STATISTICAL METHOD + MASS PRODUCTION = NEW TOOL OF RESEARCH

You will recall that my doubting Thomas friend's first question was in respect to this thing we call statistical method. What is it and how does it differ from scientific method? How often have I heard that question in the last ten years! Very often when starting to outline what a statistician does that makes him a statistician, I have been interrupted with the comment: "Why, that is what any scientist worthy of the name does every day!" As previously noted, I understand that many of you ~~at least~~ are not trained statisticians and may therefore feel much the same way. Hence in order that we may get off on the right foot, it is desirable that we consider briefly what it is that makes a method statistical.

Statistical Method - Mathematical

Some ~~20~~ twenty years ago when out of a clear sky and without previous training, I was called upon to solve what was called (and rightly so) a statistical problem, I began trying to find out what statisticians meant by statistical method and I have been trying ever since. I have collected quite an assortment of definitions over this period and of these one might say that there are no two of a kind. One finds much described in some books on statistics that can be found in books on logic or books on experimental method. Some statistical methods today are given fancy names - as for example, operation of statistical control, analysis of variance, design of experiment and the like - although much that the laymen sees in these methods he may have seen many times in different dress in books on scientific method. Again and again I have had engineers call my attention to the fact that the idea of improving the control of quality through the continuing process of detecting and eliminating causes of variability was an old, old story even before the statistician appeared on the scene. Particularly those engineers and scientists trained in physics and chemistry have likewise often pointed out to me that design of experiment is one of the oldest tricks in the scientists trade. They point out also that the laboratory trained research worker is continually studying and analyzing the observed variability (of which is one of many measures variance) in order to break it up into component parts. Such men want to know just what it is that makes the statisticians control of quality, analysis of variance, or design of experiment different from that of the successful engineer and scientist of the past. Far be it from me therefore to ever venture any statement about statistical method in the sense of the statistical method. Instead I shall simply try to indicate what I shall mean in this discussion by statistical method. Here we can learn from the dictum of the contract bridge artists: Always declare your "method of bridge" for the enlightenment of your partner. Insert #1

That which characterizes a statistician so far as we are here concerned is that he is in possession of a particular kind of mathematical model that non-statisticians do not have. The mainspring in this model is the concept of a frequency distribution from which a sample of any size n can be drawn at random.

A variable X will be said to have the frequency distribution f(X) if the frequency of occurrence of X in the range $X_1 < X < X_2$ is given by

However, by way of contrast, ~~It~~ may be helpful at this point to call attention to a procedure used in describing statistical method that is often found in the literature of statistics even today. For example, we are told that statistical method is mainly concerned with data affected by a multiplicity of causes. In this sense, statistical method is contrasted with the scientific method of the laboratory where the effects of all but one cause are supposed to be reduced to a comparatively small residuum. We are told that statistical method must accept for analysis data that are subject to the influence of a host of causes and try to discover from such data which causes are the important ones and what effect each cause produces. We are also told that statistical method is to be applied where the multiplicity of causes of variation in the observed phenomenon cannot be controlled and the observer has to take observations of things as they are and make what inferences he can from the data. Again and again statisticians have pointed to the social sciences as illustrating the difficulties of their science in that the observer of social facts cannot experiment and must study phenomena as they occur.

Several of my doubting Thomas friends - and I have had many among engineers, physicists, and chemists - have read enough about statistics to get just such a picture of the aims of statistical method as that presented in the preceding paragraph. They admit the existence of a multitude of "statistical problems" in this sense in the field of social science and elsewhere. However, many of ~~these doubting Thomas~~ them ~~friends~~ have watched the statistician perform his tricks on stock market and other kinds of "trends" and they have been none too favorably impressed, to say the least. They have asked on occasions what it profits a statistician to go through all the agony of using high-brow statistical methods in making such predictions if he cannot hit his mark more often than he has apparently done in the past. Regardless of the statistician's attitude on this question, his record in these ~~matters~~ fields where everyone has seen him work is not such as to sell many engineers, physicists, and chemists on the statistical method as a powerful tool.

To ~~do this~~ avoid the danger of having statistical method rejected upon the basis [illegible], we must look at ~~such~~ ~~this~~ statistical method a little more critically than is often done in order that we may see under what conditions it may be [illegible]. In what follows, I shall try to do this but with some trepidation, with many misgivings, and, I trust, with due humility. The outcome is a picture of the role of statistical method that is somewhat different from that just described. It is, I believe, a picture obtained primarily by adding to rather than subtracting from that just described. [illegible] also gives evidence that [illegible] be attained

What the layman sees is that many of the [illegible] predictions are not <u>valid</u> in these fields. The [illegible] (non-technical?) judges the method by its failure under [illegible] conditions. On the other hand, little is said in the literature about the conditions under which statistical method will lead to valid predictions.

$$\int_{X_1}^{X_2} f(X)\, dX. \tag{1}$$

For the sake of simplicity we shall assume that f(X) is a continuous function.

Now the phrase "drawn at random" describes the <u>operation of choosing a value of X from among the possible values of X without the choice depending in any way whatsoever upon the magnitude of the X that is chosen</u>. To my way of thinking, too much emphasis cannot be placed upon this concept of the <u>operation</u> of drawing at random.

Let us assume that

$$X_1, X_2, \ldots, X_i, \ldots, X_n \tag{2}$$

represents n such values drawn at random. Such a set of values is termed a random sample of n. Let θ_i be any specified function of these n values of the chance variable X, such as the average or standard deviation, for example. Such a function θ_i is customarily called a statistic at the suggestion of R. A. Fisher. Then the statistician conceives of the process of repeating again and again the act of drawing a sample of n from the frequency distribution. If he then computes the value of the statistic θ_i for each such sample in the order that the samples were drawn, he gets an infinite sequence,

$$\theta_{i1}, \theta_{i2}, \ldots, \theta_{ij}, \ldots \theta_{im}, \theta_{im+1}, \ldots, \theta_{im+2}, \ldots \quad (3)$$

This he calls a random sequence.

Knowing the parent distribution (1), the statistician tries to find some distribution function $\varphi(\theta_i)$ for samples of size n such that

$$\int_{\alpha}^{\beta} \varphi(\theta_i)\, d\theta_i \quad (4)$$

gives the frequency of occurrence of θ_i within the interval $\alpha < \theta_i < \beta$ to a sufficiently high order of approximation.*

There are, of course, an indefinitely large number of kinds of functions or statistics of the samples that might be computed, and for each of these, there is presumably an infinite random sequence (3) and a frequency distribution function (4).

The mathematical machine in the hands of the statistician is one into which you feed a frequency distribution and get out another frequency distribution representing the result of a random operation on the original distribution. It need not worry us here that this mathematical machine grinds slowly and has as yet only turned out comparatively few of the possible distribution functions of type (3). Sufficient for our story is the fact that the mathematical statistician has already turned out many distributions that are of great usefulness.**

---

* ~~Theoretically,~~ Of course, the function $\varphi(\theta_i)$ may be such that finite summation instead of integration must be used.

** As, for example, the distributions of K and [illegible] Students ratio, and R.A. Fisher's z, to mention three for illustration purposes.

Statistical Method as Scientific Method

So far we have mentioned the fundamental mathematical concepts behind statistical method. As in the case of any mathematical development, there is no <u>necessary</u> connection between the mathematics and observable phenomena. It is perfectly possible that the abstract mathematical model will not correspond to anything we may observe in nature and no amount of armchair philosophizing will tell us whether or not the mathematics is useful. To determine if the theory is useful, we must experiment.

So far as I see, the use of a mathematical model in statistics is not fundamentally different from the use of any mathematical model in physics or any other science. However, in drawing any scientific inference that contributes to empirical knowledge, we must start with observed data (A of Fig. 1) and end with observed data (B, Fig. 1), <u>but data different from those with which we start</u>. Mathematical

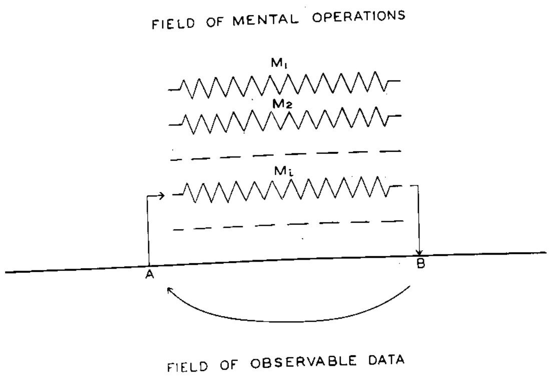


Fig. 1

and logical deductions are mental operations* that we use in the process of going from A to B. But, since there is no known necessary connection between such a deduction and observable data, we can only choose in a given case, some mental operation M_1, the use of which involves predictions as to what we may expect to find at B. We then must leave the field of mental operations and come back to the field of observable data at B to see if the predictions in this given case are valid. If, in a given case, the predictions are valid, we cannot, of course, conclude that M_1 is the operation to choose in going from A to B. ~~However, by repeating the process of choosing the mental operation M_1 whenever we have data like we had at A, having found that the predictions at B are valid, we increase our assurance that M_1 is a useful operation. That is,~~ However, we may increase our assurance that M_1 is a useful path by traversing successfully again and again the circuitous path of Fig. 1. If on the other hand, M_1 does not give valid predictions, we must choose some other path. We must learn by experience ~~that~~ what we cannot prove by deductive logic; namely, that we can often use formal mathematics as a step in the operation relating past experience to future experience.

Broadly speaking, Fig. 1 presents scientific method as an operation involving the following successive steps:

- -

*These mental operations are sometimes called constructs or conceptual models in the literature.

Now, we must note that one of the basic postulates for the usefulness of scientific method is that there must be a certain <u>uniformity</u> underlying experience. That is to say, having perceived certain data or sense impressions at A, it is necessary that we may uniformly expect more or less well-defined groups of data or sense perceptions at B every time the impressions at A are repeated. Physical and chemical theory deals with one class of phenomena of this kind. So far as I am here concerned, statistical theory must also deal with some class of phenomena of this kind if its results are to be useful. However, we must not lo~~o~~se sight of the fact that if such a uniformity does not exist, no amount of wishful thinking on our part will modify the situation so far as scientific method is concerned.

Again ~~Now~~ let us look at the mathematical machinery of the statistician ~~once more~~ <u>to see under what conditions it may reasonably ~~possibly~~ be expected to help us make valid predictions.</u> Have we any reason to expect it to ~~be useful~~ provide valid predictions when we have any old kind of complicated set of unknown or chance causes? Is it a method that will always <u>work</u> when we cannot use the so called "experimental" technique of controlling all but one cause in the laboratory? ~~Is it a kind of catch all method to be used when a multiplicity of causes affect the phenomena under observation?~~ These questions are indeed pertinent not only for those who would try to see behind the usefulness of statistical method in the sense that it is discussed in some of the older treatises in its application in the social sciences but also in the sense that it is discussed in some of the newer applications such as the analysis of variance, to mention only one. Once more we should note this idea

So far as I see, the answers to the ~~three~~ the questions ~~just raised~~ which above are ~~all~~ negative unless in all cases the complex structure of unknown chance causes is one that behaves in a very, very special manner, namely that this complex structure is one that produces variations in a manner bearing a close resemblance to what the mathematical statistician conceives of as <u>random</u>, and sad but true (at least in my own field of experience) nature does not have a habit of providing us with random data in the mathematical sense as I shall have occasion to emphasize several times as we proceed. In fact, an engineer has to do some controlling before he attains a random condition.

Of course, the statistician can always analyze a set of data at A of Fig. 1 upon the assumption that the condition of randomness exists and upon this assumption he can tell one what to expect of samples of future data. Furthermore I assume that all here would agree that these predictions would have the highest attainable validity upon the basis of the underlying assumption, but this does not mean, of course, that the basic assumption of randomness is valid and hence the predictions may not be valid. To justify th~~is~~e assumption of randomness in a given case calls for further experimentation of the kind that ~~character which~~ class I shall try to set forth in some of the simplest cases as we proceed.

the validity of predictions that is of great importance from the viewpoint of use. Hence it is not enough to know that the statistical method is useful in a given case, ~~but~~ we need to know **how** useful.

a) Taking data (A of Fig. 1).

b) Making hypotheses* (mental acts, Fig 1).

c) Using hypotheses as basis of drawing deductive inferences about measurements that can be made in the future (as B of Fig 1).

d) Testing hypotheses by making experiments to see if results deduced upon the basis of hypotheses are as predicted; modifying hypotheses if necessary, and starting again with the initial step.

From the viewpoint of our present story, I wish to emphasize three of the ~~requirements~~ characteristics of scientific method:

1. We must ~~be able to~~ start with data at A and make ~~valid~~ predictions of other data at B.

2. To validate a scientific method as a method of knowing requires that we be able to repeat at will a circular succession of steps.

3. At both A and B, an act of perception is involved.

Now let us tackle the first question that my doubting Thomas friend raised. How does statistical method differ

- -

*There are several different senses in which the word hypothesis ~~is~~ used in the literature. An hypothesis is here taken as the first ~~of~~ three stages in an induction process, the second and third stages being called deduction and verification respectively. In ~~this sense~~ an hypothesis is first propounded and then proved. However, since the process of induction can never lead to certainty in "verification", the third stage in the inductive process still leaves us with an hypothesis. At ~~In~~ going from A to B there are usually not one but several hypotheses involved.

from customary scientific method? The answer is that the statistician's mathematical model or machine simply provides the man of science with a new kind of hypothesis and a means of deducing inferences upon the basis of this hypothesis about measurements that can be made in the future but as yet have not been made. The kind of hypothesis at the basis of the mathematical deductions of the statistician is, as we have previously noted, one having to do with ~~about~~ what we may expect to get in samples drawn at random from a given population characterized by some ~~frequency~~ assumed frequency function (1).

<u>The Emergence of a Statistician as a Scientist</u>

Let me now try to paint a little more graphically just what it is that the mathematical statistician can do that no one else can do and then show what a statistician must do in order to become a scientist. Let us picture our statistician on the inside of a large black sphere so that he cannot see what is happening outside of the sphere except for numbers that appear on a small ground-glass window. Let us assume that numbers are flashed from time to time on this window. Now what the mathematical statistician can do with these numbers without ever getting outside his spherical shell that a non-mathematical statistician cannot do I shall consider to be his fundamental contribution. There are two kinds

One thing that he can obviously do is to record on a slip of paper each number as it appears on the screen. After getting a bunch of these slips, he can amuse himself by sorting them out into a frequency distribution. Then he can try to fit these discontinuous distributions by smooth curves.* He can use different kinds of functions as a starting point and can develop different criteria of fit. Thus he can build up pictures of what he has seen on his ground-glass screen. In the early days of statistics, this was the favorite pastime of the statistician.

The mathematical statistician did not long remain content to draw pictures, both graphical and analytical, of the distributions of numbers flashed on his screen. He soon had a brilliant idea: not only the distribution of the numbers but the order of their appearance should be of interest. Hence he began in effect ~~to~~ record the numbers that appeared on his ground-glass in the order that they appeared. In this way, he might get a sequence of n values such as

$$X_1, X_2, \ldots, X_i, \ldots, X_n.$$

appearing in this order. Then he hit on another idea; why not write these numbers on as many separate slips of paper, throw the slips into his hat, shake them well, and draw them out one at a time, writing the numbers down in the order that they were drawn? He could, of course, repeat this process again and again, each time getting an ordered sequence of the n numbers. What interested the mathematical statistician

* He may also try to summarize the numerical [illegible] provided by the observed distribution. See for example, [illegible] Statistical Method from the Viewpoint of Quality Control, chap. II [illegible] the United States Department of Agriculture [illegible]

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	4310	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

(a)

3875	4810	5000	4925	4375	4340	4050	4845	4615	4430	4445	4895
4065	4500	4450	4700	2965	4445	4485	4700	4570	4625	3665	4285
4615	4895	4850	4135	4850	4275	4500	3645	4335	4735	4340	5000
4915	2920	4355	4290	4335	3760	4080	4570	5000	4275	4850	4700
4450	4770	5000	4850	4640	4840	3635	4700	4570	4615	4500	4095
4170	4215	4170	4310	5200	4615	4185	4310	4700	4890	4700	3690
5000	3980	5000	5150	4125	4850	5000	4690	5000	4215	4255	4425
4700	4350	4500	4000	5075	4565	4300	4565	4325	4300	4100	4665
4640	4550	4700	4175	4410	3635	5000	3940	5075	4925	5075	4700
3650	3700	4640	3850	4560	4430	4850	4635	4790	5190	4635	4845
4450	4650	3925	3900	4560	4190	4310	5450	4895	5045	4160	4180
5250	4685	5100	4850	4840	5000	3463	4775	4570	5250	5075	3685
4090	4840	4500	4575	4095	4425	4340	5100	4855	4725	5750	4170
5000	4050	4925	4500	4375	4790	4375	4600	4790	4450	4740	5050
4685	4215	5000	4080	4770	4410	4770	4110	4255	4430	3635	4440
4325	4450	4700	4720	5000	4125	5600	3775	5150	4430	4340	4350
4575	4765	4625	5000	4700	4775	4930	3300	4080	3075	3975	2855

(b)

Can you see any fundamental difference between these two sequences?

Table 1

confined in his spherical shell was to see how the order of the n numbers as they flashed on the screen differed, if any, from the general run of orders observed when drawing the same set of n numbers out of his hat.*

In the course of the last few decades, he succeeded in finding lots of ways of comparing his "random" sequences with the ones that appeared on the screen. Some of this work he labeled distribution theory, some he called analysis of variance, some of it probability theory, some of it sampling theory, and some of it by still other names. But irrespective of what he happened to call the results of his labors, much of it was so far as we are here concerned pretty much the same: he was making it possible to ~~comparing~~ compare some property of a sequence of numbers flashed on his screen with the properties of the sequences he could construct from this given set of numbers (or from some hypothetical frequency distribution) by carrying out certain random operations. This has proved to be a fascinating game for the statistician and today he is turning out results faster than ever before and yet he can see no end to the work that he can do.**

At this point, it may be well to pause for a specific illustration. Let us assume that our statistician observed the sequence of 204 numbers shown in Table 1(a). Here the observed order starts with the first figure in the left column,

- -

*Needless to say, ~~of course,~~ our statistician being a mathematician didn't actually go to the trouble of making the hat experiment in the way I describe. Instead, he simply conceived of this kind of an operation and then figured out what he thought that he would get by such a process.

not only the results?

The practicing statistician ~~who is~~ interested in keeping abreast of the developments in this field will want to keep in touch with such journals as the Annals of Mathematical Statistics, Biometrika, the Journal of the Royal Statistical Society, and Sankyha, the Indian Journal of Statistics. The Annals of Mathematical Statistics carries ~~original~~ and research ~~articles and~~, in addition, as the official journal of the Institute of Mathematical Statistics, it is planned to publish in this journal from time to time survey articles that should be of interest to industrial statisticians. The Supplement to the Journal of the Royal Statistical Society is of particular interest in that it generally deals with applications.

As a background for anyone who wishes to read the modern literature on mathematical statistics, two comparatively recent publications will be found of great help. These are: The Theory of Statistical Inference, by S. S. Wilks, Mathematics Department of Princeton University, 1937 and Lectures and Conferences on Mathematical Statistics, by J. Neyman, published in 1938 by the Graduate School of the United States Department of Agriculture.

Many ~~other~~ important articles appear each year, scattered in various journals. However, if the journals listed above are available to the industrial statistician, they will enable him to keep pretty well in touch with the men who are making fundamental contributions.

goes down this column and then back to the first figure in the second column and thus on through the table. Suppose we watch him put these numbers in a hat and draw them out again one at a time. As a result of performing one such operation, he gets the ~~corresponding~~ order shown in Table 1(b). With a twinkle in his eye our statistician asks: "Do you as a scientist see any striking difference between the sequence in (a) and the sequence in (b)?" Most of us will probably answer no. Whereupon our statistician might show us that there are many characteristics of the sequence (a) that are unlikely to occur in "random" sequences such as (b). For example, he might show us how to construct a control chart* for averages of successive samples of four, for each of these two sequences. He will tell you <u>before</u> you construct these charts that if you construct such a chart for the data in (b) or similar sequences of these same numbers drawn at random you will seldom get a point outside the dotted control limits. If you carry out his directions for the data in Table 1, you will get the results shown in Fig. 2. The control chart for sequence (a) at the top of Fig. 2 shows points outside the dotted limits; this is something unlikely to occur in sequences drawn at random like (b) shown at the bottom of Fig. 2.

As yet, the statistician is presumably still in his shell. He is free to do only mental operations (Fig. 1). To

- -

*For the method of constructing such a control chart, see the discussion of Criterion I in my book, Economic Control of Quality of Manufactured Product, D. Van Nostrand, Co., 1931. Also see the excellent book by E. S. Pearson — The Application of Statistical Methods to Industrial Standardisation and Quality Control, British Standards Institution, 1935.

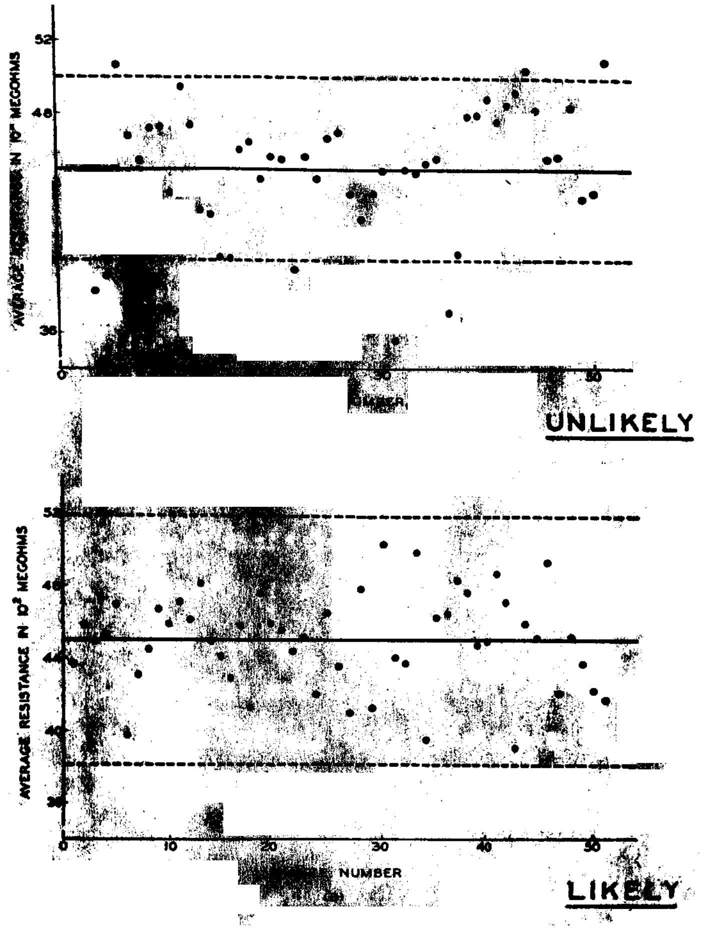


Fig. 2. Something that only a statistician can tell us.

become a scientist, as we have already seen, he must make use of evidence such as he finds in Fig. 2 as a basis for predicting something that he expects to find in the world outside his shell. ~~He~~ But in order to check this prediction, he must then get out of his shell and make contact with the world of observable data at some point B of Fig. 1, that is different from point A ~~that~~ which gave the signals flashed to him on his ground-glass window.

Let us imagine that our mathematical statistician explores the inside of his shell and finds a door alongside his ground-glass window. He suddenly becomes imbued with the idea that he can probably become a useful scientist as well as statistician.

What does our statistician do? In principle, every time he sees a number X_i appear on his screen, he sticks his head out of the door and <u>perceives</u> the physical condition C_i giving rise to the numbers X_i then appearing on his screen. The conditions associated with the X_i's provide him with a new means of arranging the numbers in the original sequence into subgroups as, for example, in a quality control chart.* Then he can determine if the observed differences between subgroups are likely to occur in subgroups selected at random from a random sequence.

However, the mathematical statistician did not stop here. He took another step and began to <u>manipulate</u>** the C's and study the corresponding effects upon the X's. Thus he entered the field of design of experiment and became pretty much a full-fledged experimental scientist.

Schematically, we may sum up the fundamental steps in the emergence of the mathematical statistician as a scientist by the ~~schematic~~ diagram in Fig. 3. Thi~~s~~e new kind of hypothesis [and the associated deductions introduced by the [illegible] fit into steps b and c of scientific method.

I find that the acts of the statistician become much more intelligible to my friends among engineers and natural scientists when these acts are viewed in relation to scientific method. This is particularly true if we go one step further

* It is in this connection that the so called Analysis of Variance Techniques originated by R.A. Fisher are of great importance. In addition to the discussion of these techniques [illegible]

R.A. Fisher, 6th Edition ~~1929~~, Oliver and Boyd.

The industrial statistician will find very ~~much~~ [illegible] helpful the book, The Methods of Statistics, by L.H.C. Tippett. Mr. Tippett will doubtless give us further illustrations in his two [illegible] beyond the [illegible].

** The outstanding text on this subject is R.A. Fisher, The Design of Experiments, Oliver and Boyd, Lon 1935

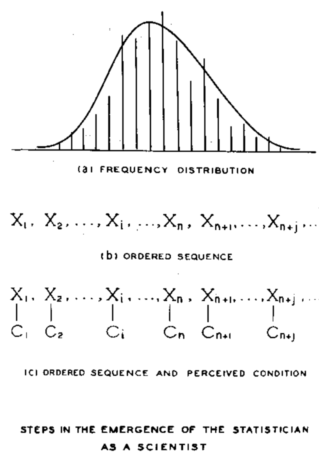


Fig. 3

and point out that no scientist claims to be able to duplicate his results <u>exactly</u>. No matter how much effort is expended, repetitions of any experiment even under presumably the same essential conditions give results that are not identical. Hence if we are to build hypotheses that are supposed to picture the behavior of observable phenomena, these must allow for the observed fact that repetitions of any experiment even under presumably the same essential conditions do not give exactly the same results. It is just this that statistical theory attempts to do for scientific method. When a statistician comes forward with great claims for his novel techniques for the control of quality of manufactured product, or for design of experiment, or the analysis of variance, engineers and scientists (particularly in physics and chemistry) may be inclined to stand aloof. What can an armchair statistician tell them about such matters? They have been controlling quality, they have been designing experiments, and they have been analyzing variances long before the statistician came on the scene. Much of what the statistician is saying is old stuff, at least to successful engineers and scientists. However, such men may fail to see that although many of the concepts employed by the statistician are stock in trade for the good scientist, the statistician does have something very important to contribute, namely more accurate hypotheses and means of testing. We shall return shortly to consider the meaning of such tests from the viewpoint of [illegible] ~~He and he alone can put a cutting edge on the scientists principle tool - scientific method.~~

<u>Why is Mass Production Ideally Suited to the Application of Statistical Method?</u>

We shall consider three characteristics of mass production that are of importance in our present story.

<u>First:</u> Mass production involves large numbers of repetitions of operations and one of the fundamental characteristics of statistical theory is that the deductions from a statistical hypothesis are not in terms of a single unique event but in terms of the characteristics of an <u>infinite</u> sequence of like events corresponding to an infinite number of repetitions of some operation in a state of statistical control.

In mass production, we ~~often~~ usually deal not with a comparatively few repetitions as we do in the field of exact science but with literally thousands upon thousands of repetitions. Hence, if the results of repetitions in mass production can be made to satisfy the conditions of randomness, statistical theory becomes the ideal method of describing the observable sequences of quality.

Second: Successful mass production involves random assembly. Mass production is much more than the process of turning out large numbers of things of a given kind: the key to successful mass production is the manufacture of interchangeable parts. Certainly the requirement of interchangeability of pieceparts is fundamental in our present discussion.*

Let us take as an example a piece of apparatus in which there are u different pieceparts. Let us take the simplest case where each of these, represented by the symbols $O_1, O_2, \ldots, O_j, \ldots, O_u$, is turned out in a continuous process so that we may write down the set of infinite sequences ordered as turned out in the manufacturing process,

- -

*It is interesting to note that, although the human race has supposedly inhabited this world ~~of ours~~ for something over a million years, it was not until about 10,000 years ago that they began to fit pieceparts together, such as in the act of fitting a handle into a stone hammer, and it was not until about 1787 or roughly 150 years ago that interchangeability was introduced!

$$\left.\begin{array}{ccccccc}
O_{11}, & O_{12}, & \ldots, & O_{1i}, & \ldots, & O_{1n}, & \ldots \\
O_{21}, & O_{22}, & \ldots, & O_{2i}, & \ldots, & O_{2n}, & \ldots \\
\cdot & \cdot & \cdots & \cdot & \cdots & \cdot & \cdots \\
O_{j1}, & O_{j2}, & \ldots, & O_{ji}, & \ldots, & O_{jn}, & \ldots \\
\cdot & \cdot & \cdots & \cdot & \cdots & \cdot & \cdots \\
O_{u1}, & O_{u2}, & \ldots, & O_{ui}, & \ldots, & O_{un}, & \ldots
\end{array}\right\} \tag{5}$$

In mass production, the assembly of parts involves the operation of taking one piecepart from each of these u sequences by some rule of choosing a piecepart that does not in any way depend upon the quality of that part. As we have seen, this rule of choosing is what the statistician calls a random one. Hence it will be observed that the assembly of pieceparts is to be carried on under what the statistician calls random sampling. Statistical theory is supposed to tell us what we may expect to get under such conditions.

Since it is not possible to make pieceparts exactly alike, the concept of a go tolerance limit was introduced about 1840. It was soon found, however, that it was more economical to try to manufacture to go, no-go tolerance limits instead of a go limit and hence these were introduced about 1870.

Let us look a little more carefully at this requirement of interchangeability. Let us think of a mechanism composed of a lot of different pieceparts. In interchangeable manufacture, these pieceparts are usually made in large quantities and then a selection of parts required to make a finished machine is, as already noted, selected at random and assembled. Each piecepart of a given kind should be so nearly alike any other of that kind that it would make no appreciable difference in the quality of a finished mechanism if one part were substituted for another. The term fit may suggest mechanical or electrical fits but as here used it implies that one piece of raw or fabricated material should be interchangeable with every other similar piece of the same material. Such a requirement of interchangeability applies not only to pieceparts but also to the parts of a given kind of food, drug and the like. Hence the requirement of interchangeability implies a certain degree of <u>uniformity</u> (or randomness) of raw materials, drugs, foods, oils, and products that are not divided into parts in the process of production.

<u>Third</u>: From the viewpoint of science and from the viewpoint of our story, the most important characteristic of mass production as a process is that it constitutes a scientific method in which we customarily recognize the following three steps:

1. Specification.
2. Manufacture.
3. Inspection.

Broadly speaking, the specification outlines a certain end result that appears feasible of attainment upon the basis of available evidence. In this sense, it represents an operationally verifiable prediction based upon certain hypotheses as to what it is believed possible to attain in the second step, manufacture. ~~This second step~~ from the viewpoint of scientific method, the second and third steps may be looked upon as a means of checking whether or not that which is predicted can be attained.

For example, in 1787 under the sway of the concept of an "exact" physical science, the specifications called for pieceparts made to exact dimensions. It was soon found, however that such parts could not be produced. As a result, it was found necessary to modify the predictions as to what was feasible of attainment. Starting with an assumption about what one can do, then making an experiment to determine whether or not the results predicted upon the basis of this assumption are attained and then modifying the assumption when necessary to take account of new data is inherently a scientific method.

With the introduction of statistical technique in to each of the three steps in the process of mass production, we have a scientific method for making the most economical and the most efficient use of raw materials and finished products.

Testing a Statistical Hypothesis

Let us consider a little more carefully than we have already done the claim that statistical theory provides the scientist with a new kind of hypothesis to be used in scientific method. One of the most important tests of this character is that for significant difference in the mean. In experimental

* It should be kept in mind that the full attainment of quality control within industry starts with raw materials and ends with finished product. Often many companies are involved and in the last analysis much depends upon the close cooperation of these ~~competitors~~ in establishing economic standards of quality involving certain statistical requirements.

As early as 1929 a committee was organized in America under the sponsorship of certain engineering and scientific societies to cooperate with representatives of industry in helping to shape such standards. This committee is now jointly sponsored by the American Society of Mechanical Engineers, the American Society for Testing Materials, the American Mathematical Society, the American Statistical Society, and the Institute of Mathematical Statistics. A similar committee was later formed in Great Britain under the auspices of the British Standards Institution. The book by E.S. Pearson previously referred to constitutes in a way a report of this committee. The American Society for Testing Materials also has a committee on the presentation and interpretation of data.

work, for example, we often have two samples that we wish to test to determine whether or not they differ significantly, as a statistician says, in their mean values. For example, Table 2 gives two sets of four observations of the breaking strength of a certain aluminum alloy made by two producers C and D. The statistician proposes to test whether the means are significantly different or, more specifically, the statistician may set out to test whether or not the two samples may be

Producer C	Producer D
25,800	26,500
25,800	27,800
25,600	25,300
26,400	27,700

Table 2

regarded as belonging to the same population. What is customarily considered as one of the most important contributions to modern statistics provides just such a test for significance. I refer, of course, to the work of Student.

As a result of his work, we may compute from these data the well-known statistic* t. In this particular instance t is equal to 5.12. Then the statistician says something like the following: "*if* sample C and sample D are both *random* samples from the same population, the chance of getting a value of t as large or larger than 5.12 is .016." Since this

---

*For details see, for example, *Statistical Methods for Reearch Workers*, by R. A. Fisher, Oliver and Boyd, 3rd Edition, 1930, pp. 106-7.

probability is greater than .01, it is customary to say that the difference is not significant. If, on the other hand, this probability had been .01 or less, it is customary to say that the difference is "clearly significant."* Neyman has recently described the traditional procedure in testing statistical hypotheses in the following general terms:**

"Having to test some specified (in early stages very vaguely specified) hypothesis H concerning the <u>random</u> variables

$$X_1, X_2, \ldots, X_n,$$

we used to choose some function T of those X_i which, for certain reasons, seemed to be suitable as a test criterion. Pearson's Chi-Square and Student's z are instances of such criteria. The next step, and a difficult one, consisted in deducing an accurate or at least an approximate probability law p(T/H), which the chosen criterion T would follow if the hypothesis H were true. The graphs of the probability laws considered usually represented curves with a single maximum at a certain point of the range, decreasing off toward the ends. This suggested a classification of possible samples into two not very distinctly divided categories, probable and improbable samples. If a sample S led to a value of the criterion T for which the value of p(T/H) is small compared with its maximum, then the sample S would be called improbable or the hypothesis H improbable, and inversely."***

<u>Significance of Such a Test of Statistical Hypothesis</u>

We must now consider the significance of this test from the viewpoint of scientific method. To begin, let us note that the immediate conclusion drawn from such a test is

- -

*This is expression used by Fisher in reference just cited.

**<u>Lectures and Conferences on Mathematical Statistics</u>, Graduate School of the Department of Agriculture, Washington, 1938, page 33. Also see "On the Use and Interpretation of Certain Test Criteria for Purposes of Statistical Inference" by E. S. Pearson and J. Neyman, <u>Biometrika</u>, XXA, pp. 175-240, and pp. 263-294 [illegible]

***Neyman used E to designate a sample. I have taken the liberty of changing it to S in the quotation, since the symbol E is used later in a different sense. Underscoring mine.

that the sample is either likely or unlikely to have arisen through the process of random sampling. Often statisticians go further, as Neyman points out, and conclude that the hypothesis is either probable or improbable, or in the case of applications of tests for significant differences, the statistician may conclude that the observed difference is clearly significant if the probability given by the test is .01 or less.* Two characteristics of such statements should be noted:

a) The immediate result of applying the test is that the statistician draws a conclusion that some event is "likely or unlikely," "significant or not significant" upon the basis of some hypothesis formulated prior to the drawing of the sample.

b) The conclusion rests upon an assumption that the data are random.

We should note, ~~however~~ moreover, that the importance of the conclusion (a) from the viewpoint of scientific method is the effect of the conclusion on our future action. The importance depends upon what one does as a direct result of the test. Very often the statistician stops with drawing the conclusion (a) and in so doing his action is such as he might carry out while still inside his shell. The crucial test of the usefulness of such a test of a statistical hypothesis is the relation he can establish between the conclusion (a) that some result

* Or more accurately, if the probability is less than previously assigned value.

is unlikely (or likely as the case may be) and something that he can validly predict upon the basis of conclusion (a) about data not yet observed that he could not as validly predict without first having applied the test to the initial set of data. Referring to ~~our~~ Fig. 1, the initial set of data corresponds to a contact* with the world of observable phenomena at A. The drawing of the conclusion that some characteristics of these data are or are not likely upon the basis of some hypothesis leaves the statistician still suspended in the field of mental ~~arts~~ acts. What is important from a scientific viewpoint is how he gets down to earth again at B in Fig. 1.

For example Fig. 2 shows us something that the statistician can tell a quality control engineer that only a statistician can tell him. Whether or not the statistician's story in this instance is of any use is another matter and depends on whether or not such results can be successfully used. We shall show in the next section that these results can be used, but to do this, we shall have to follow what the statistician does not only while he is in his shell but also what he does outside his shell and in the process of mass production.

From the viewpoint of interpreting the importance of tests for statistical hypotheses, we need also say something about the second point already noted, namely that the probability computations in such tests rest upon the assumption

*Strictly speaking, of course, the contact made [illegible] while still inside his shell is limited to the number X's and [illegible] that "samples" are random. For example, in applying Students test for significant difference to the two samples of Table 2, the assumption involved in the computation of the probability is that each sample is drawn at random.

Now we should note that the meaning of random from the viewpoint of the statistician is a meaning given to an ~~a~~rt that the statistician can carry out inside his shell. It is a man made operation.

Next we should note that scientific method must start with numbers or pointer readings that Nature gives (at point A of Fig. 1). Of course Man usually operates the experiment and some applied statisticians have assumed that, when a scientist repeats an experiment under presumably the same essential conditions, the resulting data (at A, Fig. 1) will be random. However, Nature has her finger in the pie in drawing any sample of physical measurements in a way that she does not have when the statistician within his shell conceives of a random drawing. Hence, stated in picturesque fashion, we must determine if Nature is in the habit of drawing her samples at random. Statistical method plus mass production has, I think, afforded the first critical and exhaustive investigation to find out if nature draws her samples at random.*

- -

*See for example, ~~Statistical Method from the Viewpoint of Quality Control, by W. A. Shewhart, Graduate School of the United States Department of Agriculture, Nov. 1938.~~
Shewhart, loc. cit., p. 14

My own experience, with possible exceptions that I could count on the fingers of one hand, indicates that Nature does not act in this way. The data that Nature hands the statistician at A in Fig. 1 are almost never random. ~~The fact that [illegible]~~ ~~This result is a result of the new tool of research of which I speak.~~

Let us pause for a moment to consider the importance of this result from the viewpoint of applied statistics. The experimentally established fact that Nature seldom gives us a random sample of her own accord (at least in the fields of measurement in physics, chemistry, and engineering) means that we must take all conclusions about the likelihood of observing a given difference upon the assumption that the samples are random with a grain of salt until we have first secured the condition of randomness.

Hence it is of great importance to discover if one can control his operations so as to attain a condition where repetitive operations appear to give a random sample? Can we, in other words, set up some kind of an operational procedure for detecting and eliminating some of the little demons of chance until we reach a stage where those that remain produce fluctuations that have the properties demanded of a random sample. Again we must appeal to experience for an answer. Needless to say, this cannot be done by taking a *few* data. Experience in the field of mass production shows that it cannot be done, at least there, without extensive repetition of an operation. On the other hand, it shows that through the introduction of a statistical *operation of control* soon to be considered in more detail, it is possible to attain a state of statistical control where fluctuations occur at random.

Finally we should note that in testing an hypothesis in the customary statistical sense thinking is explicitly stated in the test about the quantity of information that one should have before one can place much reliance in the results of the test. That is to say, there are what we might call different degrees of thoroughness in a test. Put in another way the degree of belief ~~to~~ that we are justified in placing in the results of a test depends upon the quantity of the data used in the test. This is a very important point from the standpoint of practical work, as I think shown in some of the illustrations.*

PROPOSITION 2:

WITH THIS NEW TOOL OF RESEARCH, SOME USEFUL THINGS CAN BE DONE THAT CANNOT BE DONE OTHERWISE

Two Objectives to be Attained

We shall consider here the following two objectives that can be attained only through the use of this new tool of research or statistical method + mass production:

a) Make possible the most efficient use of raw and fabricated materials.

b) Make possible the highest degree of quality assurance.

There is some danger that each of these objectives will be confused with the economic ones to be discussed in the next section. Here we shall be concerned with these two objectives without reference to their economic significance.

Let us first look at the matter of efficient use of material. Let us assume that we wish to build a structure for which one of the limiting factors is the quality characteristic X, such as tensile strength. Accordingly, we look about for an otherwise suitable material that has the maximum tensile strength. When found, we shall also likely find that pieces of this material show widely different values of this property as indicated schematically in Fig. 4. The lower

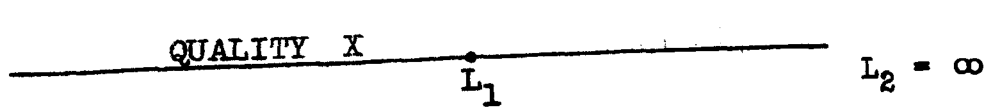


Fig. 4

Experience shows that

tolerance limit L_1 becomes the limiting factor. If we can do anything to reduce the range of variability of this material (here suggested by the frequency histogram), we automatically raise the tolerance limit L_1 and hence raise the limiting factor from the viewpoint of design. In this way, we can effectively increase the strength of the strongest material.

Now, let us look at this matter of quality assurance. The successfulness of design depends upon our being able to make <u>valid</u> predictions of certain quality characteristics within limits. It is particularly desirable to be able to do this when the test for quality is destructive, as in the case of tests for tensile strength, chemical composition, blowing time of fuses, and the like and when such properties are important from the viewpoint of safety. The engineer must set up for himself a tolerance range L_1, L_2 and

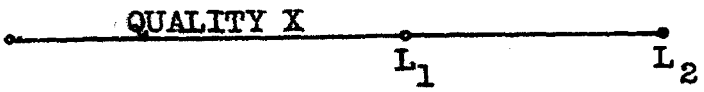


Fig. 5

stay within these limits. Let us make sure that we see what this means from the viewpoint of scientific method as illustrated schematically in Fig. 7. What the engineer must do is to start with data at A and upon the basis of analysis of these data, set up tolerance limits that will include a certain prescribed fraction (usually* .99+) of all future

* This [illegible] upon a study of [illegible] a wide field [illegible]

product. In order to maximize the assurance (based upon a given number of operations at A) that the tolerance will include the prescribed percentage,* it is necessary first to attain a state of statistical control at A, and in order to attain adequate assurance, after a state of control is reached, the engineer usually must go to samples of at least 1000 as we shall soon see. We must not, of course, confuse this problem with that of setting up what the statistician calls fiducial limits.** ~~He~~ The engineer must, as it were, set up for himself a straight and narrow path and he must be sure that he can live within it. This can only be done under random condition, and only with large samp.

__Objectives Must be Attained Under Practical Conditions__

As a background, let us recall the contrast between conditions of laboratory research and those of the process of mass production. In general, laboratory research is conducted under conditions that admittedly may be and often are different from those of mass production. In the laboratory, only a comparatively few but highly specialized people are usually engaged in carrying on a specific research problem and such work is done under conditions wherein many possible causes of variability in repetitive results have been purposely eliminated. In the process of mass production, on the other

- -

*Of course, where the data at A are not statistically controlled the engineer usually allows a much larger tolerance range than would be indicated as necessary by the observed data upon the assumption that they had been in a state of statistical control.

**The difference between these two problems is illustrated by Figs. 14 and 15 of my Washington Lectures, loc. cit., p. 14.

hand, many different men from different parts of an organization are involved. The actual production of product may be carried out in manufacturing plants located in different cities. There may be many different kinds of machines and many different machines of a kind used in the process and the raw material may be selected from many sources. In the sense that the commercial conditions of production involve many more potential sources of variability in the results of a repetitive act than are usually involved in laboratory research, the control of the process of mass production is~~, in this sense,~~ a more complicated problem than the control of an experimental procedure in the research laboratory. Yet the engineer is called upon to make, under these commercial conditions, things having quality characteristics that will meet previously specified tolerance limits. In the absence of accurate knowledge, he may be forced to set wide tolerance limits to begin with. In setting such limits, he must allow for the effects of all manner of unknown or chance causes of variability that may enter at any one or more of the stages in the production process from raw material to finished product. It is, of course, the object of the operation of control to weed out the assignable causes in the production process so that the tolerance limits may be reduced.

The Statistician Comes to the Aid of the Engineer

It was at this stage of the game of control that the engineer called for the help of the mathematical statistician.

But in order to help the engineer as a scientist, the statistician had to act as scientists act: he had to make some assumption or hypothesis about the physical situation as he saw it and then he had to try out these assumptions.

In much the same way that the scientist of old conceived of an exact science, the knowledge of which would enable him to make valid predictions with exactness, the statistical scientist conceived of an ideal attainable statistical state of control, the knowledge of which would enable him to make valid predictions within tolerance limits. He, therefore, argued with himself somewhat as follows:

If some of these unknown or chance sources of variability, or assignable causes as we shall call them, could be found and eliminated, the tolerance limits could be reduced and~~, if~~ all of the unknown or chance causes of variability that can be removed, are removed, then the resultant variability may be in a state where the fluctuations occur at random. Under these conditions, it should be possible to attain the greatest possible degree of assurance that the quality of product will lie within tolerance limits. He realized that in order to detect the presence of and eliminate the assignable causes of variability in the process of production, he would have to study the actual process of production under commercial instead of laboratory conditions.

More specifically, he conceived of a state of statistical control in the following two ways:

1) A set of unknown or chance causes of variation in the results of repeating the same operation in which no one of the causes can be found and eliminated,

2) A state wherein the observed variation in a quality characteristic takes place at random.

It is obvious that a priori a physical state might exist satisfying one of these two characterizations and not the other. What ~~Hence,~~ the control statistician did was to make ~~made~~ the assumption that a state satisfying one of these requirements would satisfy the other, and then he set about to justify this assumption by empirical methods. Let us then consider next the "operation of statistical control" whereby the control engineer has within recent years produced abundant evidence to justify this assumption.

Milestones on the Road to Control

I shall ~~Now~~ let us try to get a general picture of the operation of control. Let us think of a production process turning out a succession of pieces of a given kind of product such, for example, as a condenser, induction coil, or fuse. Let us focus our attention on some one quality characteristic X for each piece and let us assume that we can arrange the observed values of X in the same order as the pieces are produced. Since it is assumed that the operation of producing a piece of product may be repeated an indefinitely large number of times, we may represent the output of product in respect to

(The operation of control, as we shall see, is more than making and testing an hypothesis in the statistical sense; it is a method of reaching to and acting upon an ever changing environment.

the quality characteristic X by the infinite sequence

$$X_1, X_2, \ldots, X_i, \ldots, X_n, X_{n+1}, \ldots, X_{n+j}, \ldots \qquad (6)$$

<u>where the order is supposed to be that in which the pieces of product were made.</u>

Now in the repetition of the operation of production, it is customary to distinguish between the conditions under which pieces are produced. For example, one piece may be produced on one machine and another piece on another machine; one piece may be produced in one plant and another piece may be produced in a plant in some distant city. We may not have any information to indicate that the quality of an object produced under one condition is any different from that of an object produced under some other condition. However, the different conditions are potential sources of assignable differences. Let us, therefore, attach to every value of X in (6) a symbol, C, to stand for the <u>perceived</u> condition. These C's are, of course, what we have said that the mathematical statistician sees when he sticks his head out of the door of his shell. Thus we have

$$\begin{array}{ccccccc} X_1, & X_2, \ldots, & X_i, \ldots, & X_n, & X_{n+1}, \ldots, & X_{n+j}, \ldots. & \qquad (7) \\ C_1 & C_2 & C_i & C_n & C_{n+1} & C_{n+j} & \end{array}$$

Expression (7) is of outstanding importance from the viewpoint of understanding the contribution of statistical theory to quality control. In the first place, let us note that the

<u>order</u> of the X's in (7) is supposed to be that given by the production process in contrast to the indefinitely large number of "random orders" that might be established by repetitions of the operation of putting the numbers on chips, putting the chips in a hat (figuratively speaking) and drawing them one at a time in the way a statistician could do without ever getting out of his shell. In the second place, some of the C's may be perceived as being essentially the same. Hence, it may be possible to order the sequence in many different ways in terms of the C's. For example, the data in Table 2 may be thought of as a sequence of 10 observations, 5 under each of two conditions. When the C's are all perceived as being essentially the same, we still have the special order in which the pieces were made. Out of this situation, we distinguish two general ways of ordering an infinite sequence: a) by some artificial randomizing rule of procedure such as a statistician within his shell can conceive of, and b) by means of the perceived C's, including the serial order of production as a special case, such as [illegible] must crawl out of his shell in order to perceive.

Of course, any particular order of an infinite sequence established by an artificial randomizing rule is no more likely to occur than any other order. Thus the order in the original sequence (6) produced by the manufacturing process is just as likely as any other particular order that we might get by rearranging at random the terms in this sequence. Likewise, any order established by means of the C's is just as likely to occur as any other particular order established by a random process of rearranging the terms in the infinite sequence. If, however, we subdivide the infinite sequence into samples of let us say the same size for the sake of simplicity, the frequency of occurrence of a θ_i within a range θ_{i1} to θ_{i2} is presumably the same for all of the random arrangements made in terms of the C's. Statistical method, as we have already stated, gives us the means, in general, of computing the distribution function (3) for a given θ_i corresponding to a random selection and thus of setting up limits in which the observed value θ_i is likely to occur.

The job cut out for the statistician acting as a scientist in an effort to attain maximum control of quality, is to begin with, two-fold. He must set up the following two consecutive operations:

1. The operation of attaining a physical state of statistical control.

2. The operation of attaining the distribution function, $f(X, \lambda_1, \lambda_2, \ldots, \lambda_s)$ corresponding to this state of control.

We must now try to see how the efforts of the statistician, while still within his shell, contribute~~s~~ to these operations.

Looking again at Fig. 3, we see that the statistician within his shell can do at least two things: first he can construct either analytical or graphical pictures of the distribution of numbers flashed on his screen, and second he can tell whether or not certain characteristics of a sequence of numbers flashed on his screen are~~, or are not~~ likely to be given by a random process of arranging the numbers. These are his two sole implements of warfare on the little demons of chance at the time he emerges from his shell. He has made use of both of these implements in the past.

For example, the statistician acting as a scientist early tried to find some relation between the functional form of the distribution curve for the data flashed on his screen and the experimental conditions under which the observed data were obtained outside his shell. In particular, a many-peaked distribution was assumed to arise when experimental conditions were not controlled as they might be. This idea has, in fact, been extensively employed, at least since 1924, in the steel industry of Germany.*

However, the statistician's most successful implement of warfare is his means of determining whether or not the characteristics of a sequence arranged by means of the C's is a characteristic likely to occur in a random arrangement of these same numbers. Even before 1908, Student** in England was using in the brewing industry such a tool

- -

*See, for example, the writings of Karl Daeves, ~~and~~ in particular, his recent book, Praktische Grosszahlforschung, VDI-Verlag, GMBH, Berlin, 1932.

**"The Probable Error of the Mean", Biometrika, Vol. VI, 1908, pp. 1-25.

developed by him for testing the significance of observed differences between samples.

As previously noted, however, the usefulness of any such test depends upon getting a practical rule of predicting, upon the basis of the conclusion that an event is or is not likely, some other observable event such as schematically illustrated at the point B of Fig. 1. In quality control work, of course, ~~the~~ one such prediction is to the effect that a demon of chance can or cannot be found upon the basis of whether or not one or more points fall outside or inside the control limits of a quality control chart (See Fig. 2 for example). ... fundamental prediction is that by continuing ...

There are several reasons for choosing the particular procedure known as the quality control chart method, at least under the conditions of mass production. We cannot here go into these reasons nor can we cite the several different ways in which statistical tests are used in setting up what I have termed the operation of control.*

The essential thing to note is that the operation of control is essentially a rule for acting in order to attain ... . Considered in broad terms, any such rule starts with an experience A of Fig. 1. We go through some mental gyration or process of reasoning and, in effect, make a prediction - in this case that a cause can or cannot be found, depending upon the result of some computation that in itself has no necessary connection with the initial experience or the experience predicted. The next stage in this process is to see if the prediction is valid.

...[illegible] limits become action limits located, in general, within the tolerance limits $L_1 L_2$. For a more detailed discussion of these and other facts about the operation of control, see Shewhart, loc. cit. p. 12. and "Some Application of Statistical Methods to Manufacturing Problems", Journ. of the Franklin Institute Vol. 226, No. 2, August, 1938, pp. 163-186.

* For example, the operation of control involves, among other things, [illegible] elimination of an assignable cause [illegible] [illegible] that a point [illegible] outside the control chart was taken as evidence of the existence of an assignable cause. In order to [illegible] the cause, it may be necessary or at least advantageous [illegible] [illegible] for the analysis of variance.

Furthermore, in the process of finding the causes, one may and usually does employ some of the principles of design of experiment. I like to think of the control chart as a practical means of indicating when it is desirable to look for trouble whereas the [illegible] of variance is what ~~[illegible]~~ one might [illegible] [illegible] method of a "[illegible]" to use in looking for the causes.

The [illegible] control in mass production, however, [illegible] [illegible] [illegible] with the use of the control chart simply in indicating an assignable cause. Instead it involves a series of steps or operations by which one attains a state of statistical control through [illegible] specification, production, and [illegible]. For example [illegible] this [illegible] the

However, in order to justify the theory of control, we must show that we can follow through such a process again and again continuing to eliminate causes of variation and thereby reducing the resultant variability until we have attained a state where the points of a control chart remain within the boundaries in the same way that they do when drawings are made from an experimental bowl universe. That such a state of control can be attained under conditions of mass production, is illustrated by Fig. 6 which shows a control chart for the blowing time of fuses.

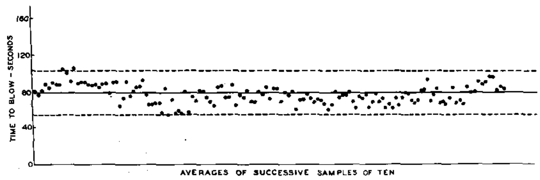


Fig. 6

In general, experience shows that such a state of statistical control is attained only at the end of a comparatively long succession of repetitive observations such as one may obtain in mass production. Furthermore, experience indicates that one is not justified in believing that such a state of control exists until he has attained a condition in which the scientist or engineer himself not only believes the conditions under which the repetitive process is carried out to be essentially the same but also that at least twenty-five samples of four satisfy what I have already referred to as Criterion I of control. This is an important criterion on the quantity of data.

Thus far we have considered simply the problem of attaining what we term a physical state of statistical control. But this is not the end. Let us suppose that our experience, instead of indicating as it usually does, that we have to eliminate assignable causes of variability before reaching this state, had indicated that in almost every instance, sequences of observations observed under presumably the same conditions satisfy the criteria of control. It should be remembered that even under this condition it would be necessary in a manufacturing process to establish tolerance limits that would include roughly 99.7% or more of the product. Furthermore, it is necessary for economic reasons that these limits be set with a comparatively high degree of accuracy.

Suppose we take a very simple case in which we know that the quality of a given product is controlled and what is more, we know that it is controlled in accord with the normal law. How large a sample would it require before we could set 99.7% tolerance limits with a ~~reasonable~~ degree of accuracy usually demanded in engineering work. Fig. 7 shows one hundred such tolerance ranges derived from one hundred samples of four; forty ranges of samples of one hundred; and four ranges of samples of one thousand. The dotted limits show the ~~true~~ 99.7% limits.

This point has already been emphasized on page 33.

This chart shows* ~~It~~ would be necessary in order to make the most efficient use of materials to attain something like the degree of accuracy corresponding to a sample of one thousand even under these ideal conditions.

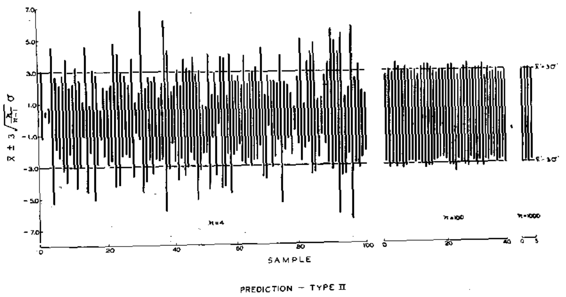
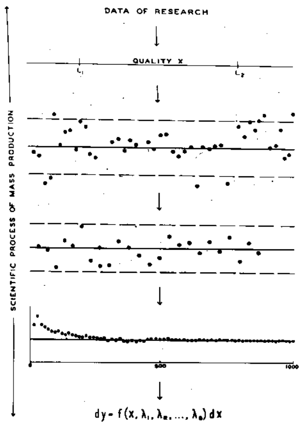


Fig. 7

Enough has been said, I think, to justify the claim that in order to make the most efficient use of materials in the sense here defined by minimizing the tolerance, and in order to obtain the state in which predictions can be made with the greatest degree of validity, it is necessary to use the technique of statistical control in mass production.

Fig. 8 shows in summary fashion the milestones in the ~~step~~ road to a state of control. We start with the data of research and set tolerance limits L_1, L_2. Then applying the operation of

That is to say, the fraction allowed outside tolerance limits [illegible] sample, [illegible] less than one [illegible], and hence [illegible] error [illegible] within the tolerance [illegible] control, we look for assignable causes whenever we find points outside of control limits. So soon as we have gotten a run of at least 25 samples of four that stay within the limits, we assume that we have attained a state of statistical control.



MILESTONES ON THE ROAD TO CONTROL

Fig. 8

The next step is to find the parameters in the frequency distribution representing the state of control by means of the statistical limit. To secure adequate accuracy, it is usually necessary to take at least 1000 observations. We then are in a position to set up the distribution function λ_s

$$dy = f\ (X,\ \lambda_1,\ \lambda_2, \ldots, \lambda_5)\ dX.$$

<u>In attaining the desired goal, it is not simply the characteristics of statistical method but it is the combination of these characteristics with those of mass production that makes the attainment of these objectives a possibility.</u>

PROPOSITION 3:

USE OF STATISTICAL METHOD IN MASS PRODUCTION PROVIDES SCIENTIFIC METHOD OF MINIMIZING THE COST OF PRODUCING THINGS TO SATISFY HUMAN WANTS

What Economies Can be Made?*

Three economies can be made. They are:

1. Reduction in the cost of inspection.
2. Reduction in the cost of rejection.
3. Saving through efficient use of material.

By weeding assignable causes of variability out of the production process, an engineer can effectively reduce the cost of inspection at every stage of the process from the inspection of raw material to finished product. By attaining a state of statistical control, the engineer can minimize his cost of inspection.

In respect to reduction in the cost of rejection, suppose we let the results of the first large scale experiment to determine the effect of weeding out assignable causes revealed by the control chart, speak for itself. About thirty typical items used in the telephone plant and produced in lots running into the millions per year were made the basis of this study. Fig. 9 shows that during 1923-24 these items showed

- -

*The economic aspects of quality control have been extensively discussed in the literature. Here I shall simply present a graphic summary of what can be done.

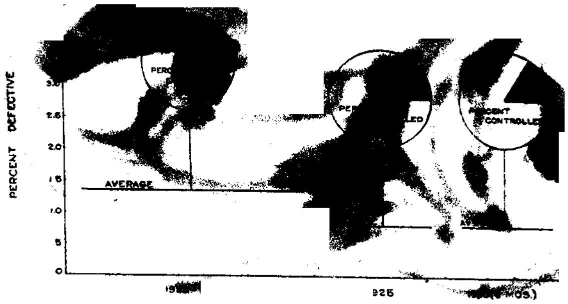


Fig. 9

about 68 percent control about ~~a~~ the relatively low average of 1.4 percent defective.* However, as the assignable causes - the little demons of chance - were detected by points falling outside the control charts for fraction defective and then found and removed, the quality of product approached the state of control as indicated by an increase of from 68 to 84 percent control by the latter part of 1926. At the same time, the percent defective dropped from 1.4 in 1923-24 to 0.8 in 1926. If we let such typical experience speak for itself as to what happens to rejections as we approach a state of statistical control, we get this answer: "Rejections drop,

- -

*R. L. Jones, "Quality of Telephone Materials", Bell Telephone Quarterly, June 1927.

drop, drop, with successive steps toward the state of control and approach asymptotically some minimum value."

Now in what way can we save through efficient use of material? In general, by making it possible to stay with narrower tolerances, we can often reduce the amount of material required in a given design. This is particularly important in complicated designs where the overall tolerance is a function of the tolerances on many parts. Another field where an improvement in the homogeneity of quality may make possible appreciable savings is that of protective coatings of one kind or another. In respect to the economic importance of attaining homogeneity in the production of steel, John Johnston, Director of Research of the U. S. Steel Corporation has recently said: "The possibility of improving the economy of steel to the consumer is therefore largely a matter of improving its uniformity of quality, of fitting steels better for each of the multifarious uses, rather than of any direct lessening of its cost of production."*

How Can Economies be Made?

Let us look once more at the three major steps in mass production, namely specification, production (or manufacture) and inspection. The use of statistics in a given industry often has the habit of coming in at the back door,

- -

*"The Applications of Science to the Making and Finishing of Steel", Mechanical Engineering, Feb. 1935.

inspection. Very often inspection is looked upon as pretty much a routine for screening defects from the product. Since this screening must often be done on a sampling basis, there arises the question: How large a sample shall be taken? Almost everyone knows that a statistician is supposed to know something about sampling and hence the first contact a statistician often gets with the problems of mass production is in answer to a call for assistance in setting up a sampling plan. After the demons of chance have been given an opportunity to do their dirty work, the statistician is called in to help cull out the defects! This is much as it was in the field of medicine before the days of prevention. We let the mosquitos bite the millions and then called in the doctor to cure the sick.

In much the same way that the doctors who simply cured the sick earned much more than their keep, so too the statistician ~~who~~ simply draws up sampling plans to help relieve the economic pains in a production alive with the little demons of chance much more than earns his keep. But why keep living with sickness if it can be prevented? Why keep living with those demons of chance that we can get rid of?

The application of statistical method in the sense of the present discussion is preven~~ta~~tive, not curative. It is not simply a technique for handling the ~~1001~~ little sampling problems that may occur here and there in every industry, important though that might be; it is not simply a technique of interest to one department such as that of inspection in a large industry; it is not one that simply affects the producer. Instead, an understanding of the potential possibilities of the combination - statistical method + mass production - calls for a complete change of view of the mass production process from that view on which the process was founded some hundred and fifty years ago.

Then we were living under the sway of the concept of an "exact" science. We conceived of the three steps, specification, production, and inspection as independent, Fig. 10. We conceived of a knowledge of scientific facts and principles adequate for setting economic tolerance limits and looked upon the process of mass production pretty much as an <u>application</u> of scientific knowledge. There is little evidence that we thought of the mass production process as <u>a scientific method of acquiring more knowledge</u>, let alone as being the only method for acquiring the knowledge necessary for minimizing tolerances and maximizing quality assurance as well as for effecting the three economies just considered.

Today, I have tried to picture the three steps in the mass production process as steps in a scientific method of continually acquiring new knowledge to be fed back into the process until we know enough to stay within a straight and narrow. At the start of production of a new design,

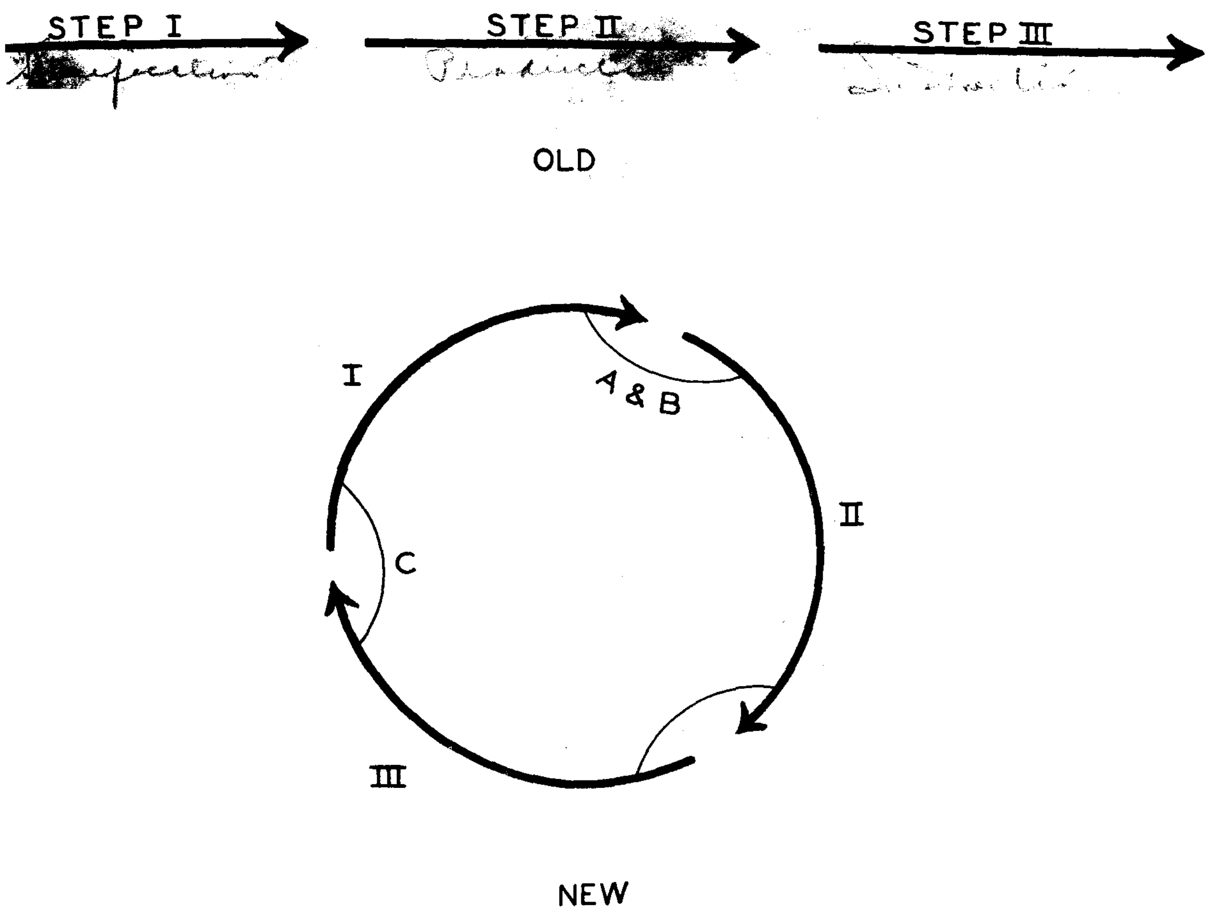


Fig. 10

comes the step of specification of tolerances - an act that involves the use of available data and the use of hypotheses as a basis for predicting that these tolerances can be met. Then comes the acts of production and inspection to test the predictions. Information attained in steps II and III must be fed back into step ~~one~~ I and the circle completed again and again before we pass all the milestones on the road to control, Fig. 8. Statistical concepts and techniques must form a background for viewing this whole process.

Viewed in this way the three steps are not independent; ~~instead~~ they are correlated. They follow one another in a circle much after the manner of Fig. 7. We have seen that, in order to take Step I economically, we need action limits A and B of the control chart. The basis for setting these limits must come from Steps II and III. ~~Also~~ We also need the aimed-at value C to be used in design formulae, but this must come from Step III. ~~However,~~ We cannot take the third step, however, and attain the expected value C unless we first attain a state of statistical control, Step II, such that C is reached as a statistical limit. But this state of control can be attained only as a statistical limit in the continuing process of mass production.

Hence we may conclude that statistical method + mass production makes possible the most efficient use of raw materials and manufacturing processes, effects economies in production, and makes possible the highest economic standards of quality for the manufactured goods used by all of us.