# APPLICATION OF STATISTICAL METHODS TO MANUFACTURING PROBLEMS.*

BY

W. A. SHEWHART, Ph.D.,

Bell Telephone Laboratories.

ABSTRACT.

The application of statistical methods in 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. The story of the application, however, is of much broader interest. The economic control of quality of manufactured goods is perhaps the simplest type of scientific control. Recent studies in this field throw light on such broad questions as: How far can Man go in controlling his physical environment? How does this depend upon the human factor of intelligence and how upon the element of chance?

## INTRODUCTION.

I sincerely appreciate the honor of the invitation to present to you this evening some of the recent developments in the application of statistical techniques in mass production.

What is there about the application of statistical methods in manufacturing of common interest to all of you? I assume that most of you are not statisticians looking for new worlds to conquer and therefore interested in the problems of production. Likewise I assume that many of you, at least, are not manufacturers looking for new tools with which to solve your problems. However, all of us are affected in one way or another by developments in the field of applied science. In fact the application of science is for the very purpose of satisfying the human wants of each and every one of us. Without such wants, there would be no applied science. Therefore each of us has an interest in industry's improvement of the technique of giving us what we want.

Furthermore the new statistical techniques developed and used by industry in the control of quality are of interest in

* Presented at a meeting held Thursday, February 28, 1937.

themselves as constituting extensions of scientific methodology that have now been tested and not found wanting—methods which make possible the most efficient use of raw materials, the reduction of the cost of production, and the maintenance of economic standards of quality.

The story of the application of statistical methods in manufacturing is, however, of much broader interest. All of us are interested directly or indirectly in the human control of some one or more of the physical aspects of the world in which we live. Now, the control of quality of manufactured product is perhaps the simplest type of control problem. Hence studies in this field may be expected to throw light on such broad questions as: How far can man go in controlling his physical environment? How does this depend upon the human factor of intelligence and how upon the element of chance?

### The Problem of Quality Control.

Without more ado, let us sketch in broad outline the problem of making things to satisfy human wants. This problem is twofold: (*a*) discovery of what is wanted and of physical facts and principles, and (*b*) efficient use of the results of discovery. Engaged in determining what is wanted we have literally hundreds of consumer research agencies, trade associations, and the like, and engaged in determining ways and means of making use of physical facts and laws in producing things to satisfy such wants, we have more than 1600 industrial research laboratories in America alone. To sketch some of the potential advantages of the application of the recently developed statistical methods for research workers in all such work would constitute a story in itself. For lack of time we must pass it by tonight.

However, given the results of research, there remains the problem of using such results—the problem of development and design, then production, and last but not least inspection. Tonight we shall limit our consideration to mass production and consider primarily the three steps: specification, production, and inspection.

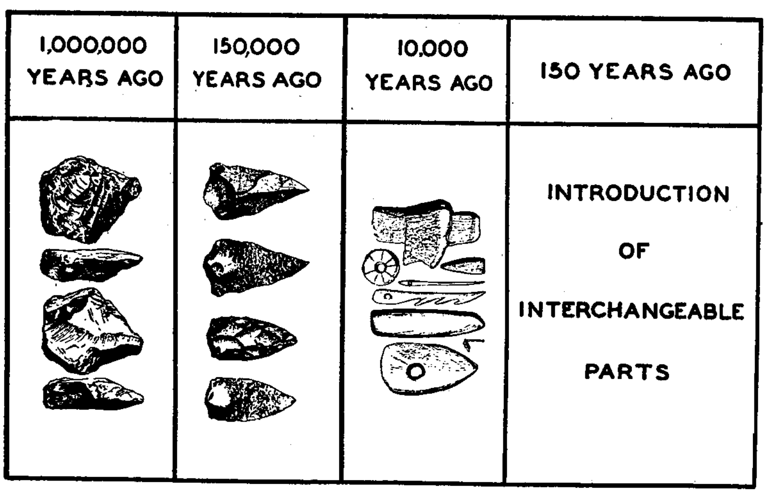
Now let us see what characterizes these three steps. What for example is a specification? In general, it is a technical description in words and engineering drawings of the thing that is wanted. The act of specification leads to a symbolic representation of a physical thing to be produced—a thing that exists as an idea in the mind of the engineer. Production on the other hand is an act leading to physical things, not symbols. Inspection as usually considered is the act of comparing the things produced with that symbolized by the specification.

We shall see how within the past twelve years fundamental changes have been brought about in the techniques underlying these three steps. We shall see that these changes are based upon fundamental changes in ideas as to the nature of the inspection process and as to what constitutes an adequate specification, what constitutes a standard of quality, and how much variation in the quality of things produced should from an economic viewpoint be left to chance. These changes in ideology, after all, are the most important elements in our story because ideas, concepts, and beliefs rule our actions.

### SOME IMPORTANT HISTORICAL STAGES IN CONTROL OF QUALITY.

To give us a perspective from which to view recent developments, let us look at Fig. 1. That which to a large extent

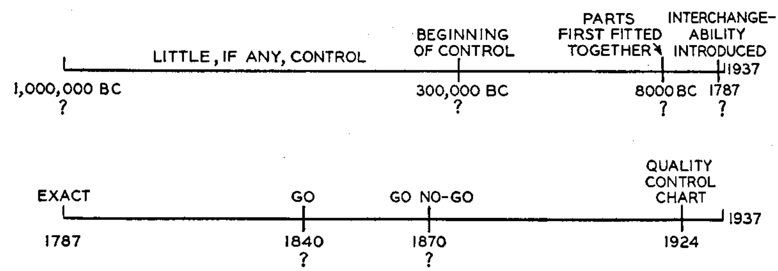
FIG. 1.




differentiates man from animals is his control of his surroundings and particularly his production and use of tools. Apparently the human race began the fashioning and use of stone implements about a million years ago as is evidenced by the crude stone implements shown to the left of Fig. 1 which were recently (1921) discovered[1] just north of London. Little progress in control was made, however, until 10,000 years ago or thereabouts when man first began to fit parts together as evidenced by the holes in the instruments of that day illustrated in Fig. 1.

Throughout this long period apparently each man made his own tools, such as they were. As far back as 5000 years ago the Egyptians are supposed to have made and used interchangeable bows and arrows to a limited extent. It was not, however, until about 1787 or a hundred and fifty years ago

FIG. 2.




that we had the first real introduction of the concept of interchangeable parts. Only yesterday, as it were, did man first begin to study the technique of mass production!

From the viewpoint of ideology it is significant that at the time this first step was taken science was thought to be exact. Accordingly an attempt was made to produce pieceparts to exact dimensions. How strange such a procedure appears to us today, accustomed as we are to the concept of tolerances. But as shown schematically in Fig. 2, it was not until about

[1] This discovery is reported in "Man Rises to Parnassus" by Henry Fairfield Osborne, Princeton University Press, 1928. The photograph of the stone implements of a million years ago has been reproduced by permission from this most interesting book. The implements of 150,000 and 10,000 years ago as shown in Fig. 1 have been reproduced by permission from the fascinating story told in "Early Steps in Human Progress" by Harold J. Peake, J. B. Lippincott and Company, 1933.

1840 that the concept of a "go" tolerance was introduced and not until about 1870 that we find the "go no-go" tolerance.[2]

Why these three steps: exact, go, go no-go? The answer is quite simple. Manufacturers soon found that they could not make things exactly alike in respect to a given quality, it was not necessary that they be exactly alike, and it was too costly to try to make them exactly alike. Hence by about 1840 they had eased away from the requirement of exactness to the go tolerance. Still too much time was wasted unnecessarily in trying to stay reasonably close to the tolerance. Then came the idea of specifying the go no-go tolerance or the range within which the quality characteristic might vary and still be satisfactory. This was a big forward step because it gave the production man more freedom and brought a still greater reduction in cost. All he had to do was to stay within the tolerance range—he didn't have to waste time trying to be unnecessarily exact.

Though this step was of great importance something else remained to be done. The way the limits are necessarily set is such that every now and then pieces of product are produced with a quality characteristic falling outside the specified range—in other words, defective. To junk or modify such pieces adds to the cost of production. But to find the unknown or chance causes of defectives and try to remove them also costs money. Hence after the introduction of the go no-go tolerance there remained the problem of trying to reduce the fraction $p$ of defectives to a point where the rate of increase in cost of control equals the rate of increase in the savings brought about through the decrease in the number of rejects.

For example, in the production of the apparatus going into the telephone plant, raw materials are gathered literally from the four corners of the earth. More than 110,000 different kinds of pieceparts are produced. At the various stages of production inspections are instituted to catch defective parts before they reach the place of final assembly and are thrown

---

[2] It will be noted that the first six dates shown in Fig. 2 are given with a question mark—authorities are not in unanimous agreement as to the exact dates. I think, however, that the dates here shown will be admitted by all to be approximately correct.

out Here one faces the problem of finding the economic minima for the sizes of the piles of defects thus thrown out.

This problem of minimizing the per cent defective, however, was not the only one that remained to be solved. Tests for many quality characteristics—strength, chemical composition, blowing time of fuse, and so on—are destructive. Hence every piece of product cannot be tested for such a characteristic to see if it falls within the specified tolerances. Engineers must appeal to the use of a sample. But how large a sample should be taken in a given case in order to give adequate quality assurance?

The attempt to solve these two problems, giving rise to the introduction of the quality control chart technique in 1924, may therefore be taken as the starting point of the contributions of statistical technique to the control of quality of manufactured product in the sense here considered.

#### Why After 1900?

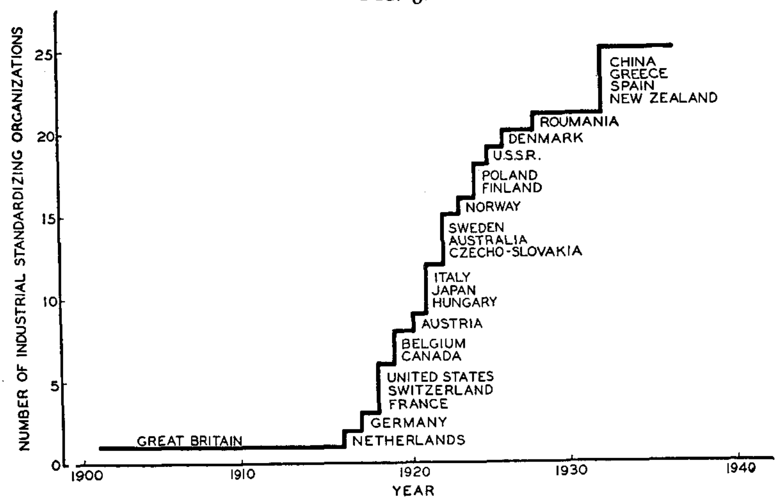
Why, you may ask, was it something like one hundred fifty years from the start of mass production of interchangeable parts to the time of the more or less intensive study of the application of statistical methods in this field? There are at least two important reasons.

First, there was the rapid growth in standardization. Fig. 3 shows the rate of growth in the number of industrial standardization organizations both here and abroad. The first one was organized in Great Britain in 1901. Then beginning in 1917 we get a rapid spread of the realization of the importance of national and even international standards. Now, fundamentally the output of such standardization organizations is specifications of the aimed-at quality and in certain instances of methods of measuring this quality. But when one comes to write such a specification, he runs into the two problems discussed in the previous section—minimizing the number of rejections and minimizing the cost of inspection to give an adequate degree of quality assurance. Hence the growth in standardization spread the realization of the importance of such problems in industry.

Second, there was a more or less radical change in ideology. We passed from the concept of the exactness of science in 1787, when interchangeability was introduced, to the concepts of probability and statistics which came into their own in almost every field of science after 1900. Whereas the concept of mass production of 1787 was born of an exact science, the concept underlying the quality control chart technique of 1924 was born of a probable science.

We may for simplicity think of a manufacturer trying to produce a piece of product with a quality characteristic falling within a given tolerance range as being analogous to shooting at a mark. Now, if one of us were shooting at a mark and

FIG. 3.



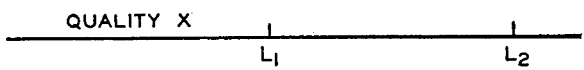

failed to hit the bull's-eye, and some one asked us why, we would likely give as our alibi, CHANCE. Had some one asked the same question of our earliest known ancestors, they might have attributed their lack of success to the dictates of fate or to the will of the gods. I am inclined to think that in many ways one of these alibis is just about as good as another. Perhaps we are not much wiser in blaming our failures on chance than our ancestors were in blaming theirs on fate or the gods. One element of human interest in our story tonight is that our William Tell, the manufacturer, has proved his unwillingness since 1900 to attribute all such failures to chance.

This represents a remarkable change in ideology which characterizes the developments in the application of statistics in the control of quality.

### APPLICATION TO SPECIFICATION.

Step I in quality control, as we have defined it, is specification. Let us now consider the application of statistical methods to this step. With the introduction of the go no-go tolerances of 1870, it became more or less generally accepted practice to specify for any given quality characteristic $X$ that this quality should lie within specified limits $L_1$ and $L_2$, represented schematically in Fig. 4. Such a specification is of the nature of an end requirement on the specified quality characteristic $X$ of a finished piece of product. It provides a basis on which the quality of a given product may be gauged to determine whether or not it meets the specification. From

FIG. 4.




this viewpoint, the process of specification is simple indeed. Knowing the limits $L_1$ and $L_2$ within which it is desirable that a given quality characteristic $X$ shall lie, all we need to do is to put these limits in writing as a requirement on the quality of a finished product. With such a specification at hand, it is presumed to be possible through measurement to classify a piece of product as conforming or non-conforming to specification.

As we have seen, however, two difficulties arise with this form of specification. Suppose that the quality in question, the blowing time of a fuse for example, is one that can be determined only by destructive tests. How can one give assurance that the quality of a given piece of product will meet the specification without first destroying it? Or again, if we concern ourselves with the fact that even where the quality characteristic may be measured, there is always a certain expected fraction $p$ falling outside the tolerance limits, how can we go about attaining an economic minimum to this fraction non-conforming? A little reflection shows that the simple specification of a go no-go tolerance is not satisfactory in such instances from the viewpoint of: (1) Economy, and (2) Quality Assurance.

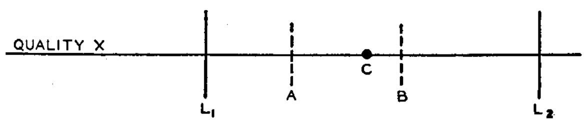
At this juncture statistical theory steps in with the concept of two limits $A$ and $B$ which we shall term action limits, and which lie, in general, within $L_1$ and $L_2$. These limits are such that when the observed quality of a piece of product falls outside of them, even though it be still within the limits $L_1$ and $L_2$, it is desirable to look at the manufacturing process in order to discover and remove, if possible, a cause of variation which need not be left to chance. In other words, whereas limits $L_1$ and $L_2$ provide a means of gauging product already made, action limits $A$ and $B$ provide a means of directing action toward the process in order that the quality of product not yet made may be less variable on the average.

Furthermore, the statistical theory of quality control introduces the concept of another point $C$ lying somewhere between the action limits $A$ and $B$ which is the expected or the aimed-at average quality in an economically controlled state. We should perhaps pause a moment to note the significance of the point $C$ from the viewpoint of design or the use of material that has already been made. Let us take, for example, a very simple case of setting over-all tolerances. Suppose we start with the concept of the go no-go tolerance of 1870 and wish to fix the over-all tolerance for $n$ pieceparts assembled in such a way that the resultant quality of the $n$ parts is the arithmetic sum of the qualities of the parts. An extremely simple example would be the thickness of a pile of $n$ washers. The older method of fixing such a tolerance is to take the sum of the tolerances on the pieceparts. This is generally many times too large from the viewpoint of economy. The efficient way of setting such tolerances is in terms of the concept of the expected value and the expected standard deviation about this value. In other words, the concept of expected value is of fundamental importance in all design work in which an attempt is made to fix over-all tolerances in terms of those of pieceparts.

Thus we see how, starting with the simple concept of a go no-go tolerance in a specification as illustrated in Fig. 4, it is necessary in many cases [3] for economy and quality assurance reasons to introduce certain action limits $A$ and $B$ and also a certain expected value $C$ to be used in design formulae. The situation corresponding to the simplest case is shown schematically in Fig. 5. Statistical theory alone is responsible for the introduction of the concept of action limits $A$ and $B$ and the expected value $C$.

The next question to be considered is that of determining the points $A$, $B$, and $C$. It is extremely important to note that whereas $L_1$ and $L_2$ can for the most part be set a priori, the other three points cannot be thus set because they depend upon the results economically attainable in the process of production and in the process of inspection. In particular, the action limits $A$ and $B$ call for action directed at the production process. Let us therefore consider next the second step in quality control, production.

FIG. 5.




## APPLICATION TO PRODUCTION.

As already noted, if science were really exact and physical laws were such that an engineer could figure out beforehand just how to make exactly what he wanted, there would be no story to tell tonight. Though scientists in the past may at times have entertained this fair dream, such hopes have long ago been blasted. Seemingly we are foreordained to live in a world where chance or unknown causes play their part. Even in the simplest case where we try to do the same thing, or go through the same kind of operation, again and again under what appear to us to be the same essential conditions, our

[3] It should be noted, of course, that if there is no economic or quality assurance reason for going beyond the concept of the go no-go tolerance, statistical theory has nothing to add. Likewise, it should be noted that, although the action limits $A$ and $B$ may lie within the tolerance limits $L_1$ and $L_2$, product already produced and found within the limits $L_1$ and $L_2$ is still considered as conforming although outside $A$ and $B$. In other words, the action limits $A$ and $B$ do not apply as a gauge for product already made.

results are marked by certain variations which we cannot explain in terms of known causes. The real economic problem in such a case is to have some kind of practical technique for determining when we have gone as far as it is economically feasible to go in finding and eliminating causes of variation.

#### Control Chart Technique—How it Works.

Now the control chart technique referred to in Fig. 2 is the practical contribution of statistical theory to this problem of indicating when one who is trying to do the same thing again and again should look for assignable causes. Let us think of any group of chance causes of variation as consisting of a subgroup that can be found and controlled and of a remainder which it is perhaps beyond the power of man to find and control. For the purpose of picturesqueness, let us think of these assignable or findable causes as little demons that one is trying to remove from the production process. The control chart technique may then be thought of as a trap to catch these demons. Let us see how it works.

I take as an example a problem that arose in the early testing out of this technique—the development of a high insulation material with minimum variability. Two hundred four different pieces were made and tested and gave the results shown in Table I, where the numbers are in megohms. The

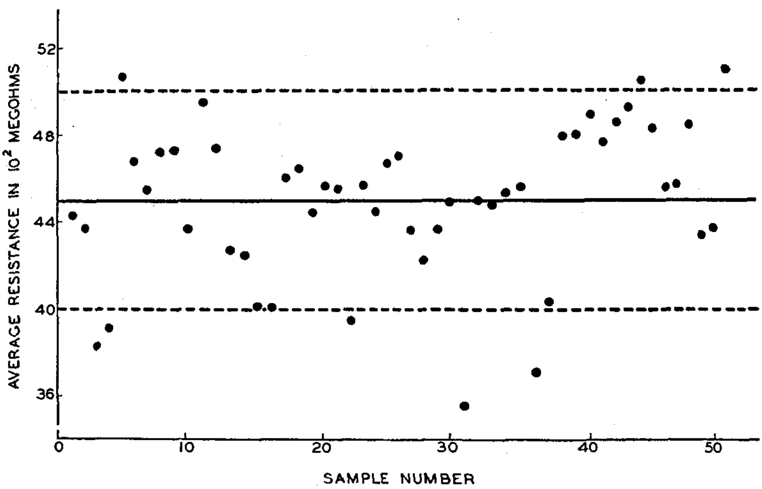
TABLE I.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 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 | 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 |

data are given here in the order that the pieces were made, beginning at the top of the first column, reading down that column, then down the second column, and so on. The practical problem is simply this: Have all the little demons been removed from the production process? In other words, are the resultant variations such that economically they should be left to chance?

We cannot take time here to describe the details of the technique other than to say that it consists in this case of dividing the 204 observed values into fifty-one sets of four in the order in which they were taken, finding the average for

FIG. 6.



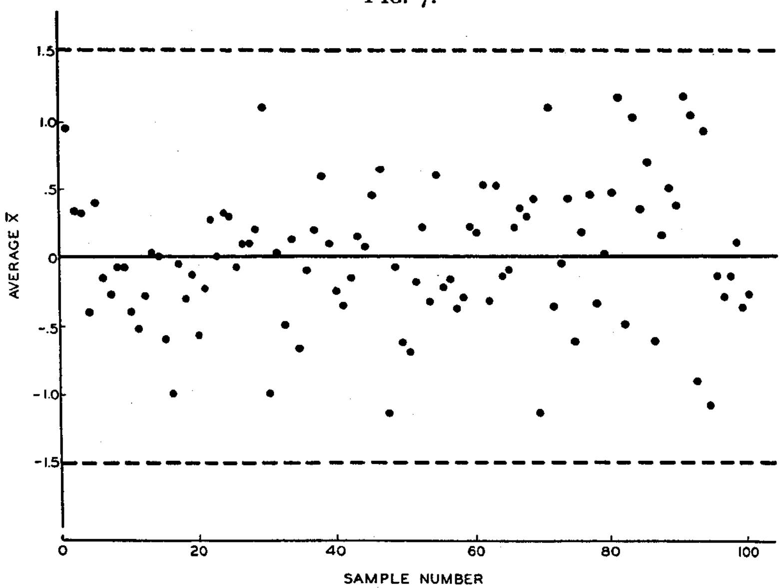

each set, and plotting these averages on a vertical scale as shown in Fig. 6. By going through a little computation, we get the two limits which are shown by the dotted lines in this figure. If any point falls outside these limits, the technique says to look for a little demon. The rest of the story in this case is simple indeed. They looked, they found, they removed some demons. As a result they succeeded in reducing to a minimum the variability which should be left to chance.

Years of experience in the application of such techniques have met with success. The technique itself is one which could be attained only through the use of statistical theory.

**How the Technique Works in a Case Where There Is No Demon of Chance.**

It is somewhat illuminating to see just how this test works in a case where I think all of you will agree that one has gone as far as he can in removing the causes of variation. Suppose I had a bowl of chips before me like the one I have in my hand and that we write numbers on these chips distributed, let us say, in the well-known bell-shaped normal distribution. Suppose one of you were to step up here and, after being blindfolded, you mixed the chips in the bowl and drew out one and we recorded the number here on the board. Suppose you put

FIG. 7.

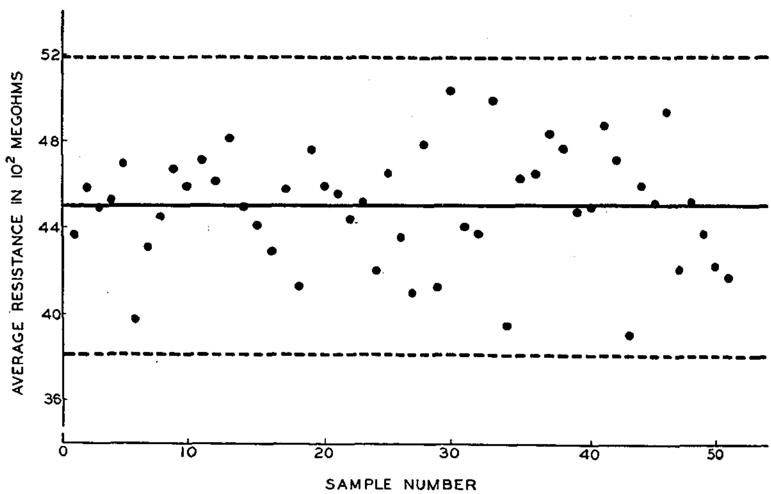



the chip back, stirred them up, and drew out another number, continuing this process until we had four hundred numbers on the board. Now, is there any way that any of you know in which you might under such circumstances control the numbers you draw, assuming, of course, that the chips are of the same size and feel the same? I take it that you would agree that the kind of variations you get in the numbers through such an operation is beyond your control. Well, let us see what the criterion shows when tried on such a set of four hundred numbers grouped into samples of four. Fig. 7 shows the averages of such a series of samples—you would not have occasion to look for any little demon or findable cause. The chart technique would not lead you astray.

#### The Human Element in the Technique.

I said something in the introduction about trying to indicate the role of intelligence in the application of such techniques. It is perhaps trite to note that as the basis of any inference there is some hypothesis. There could be no facts without a theory and there could be no theory without a fact. Take, for example, the set of 204 observations shown in Table

FIG. 8.

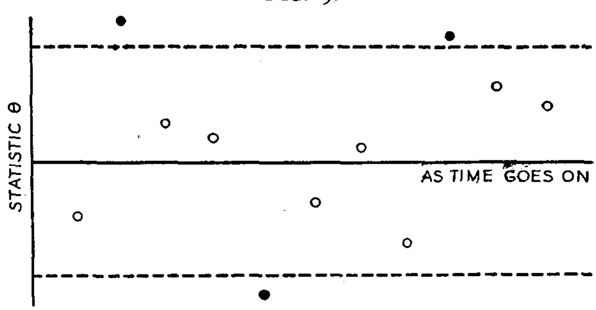



1. So far as the data are concerned, they are simply a set of 204 numbers. It is the way you use the numbers that counts. Note, for example, that we took the sets of four in the order in which the pieces were made. Suppose now that instead of doing this, we were to put the 204 numbers on a set of as many similar chips, put the chips in a bowl, and draw the numbers out, one at a time. Suppose then that we were to apply this same criterion, taking the first four, the second four, and the like, and plotting the averages of the sets, what would we get? One such experiment gave the results shown in Fig. 8; all the points are within the action limits. By mixing these

numbers one has lost some essential information that all the King's horses and all the King's men could not put back into the data.

There is in the literature on the application of such tests a very important statement, namely, that the one who takes the data must first divide them into what is termed rational subgroups.[4] I do not care to worry you about the technical meaning of this term except to point out that before you can successfully apply a test to catch one of these little demons of chance, you have to use your head in this process of dividing the data into rational subgroups. In other words, human intelligence, the ability to make hypotheses, is still the greatest power given to man in the control of his surroundings, even in

FIG. 9.




this simple case. But given intelligence to make hypotheses, there is something that statistical theory contributes through the control chart technique that is essential to the job of going to the economic limit in eliminating causes of variation and thereby getting control of a process to the limiting extent to which one can hope to go.

That is to say, given a situation in which you have an intelligent grouping of the data, statistical theory contributes a definite operational technique whereby you can calculate a statistic $\theta$ such as an average, fraction defective, or the like, which when plotted in a control chart schematically represented in Fig. 9 tells one when to look for trouble. When a point falls outside, look for trouble in the process; when it falls

[4] See, for example, Shewhart, W. A., "Economic Control of Quality of Manufactured Product," D. Van Nostrand and Company, New York, 1931.

inside, assume that ALL'S WELL. The dotted lines are the practical action limits $A$ and $B$ discussed under Step I—specification.

## APPLICATION TO INSPECTION.

### Some Early Perplexing Problems.

Production is followed by inspection, Step III in the control process; someone must look upon that produced and say whether or not its quality is what it was supposed to be. In principle a manufacturer first specifies, then produces, and then inspects. Classic practice is to assign these jobs to three separate departments or at least to different personnel. Ever since there was production there has been inspection; so let us look at the fundamental inspection problem.

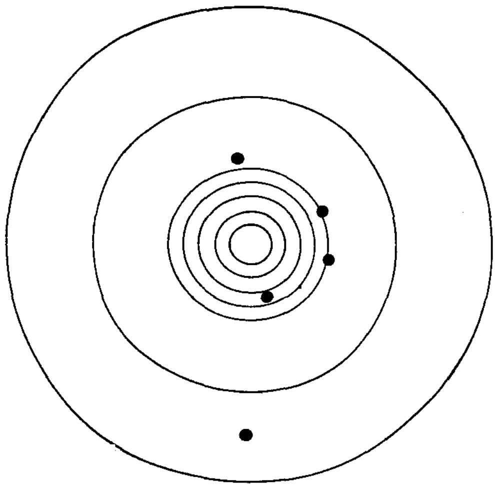
The production department turns out a lot of $N$ pieceparts or units. Two conditions may arise: (*a*) it may be too expensive to look at all $N$ units even when the test is non-destructive, (*b*) the test may be destructive, as tests for chemical composition, breaking strength, blowing time of fuse, etc., and we have no other practical choice than to appeal to a sample. In fact this problem constitutes the door wherein statistics made its entry into the field of manufacturing in the sense of quality control as considered here tonight. The intricacies of this problem are legion, but we do not have to consider this mountain of detail to see clearly a fundamental contribution of the statistical theory of quality control.

I take it that all of you have either shot at a mark or watched someone else doing so. Hence I shall start with a question to which some or all of you may be willing to risk an answer. Fig. 10 shows the placement of five shots fired by the same person at the bull's-eye. The question I want to raise is: Where will the next shot fall if fired in the same way by the same person? To make the question a little more obviously similar to the inspection problem, I might ask: Suppose the person is to fire a series of $N$ shots of which the first five are shown in Fig. 10. What fraction $p$ of these $N-5$ shots will fall within a specified circle?

Possibly some of you at least will be cautious and say that you wouldn't attempt a prediction on such meagre information. But suppose that instead of shots we have 25 fuses and the blowing times for five of them. We may think of the chance cause system in the production process as producing variation in the blowing times. Having seen by test where the chance causes place the blowing times of the five, what can you say about the blowing times of the twenty as yet untested?

In respect to the question about the five shots, you may suggest that you ought to take a few more before risking an answer. However, in the case of the fuses, you will pretty likely think of your pocketbook before you answer. Possibly you now begin to feel, even if you have not previously done so, the way in which economics enters such questions.

FIG. 10.



But let us not worry over these two comparisons too much. Let us take at its face value the suggestion that we need more than a sample of five. Let us think back to the eve of November 3, 1936. One of the questions of the day was whether or not the published results of the Literary Digest poll of more than a million was sufficient to tell the story of the morrow. It was not. Size alone is not enough. This is one of the important facts which an inspection engineer learns early in the game.

There are lots of other knotty situations for the inspection engineer to face. For example, let us refer again to the placement of shots in Fig. 10. You will note that three of the shots are to the right of a perpendicular to the base of this figure drawn through the bull's-eye. Suppose this were your target, would you pull a little to the left the next time you shot? That is to say, would you infer from the placement of the five shots, even though you had no other evidence, that there were some constant or assignable causes—some little demons—tending to place your shots to the right of this perpendicular? Everyone to whom I have put this question has answered yes. They have said that they would aim a little to the left to "correct" for this apparent effect.

Well, now I will let you in on a little secret as to how the shots on that target were fired. We started with a bowl of 1000 chips on which numbers were marked in such a way that there was a normal distribution in the bowl; a chip was drawn by one of us blindfolded, the number recorded, the chip returned to the bowl, and this number was taken as the $X$ displacement of a shot; the $Y$ displacement was obtained in a similar way by another drawing. With this information at hand, I think you will agree that your tendency to pull to the left in this case would not have been justified. In fact, such an effort to "correct" would tend to increase the spread of your shots on the target beyond what would in such a case be necessary. In other words, if we followed such a procedure in this particular case, we could not reduce the variability to a minimum.

To discuss the separate aspects of this problem is enough for an evening in itself. All I wish to do here, however, is to get all of us to feel some of the uncertainties in the interpretation of a sample. In so doing I will have gotten you to sense some of the stumbling blocks that lay in the path of early attempts to develop a rational theory for inspection engineering. In other words, you get a little of the atmosphere existing at the beginning of the application of statistical theory and technique to inspection in the sense considered in this paper.

Let us get back to the fact that some of you may have felt that the sample of five shots was too small to tell you much about where the next ones would fall and the fact that some of you may have wondered why a sample of more than a million failed to indicate the landslide of November 3, 1936.

Of course, one could find lots of advice to the effect that statistical sampling theory applied only to certain kinds of samples such as random and representative ones. But how is an inspection engineer to know a random or representative sample when he meets one face to face? That was the rub! The inspection engineer had to answer that one before he could taken on another.

This was one side of the story. The other was that it was a pretty well established fact that conditions exist in certain realms of nature where experience had justified predictions made on the basis of probability theory. Witness, for example, the insurance business, many molecular phenomena, the throws of a symmetrical die, drawings from a bowl of similar chips, or the distribution of suits in a well-shuffled deck of cards.

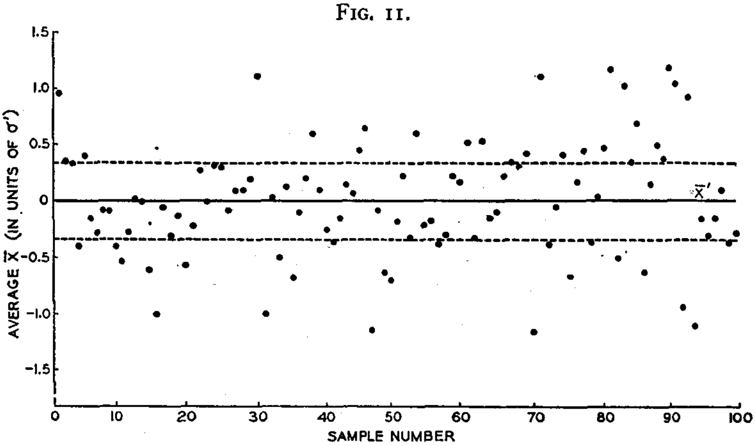
*What Was Done.*—Here then was the setting. What was done? Well, it was to try to take account of what we know about a sample in addition to its complexion and size. For example, as soon as I told you how the sample of five shots, Fig. 10, were made, you probably agreed that such variations are the kind that we must leave to chance in conducting such a drawing. Let us then just get a wee picture of the importance of what we know about the origin of a sample as a factor influencing our ability to predict.

I have here a bowl containing 1000 chips physically similar to the one I hold in my hand. On each chip is a number. The distribution of numbers is in accord with the normal law. The arithmetic mean of these numbers is zero and the root mean square deviation is 1. With this information, I can make a lot of verifiable predictions about samples of any size $n$ drawn with replacement from this bowl, without first drawing a sample.

For example one can say that in, let us say, $N$ drawings of samples of size $n$, approximately 50 per cent of the averages of these samples will fall within the limits $0 \pm \frac{.6745}{\sqrt{n}}$. Fig. 11 shows the results for one such experiment of 100 samples of four. It worked pretty well. The number of averages within is 52 per cent as compared to an expected 50 per cent. This condition of predictability exhibited by the bowl represents an ideal. If in some way one could come to know a process of production well enough that he could predict the quality of future output of that process with as much justified assurance as he can predict the results of future drawings from the bowl of chips here on the table, he would have gone as far as one may reasonably hope to go.

Now, let us assume that we don't know the average and standard deviation even though we know that the distribution of numbers has the same normal shape. Have we any way of drawing the limits such as shown in Fig. 11 that will include 50 per cent of the averages? The answer is No. First

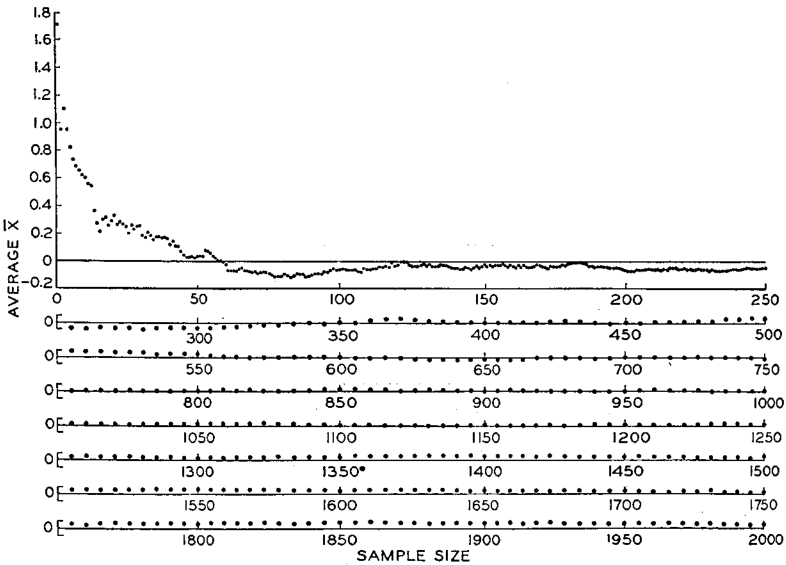
FIG. 11.




we must find the average and standard deviation. Suppose you can find these only by sampling. This corresponds to the practical case, where instead of drawing from a bowl we have the fluctuations in some quality characteristic of pieces of a given kind of product produced by a given kind of process.

How large a sample would you have to take from the bowl before you could predict as well as in the case above about the 50 per cent range for future samples of four? So that you may have some basis for risking an answer, assuming that you are not already professional statisticians, let us see how rapidly and in what sense the average of a sample of $n$ approaches the true average in this bowl. Fig. 12 shows one such experimental determination. The ordinate of a point is the observed average for the sample of size corresponding to the abscissa of that point. Of course, if you took another series of 2000 drawings and plotted the points for the same set of sample sizes the set of points would almost certainly not be the same. In fact they might start any place from + 3 to − 3 (these being the maximum and minimum numbers in the bowl) but in general the points would tend to hug the zero line progressively closer with increase in sample size much as shown in this figure.

FIG. 12.

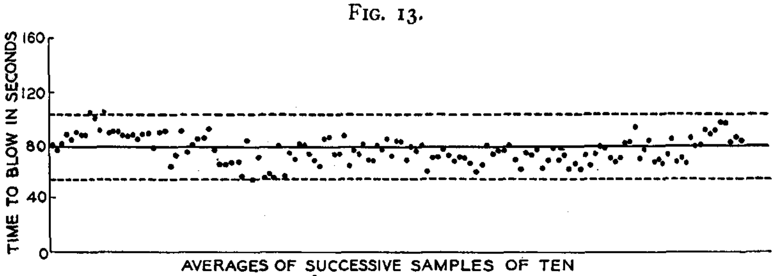



The first point I wish to make is that the rate of approach is slow. I don't believe you would think much in general of prediction from samples of 25 or even 50. This is particularly true when you learn that the rate of approach of an observed standard deviation is much slower, and you need that too in order to set the limits. In fact, you might want to go to a sample of 500 or 1000 before you would be satisfied with the likely reduction in your error of prediction.

Now, the second and even more important point I wish to make is that there is no royal road to derive such limits from a sample so as to reduce the required sample size. Be he king or peasant every man must travel this same road.

But this is not the crux of the practical case. Here we have assumed sampling from a bowl. In practice we must first convince ourselves that a sample behaves as though it were drawn from a bowl. How can an engineer do this? The answer given by the theory and practice of the use of the control chart shows that we must first get rid of assignable causes revealed by such a method. In other words, in order to attain desirable predictability in the probability sense, we must first apply the quality control chart technique to give assurance that the causes of variability are such as must from a practical viewpoint be left to chance. In general, to do this requires quite a large sample as I have shown in detail elsewhere.

FIG. 13.




All I can do here is to show you a sample of the proof of the pudding. Fig. 13 shows certain practical limits for the quality of 99 per cent of the product of a certain kind of fuse in respect to the blowing time (a destructive test). The points are averages of 10. They stay within almost as well as the data from a bowl!

**Practical Significance.**

What has statistical theory and method contributed to the third step in control, inspection? Well, first, it has shown that in the interpretation of the sample we need to go behind the sample. We need to know how it was taken and most certainly we need to consider the available evidence as to whether or not the sample came from a controlled process. That is, we are thrown back on the evidence as to the degree of economic control obtained in the second step, production. It has shown very definitely that in order to minimize the cost of inspection requisite for giving a specified degree of quality assurance, it is necessary first to attain a state of statistical control as evidenced by the satisfaction of certain quality control chart criteria.

This has made necessary two new techniques in the field of inspection.

(*a*) *Quality Report.*—A continuing quality report providing a record of the evidence of the state of control of the quality.

(*b*) *Inspection Practices.*—The preparation and use of inspection practices that take into account the state of control of the quality.

The quality report not only furnishes a background upon which to interpret the observed quality in a current sample of product but the central line in the chart approaches as a statistical limit the expected value $C$ to be used in the specification of currently economic standard quality. It should be recalled that this $C$ is the expected quality that enters into the computation of the overall quality of a design in terms of the qualities of the parts and is therefore a necessary factor. A third use of the quality report is to provide for design and production engineers evidence of lack of control or the presence of assignable causes of variability which should not be left to chance.

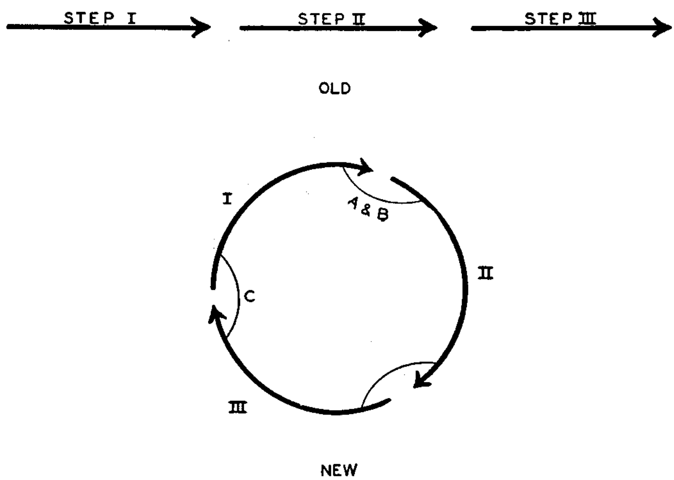
Now we are in a place to consider a very important potential contribution of mass production to scientific industrial progress. You have seen how in order to remove the assignable causes of variability in quality of product, it is possible to apply the quality control technique and you have seen how essential it is to have a high degree of control in order to give the highest quality assurance. But you have also seen that this potential state of economic control can only be approached slowly as a statistical limit. Control of this kind cannot be reached in a day. It cannot be reached in the production of a product in which only a few pieces are manufactured. It can, however, be approached scientifically in a continuing mass production.

## CONCLUSION.

We have seen clearly that statistical theory and technique has contributed something very definite to each of the three steps in control. But it has done something much more than this. It has changed our whole concept of control.

As previously conceived, Steps I, II, and III were more or less independent. Of course, if we had *certain* knowledge they might be independent, but in science we can only have *probable* knowledge. What then has happened to our concept of three independent steps? The answer is shown schematically in Fig. 14. We see that the steps must go in a circle. They are no longer conceived of as being independent; they are *correlated*. We have seen that in order to take Step I we need action limits *A* and *B*. These must come from Step II.

FIG. 14.




Also we 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 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. Finally we should note that both statistical theory and practice show that this state of control can be attained as a statistical limit only in the continuing process of mass production.

Thus you see how statistical theory has helped the William Tell of our story, the manufacturer, to hit his mark—the highest standard of quality and the greatest quality assurance at a given cost.