R331
WAS

I. E. B. 4

BASIS FOR ECONOMIC CONTROL OF QUALITY

BY

W. A. SHEWHART

SCIENTIFIC BASIS FOR INTERPRETING THE SIGNIFICANCE
OF CHANCE VARIATIONS IN QUALITY OF PRODUCT AND
FOR ELIMINATING CAUSES OF VARIABILITY WHICH NEED NOT
BE LEFT TO CHANCE MAKING POSSIBLE MORE UNIFORM
QUALITY AND THEREBY EFFECTING CERTAIN ECONOMIES

BELL TELEPHONE LABORATORIES, INC.
INSPECTION ENGINEERING DEPARTMENT

ISSUED FOR USE OF MEMBERS OF THE DEPARTMENT
OCTOBER 1929

FOREWORD

This is the fourth of a series of bulletins issued primarily for the use of members of the Inspection Engineering Department. Each of these bulletins treats of a particular phase of the general subject, "Control of Quality of Manufactured Product". An attempt has been made to make the discussion in each bulletin as nearly as possible a complete and independent unit so that the material contained therein may be used independently of that contained in other bulletins of the series. On the other hand, however, it is hoped that when all of the bulletins in the series have been issued, they will constitute a unified treatment of the above subject divided into the following parts:

I - Introduction.
II - Presentation of Data by Means of Simple Statistics.
III - Basis for Economic Control of Quality.
IV - Detection of Quality Variation which Should not Be Left to Chance.
V - Measurement of Quality.
VI - Quality Standards for Raw Materials.
VII - Economic Control of Quality through Inspection.
VIII - Economic Control of Quality through Design.
IX - Tables and Nomograms with a Discussion of Nomographic Treatment of Data.

The order of presentation of these parts has been governed by the immediate needs of the Department. For example, I.E.B. 1 and I.E.B. 2 constitute as it were Parts IV and V of the complete story. The present bulletin I.E.B. 4 constitutes Part III and treats of the subject "Basis for Economic Control of Quality".

This bulletin presents a rational basis for making more efficient use of experience through generalizations having to do with phenomena controlled by chance, and it also provides a practical basis for knowing when the limit in this direction has been reached.

It raises certain fundamental questions which cannot be explicitly answered but in every case a practical answer has been given through an empirical method which experience has shown to work. Engineering experience has shown that there exist in practice systems of causes which give rise to product whose quality distribution obeys the Law of Large Numbers. Of still greater practical significance is the fact that for such distributions empirical information has shown what limits associated with a given probability are to be used in detecting lack of control even though it is not possible to establish such

SHEWHART'S COLLECTION

STATISTICAL INSTITUTE LIBRARY No. R 331

FOREWORD

This is the fourth of a series of bulletins issued primarily for the use of members of the Inspection Engineering Department. Each of these bulletins treats of a particular phase of the general subject, "Control of Quality of Manufactured Product ". An attempt has been made to make the discussion in each bulletin as nearly as possible a complete and independent unit so that the material contained therein may be used independently of that contained in other bulletins of the series. On the other hand, however, it is hoped that when all of the bulletins in the series have been issued, they will constitute a unified treatment of the above subject divided into the following parts:

I - Introduction.
II - Presentation of Data by Means of Simple Statistics.
III - Basis for Economic Control of Quality.
IV - Detection of Quality Variation which Should not Be Left to Chance.
V - Measurement of Quality.
VI - Quality Standards for Raw Materials.
VII - Economic Control of Quality through Inspection.
VIII - Economic Control of Quality through Design.
IX - Tables and Nomograms with a Discussion of Nomographic Treatment of Data.

The order of presentation of these parts has been governed by the immediate needs of the Department. For example, I.E.B. 1 and I.E.B. 2 constitute as it were Parts IV and V of the complete story. The present bulletin I.E.B. 4 constitutes Part III and treats of the subject "Basis for Economic Control of Quality".

This bulletin presents a rational basis for making more efficient use of experience through generalizations having to do with phenomena controlled by chance, and it also provides a practical basis for knowing when the limit in this direction has been reached.

It raises certain fundamental questions which cannot be explicitly answered but in every case a practical answer has been given through an empirical method which experience has shown to work. Engineering experience has shown that there exist in practice systems of causes which give rise to product whose quality distribution obeys the Law of Large Numbers. Of still greater practical significance is the fact that for such distributions empirical information has shown what limits associated with a given probability are to be used in detecting lack of control even though it is not possible to establish such

SHEWHART'S COLLECTION

STATISTICAL INSTITUTE LIBR.
No. R 331

limits upon the basis of a priori reasoning.

In general each chapter starts with a simple statement of the problem or outline of the points to be considered and the last paragraph gives a statement of definite conclusions reached. As far as possible the statement has been made free of mathematics and it is believed that the reader may successfully obtain a logical basis for the use of the principles of control without going into all of the details of even the small amount of mathematics given. Quotations are cited throughout the bulletin to show that the principles of control here discussed form a basis for the solution of similar problems arising in many fields outside of engineering.

The following is an outline of the material by chapters.

Chapter I presents a simple qualitative picture of what we mean by a state of control and how this condition of control is used to advantage in certain fields.

Chapter II gives five specific reasons illustrated by simple problems showing why control is desirable in engineering.

Chapter III presents a causal interpretation of chance phenomena forming a basis for the specification of the necessary and sufficient conditions for control in terms of causes.

Chapter IV translates these necessary and sufficient conditions into terms of the characteristics of the distributions of effects.

Chapter V introduces the law of large numbers which enables us to connect the distribution of effects of a hypothetical chance cause system to that of a natural chance cause system.

Chapter VI shows how the principle of control is invaluable in all scientific endeavor. First, because it is necessary in the interpretation of any set of observed data and, second, because many of the properties of matter and even many natural laws are but statements of certain conditions arising under controlled states.

Chapter VII takes up the important problem of making control practical by showing how we can establish criteria for detecting lack of control and presents information based upon the experience of the Laboratories to show that the probability associated with any one of these criteria can be and has been satisfactorily determined upon an empirical basis.

Chapter VIII closes the study with an outline of the ways in which it should be possible to make more efficient use of experience in connection with every research program both in pure and applied science. This stresses the necessity for testing all data for lack of control before they are used in engineering formulas. This is particularly necessary where the data represent the physical properties of raw materials. This chapter also indicates the nature of advantages to be obtained by testing for lack of control at each of the five stages of development and use from raw material to the end of the life of the finished product. Such advantages are obtained by a close study of the results of tests made at each of the steps and by cooperation with the laboratory research organization in the elimination of assignable causes of variability.

TABLE OF CONTENTS

"To-day the mathematical physicist seems more and more inclined to the opinion that each of the so-called laws of nature is essentially statistical, and that all our equations and theories can do, is to provide us with a series of orbits of varying probabilities."

Engineering - July, 1927

"A large amount of work has been done in developing statistical methods on the scientific side, and it is natural for any one interested in science to hope that all this work may be utilized in commerce and industry. There are signs that such a movement has started, and it would be unfortunate indeed if those responsible in practical affairs fail to take advantage of the improved statistical machinery now available."

Nature - January, 1926

CHAPTER I

NATURE OF CONTROL

> "Perfect knowledge alone can give certainty, and in nature perfect knowledge would be infinite knowledge, which is clearly beyond our capacities. We have, therefore, to content ourselves with partial knowledge - knowledge mingled with ignorance, producing doubt."
>
> W. Stanley Jevons
> Principles of Science

1. Why Controlled Quality is a Variable Quality

To make a thing the way you want to make it is one popular conception of control. The kernel of this conception is as old as the human race, for throughout the ages man has been trying to gain control of his surroundings through acquired knowledge of the physical world. Today we see the fruition of this attempt in marvelous industrial development attributable for the most part to applied science. In other words, we have achieved a certain amount of success by formulating principles which experience has proved to be serviceable in gaining control of our surroundings. The more scientific knowledge we have gained, the more able have we become to do what we want to do. Once we had to walk to get from one place to another; today we can fly. Thus progress has been made because today we can do some things we want to do that we could not do before. But even today we are far from the goal of being able to do <u>exactly</u> what we want to do.

Let us consider a very simple illustration of our inability to do exactly what we want to do. In this way we come upon the first essential principle in our understanding of the control of product. We come to see that control must mean doing what we want to do <u>within limits</u>.

Write the letter *a* on a piece of paper. Now make another *a* just like the first one; then another and another until you have a series of *a*'s, *a*, *a*, *a*, *a*,.... You try to make all the *a*'s alike but you don't; you can't. You are willing to accept this as an empirically established fact. But what of it? Let us see just what this means in respect to control. Why can we not do a simple thing like making all the *a*'s just alike? Your answer leads to a generalization which all of us are perhaps willing to accept. It is that there are many causes of variability among the *a*'s: the paper was not smooth, the lead in the pencil was not uniform and the unavoidable variability in your external surroundings reacted upon you to introduce variations in the a's. But are these the only causes of variability in the a's?

Probability not. We accept our human limitations and say that likely there are many other factors. If we could but name all the reasons why we cannot make the a's alike, we would most assuredly have a better understanding of a certain part of nature than we now have. Of course this conception of what it means to be able to do what we want to do is not new; it does not belong exclusively to any one field of human thought; it is a commonly accepted conception. Thus, Tennyson says,

"Little flower, but <u>if</u> I could understand,
What you are, root and all, and all in all,
I should know what God and man is."

The point to be made in this simple illustration is that we are limited in doing what we want to do; that to do what we set out to do, even in so simple a thing as making a's that are alike requires almost infinite knowledge compared with that which we now possess. It follows, therefore, since we are thus willing to accept as axiomatic that we cannot do what we want to do and that we cannot hope to understand why we cannot, that we must also accept as axiomatic that a controlled quality will not be a constant quality. <u>Instead a controlled quality must be a variable quality</u>.

But go back to the results of the experiment on the a's and we shall find out something more about control. Your a's are different from my a's; there is something about your a's which makes them yours and something about my a's that makes them mine. True, not all of your a's are alike. Neither are all of my a's alike. Each group of a's varies within a certain range and yet each group is distinguishable from the others. This distinguishable and, as it were, constant variability is something which on the face of it gives us a basis for extending our conception of control.

Let us amplify this point. We try to make several things alike but they come out different. We tried to make the a's alike and they came out different. Why the variability we do not know, but we postulate that there are numerous unknown causes of variation. Without the guiding hand of experience in such a situation where there are so many unknown causes of variability, we might give it up as hopeless to be able even to set certain limits within which variations might lie. However, experience saves us from such a hopeless state. It sets us trying to understand why your a's are yours and mine are mine and yet the two are different. It sets us trying to find out how unknown groups of causes may and do act in nature. It leads us to see that the distribution of effects of an unknown group of causes has in it very definite and important information for him who is but willing to see and understand.

This bulletin is a report of the findings in such a quest for a better understanding of how unknown groups of natural causes of variability appear to act, so that we may better be able to do what we want to do in the production of uniform quality of manufactured product. The course followed is a natural one. We glean what we can about the workings of these unknown chance causes and use this information to show how a basis can be established first for control within certain fixed limits and then for control within the narrowest attainable or minimum limits.

Let us take some very simple illustrative phenomena admittedly under the influence of chance, say length of life and molecular motion. Perhaps nothing is more uncertain than life itself unless it be molecular motion. Yet there is something certain about these uncertainties. In the law of mortality we find some of the essential characteristics of control within limits, and in the law of distribution of molecular displacement, similar characteristics of maximum control or control within minimum limits.

2. Control within Limits

The date of death always has seemed to be fixed by chance even though great human effort has been expended in trying to rob chance of this prerogative. We come into this world and from that very instant on are surrounded with causes of death seeking our life. Who knows whether or not death will overtake us within the next year? If so, what will be the cause? These questions we cannot answer. Some of us are to fall at one time from one cause, others at another time from another cause. In this fight for life we see then the element of uncertainty and the interplay of numerous unknown or chance causes.

Now, however, when we study the effect of these chance causes in producing deaths in large groups of individuals, we find some indication of a controlled condition. We find that this hidden host of causes produce deaths at an average rate which does not differ much over long periods of time. Even comparative life curves, Fig. 1, reveal an interesting similarity. They seem to show the effects of these causes to be much the same, say, in Germany as they are in England or here at home. Of course, upon closer observation the life curves do show marked differences which we attribute to the general effects of such factors as climate, sanitary conditions, medical service and habits of life in the different countries.

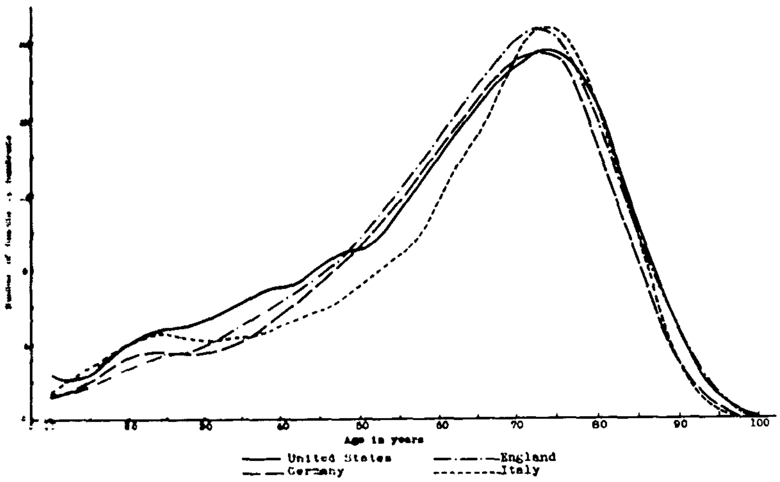


FIG. 1 - SOMETHING CERTAIN ABOUT THE UNCERTAINTIES OF LIFE
The similarity of life curves. The kind of evidence giving rise to a faith in generalized effects of chance causes.

However, even when we take great care in selecting a homogeneous group living under supposedly the same essential conditions, we still find a life curve quite similar to those already shown in Fig. 1, even indicating almost as wide a range in the rate of death at the different ages. From such observations we are led to believe that, as we approach the condition of homogeneity of population and surroundings, we approach what is customarily termed a "Law of mortality" such as indicated schematically in Fig. 2. In such a case we believe that the causes of death function so as to make the probability, let us call it p', of dying within given age limits, such as forty-five to fifty, constant.

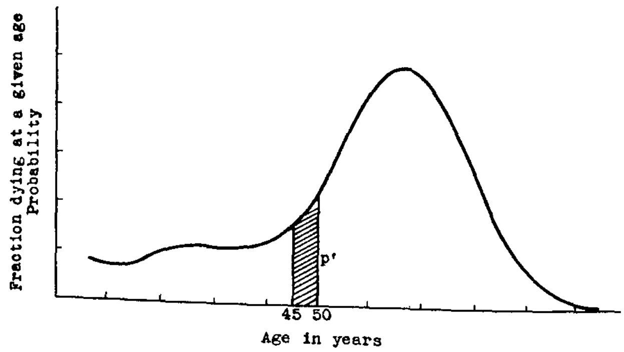


FIG. 2 - LAW OF MORTALITY - LAW OF FLUCTUATIONS CONTROLLED WITHIN LIMITS

Now, if an actuary can assure himself that the deaths in a population are produced by such a system of chance causes, he can set up limits within which the observed fraction dying in a group from this population may be expected to lie with any given degree of probability. In this sense death becomes a controlled phenomenon within limits. This makes possible the prediction that the observed fluctuations in the number of deaths in groups of n individuals from a given population should not exceed limits set by probability theory with the expectation of a check as close as the experimental one given in Fig. 3.

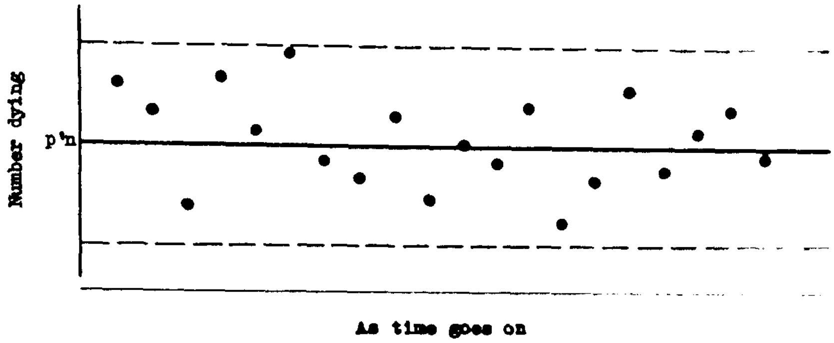


FIG. 3 - CONTROL WITHIN LIMITS EXHIBITED BY THE SYSTEM OF CHANCE CAUSES OF DEATH

As is well-known the insurance business is founded on just such an application of probability theory, and life insurance is a big business. The income from premiums alone in 1928 was over three billion dollars. In turn this application rests upon the assumption of the existence of a kind of statistical equilibrium among the effects of an unknown system of chance causes expressible in the assumption that the probability of dying within a given age limit is, under the assumed conditions, an objective and constant reality.

Thus briefly we have seen that the insurance business is founded upon an assumed constancy of the probability of dying within any prescribed age limits which, in turn, depends upon the assumed constancy of a chance system of causes in the sense of control within limits.

If, now, the quality of a given product is controlled by such a system of chance causes, may we not reasonably expect to make use of this fact in setting limits within which this quality may be expected to vary and thus gain certain economic advantages soon to be outlined? Obviously the crucial test as to the possibility of making such applications is to show that the probability of the cause system producing a quality within a given range is constant.

3. Control within Minimum Limits - Maximum Control

Just about a century ago, in 1827 to be exact, an English botanist, Brown, saw something through his microscope that caught his interest. It was motion going on among the suspended particles almost as though they were alive. In a way it resembled the dance of dust particles in sunlight, so familiar to us, but this dance differed from that of the dust particles in important respects, - for example, adjacent particles seen under the microscope did not necessarily move in even approximately the same direction, as do adjacent dust particles suspended in the air.

A photograph (after Compton) of the kind of motion seen by Brown is reproduced in Fig. 4. What we see here represents the traces of particles, far too small to be seen with the naked eye and yet millions of times larger than an atom, as they take part in their own characteristic motion.

FIG. 4 - MOLECULAR JAZZ

Watch such motion for several minutes. So long as the temperature remains constant, there is no change. Watch it for hours, the motion remains characteristically the same. Watch it for days, we see no difference. Even particles suspended in liquids enclosed in quartz crystals for thousands of years show exactly the same kind of motion. Therefore, to the best of our knowledge there is remarkable permanence to this motion. Its characteristics remain constant. Here we certainly find a remarkable degree of constancy exhibited by a chance system of causes.

Suppose we follow the motion of one particle to get a better picture of this constancy. This has been done for us by several investigators, notably Perrin. In such an experiment he noted the position of a particle at the end of equal intervals of time, Fig. 5. He found that the direction of this motion observed in one interval differed in general from that in the next succeeding interval. He found that the direction of the motion presents what we instinctively call absolute irregularity. Let us ask ourselves certain questions about this motion.

Suppose we fix our attention on the particle at the point A. What made it move to B in the next interval of time? Of course we answer by saying that a particle moves at a given instant in a given direction, say AB, because the resultant force of the molecules hitting it in a plane perpendicular to this direction from the side away from B is greater than that on the side toward B; but at any given instant of time there is no way of telling what molecules are engaged in giving it such motion. We do not even know how many molecules are taking part. Do what we will, so long as the temperature is kept constant, we cannot change this motion in a given system. It cannot be said, for example, when the particle is at the point B that during the next interval of time it will move to C. We can do nothing to control the motion in the matter of displacement or in the matter of the direction of this displacement.

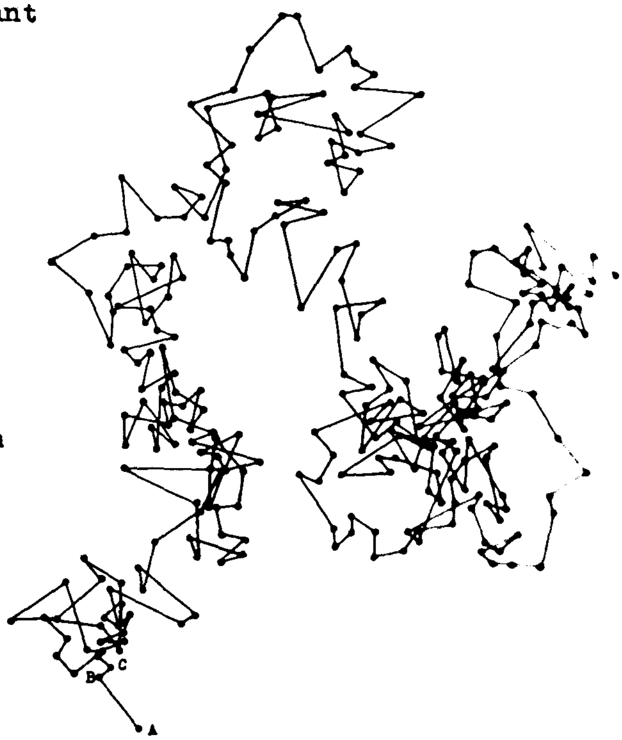


FIG. 5 - A CLOSE-UP OF MOTION CONTROLLED WITHIN MINIMUM LIMITS
Molecular Displacements in Successive Intervals

Since this motion cannot be changed without changing the system itself we are going to call it a state of <u>maximum control</u>. It is control because such characteristics as the length of displacement in a given interval of time can be foretold within limits; it is maximum control because we cannot reduce these limits.

Is there a state reached in the production of quality of product where the unknown chance causes produce fluctuations such that it would be just as difficult to find the causes thereof as it would be to find out just which molecules took part in producing motion, let us say from A to B? If there is such a state, it is reasonable to believe that <u>it would be just as foolish to waste time trying to control product beyond this point as it would be to try to control the motion of the molecules</u>. If there is, wouldn't it be reasonable to expect that the fluctuations in this quality might in some way resemble this random motion of the molecule? At least it seems a reasonable hunch. Later we shall find that it is theoretically justified. But we are getting ahead of our story. For the present we should learn more about this characteristic motion under maximum control.

Let us examine first the distribution of the x components of the segments of the paths. Within recent years we find an abundance of evidence gathered together indicating that these displacements appear to be distributed about zero in accord with what is called the normal law, Fig. 6.

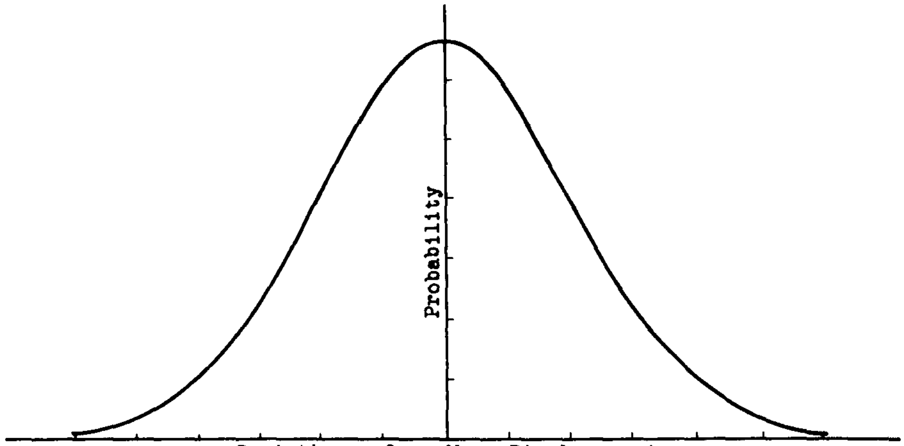


FIG. 6 - DISTRIBUTION OF MOLECULAR DISPLACEMENTS UNDER MAXIMUM CONTROL
Can we expect the same type of distribution to be characteristic of quality controlled within minimum limits?

That is to say, if x represents the deviation from the mean displacement, zero in this case, the probability dy of x lying within the range x to x + dx is given by

$$dy = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx \qquad (1)$$

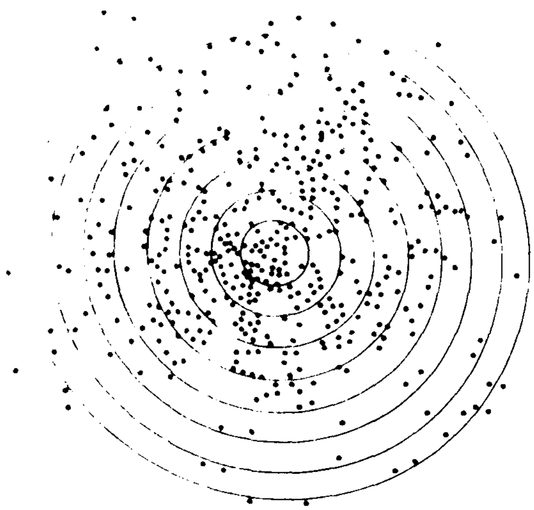


FIG. 7

where σ is the root mean square deviation.

In a similar way, if instead of measuring the x displacement we take a common origin for the vectors representing the motion through successive intervals of time, we get a distribution in the X, Y plane resembling that of Fig. 7. Experimental evidence indicates that the frequency distribution of the simultaneous occurrence of the deviation x with the deviation y, where the deviations are measured from the point corresponding to the center of the circles in Fig. 7, is given by the normal law for two variables and is represented schematically in Fig. 8.

In other words, it appears that the probability dz of the occurrence of a deviation within the interval x to x + dx simultaneously with a deviation y within the interval y to y + dy is given by the expression

$$dz = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}}\, e^{-\frac{1}{2(1-r^2)}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} - 2r\frac{xy}{\sigma_x\sigma_y}\right)}\, dx\, dy \qquad (2)$$

where σ_x and σ_y are the respective root mean square deviations in x and y and r is the correlation coefficient between x and y.[1]

Naturally the fact that this molecular motion appears to be distributed normally in accordance with Equations (1) and (2) does not necessarily

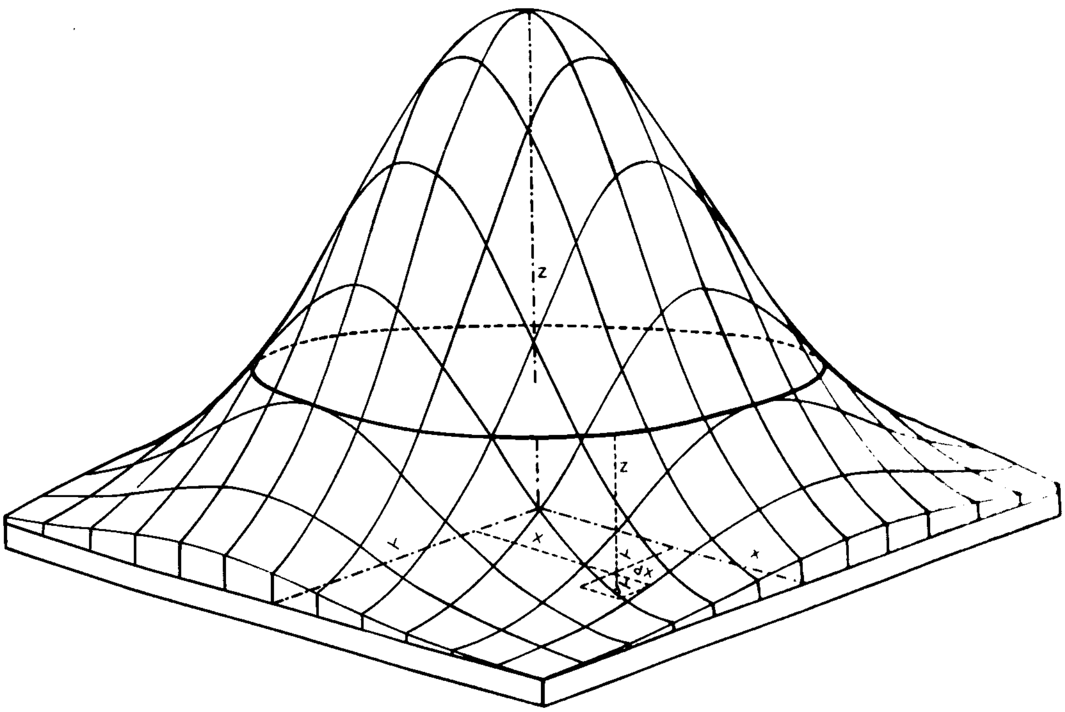


FIG. 8 - DISTRIBUTION OF THE SIMULTANEOUS OCCURRENCE OF THE DEVIATION x WITH THE DEVIATION y UNDER MAXIMUM CONTROL.

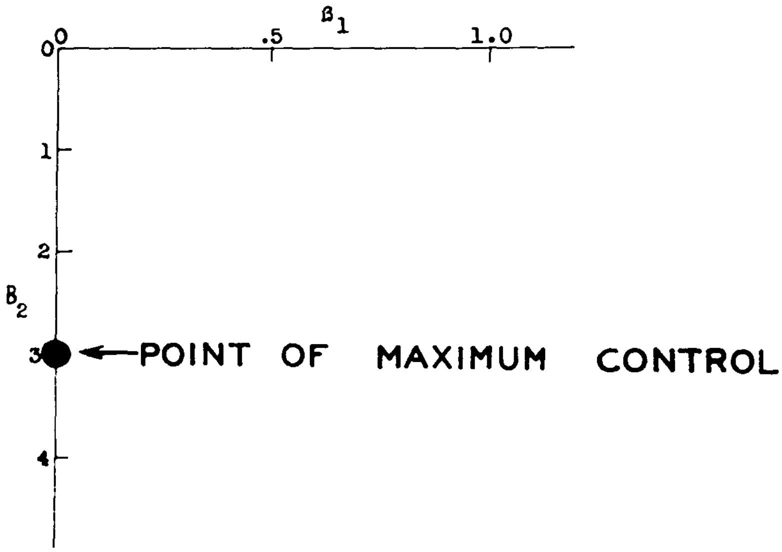
indicate that all the distributions under maximum control should bear even any resemblance to normality. It remains for us to present evidence later in this

- -

1. Those not familiar with these terms will find them discussed in I.E.B. 3.

bulletin indicating that this normal distribution is a limiting condition for maximum control.

Keeping in mind then that we are to come back again and again to the normal distribution as characterizing the limiting condition of maximum control, it will be interesting to see what this really means when expressed in terms of the skewness and flatness or kurtosis[1] of a distribution, or, in other words, β_1 and β_2. As is well-known, the normal law for one variable is characterized by the point 0,3 in the $\beta_1\beta_2$ plane. Looked at in this way, we see that in molecular motion at least there is an objective point of maximum control.



We shall see later that this same point is the objective point of maximum control of manufactured product.

4. Summary Statement

To summarize, we have seen that:

A. It is not possible to make pieces of product identical one with another and therefore that a controlled product must be a variable product.

B. We have reviewed evidence indicating that certain phenomena in life are controlled within limits in the sense that the probability of the unknown cause system producing a deviation within a given range appears to remain constant.

---

1. Those not familiar with these terms will find them discussed in I.E.B. 3.

C. In certain molecular phenomena where it is admitted that the condition of maximum control exists in the sense that it is humanly impossible to modify the action of the unknown cause system without changing the system itself, we have seen that there is a characteristic distribution of effects which is called normal.

In the light of these results, we take as our thesis that a controlled product must be variable, that it is reasonable to expect that the condition of control can be attained in production, and that there is a condition of maximum control of the production process beyond which it is not feasible to go unless we change the entire process.

INDIAN STATISTICAL INSTITUTE LIBRARY
No. R 331
Date.
CALCUTTA.

SHEWHART'S COLLECTION

CHAPTER II

WHY CONTROL QUALITY?

"It will help the builder of economic theory to have mastered the principle of movement towards equilibrium."

F. Y. Edgeworth
Application of Mathematics
to Political Economy

1. Basic Considerations

We have considered in some detail the kind of available evidence for the existence in nature of unknown or chance cause systems which act as though the probability of their producing a resultant effect lying within a given range is approximately constant. Let us now consider ways and means of making use of the knowledge that a product is controlled by such a system of causes. To begin with we shall say, in the light of the kind of evidence already reviewed, that:

A. If it can be shown that the deviations in the quality of product are not attributable to a controlled system of unknown or chance causes, then it is not possible to predict how the quality will vary in the future.

B. If it can be shown that the deviations in the quality of product are attributable to a controlled system of chance causes, then it is possible to set down limits within which the quality may be expected to vary corresponding to any given probability.

So long as the future quality lies within these limits, the producer may rest assured of its uniformity. In other words, he may be sure that the quality of his product is not changing from period to period by more than an amount attributable to sampling fluctuations. It will be shown later that it may often be possible in such cases to find and eliminate some of the causes of variability, thus reducing the limits.

C. If it can be shown that the quality of product deviates in a way attributable to a system of unknown or chance causes exhibiting maximum control, then it is possible to set down minimum limits within which the quality may be expected to vary corresponding to any given probability.

It will also be shown later that under these conditions it is doubtful whether or not future research will reveal ways and means of decreasing the variability of the quality without effectively changing the manufacturing process.

At this point you may ask: How are we to show that the quality of product satisfies one of the three conditions mentioned? Also, granted for the time being that control has existed in the past, how are we to justify the assumption that it will continue to do so in the future.

The answer to the first question involves the use of sampling theory considered in other bulletins of this series[1]. All that we are to do in the present bulletin is to set up certain necessary and sufficient conditions for control expressed in terms of the effects of the cause system. The sampling theory just mentioned will establish certain limits on the variability of these effects such that, if the observed fluctuation in quality lies within these limits, it will be assumed that the quality is controlled. In the discussion that follows, some such limits may be introduced as giving evidence of control.

In respect to the second question, evidence will be presented as we go along to indicate that quality once controlled may in general be expected to remain controlled in just the same way that we expect the continuance of control in the mortality based upon the knowledge that it has been approximately controlled in the past. In this chapter, let us consider in particular some of the specific advantages that accrue from control of manufactured product.

2. Reasons for Control Within Limits

A. To reduce the cost of inspection

If we can assure ourselves that something we use is produced under controlled conditions, then likely we will have confidence in the product and not demand much inspection. If, however, we cannot so assure ourselves and find that the previously observed variations in quality are greater than can reasonably be attributed to random fluctuations of a controlled system of chance causes, quite likely we will have some concern about what may happen in the future and therefore will demand a very extensive inspection of the product before buying it. In other words, if we have been using some particular product for several years and have found its quality to differ not more than we are

---

1. I.E.B. 1 and I.E.B. 5.

willing to attribute to chance, we gradually become more and more willing to accept the product without inspection. For example - we do not waste our money on doctors' bills so long as we are willing to attribute the variability in our health to the fluctuations of a controlled system of chance causes.

So it is in engineering. In the early stages of production there are usually causes of variability that must be weeded out. As we proceed in this way, product approaches a more stable condition, one in which it appears that the causes of variability are nearing the state of equilibrium or control. In the initial stages of production the presence of these causes makes necessary a comparatively large amount of inspection but as these causes are weeded out one by one, the need for such inspection becomes less.

Fig. 9 illustrates the way in which observed quality is often found to approach the state of equilibrium or control. Here we see how the quality of one product measured in terms of per cent defective fluctuates quite widely for a certain period of time and then as the assignable causes are weeded out, the fluctuation in this percentage becomes less and less until finally it comes within the dotted control limits shown in the figure. The same type of phenomenon is shown by the lower chart in this figure which represents the approach to control by the quality of another product measured as a variable.

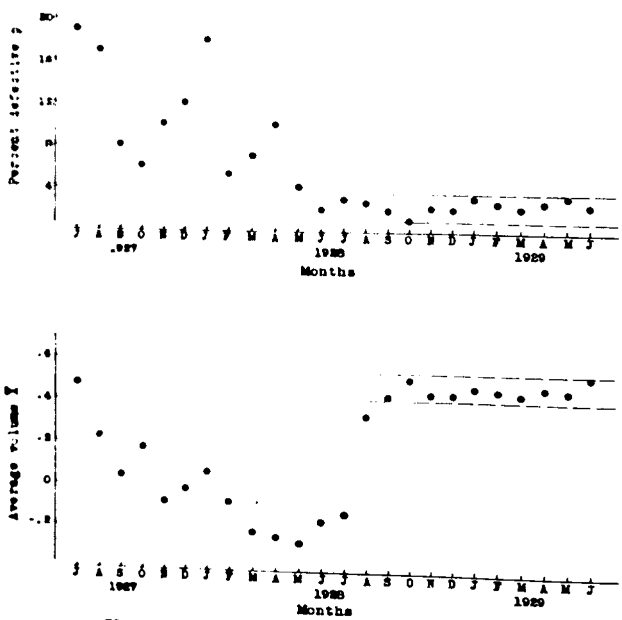


FIG. 9 - EVIDENCE OF APPROACH TO CONTROL UPON WEEDING OUT OF ASSIGNABLE CAUSES

Fig. 10 shows the results of one of the first large-scale experiments to determine whether or not we can reasonably apply the principles of control in production. It shows that we can. About thirty typical items used in the telephone plant and produced in lots running into the millions per year were made the basis for this study. As shown in this figure during 1923-24 these items showed 68% control about a relatively low average of 1.4%

defective.[1] However, as the causes of erratic deviations, that is deviations in the observed monthly per cent defective falling outside of control limits were found and eliminated, quality of product approached the state of control as indicated by an increase of from 68% to 84% control by the latter part of 1926. At the same time the quality improved; in 1923-24 the average per cent defective was 1.4% whereas by 1926 this had been reduced to .8%.

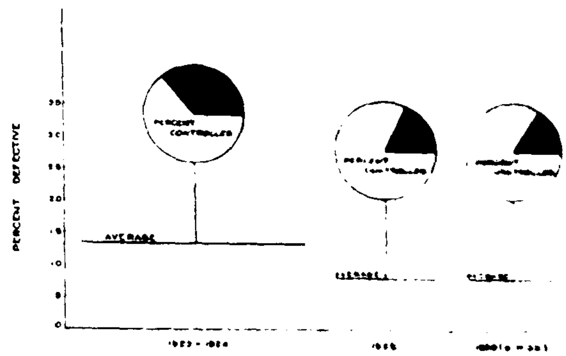


FIG. 10 - EVIDENCE OF IMPROVEMENT IN QUALITY WITH APPROACH TO CONTROL

Here we get some typical evidence that, in general, as the causes of erratic fluctuation are removed, no others enter to take their place. Instead there is evidence that the chance cause system approaches a state of equilibrium or control. This kind of definite evidence could be extended at will to indicate that when we obtain control of a manufacturing process, the product produced thereafter may be expected to fall within the control limits as determined by probability theory. In other words, having the assurance that a product is controlled at one time, we can usually rest assured that it will remain so, thus decreasing the need for inspection.[2]

One of the very common instances where advantages can be derived from a knowledge of control within limits is the inspection of raw materials. Such materials are usually secured through different suppliers. It is more or less accepted practice in many organizations to inspect more product coming from a source which has previously shown erratic fluctuations than product from a source which has not shown such erratic fluctuations. Now, our judgment of what may or may not be an erratic fluctuation is open to serious question unless

---

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

2. Of course final product is usually given some kind of service test at the time it goes into the hands of the consumer. However, a large part of the cost of inspection arises from inspections of raw material, piece-parts and assembly operations at the various stages of production.

we make use of the principles to be laid down in this and other bulletins, showing how tests can be applied to indicate whether or not a product is controlled.[1]

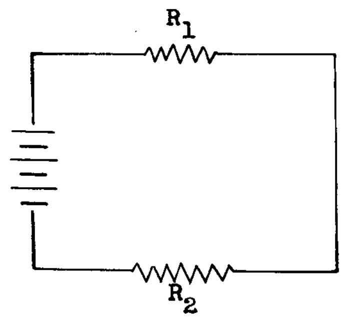
We shall now consider the second reason for control within limits.

B. <u>To insure maximum benefits from quantity production</u>

The quality of the finished product depends upon the qualities of the raw materials, upon the qualities of the piece-parts produced from these materials and upon the qualities of the assembly of piece-parts. Now, it can easily be shown[2] that so long as the above three factors influencing the resultant quality are controlled, the quality of the assembled unit or instrument will be controlled.

We have seen in the previous paragraphs how control tends to produce fewer rejections of piece-parts and it follows, along the same line of reasoning, that control of the qualities of piece-parts should lead to lower rejections of finished product. Hence we see that in a factory where each of the elements affecting the resultant quality is controlled, the resultant quality itself should also be controlled.

Let us get a little closer picture of this process by means of an example chosen because of its idealized simplicity and not because of its great economic importance. The actual problems of this nature are of greater commercial importance but they are at the same time very much more complicated. Assume that we have the simplest kind of electrical circuit involving two types of resistance coils and a battery such as shown in the adjacent sketch. Assume that it is essential that the total resistance of this circuit shall lie within prescribed limits.



It follows from what has just been said that, if the resistances of these two types are controlled, then the combined resistance will also be controlled within limits that we can determine from the known distributions of the resistances of the two types. To see how closely this can be done let us consider some experimental results. One hundred circuits

---

1. See, for example, I.E.B. 1 for certain tests of this nature.

2. See I.E.B. 6.

were to be made up from resistances selected from controlled product. Columns 1 and 2, Table 1, present the results of one such experiment, and the third column shows the combined resistances as they appeared in the completed circuits. From a knowledge of the fact that the production of each type of resistance was under control, it follows that the combined resistance should be distributed as indicated in Fig. 11. The points in this figure show how closely the observed distribution, Column 3, checks the one forecast upon the basis of the assumption of control. Here we see how control of piece-parts lead to control of the more complicated unit composed of several piece-parts.

R_1	R_2	$R_1 + R_2$
.5	10.3	10.8
.8	10.4	11.2
1.6	10.8	12.4
3.0	10.4	13.4
1.9	10.9	12.8
.	.	.
.	.	.
.	.	.
.	.	.
1.2	10.6	11.8
1.5	10.7	12.2
1.8	10.9	12.7
1.9	11.0	12.9
2.3	11.1	13.4

TABLE 1

We may look at the advantages to be gained by control of the separate steps in the fabrication process from two different angles, - that of the engineer writing the specifications and that of the producer. It is up to the specification engineer to assure himself that the specifications which he sets on the quality can economically be lived up to. It is up to the production engineer to produce the specified quality at a minimum manufacturing cost. Looked at from the viewpoint of control, the interests of the specification and production engineer are quite similar.

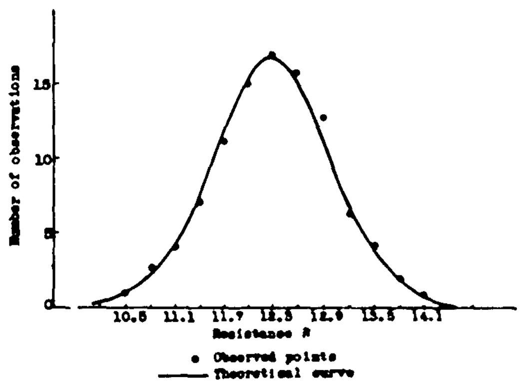


FIG. 11

From the specification viewpoint we wish to set standards and tolerances which can be met economically. Obviously, if the tolerances are too wide, unnecessary variability in product is permitted even to such an extent that trends in the quality may be overlooked. On the other hand, if the tolerances are decreased, the percentage of rejections is raised. What the design engineer usually does, therefore, is to make preliminary tests on tool-made samples and from these data try to derive satisfactory tolerances. If, however, the production process involved in making the tool-made samples is not controlled, it follows from reasons already outlined that the tolerances derived from tests made thereon will likely not prove to be satisfactory even under the assumption that the manufacturing process can be made to duplicate that used in producing the tool-made samples. It also follows that we may not expect such tolerances to be the most economical. If, however, the test results obtained on the tool-made samples indicate the experimental manufacturing process to be controlled, it is reasonable to believe that the tolerances set upon the basis of such data will prove to be economically satisfactory.

Looked at now from the production engineer's viewpoint he has only to duplicate the manufacturing process used in producing the tool-made samples provided the limits or tolerances have been established upon the basis of a controlled process of manufacture in obtaining the tool-made samples. If, however, the tolerances on resultant quality as given to the production engineer have not been obtained from a controlled experimental process, then it follows that it is to his advantage to attain control by weeding out the causes of erratic fluctuations until he has secured a controlled product.

As already stated, the above problem is ridiculously simple compared with those usually arising in practice. Thus, instead of resistances usually we must consider impedances and instead of two elements in the circuit we must consider several, including condensers, induction coils, relays, vacuum tubes, etc. From the theoretical viewpoint, however, the method of attack is identical with that employed in this very simple illustration. If we keep the qualities of each of these parts within control, including the element of assembly, the resultant quality of the complicated circuit will also be under control.

It is perhaps of interest to consider one typical question that arose in setting quality control standards for the quality of one of these complicated circuits. Here there were at least eight characteristics of the completed circuit for which it was necessary to set standards of quality, including tolerances. We shall, however, consider only one of these quality characteristics, namely, the peak voltage at a certain point in the circuit under test conditions.

Fig. 12 shows the distribution of observed peak voltages in the first 198 circuits that were set up. With this information on hand the following question was raised. What limits shall we set on peak voltage of circuits of this character and what is the probability that future circuits will have peak voltages lying within these limits?

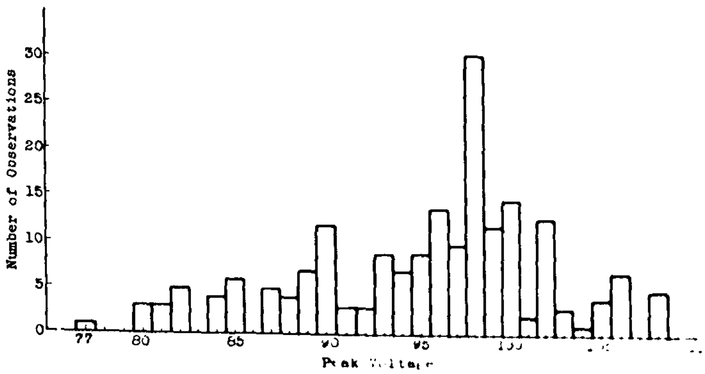


Fig. 12

A few simple tests applied to these results gave very definite indication that the quality of the first 198 circuits had not been under control for some unknown reason. In other words, these data furnish evidence of lack of constancy in the probability of the production of a circuit with peak voltage lying within a given range. Hence it is not feasible to set limits associated with a given probability because as far as we know there may have been a very definite trend present in the production process during the time that the circuits were being made. If in the absence of any further information we were to set limits on the basis of such data and use them as we would use similar limits derived from a controlled process, we might expect to find at times that there will be an unexpected number of rejections.

Both of the advantages of control thus far discussed have been economic in character. Now we come to a reason for control which is even more important than either of the previous two.

C. <u>To help insure[1] quality even though inspection test is destructive</u>

So often the quality of a material of the greatest importance to the individual is one which cannot be measured directly without destroying the material itself. So it is with the fuse that protects your home; with the steering rod on your car; with the rails that hold the locomotive in its course; with the propeller of an aeroplane, and so on indefinitely. How are we to know that a product which cannot be tested in respect to a given quality is satisfactory in respect to this same quality? How are we to know that the fuse will blow at a given current; that the steering rod of your car will not break under maximum load placed upon it? To answer such questions we have to rely upon the

---

1. Other phases of this problem will be discussed in I.E.B. 6 and I.E.B. 7.

fact that fuses of this kind previously blew at or near their rated current and that other steering rods similar to the one in your car have stood the strain placed on them in service.

In other words, we rely upon our experience under what we assume to be similar conditions to give us an idea of what we may expect to find in the future even though we appreciate full well that we do not even know the causes of variability. Thus we make certain observations of the quality of a thing under a given set of conditions and, so long as we feel that these conditions remain essentially the same, we assume that the quality of the thing tested will be approximately the same as previously observed. The theory of control makes possible a better use of our experience to determine when the fluctuations previously observed indicate lack of constancy of the chance cause system.

Imagine ourselves in the position of an engineer charged with the purchase of raw materials from different sources. Suppose, for example, that he is buying ordinary manila rope, one of the most important properties of which is tensile strength. Naturally he would be guided by some specification which may state, for example, that the tensile strength shall not be less than some predetermined value. Now, all of the rope cannot be tested and he must rely upon tests of a comparatively few samples although he knows that there is wide variability in the observed tensile strengths. He faces the responsibility of buying only satisfactory material although he knows that the raw material from one source may be entirely different from that from another. In a given case it may happen that the average of all tests on material from one source may be approximately the same as that from another, although the nature of the variability about these averages may be entirely different.

As an illustration, three characteristic conditions often observed in such studies are presented schematically in Fig. 13. The dots show the observed fluctuations in quality for successive tests on material from three different sources. If you were a purchasing engineer, from which source would you prefer to secure this particular kind of material assuming the costs to be the same? In each case the maximum observed range in quality is approximately the same. However, the quality of material from the first source appears to fluctuate in a somewhat random manner about the mean observed quality. In each of the other two cases there is a definite indication of lack of control. In the second one there appears to be a trend. In the third one there appear to be erratic fluctuations. In no one of the three cases do we know why the material fluctuates as it does but, so far as the average quality is concerned, material from all three sources is satisfactory. Which material would you select, assuming the cost to be the same?

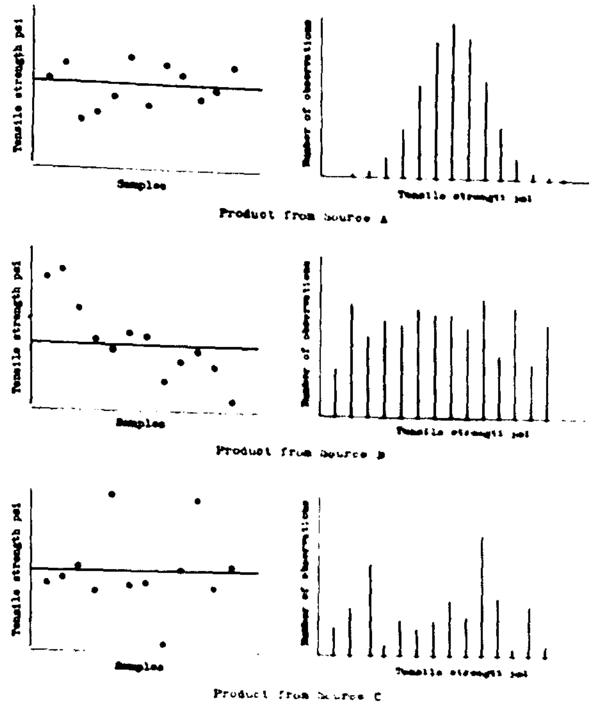


FIG. 13 - WHICH ONE WILL YOU CHOOSE?

As already stated, the quality of the material from the first source appears to deviate in a random manner about the mean value whereas the others do not. Hence if the material from the first source can be shown to give evidence of coming from a controlled system of causes and if we are willing to assume that this material will not fluctuate beyond certain predictable limits so long as we have no evidence of lack of control, then it follows that the material from the first source is perhaps the best. That is to say, although the quality of this material was observed to vary over the same range as that of the material from each of the other two sources, nevertheless, we may expect it to keep within this range in the future whereas we cannot reasonably expect this condition to be fulfilled in respect to the other two materials.

Of course, we have chosen the three illustrations, Fig. 13, such that one might be led to accept the first source of material without much further consideration. In practice, as might be expected, it is often far from being so easy to choose the sources of supply and we must have recourse to tests based upon the application of probability theory to assist us in arriving at the most reasonable choice. As we shall see in other bulletins, there are a few very simple and very useful tools based upon some of the most recent developments in the theory of mathematical statistics to help the purchasing engineer make the best choice of raw materials.[1]

So far we have looked at this problem from the viewpoint of the purchasing engineer previously supplied with tolerances on raw materials. Let us look at the same problem from the viewpoint of the specification engineer. Naturally what he would like to do is to specify that the quality of a given material should have some fixed value but, of course, he knows that such a specification cannot be met. He realizes that it is impossible to get around the fact that the qualities of most materials vary over large ranges compared with their average values. Take strength of material for example. Maximum ranges of variation are often greater than 50% of the average or expected strengths.

From what has already been said it logically follows that the specification engineer should base his specifications upon the results of tests made on the given material produced under what is apparently a controlled process. Having thus determined the process to be controlled, estimates of the average quality and the root mean square or standard deviation of this quality are the most useful statistics for specifying the characteristics of the distribution of any particular quality under controlled conditions.[2] Knowing this average and standard deviation, it is only necessary to say that the quality of the material secured from any given source shall constitute a controlled product about the specified average and standard deviation.

Now, of course, certain instances often arise where for one reason or another it is not economically feasible to secure controlled product. Then the practice to be followed is to show why the material from a given source is not controlled and then to show that to regulate these particular factors so as to secure controlled product would increase the cost more than is warranted. Under these conditions material can be accepted provided it deviates from the standards for a controlled product because of one or more kinds of reasons mentioned above. This case will be considered in greater detail later in the discussion of economic control.

---

1. I.E.B. 1 and I.E.B. 5 for example.
2. See I.E.B. 3 for a consideration of the usefulness of the average and standard deviation.

3. Reasons for Control within Minimum Limits

So far we have considered the advantages of control within limits. We have assumed that all that is wanted is reasonable assurance that future fluctuations in quality will lie within limits specified in terms of available information. But to stop here would be to overlook a very important human want, namely minimum variability whenever economically attainable. The car that ran all right yesterday but won't run today; the fluctuating voltage on your radio when you are listening to something particularly interesting; the oil burner that needs constant adjustment, are typical everyday reminders of this human want. In other words, we do not necessarily stop at merely wanting to be able to predict in terms of the past what variations may be expected in the future but further we want to minimize these variations.

However, as already stated, there is a limit to what we are willing to pay for reduction in variability. What we really want is minimum variability at a given cost. When we buy something, we want the assurance that everything feasible has been done to eliminate causes of variability. We want to be assured that the producer has done the best he can so that in popular phraseology, "Angels could do no more". In technical terms, we want to be reasonably sure of the removal of all causes of variability which can be found and removed without adding to the cost at a rate greater than that of increase in value. Let us consider a few specific objects of control within minimum limits.

A. To reduce the variability in product to economic minimum

First, put yourself in the place of the producer. Suppose you were developing an insulating material and that the resistances in megohms of the first 204 specimens were as indicated in Fig. 14. Assume that you had no a priori reason for believing that you could find and remove any one cause or group of causes of variability. Would you rest assured that you had gone as

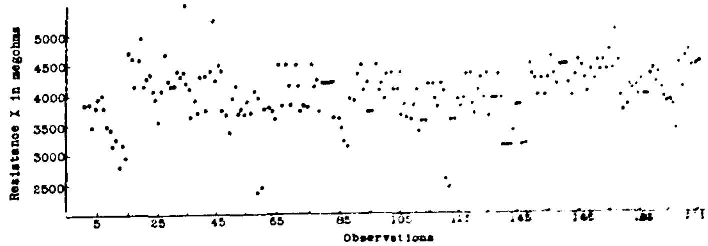
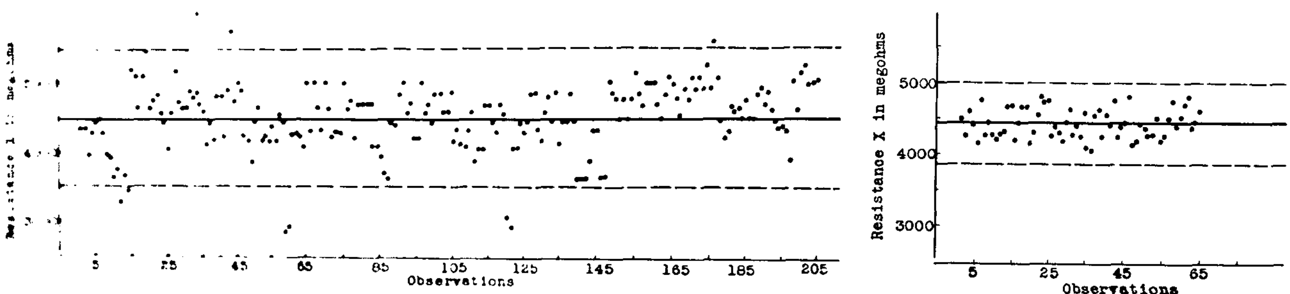


FIG. 14

far as you should try to go in reducing variability? Next put yourself in the place of the consumer. Would you be satisfied that the deviations in the observed resistances of the material were no greater than should be left to chance? In other words, would you accept this product as having been produced under maximum control?

The above questions are but typical of those arising every day in industrial research. Thus far we have not considered any detailed tests for maximum control. Nevertheless, it may not be out of place to review briefly the way one such test was applied in this particular case. In this way we shall see how it worked even before we consider all the reasons why it worked. It turns out that if these points do not fall within certain limits,[1] we as a producer should continue to look for causes of variability and as a consumer we should not be willing to accept the variability of product as being no greater than that which must be left to chance.[2] Applying this test we found points falling outside the



15-a - Indicated some cause could be found.

15-b - After causes were found and removed

FIG. 15 - THE KIND OF EVIDENCE THAT SHOWS THAT CAUSES OF TROUBLE CAN BE DETECTED

limits, Fig. 15-a. Hence we could reasonably expect to be able to find and remove some of the causes of variability.

The engineer in charge of this experimental work continued his research and six months later came back with the data shown in Fig. 15-b. In the meantime he had found and removed certain causes of variability, thus confirming the indication that such causes could be found. In fact the removal of these causes reduced the standard deviation to approximately 50% of its previous value. Apparently the engineer had gone about as far as he could expect to go for now the points fall within the limits, indicating maximum control.

1. See Modified Criterion 1 of I.E.B. 1.

2. Another way of looking at this in terms of tolerances is discussed in I.E.B. 7.

As a producer in this situation we should not expect to go much further in reducing variability without changing the whole production process. As a consumer we should be willing to leave the resultant variability to chance. Incidentally, this illustration not only shows us the advantages of maximum control but the illustration itself is typical evidence that we can approach a state of maximum control when we will to do so.

Up to this point we have been considering the advantages of control of quality of product primarily at the time of delivery to the consumer or at some intermediate stage in the production process. Important though these applications may be, we should not overlook the perhaps even greater importance of control of quality throughout the life of the product.[1] Obviously we cannot apply a life test (either artificial or field) to all product because such a test is destructive. Nevertheless, we as consumers are just about as interested in knowing that the product does not vary in respect to quality throughout life by more than can reasonably be left to chance as we are in knowing that the product at the time it reaches us satisfies these requirements.

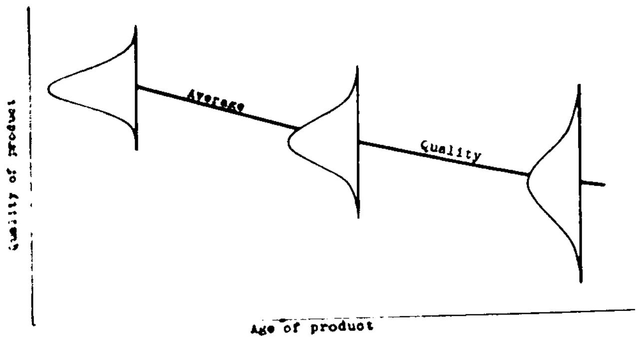
A very homely illustration will serve to illustrate this point although others could be taken from almost any field of industry. Not long ago an inquiry came from an engineer interested in the production of sheet metal watering troughs. He was very much disturbed at that particular moment because of what appeared to be an excessive number of customers' complaints. It seems that many of the customers felt that the lives of tanks which they had recently bought were far below what they should be when compared with those of other tanks which they assumed to be comparable.

Naturally in such a case a producer is very much interested in being able to detect indications of the presence of causes of erratic deviations long before it is noted on the part of the consumer. His object is, of course, to find and eliminate these causes wherever possible.

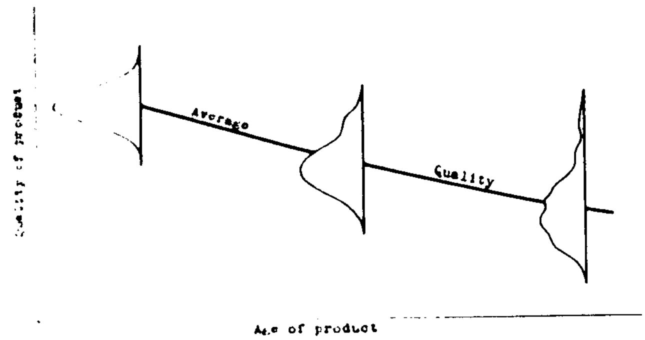
To get a little clearer picture of this kind of problem let us consider one case which may arise. Suppose the life histories appeared like those shown in Fig. 16.

- -

1. Here we may be testing lack of similarity in the conditions under which product is being used.



16-a - Product controlled throughout life



16-b - Product controlled only at time of manufacture

FIG. 16

Here the average quality and the variability of the quality (the variability being measured by the standard deviation) are supposed to change at about the same rate throughout the life of the instruments in the two conditions. The frequency curves are supposed to be the observed distributions in quality at the same time of measurement and it is assumed that each of the frequency distributions under Condition A indicates control within minimum limits where those under Condition B indicate lack of control.

In such a case the history of the two samples of the product taken under different conditions reveals that there is little to be done by way of improving the quality throughout life under Condition A whereas under Condition B there appear to be discoverable causes of variability which it should be possible to eliminate.

We have briefly touched upon some of the advantages of control. In all these cases, however, sampling inspection is assumed. It will be interesting therefore to consider one case where control may be desirable even though all product can be and is submitted to a go-no-go inspection test.

B. To cut down rejections to economic minimum

Let us consider the production of a simple mica insulating washer like one so often sees used in an assembly of electrical devices. We take this extremely simple illustration because it is so easy to visualize how the principle of control assists in reaching an economic minimum in rejections and not necessarily because the amount of money to be saved in this instance is of any important magnitude. Of course, the saving depends upon decreasing the number defective. In the present illustration, however, it is very easy to see that good or satisfactory product may be sifted from the bad or defective product simply by means of a go-no-go gauge.

Suppose that the thickness called for in the specifications is .010" ± .001" and that the first hundred washers produced under this specification gave the distribution of thickness shown in Table 2.

As was to be expected the thickness of the washers varied over quite a range. In fact, 2% of them fell below the lower limit and 27% above the upper limit whereas the average thickness was nearly a mil (.0007") greater than the design standard of ten mils. Now, if the first hundred washers are to be taken as indicative of what the future product will be, it follows that something like 29% of the product must be either junked or reclaimed. This 29% rejection, however, is high,- nearly 1/3 of the product.

Now, if we could modify the distribution of the thickness of the washers by finding and eliminating some particular cause of variability, it would be possible to cut down percentage rejection and thus cut down the cost involved in junking or reclaiming the defective material. If in such a situation we know a priori certain important causes of variability, we must consider the cost of controlling these known causes so as to reduce the number rejected. This cost must be balanced against the increased value of the product through a reduction in the per cent rejection. In the particular case in hand, however, no important cause of variability was known a priori. Therefore it was necessary to consider whether or not it was likely that such a cause could be found. In such a case two courses of action are open. One is to set about in an experimental way to find such causes of variation in the observed distribution of product. The other is to make use of the principle of maximum control to indicate whether or not it is likely that such an action on our part will lead to finding the causes sought for.

Thickness in Inches	Number of Washers
.0088	1
.0089	1
.0090	0
.0091	0
.0092	1
.0093	1
.0094	1
.0095	1
.0096	0
.0097	0
.0098	2
.0099	1
.0100	2
.0101	5
.0102	2
.0103	3
.0104	7
.0105	5
.0106	8
.0107	10
.0108	10
.0109	7
.0110	5
.0111	3
.0112	5
.0113	6
.0114	6
.0115	3
.0116	3
.0117	0
.0118	0
.0119	1

TABLE 2 - DISTRIBUTION OF THE THICKNESS OF MICA WASHERS.

To find out whether or not certain important causes of variability exist is almost always expensive and the principle of maximum control simply makes it possible for us to eliminate this kind of expense. If, through application of the principle, it appears that there are causes of this nature which can be found, the ordinary method of procedure is to find and control these so as to decrease the percentage rejected. Quite naturally where important or significant causes of variability exist they are modified only to that point where the rate of increase in cost involved in controlling the causes is just equal to the rate of increase in value of the product produced by decreasing the percentage rejection.

One may argue, however, that in practically all instances in his experience, he has been in a place where he knew apriori certain important or assignable causes of variability so that he did not need the principle of maximum control. This may be true but the point should be kept in mind, however, that even after the effects of causes known a priori to be assignable causes of variability have been eliminated there still remains a variable quality. The question is still present as to whether or not we have overlooked something that is assignable, - we cannot get away from the necessity of asking ourselves whether or not <u>we have gone as far as we can unless we have assurance that we have reached the state of maximum control</u>.

4. <u>Summary Statement of Advantages of Control</u>

We have seen that to the degree with which we can attain control we can secure the following advantages:

A. Reduction in the cost of inspection of piece-parts.

B. Automatic control of final product through control of piece-parts in assembly, thus securing advantages of quantity production and a reduction in the amount of inspection.

C. Best assurance of satisfactory quality even though inspection test is destructive, such as it must necessarily be in testing many raw materials and in all kinds of life tests.

To the extent that we can attain maximum control, we can secure the following advantages:

A. Reduction of the variability of product to economic minimum.

B. Reduction of the number of rejections to an economic minimum.

CHAPTER III

NECESSARY AND SUFFICIENT CONDITIONS FOR CONTROL

> "In contrast to the sharply defined causality which is evident in macroscopic physics, the latest theories have emphasized the indeterminate nature of atomic processes, they assume that the only determinate magnitudes are the statistical magnitudes which result from the elementary processes of physics."
>
> Arthur Haas
> Wave Mechanics and the
> New Quantum Theory

1. Statement of Conditions in General Terms

The necessary and sufficient conditions for control and maximum control are taken to be respectively:

A. The probability of the unknown chance cause system producing an item with a quality characteristic lying within a given range must be constant for the quality of each item produced. In this sense the unknown chance cause system must be constant.

B. The unknown chance cause system must be constant and must be such that it is not feasible to single out or assign any one of the component causes or groups of causes.

For example, we have seen that the constancy of the probability of death within given age limits characterizes an ideal law of mortality. Here, however, we are not willing to assume that some of the causes of death cannot be found and modified so as to change this probability. Similarly, we have considered molecular motion as an illustration of maximum control or of a phenomenon controlled by a constant system of chance causes which cannot be found and modified by human effort in the sense previously indicated.

But how are we to know that the unknown chance causes satisfy either of these conditions? To know that a cause system is constant means to know that at all times the ratio of the number of ways the system of causes could produce a unit of product with a quality lying within a specified range to the total number of ways that the cause system could produce a unit of product with a quality lying between plus and minus infinity is constant. To know that a cause system represents maximum control means to know that the previously mentioned ratio is constant and that the causes are not findable and modifiable. Now, a little consideration shows that likely we can never hope to know that a phenomenon is controlled in this rigorous fashion because we can never hope to know enough about the cause system. What we must try to do, therefore, is to express these conditions in terms of the possible effects of controlled cause systems and then set up certain ways and means of deciding whether an observed distribution of effects of some cause system resembles the distribution of effects of some controlled cause system closely enough to indicate control of one or the other of the two types.

Thus we have two problems: One is the establishment of standard theoretical distributions representing control. The other is the establishment of ways and means of determining when an observed distribution of effects looks enough like one of the standard types to indicate control. The first problem forms the subject of the remainder of this Bulletin. The second problem is treated in other[1] Bulletins of this series.

2. Specifications of Controlled Distributions of Effects

At least two ways are open to us for specifying distributions of effects to be taken as representing one or the other types of control. These are:

A. Accept as standards observed distributions resulting from cause systems previously believed to be controlled.

B. Postulate rational, controlled systems of causes and specify the distributions of effects given by these as standards.

Let us start with a consideration of the first of these two methods.

Already we have introduced the concept of distributions of mortality and molecular displacements as illustrating what we believe may happen under control and maximum control respectively. You will note, of course, that we speak of the concept of distributions. What we really mean is that this method involves the assumption that there are one or more definite laws of mortality and one or more laws of molecular displacement. Here the term law is used in the same sense as it is used in speaking of Ohm's law or any other law of physics.

---

1. In particular I.E.B. 1 and I.E.B. 5.

Quite naturally any observed distribution given by a controlled system of causes can be expected to differ from the law of distribution for this same system because of errors of sampling. In a similar way any observed set of values of resistance, voltage and current in a circuit can be expected to differ from Ohm's law because of errors of measurement. Of course then in all such cases a law is a conceptional reality. It is an abstract generalization assumed to represent the conditions existing in the world about us.

Starting on this line of attack, therefore, we as engineers must first provide ourselves with a list of accepted or pedigreed distribution laws representing the two types of control. Naturally then we turn to the literature of the subject of chance. We find that beginning way back two centuries ago, a few people began talking about laws of chance, that is to say, laws of distribution of effects of chance systems of causes. For example Simpson and Lagrange assumed a very simple distribution. Then came more and more complicated ones until today the literature is full of proposed laws, some of which are so general as to include practically every possible kind of distribution. This state of affairs is outlined in Fig. 17.

How then can an engineer use this first method of specifying distributions to be taken as standards of comparison or as laws typical of controlled conditions? The answer is almost obvious. He cannot logically do so although many references could be cited where this method has been applied. Furthermore this method does not lead to a differentiation between control and maximum control. To follow this method alone leads to the acceptance of almost any kind of observed distribution as indicating a controlled condition. Let us therefore turn to a consideration of the second method in the hope that it will give us a rational basis for establishing standards of control. Let us postulate a particular constant system of chance causes exhibiting control, and study the types of possible distributions of effects of such a system. Later we shall see how such a system can be modified to make it satisfy the conditions of maximum control.

3. Controlled or Constant System of Chance Causes

We shall assume that constant chance cause systems are of two classes - continuous and discontinuous. In either case we assume that there is

HOW CHANCE EFFECTS HAVE BEEN PORTRAYED.

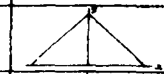
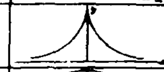
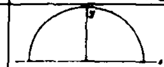
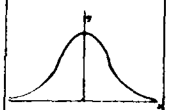
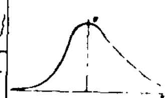
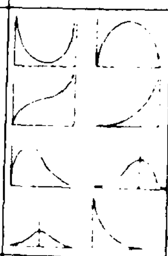
DATE	NAME	GENERAL CHARACTERIZATION OF CHANCE	PARTICULAR CHARACTERIZATION OF CHANCE	NATURE OF DISTRIBUTION OF CHANCE EFFECTS	EXPRESSION OF THE LAW OF CHANCE: ANALYTICAL FORM	EXPRESSION OF THE LAW OF CHANCE: GRAPHICAL FORM	REFERENCES
1756 1778	SIMPSON AND LAGRANGE	[illegible] ATIONS AROUND THE ARITHMETIC MEAN	1. POSITIVE AND NEGATIVE DEVIATIONS CAUSED BY CHANCE ARE EQUALLY LIKELY. 2. THE MEAN VALUE IS NEARER TO THE TRUTH THAN ANY OTHER SINGLE OBSERVATION.	SYMMETRY IN DISTRIBUTION. LIMITED RANGE OF VARIATION.	$y = \pm mx + c$		E. WHITTAKER & ROBINSON; J. KEYNES; I. TODHUNTER
1774	LAPLACE	DITTO	DITTO	SYMMETRY IN DISTRIBUTION. UNLIMITED RANGE OF VARIATION.	$y = \frac{m}{2} e^{-mx}$		[illegible]; I. TODHUNTER
1778	DANIEL BERNOULLI	DITTO	DITTO	SYMMETRY IN DISTRIBUTION LIMITED RANGE OF VARIATION	$y = \sqrt{r^2 - x^2}$		I. TODHUNTER; ARNE FISHER
1809	GAUSS	DITTO	1. PROBABILITY OF A DEVIATION FROM THE AVERAGE IS A FUNCTION ONLY OF THE MAGNITUDE OF DEVIATION. 2. THE MEAN VALUE IS MOST PROBABLE VALUE 3. DEVIATIONS ARE SMALL (THEIR CUBES ARE NEGLIGEABLE) 4. THE NUMBER OF OBSERVATIONS IS LARGE	1. SYMMETRY IN DISTRIBUTION. 2. UNLIMITED RANGE IN VARIATION. 3. OBSERVED CHANCE EFFECTS DISTRIBUTIONS IN ALL EXPERIMENTAL DATA ARE GOVERNED BY THE LAW OF CHANCE EXPRESSED IN GAUSSIAN NORMAL LAW OF ERRORS.	$y = \frac{h}{\sqrt{\pi}} e^{-h^2x^2}$		E. WHITTAKER & ROBINSON; J. BERTRAND; ARNE FISHER
1812	LAPLACE	DITTO AND CHANCE DISTRIBUTIONS OF ELEMENTS OBSERVABLE IN NATURE ARE IDENTICAL WITH APRIORI DISTRIBUTIONS OF SUCCESSES IN REPEATED INDEPENDENT TRIALS IN ARTIFICIALLY CONSTRUCTED [illegible] (URNS, COINS, DICE ETC.)	1. THE NUMBER OF ELEMENTS CONSTITUTING THE CHANCE EFFECT IS LARGE. 2. VARIATIONS IN CONSTITUENTS ARE INDEPENDENT AND OBEY THE SAME LAWS OF FREQUENCY. 3. CHANCE EFFECTS ARE AGGREGATED BY SIMPLE SUMMATION.	1. SYMMETRY IN DISTRIBUTION. 2. UNLIMITED RANGE IN VARIATION. 3. OBSERVED CHANCE EFFECTS DISTRIBUTIONS REPRESENT AN APPROXIMATION TO THE LAW OF CHANCE.	$y = \frac{1}{\sqrt{2\pi P_2}} e^{-\frac{x^2}{2P_2}}$		E. WHITTAKER & ROBINSON; F. EDGEWORTH; J. BERTRAND; [illegible]
1837 [illegible] 1849	POISSON AND QUETELET	CHANCE EFFECTS INFLUENCE ALL VARIATIONS AND CONSTANCIES IN THE UNIVERSE.	1. STABILITY OF CHANCE EFFECTS IS MANIFESTED ONLY THROUGH A LARGE NUMBER OF OBSERVATIONS. 2. SERIES OF OBSERVATIONS ARE INDEPENDENT. 3. AN ARITHMETIC MEAN IS AN ESTIMATE OF TRUE VALUE. 4. EMPIRICAL FREQUENCY IS AN ESTIMATE OF APRIORI PROBABILITY. 5. THE LARGER IS A NUMBER OF OBSERVATIONS THE BETTER ARE THESE ESTIMATES.	1. AN APPARENT CAUSAL IRREGULARITY IN OBSERVED EVENTS IS GOVERNED BY THE UNIVERSAL LAW OF LARGE NUMBERS. 2. SYMMETRICAL AS WELL AS ASSYMETRICAL DISTRIBUTIONS OF FINAL EFFECTS OF CHANCE.	$y = \frac{1}{\sqrt{2\pi P_2}} e^{-\frac{x^2}{2P_2}} = \varphi(x)$ $y = \frac{e^{-m} m^x}{x!} = \psi(x)$		ARNE FISHER; J. KEYNES; [illegible]; [illegible]
1879	GRAM	THE LAWS OF CHANCE EFFECTS DISTRIBUTIONS MAY BE IDENTIFIED WITH [illegible] REPRESENTABLE BY SERIES OF TERMS DERIVED FROM CERTAIN ASSUMED PRIMARY FUNCTION (GENERATING FUNCTIONS)	GENERATING FUNCTION IS NORMAL LAPLACIAN FUNCTION.	1. SYMMETRICAL } 2. ASSYMETRICAL } DISTRIBUTIONS 1. LIMITED } 2. UNLIMITED } RANGES OF VARIATIONS	$y = a_0\varphi^0(x) + a_1\varphi^1(x) + a_2\varphi^2(x) \dots$ $\varphi^0 = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$; $\varphi^i = \frac{d^i}{dx^i}(\varphi^0)$		ARNE FISHER
1889	THIELE		DITTO		[illegible]		THIELE; ARNE FISHER
19[illegible]	CHARLIER		GENERATING FUNCTIONS ARE NORMAL LAPLACIAN FUNCTION OR POISSON'S EXPONENTIAL.		$y = C_0\psi(x) + C_1\Delta'\psi(x) + C_2\Delta^2\psi(x)$ $\Delta\psi(x) = \psi(x) - \psi(x-1)$ $\Delta^2\psi(x) = \Delta\psi(x) - \Delta\psi(x-1)$		CHARLIER; ARN. FISHER
1895 1916	K. PEARSON	CHANCE EFFECTS ARE REPRESENTABLE BY FREQUENCY FUNCTIONS AND DO NOT FOLLOW IN THEIR MODE OF OPERATIONS THE ARTIFICIAL SCHEMES ASSOCIATED WITH EXPERIMENTS WITH COINS, DICE, OR URNS".	1. POSITIVE AND NEGATIVE CHANCE EFFECTS ARE NOT EQUALLY PROBABLE. 2. THE NUMBER OF ELEMENTARY CAUSE GROUPS IS NOT INFINITE. 3. ELEMENTARY CAUSE GROUPS DO NOT CONTRIBUTE INDEPENDENT ELEMENTS BUT CORRELATED ONES.	1. ASSYMETRY IN DISTRIBUTION. 2. FINITE RANGE OF VARIATION. 3. NON EXISTANCE OF ONE UNIVERSAL LAW OF CHANCE.	$\frac{1}{y}\frac{dy}{dx} = \frac{a_0 + x}{c_0 + c_1x + c_2x^2}$		K. PEARSON
1914-1935	EDGEWORTH	CHANCE EFFECTS IN A MAGNITUDE ARE PRODUCED IN FORM OF FREQUENCY WITH WHICH DIFFERENT VALUES ARE ASSUMED BY THIS MAGNITUDE DEPENDING ON A NUMBER OF VARYING ELEMENTS.	1. THE LARGER IS A NUMBER OF ELEMENTS THE MORE FREELY THEIR FREQUENCY LAWS MAY DIFFER FROM NORMAL LAW. 2. AGGREGATION OF ELEMENTS IS REPRESENTABLE BY A LINEAR FUNCTION OF THEIR VALUES. 3. VARIATIONS IN VALUES OF ELEMENTS MAY BE INDEPENDENT AS WELL AS CORRELATED.	DITTO	[illegible]		EDGEWORTH
1903-1916	J. KAPTEYN	CHANCE CAUSES MUST BE CONSIDERED IN CONNECTION WITH THE MAGNITUDE OF ELEMENT AFFECTED BY CAUSES.	THE NUMBER OF CHANCE CAUSES IS FINITE. CHANCE CAUSES ARE NOT INDEPENDENT.	1. FREQUENCY CURVE REPRESENTS EVIDENCE AS TO THE CHARACTER OF CAUSES WORKING IN A GIVEN CASE. 2. AN ASSYMETRICAL FREQUENCY DISTRIBUTION ALWAYS IS GENERATED BY CERTAIN SYMMETRICAL FUNCTION.	$y = \frac{F'(x)}{\varepsilon\sqrt{\pi}} e^{-[illegible][F(x) - F(\omega) - M]^2}$		J. KAPTEYN
1910-1916	TSCHUPROW	"CHANCE EFFECTS ARE TREATED AS CHANCE VARIABLES DETERMINED BY THEIR LAWS OF DISTRIBUTIONS AND LAWS OF STOCHASTIC CONSTRAINTS.	A FUNCTION OF CHANCE EMPIRICAL DATA APPROACHES (STOCHASTICALLY) THE LIMIT SET BY A CORRESPONDING FUNCTION OF A APRIOR CHARACTERISTICS OF A GIVEN SET OF DATA.	STATISTICAL METHOD GIVES A BEST ESTIMATE OF THE COMPLEX [illegible] EFFECTS OF CHANCE CAUSES IN FORM OF APRIORI VALUES OF PARAMETERS DETERMINED FROM EMPIRICAL RESULTS.	$y = F(x, \lambda_1, \lambda_2 \dots \lambda_n)$		TSCHUPROW

FIG. 17 - PROPOSED LAWS OF CHANCE

a finite number m of independent causes,

$$C_1, C_2, \ldots, C_i, \ldots, C_m,$$

and that the resultant effect of the system is the sum of the effects of the independent causes.

For the discontinuous cause system we assume that these m causes produce effects

$$x_1, x_2, \ldots, x_i, \ldots, x_m$$

respectively, with corresponding probabilities

$$p_1, p_2, \ldots p_i, \ldots, p_m.$$

For the continuous cause system we assume that the probability of the ith cause ($i = 1,2,\ldots,m$) producing a contribution x in the interval x to x + dx is

$$f_i(x)dx.$$

But what do we mean by cause as here used? The engineer usually wants this information before he goes ahead. We shall find that there is the same kind of indefiniteness about the answer to this question here as is present in other fields of science, but here as in the other fields the concept of cause in a specific case has practical significance.

4. Meaning of Cause

Let us see just what we mean by cause in other fields. Suppose you had been sitting alongside of Newton watching the apple fall to the ground. Suppose Newton had said, - There is a cause, gravity, which makes the apple fall such that its velocity after any interval of t seconds is equal to gt where g is a constant for any particular point on the earth. Suppose you had then asked Newton, What is this cause? He might have discoursed at length on gravity, but he probably wouldn't have told you to your complete satisfaction how that thing he called gravity could reach out through what appears to be intervening space and pull that apple to the ground. No one knows even today what this thing gravity really is in this sense.

As human beings we always want a cause for everything. But nothing is more elusive than this thing we call a cause. Every cause has its cause and so on ad infinitum. We never get quite to the infinitum. In this sense there must always exist a certain amount of topsi-turviness about the world as we perceive it. All that we can do is to find certain practical rules or relationships among things we observe. In doing this we introduce a lot of terms which we can't explain in the fundamental sense, but which we use to great advantage as, for example, mass, energy, electron and so on. Nevertheless undaunted we go ahead and introduce theories as to how these things are related. But, if the truth were known, we do not even know what these things are that we talk about. We have theories of light but we can't answer, — What is light? In some ways it looks like a wave but in others, like a corpuscle. Nevertheless we make use as best we can of these theories. Without them we would be at a great loss. From the engineer's viewpoint the justification of the use of either the wave theory or the modern corpuscular theory of light is that it helps him to attain a desired end.

So too in the simple theory of control we talk about causes, but we don't know what a cause really is any more than we know what light really is. Nevertheless, when we apply control theory, it is just as easy to get a "feeling" for what we mean by cause in a specific case as it is to get a feeling for what we mean by light when we apply a theory of light.

Next let us consider the restrictions to be placed upon this constant system of chance causes to make it one exhibiting maximum control.

5. Necessary and Sufficient Conditions for Control and Maximum Control

Control: As already seen, the necessary and sufficient condition for the existence of control is that the chance cause system be constant.

Maximum Control: Now, if we have reasons to believe that there are only two or three or at most a few causes of variability, we usually expect to find them. If, however, we have reason to believe that there are many causes of variability, we are not so certain about being able to find them. For example, if we were making a series of physical measurements and found them to show wide deviations from their mean value presumably because of chance errors of measurement, it would give us considerable satisfaction to know that there were only a very few causes.

But it is not only the number of causes that is an important factor in determining whether or not a cause is findable. Obviously we must consider the comparative magnitudes of the effects of the causes. In general, we should expect that when there are a comparatively large number of causes and when no cause produces a predominating influence there should be great difficulty in finding and singling out a separate cause.

For this reason, we shall take as the necessary and sufficient conditions for maximum control for the discontinuous system

$$\text{either}\begin{cases} p_i = p_j \\ x_i = x_j \\ m \text{ large} \end{cases} \qquad \text{or} \begin{cases} p_i \neq p_j \\ x_i \neq x_j \\ m \text{ very large,} \end{cases}$$

and for the continuous system

$$\text{either}\begin{cases} f_i(x) = f_j(x) \\ m \text{ large} \end{cases} \qquad \text{or} \begin{cases} f_i(x) \neq f_j(x) \\ m \text{ very large} \end{cases}$$

Thus we have arrived through a rational process at a statement of the necessary and sufficient conditions for the two types of control which we shall now use analytically in the establishment of necessary and sufficient conditions expressable in terms of effects.

CHAPTER IV

NECESSARY AND SUFFICIENT CONDITIONS FOR CONTROL IN TERMS OF EFFECTS

> "For the physicist the definition of causality or determinism means the specification of conditions by which its existence may be experimentally established."
>
> P. Jordan
> Nature, April 16, 1927.

1. The Problem

Let us start with the specification of conditions for control within limits. Quite obviously a system of causes exhibiting this type of control may produce any kind of a distribution of effects. Hence, we cannot specify necessary and sufficient conditions in terms of the characteristics of the distribution of possible effects but, as the name of this type of control indicates, we can specify that the variation of the distribution of effects of the cause system from one time to another must remain within limits.

In other words, suppose that such a cause system produces a series of, say, n things with quality characteristics $X_1, X_2, \ldots, X_i, \ldots, X_n$, at one time and another series of the same number of objects having a similar set of values of quality characteristics at another time. These two groups of quality characteristics will in general not be the same, but so long as the cause system is constant, the differences between the two groups must remain within certain limits. By differences here, we include differences in all kinds of functions used to characterize the distribution of quality characteristics.

Now we are ready to consider the necessary and sufficient conditions for maximum control. It is obvious that the first requirement is the same as that for control within limits, but in addition to this we can now specify certain requirements to be satisfied by the possible distribution of effects of the cause system.

To make our problem specific let us look at it in this way. Any set of n observed values of some quality may be represented graphically by n points along a line, Fig. 18-a. Any set of n pairs of observed qualities may be represented by n points in a plane and so on for three or more qualities, Fig. 18. Our question now becomes: "Is there any particular spacing of these points in one, two or more dimensions which is indicative of the state of maximum control?".

We shall find that the answer to this question is: "Yes". We shall limit our discussion primarily to distributions in only one or two dimensions. We shall see to what extent smoothness, unimodality and normality of distribution of effects are necessary and sufficient conditions for maximum control.

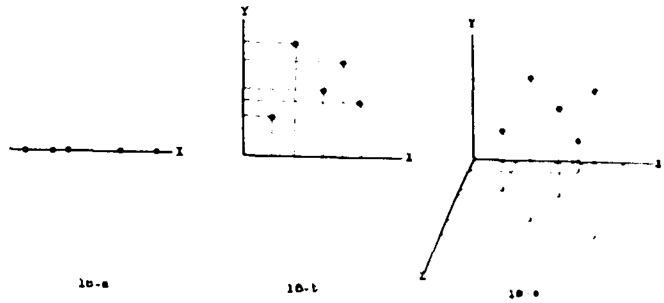


FIG. 18 - WHAT PARTICULAR SPACING BETWEEN POINTS WILL CHANCE GIVE?

2. Smoothness and Unimodality as Necessary Conditions - Simple Cause System

The simplest of the previously given cause systems exhibiting maximum control is the discontinuous one satisfying the conditions:

$$p_i = p_j$$
$$x_i = x_j$$
$$m \text{ is large},$$

where as before m is the number of causes and p_i is the probability that the ith cause will produce an effect x_i. It follows as we shall now see that the probabilities of the occurrence of a resultant magnitude X equal to 0, x, 2x, 3x,..., rx,..., m x, are given by the terms of the point binomial $(q + p)^m$.

Since X is the sum of the effects of the individual causes, its value depends upon the number of causes which produce effects or which "operate" as we shall say. Now, the probability that any one of the m causes fails to operate is $q = 1-p$.

Since the causes act independently one of another, the probability that $X = 0$ is equivalent to the probability that no one of the causes operates, and this by the multiplication theorem of independent probabilities is equal to q^m. The probability that $X = x$ is equal to the probability that any one of the m causes produces its effect and the other $m-1$ do not. This probability is evidently $mq^{m-1}p$, because we have m mutually exclusive ways in which the effect x can be produced and the probability that any one of the ways occurs is $q^{m-1}p$. In a similar way, we may proceed to find the probability that $X = 2x, 3x, \ldots, mx$.

In general, we are interested in the probability that $X = rx$ since this includes all the others as special cases. Now, the probability that any particular r of the causes (say the first r) produce their effects and the remaining m-r do not, is evidently $q^{m-r}\, p^r$. But X would have this same value no matter which r of the causes operated. In other words, to get the probability that $X = rx$, we should multiply the $q^{m-r}\, p^r$ by the number of ways in which r causes could produce their effects, the remaining m-r failing to do so. This number of ways is the total number of combinations of m things taken r at a time which is

$$C_r^m = \frac{m!}{r!\;(m-4)!}\,.$$

Hence the probability that $X = rx$ is $C_r^m\, q^{m-r}\, p^r$ which, when r varies from 0 to m, gives us the probabilities corresponding to all possible effects of the cause system and these are precisely the terms of the binomial expansion $(q + p)^m$.

Now, it becomes a simple matter theoretically to compute the terms of this expansion and plot ordinates proportional to these terms at the points 0, x, 2x,..., mx of the horizontal axis. Such a distribution is always a smooth unimodal one.

Hence smoothness and unimodality of the distribution of effects must be taken as a necessary characteristic of maximum control for a discontinuous constant system of causes.

We shall now see that the closer we approach the state of maximum control through increasing the number of causes, the closer we approach normality in the distribution of effects characterized in terms of the skewness β_1 and kurtosis β_2. In other words, the closer we approach maximum control, the closer we approach in the $\beta_1\beta_2$ plane the point previously characterized as that of maximum control, subject to certain limitations to be noted as we proceed. What we shall try to do then is to express the probability of occurrence of the resultant effect of the m causes by means of a continuous function. Next we shall consider the way in which the form of this function approaches that of the normal law under certain limiting conditions.

3. Normality as a Necessary Condition - Simple Cause System

To simplify the calculations we shall consider the effect of each cause to be unity. Then the resultant effect X of the cause system can be expressed in the form pm + x, where x now expresses the deviation of the resultant effect X from its mean value[1] pm. Then, wherever the quantity pm + x is an integer lying between zero and m inclusive, the probability P_x of getting an effect of this magnitude is equivalent to the probability that pm + x of the causes operate and the remaining qm - x do not operate.

Thus

$$P_x = C^m_{pm+x} p^{pm+x} q^{qm-x},$$

$$= \frac{m!}{(pm+x)!\,(qm-x)!}\, p^{pm+x}\, q^{qm-x} .$$

Let us express the probability P_x approximately in terms of a continuous function of x. This can be done with a sufficient degree of accuracy for most practical purposes through the use of a series expansion involving the number of causes m and the probabilities p and q.

Now

$$(pm+x)! = (pm + x)(pm + x - 1)\ldots\ldots(pm + 1)\, (pm)!,$$

and

$$(qm-x)! = \frac{(qm)!}{qm(qm-1)(qm-2)\ldots\ldots(qm-x+1)} .$$

Hence

$$P_x = \frac{m!}{(pm)!\,(qm)!}\; \frac{qm(qm-1)\ldots\ldots(qm-x+1)}{(pm+x)(pm+x-1)\ldots\ldots(pm+1)}\; q^{qm-x}\, p^{pm+x}$$

$$= \frac{m!}{(pm)!\,(qm)!}\; \frac{(1-\frac{1}{qm})(1-\frac{2}{qm})\ldots(1-\frac{x-1}{qm})(qm)^x}{(1+\frac{1}{pm})(1+\frac{2}{pm})\ldots(1+\frac{x}{pm})(pm)^x}\; q^{qm-x}\, p^{pm+x}$$

$$= C\, \frac{(1-\frac{1}{qm})(1-\frac{2}{qm})\ldots(1-\frac{x-1}{qm})(1-\frac{x}{qm})}{(1+\frac{1}{pm})(1+\frac{2}{pm})\ldots(1+\frac{x}{pm})(1-\frac{x}{qm})} ,$$

---

1. The method of obtaining the mean value pm and other moments of the binomial distribution is given in almost any elementary text in Statistics.

where
$$C = \frac{\underline{|m}}{\underline{|pm}\;\underline{|qm}}\; q^{qm}\, p^{pm}.$$

Therefore

$$\log_e\left(\frac{P_x}{C}\right) = \sum_{i=1}^{x} \log_e\left(1 - \frac{i}{qm}\right) - \sum_{i=1}^{x} \log_e\left(1 + \frac{i}{pm}\right) - \log_e\left(1 - \frac{x}{qm}\right)$$

$$= \sum_{i=1}^{x}\left(-\frac{i}{qm} - \frac{1}{2}\frac{i^2}{q^2m^2} - \frac{1}{3}\frac{i^3}{q^3m^3} - \cdots\right)$$

$$+ \sum_{i=1}^{x}\left(-\frac{i}{pm} + \frac{1}{2}\frac{i^2}{p^2m^2} - \frac{1}{3}\frac{i^3}{p^3m^3} + \cdots\right)$$

$$+ \frac{x}{qm} + \frac{1}{2}\frac{x^2}{q^2m^2} + \frac{1}{3}\frac{x^3}{q^3m^3} + \cdots,$$

provided x is less than the smaller of pm and qm.

Summing these series by columns, we have

$$\log\left(\frac{P_x}{C}\right) = -\left(\frac{1}{pm} + \frac{1}{qm}\right)\sum_{i=1}^{x} i - \frac{1}{2}\left(\frac{1}{q^2m^2} - \frac{1}{q^2m^2}\right)\sum_{i=1}^{x} i^2$$

$$- \frac{1}{3}\left(\frac{1}{q^3m^3} + \frac{1}{p^3m^3}\right)\sum_{i=1}^{x} i^3 - \cdots + \frac{x}{qm} + \frac{1}{2}\frac{x^2}{q^2m^2} + \frac{1}{3}\frac{x^3}{q^3m^3} + \cdots$$

$$= -\frac{1}{pqm}\,\frac{x(x+1)}{2} - \frac{(1-2q)}{2p^2q^2m^2}\,\frac{x(x+1)(2x+1)}{6}$$

$$- \frac{1}{3}\,\frac{(1-3pq)}{p^3q^3m^3}\,\frac{x^2(x+1)^2}{4} - \cdots + \frac{x}{qm} + \frac{1}{2}\frac{x^2}{q^2m^2} + \frac{1}{3}\frac{x^3}{q^3m^3} + \cdots,$$

after putting in the values of the sums and making use of the fact that $p + q = 1$.

Terms of the same order of magnitude may be more easily collected if we measure x in units of σ, i.e., if we let $x = z\sigma$.

Then

$$\log\left(\frac{P_x}{C}\right) = -\frac{z^2\sigma^2 + z\sigma}{2\sigma^2} - \frac{1-2q}{2\sigma^4}\,\frac{2z^3\sigma^3 + 3z^2\sigma^2 + z\sigma}{6}$$

$$- \frac{1}{3}\,\frac{1-3pq}{\sigma^6}\,\frac{z^4\sigma^4 + 2z^3\sigma^3 + z^2\sigma^2}{4} - \cdots + \frac{pz}{\sigma} + \frac{1}{2}\frac{p^2z^2}{\sigma^2} + \cdots$$

Now, we have already noted that as the number m of causes increases, the difficulty of singling out one of the causes becomes greater. In other words, the greater the number m of causes the closer we approach a characteristic condition of maximum control in terms of the distribution of effects, it being assumed throughout this discussion that the effects of the causes are equal. What we need to do then is to see what happens to our approximation as m approaches infinity.

If in the above expression, we neglect terms involving $\frac{1}{\sigma^2} = \frac{1}{pqm}$, we have

$$\log \frac{P_x}{C} = -\frac{z^2}{2} - \frac{z}{2\sigma} - z^3 \frac{(1-2q)}{6\sigma} + p\frac{z}{\sigma}$$

or

$$= -\frac{z^2}{2} - \frac{q-p}{2\sigma}\left(z - \frac{z^3}{3}\right)$$

$$= -\frac{x^2}{2\sigma^2} - \frac{k}{2}\left(\frac{x}{\sigma} - \frac{x^3}{3\sigma^3}\right).$$

Whence

$$P_x = \frac{1}{\sigma\sqrt{2\Pi}} e^{-\frac{x^2}{2\sigma^2}} \left\{1 - \frac{k}{2}\left(\frac{x}{\sigma} - \frac{x^3}{3\sigma^3}\right)\right\}, \qquad (3)$$

after putting in as before the approximate value of C. This expression we recognize as the second approximation of the previous bulletin[1].

If now for a given value of p and q, m is further increased so that we may neglect terms in the expansion involving $\frac{1}{\sigma} = \frac{1}{\sqrt{pqm}}$, then we have

$$\log \frac{P_x}{C} = -\frac{z^2}{2} = -\frac{x^2}{2\sigma^2},$$

or

$$P_x = \frac{1}{\sigma\sqrt{2\Pi}} e^{-\frac{x^2}{2\sigma^2}}$$

which is our old familiar friend, the normal law.

Hence we see that the nearer we approach the state of maximum control the nearer we approach normality. However, we need to examine this statement a little further.

Going back to our homely illustration of the variability in one's

- -

1. I.E.B. 3.

temperature, we readily admit, I think, that it would be hard to find - at least to prove beyond reasonable doubt that we have found - one of the causes when there are more than, say, a dozen. However, if when $m = 12$, p or q is very small, neither $\frac{1}{pqm}$ nor $\frac{1}{\sqrt{pqm}}$ is small and hence cannot be neglected. Hence we see that for all practical purposes we may have reached maximum control long before the condition of normality has been satisfied. This need not bother us, however, because we use normality as a criterion of maximum control simply as a limit beyond which we need not go. If we were to adopt normality alone as a criterion, we might sometimes go further than we needed to, but upon the assumptions made above we would certainly never quit looking for causes of variability before we had just reason to do so.

In this connection it may be of interest to get some practical idea as to the approach to normality. Fig. 19-a shows how closely the distribution given by a system of 16 causes producing equal effects and with p equal to q approaches normality. Similarly Fig. 19-b shows how closely the distribution given by 100 causes for which $p = 0.9$ and $q = 0.1$, approaches the second approximation.

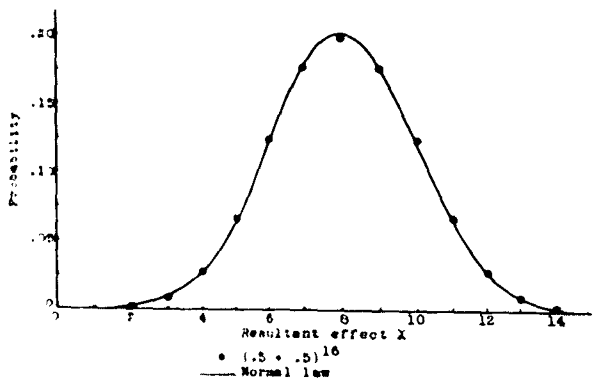


Fig. 19-a

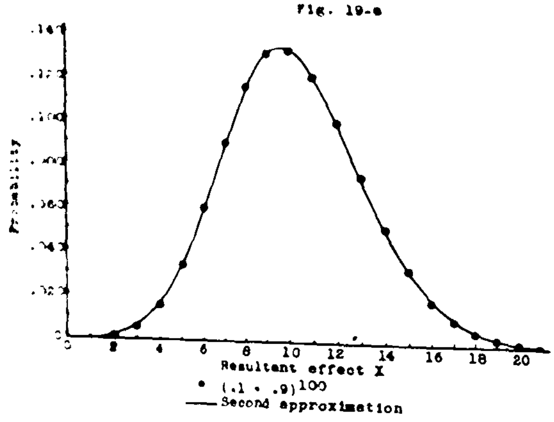


Fig. 19-b

FIG. 19

We may get at this approach to normality in a slightly different manner through the use of β_1 and β_2. It may easily be shown that

$$\sqrt{\beta_1} = k = \frac{q-p}{\sqrt{pqm}},$$

and

$$\beta_2 = 3 - \frac{1 - 6\,pq}{pqm}.$$

Fig. 20 shows how these factors approach the value 0 and 3 respectively as m approaches infinity for various values of p and q.

From what has been said we conclude that the approach to normality is quite rapid when the probability p_i of producing an effect x_i is equal to the probability q_i of not producing any

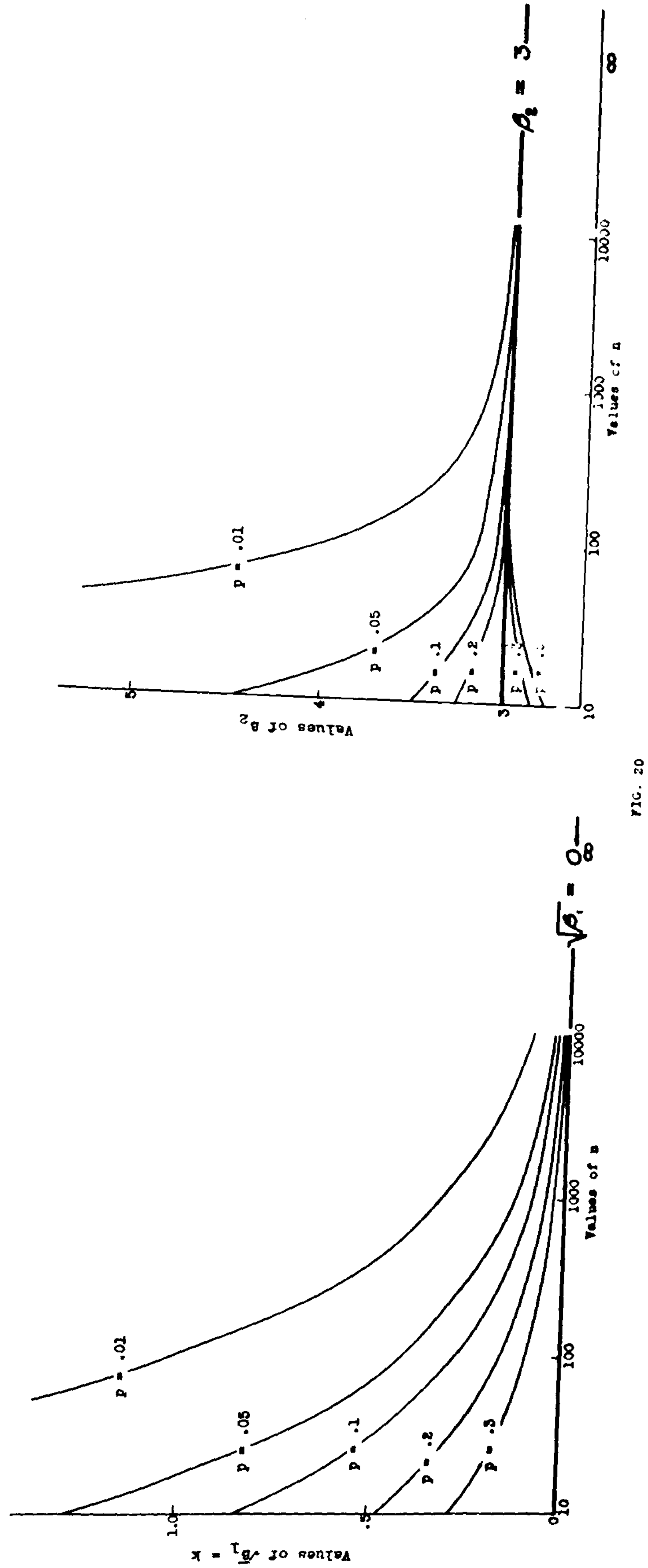

Values of B_2
Values of n
p = .01
p = .05
p = .1
p = .2
10
100
1000
10000
∞
$\beta_2 = 3$
Values of $\sqrt{B_1}$ = k
$\sqrt{\beta_1} = 0$
p = .3
0
.5
1.0

FIG. 20

effect. In this case even for a number of causes no greater than 16, the distribution of possible effects is approximately normal. When $p_i \neq q_i$ the number of causes m must be greater than in the previous case to attain the same approach to normality.

So far we have seen that, for the single discontinuous system of chance causes, smoothness and unimodality of the distribution of effects is a necessary requirement for maximum control. Furthermore we have seen that, as the number m of the causes increases, the distribution of effects approaches normality. Now, we shall consider the problem of specifying the characteristics of the distribution of effects of a continuous cause system necessary for the state of maximum control.

4. <u>Normality as a Necessary Condition - Continuous Cause System</u>

We start with the previously given necessary and sufficient conditions for maximum control in terms of causes; viz.,

$$\underset{m \text{ large}}{f_i(x) = f_j(x)} \qquad i, j = (1,2,3,\ldots, m).$$

Now, in practice, even though a cause gives a continuous distribution we can only observe a finite number of values of X differing by not less than the minimum error of measurement of the given quality characteristic. Hence in the beginning of our discussion we shall consider that any cause gives a discrete distribution of effects.

Let us assume that the resultant effect X produced by the set of m causes is the sum

$$X = x_1 + x_2 + \ldots + x_i + \ldots + x_m$$

where x_i is the component effect of the ith cause.

Our problem now becomes the specification of the distribution of resultant effects of such a system of chance causes exhibiting maximum control. We shall show that this distribution approaches normality with increase in the number m of causes irrespective of the functional form $f_i(x)$, so long as the moments of $f_i(x)$ about the expected value for this cause are finite.

To do this we shall first show that the distribution of $\frac{X}{m}$ approaches normality as m approaches infinity, but it follows from this that the distribution of X also approaches normality.

Hence, we come out with the result that the distribution of X necessary for maximum control is normal.

Now, in practice the number m of causes may be only moderately large and yet the cause system may be in a state of maximum control in the sense that it is not feasible to find and eliminate one of the causes. For this reason, a few experimental results are later introduced to show that even though we consider normality to be a necessary condition for maximum control, quite rigorously true when m is large, we can still count on this requirement to indicate many cases where maximum control is practically reached even though the number m of causes is not great. We are now ready to show that the distribution of the mean $\frac{X}{m}$ approaches normality as m approaches infinity.

In accordance with what was said at the beginning of this section in respect to practical continuity, we assume that a cause $f_i(x)$ can give rise to s different values of x

$$x_1, x_2, \ldots, x_i, \ldots, x_s$$

with corresponding probabilities

$$p_1, p_2, \ldots, p_i, \ldots, p_s,$$

where of course

$$\sum_{i=1}^{s} p_i = 1.$$

Now, clearly the distribution of the resultant effect X of the m identical causes of the system is the same from a probability viewpoint as that of the sum of m effects from one cause. Hence we shall first show that the distribution of the mean $\bar{x}$ of m effects from one cause approaches normality and then that the distribution of the sum of these m effects also approaches normality.

To this end we show that $\beta_{1_{\bar{X}}}$ and $\beta_{2_{\bar{X}}}$ of averages $\bar{X}$ produced by any one of the causes $f_i(x)$ approach, as m becomes large, 0 and 3 respectively or the values of these β's for the normal law.

Let the absolute frequency of

$$x_1, x_2, \ldots, x_s$$

obtained in a set of m effects of the cause system be

$$f_1, f_2, \ldots, f_s,$$

where of course some of the f's may be zero.

Then our problem is to investigate the distribution of

$$\bar{x} = \sum_{i=1}^{m} \frac{f_i x_i}{m}$$

in different possible sets of m effects.

Denote the average and higher moments of the distribution of effects x of the cause system by

$$\bar{x}', \mu_2, \mu_3, \ldots\ldots, \mu_i, \ldots$$

and of the average $\bar{x}$ by

$$M_1, M_2, M_3, \ldots\ldots, M_i, \ldots$$

As stated above we are concerned with β_1 and β_2 of $\bar{x}$, and since these quantities are certain functions of the moments of $\bar{x}$, our first task is to find an expression for these moments. What we shall do is to express the M's in terms of the μ's, which for the given postulated system of causes are constant.

Denote the expected value of $\bar{x}$ by the symbol $E(\bar{x})$. Then by definition

$$M_1 = E(\bar{x}) = E\left(\frac{x_1+x_2+\ldots\ldots+x_m}{m}\right) = \frac{1}{m}\left[E(x_1)+E(x_2)+\ldots+E(x_m)\right] = \bar{x}',$$

since the expected value of each x is $\bar{x}'$. Hence in finding the higher moments of $\bar{x}$ we may replace M_1 by $\bar{x}'$.

Romanovsky has developed a very elegant and simple way of obtaining these moments, as follows:

Consider the function of t defined by

$$U = \left(\sum_{i=1}^{s} p_i \; e^{\frac{t}{m}(x_i-\bar{x}')}\right)^m$$

$$= \left(p_1 e^{\frac{t}{m}(x_1-\bar{x}')} + p_2 e^{\frac{t}{m}(x_2-\bar{x}')} + \ldots + p_s e^{\frac{t}{m}(x_s-\bar{x}')}\right)^m$$

By the multinomial theorem we have

$$U = \frac{\Sigma \lfloor m}{\lfloor f_1 \lfloor f_2 \ldots\ldots \lfloor f_s} \left(p_1 e^{\frac{t}{m}(x_1-\bar{x}')}\right)^{f_1} \left(p_2 e^{\frac{t}{m}(x_2-\bar{x}')}\right)^{f_2} \ldots\ldots \left(p_s e^{\frac{t}{m}(x_s-\bar{x}')}\right)^{f_s}$$

$$= \frac{\Sigma \lfloor m}{\lfloor f_1 \lfloor f_2 \ldots\ldots \lfloor f_s} p_1^{f_1} \; p_2^{f_2} \ldots\ldots p_s^{f_s} \; e^{\frac{t}{m}\Sigma f_i (x_i-\bar{x}')}, \qquad (4)$$

the summation being extended to all f's whose sum is m.

Now the factor

$$\frac{\lfloor m}{\lfloor f_1 \lfloor f_2 \ldots\ldots \lfloor f_s} p_1^{f_1} p_2^{f_2} \ldots\ldots p_s^{1_s},$$

is the probability of getting, in m effects of the given cause $f_1(x)$, f_1 x_1's, f_2 x_2's, ..., f_s x_s's. Or, in other words, this factor is the probability of getting an $\overline{x}$ constructed in a particular way. Also for a particular construction of $\overline{x}$, the exponent of e in (4) becomes

$$\frac{t}{m} (\Sigma f_i x_i - \overline{x}' \Sigma f_i) = t(\overline{x}-\overline{x}').$$

Hence

$$U = \frac{\Sigma \lfloor m}{\lfloor f_1 \lfloor f_2 \ldots \lfloor f_s} p_1^{f_1} p_2^{f_2} \ldots\ldots p_s^{f_s} e^{t(\overline{x}-\overline{x}')}. \qquad (4')$$

Differentiating r times with respect to t and afterwards setting t=o, we have

$$\left(\frac{d^r U}{dt^r}\right)_{t=o} = \frac{\Sigma \lfloor m}{\lfloor f_1 \lfloor f_2 \ldots \lfloor f_s} p_1^{f_1} p_2^{f_2} \ldots\ldots p_s^{f_s} (\overline{x}-\overline{x}')^r, \qquad (5)$$

since each differentiation of a particular term in the sum (4') merely multiplies this term by the constant $(\overline{x}-\overline{x}')$.

By virtue of the way in which the right side of (5) has been built up, it is clear that this sum is precisely the rth moment M_r of the mean $\overline{x}$ about its mean value. The method of obtaining any moment of $\overline{x}$ is then a very simple one. Most of the details are given below.

To facilitate the work, we shall set

$$u = \sum_{i=1}^{s} p_i e^{\frac{t}{m}(x_i-\overline{x}')},$$

or

$$U = u^m.$$

The area or zero moment of $\overline{x}$ is then

$$(U)_{t=o} = (p_1+p_2+ \ldots\ldots + p_s)^m = 1.$$

since $$\sum_{i=1}^{s} p_i = 1.$$

$$M_1 = \left(\frac{dU}{dt}\right)_{t=o} = \left(mu^{m-1}\frac{du}{dt}\right)_{t=o} = \left(mu^{m-1}\frac{1}{m}\sum_{i=1}^{s} p_i(x_i-\bar{x}')e^{\frac{t}{m}(x_i-\bar{x}')}\right)_{t=o}$$

$$= \sum_{i=1}^{s} p_i\,(x_i-\bar{x}') = 0,$$

since this last sum is the first moment of the effects of the given cause system about its mean value.

In what follows it will be understood that Σ stands for the summation from 1 to s.

$$M_2 = \left(\frac{d^2U}{dt^2}\right)_{t=o} = \Bigg[mu^{m-1}\frac{1}{m^2}\Sigma\, p_i\,(x_i-\bar{x}')^2\, e^{\frac{t}{m}(x_i-\bar{x}')}$$

$$+ m(m-1)u^{m-2}\left(\frac{1}{m}\sum_{i=1}^{s} p_i(x_i-\bar{x}')e^{\frac{t}{m}(x_i-\bar{x}')}\right)^2\Bigg]_{t=o}$$

$$= \frac{1}{m}\mu_2 .$$

$$M_3 = \Bigg[\frac{u^{m-1}}{m^2}\Sigma\, p_i(x_i-\bar{x}')^3 e^{\frac{t}{m}(x_i-\bar{x}')} + \frac{3(m-1)}{m^2}u^{m-2}\left(\Sigma p_i(x_i-\bar{x}')e^{\frac{t}{m}(x_i-\bar{x}')}\right)$$

$$\left(\Sigma p_i(x_i-\bar{x}')^2 e^{\frac{t}{m}(x_i-\bar{x}')}\right) + \frac{(m-1)(m-2)}{m^2}u^{m-3}\left(\Sigma p_i(x_i-\bar{x}')e^{\frac{t}{m}(x_i-\bar{x}')}\right)^3\Bigg]_{t=o}$$

$$= \frac{\mu_3}{m^2} .$$

In an exactly similar way we shall find

$$M_4 = \frac{3(m-1)}{m^3}\mu_2^{\,2} + \frac{\mu_4}{m^3} .$$

Denoting by $\beta_{1_{\bar{x}}}$ and $\beta_{2_{\bar{x}}}$ the skewness and kurtosis, respectively, of the average, we have by definition

$$\beta_{1_{\bar{X}}} = \frac{M_3^{\,2}}{M_2^{\,3}} = \frac{\mu_3^{\,2}}{m^4}\,\frac{m^3}{\mu_2^{\,3}} = \frac{1}{m}\,\frac{\mu_3^{\,2}}{\mu_2^{\,3}}$$

$$= \frac{\beta_1}{m},$$

and

$$\beta_{2_{\bar{X}}} = \frac{M_4}{M_2^{\,2}} = \left[\frac{3(m-1)}{m^3}\,\mu_2^{\,2} + \frac{\mu_4}{m^3}\right]\frac{m^2}{\mu_2^{\,2}}$$

$$= \frac{3(m-1)}{m} + \frac{\mu_4}{m\,\mu_2^{\,2}}$$

$$= \frac{\beta_2 - 3}{m} + 3,$$

where β_1 and β_2 are the skewness and kurtosis of the distribution of effects of the given cause. Hence we see that as m becomes large, $\beta_{1_{\bar{X}}}$ and $\beta_{2_{\bar{X}}}$ approach 0 and 3, respectively, or the value of these quantities for the normal law.

Now we may go much further than this and show that as m approaches infinity the ratio of any moment of $\bar{X}$, say the hth, to the hth power of the standard deviation of $\bar{X}$ approaches, as m approaches infinity, the same value as does this ratio for the normal law[1].

It remains for us to consider the empirical information showing that the approach to normality is quite rapid. Fig. 21 shows that, at least when $f_i(x)$ is normal, rectangular or right triangular, the distribution of resultant X is approximately normal even when the number m of causes is only 4. The black dots are the observed frequencies in 1000 drawings and the solid curves are normal distributions[2]. Naturally some of the deviations from normality in this case are introduced through sampling and hence from our present viewpoint are most likely greater than they would be if we had the distribution of all possible resultant effects of 4 causes produced by these three different cause systems. Hence we have good reason to believe that, <u>so long as the condition</u>

---

1. See V. J. Romanovsky: On the Distribution of an Arithmetic Mean in a Series of Independent Trials. Bulletin of the Russian Academy of Science, 1926.

2. The drawings were made from discrete universes approximately normal, rectangular and right triangular.

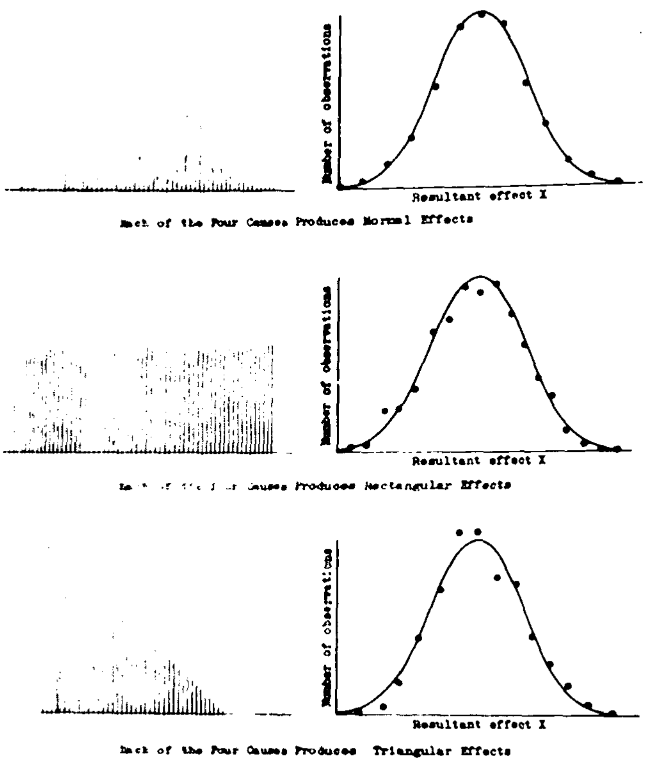


FIG. 81 - MANY CAUSES PRODUCING EQUAL EFFECTS TEND TO EXHIBIT CHARACTERISTICS OF MAXIMUM CONTROL

<u>of equal effects of the cause systems is satisfied, that is, so long as $f_i(x) = f_j(x)$, the distribution of resultant effects of m such causes will be approximately normal even for moderate values of m.</u>

Next, however, we must investigate the situation where the number of causes m is large but the causes do not produce equal effects.

5. <u>Necessary Conditions - General Cause System</u>[1]

So far we have been quite successful in translating the necessary and sufficient requirements expressed in terms of causes themselves into necessary conditions expressed in terms of effects. We have seen, in general, that smoothness, unimodality and normality play an important role in characterizing the condition of maximum control in terms of effects for those particular cause systems so far considered.

We are soon to find, however, that the problem of expressing necessary conditions for maximum control in terms of effects is far more difficult when we consider the general cause system exhibiting maximum control characterized by the conditions:

- -

1. It is realized of course that in attempting to establish a necessary and sufficient condition for maximum control in terms of effects, we have considered only the simplest hypothesis concerning the way in which the resultant effect X could be built up.

   Specifically, we have assumed that, in all cases where we are dealing with the independent causes of variation,

   $$X = x_1 + x_2 + \dots + x_i + \dots + x_m,$$

   where x_i is the effect of the ith cause. Now, of course, we may argue as follows:

   Suppose, for example, that the oranges on a tree are produced by a constant system of chance causes exhibiting maximum control and further that the distribution of diameters of these oranges is normal. Now consider the volumes of these same oranges. From one point of view, it is quite

$$x_i \neq x_j$$

$$p_i \neq p_j$$

m very large.

Now, this condition is often both necessary and sufficient; but, on the other hand, as already mentioned in the early stages of this development, the relative effect of the causes as well as the number of causes, is important when we are specifying conditions for maximum control.

In fact, even if m is very large, if one of the causes produces an effect which is larger than the resultant of all the remaining causes, it is reasonable to assume that we could find it. In such cases, then, the system cannot be thought of as exhibiting maximum control. Hence, it seems reasonable that we should add to the above conditions for maximum control the restriction that any one cause is small in comparison to the resultant of all the other causes.

With this understanding it will be seen that we may reasonably take as a necessary condition for maximum control in terms of effects that the distribution approaches normality.

We are now in a position to examine in some detail the significance of these statements.

Let us start with the above mentioned cause system. Under these conditions there is, unfortunately, no general way of finding the distribution of effects but we can derive considerable information from the results obtained in

---

reasonable to say that the same causes that produced changes in diameters should naturally produce the corresponding changes in volume. In other words, if we think of the system of causes producing a certain diameter X, then it is natural to think that these same causes produced a volume $\frac{\pi X^3}{6}$. Hence, the distribution of volumes, say X_1, will be obtained from that of the diameters by means of the transformation

$$X_1 = \frac{\pi}{6} X^3.$$

It can now be shown that the distribution of volumes X_1 is definitely not normal although it may approach normality if certain restrictions are placed on the distribution of diameters.

In a similar way, the particular characteristic in question may vary as a logarithm of X. Here also, if the distribution of X is normal, the distribution of log X will not be normal. Of course in the perfectly general case, we would have to consider the distributions of f(X) where f is a very arbitrary kind of function and we would expect to get under these conditions a very arbitrary type of distribution.

certain special cases. Fig. 22 shows the distributions of effects of typical systems of this type.[1]

The possible effects that a cause system can produce are indicated relatively along the horizontal axis. The ordinates in the figures indicate the probabilities of occurrence of the effects of the corresponding magnitudes.

The method of indicating the cause system may be illustrated for the case of Fig. 22-a. Here we have

$$m = 5$$

$$x: \; 1, 1, 1, 1, 1$$

$$p: \; \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}, \tfrac{1}{6}$$

and this means that there are five causes, each producing an effect of unit magnitude with a probability of $\frac{1}{6}$, it being understood that the probability of producing no effect is in each case $1 - \frac{1}{6} = \frac{5}{6}$. This particular system is of the type $(q + p)^m$ previously described and hence gives a smooth, unimodal distribution of effects. It is introduced here simply to serve as a basis for comparison with the five other cause systems shown in this figure.

First let us contrast the smooth distribution of effects for this simple system with the irregular one given in Fig. 22-b where the maximum effect produced by a single cause is five times that of the cause producing minimum effect. In the same way c, d, and e illustrate the way in which irregularities appear in the distribution of effects when $x_i \neq x_j$ and $p_i \neq p_j$. We should note, however, that as m becomes large, the distribution appears to smooth out and approach unimodality. We also observe that in these cases the effect of any one cause is not as large as the resultant of all the remaining causes. However, as already stated, the condition that m is very large is not a sufficient condition for maximum control in terms of causes for the case now studied.

Consider for example the cause system defined by

$$x: \; 1, 2, 4, 8, 16, \ldots, 2^{m-1}$$

$$p: \; \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \ldots, \tfrac{1}{2}$$

m very large.

---

1. The details of the method of arriving at any one of these distributions are presented in Appendix 1.

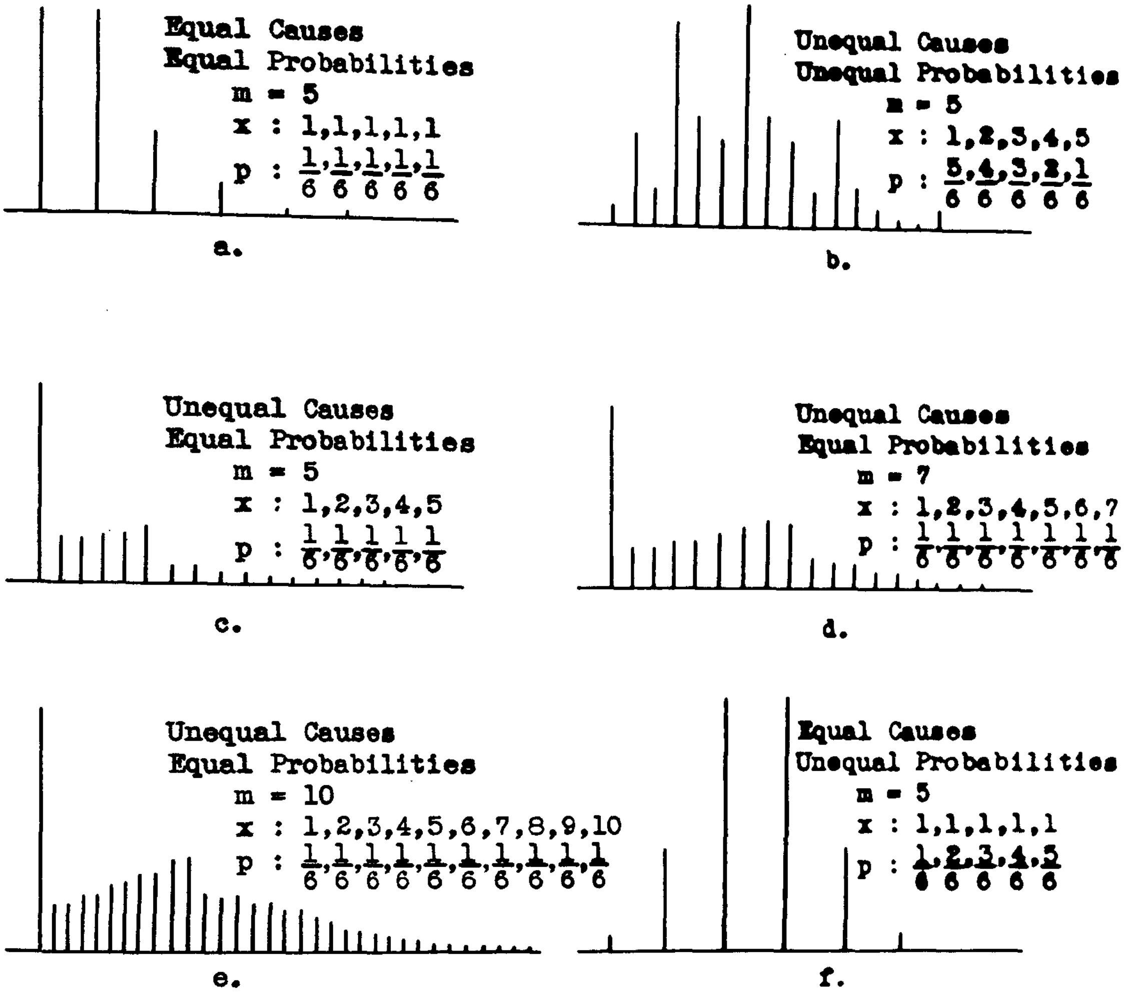


FIG. 22

The distribution of effects remains rectangular no matter how large m may be. Fig. 23 shows the distribution of effects for the special case m=5. First observe that the fifth cause is larger than the resultant of the four remaining causes. This is also true for any larger value of m. But, under these conditions, it is reasonable to believe that at least the cause

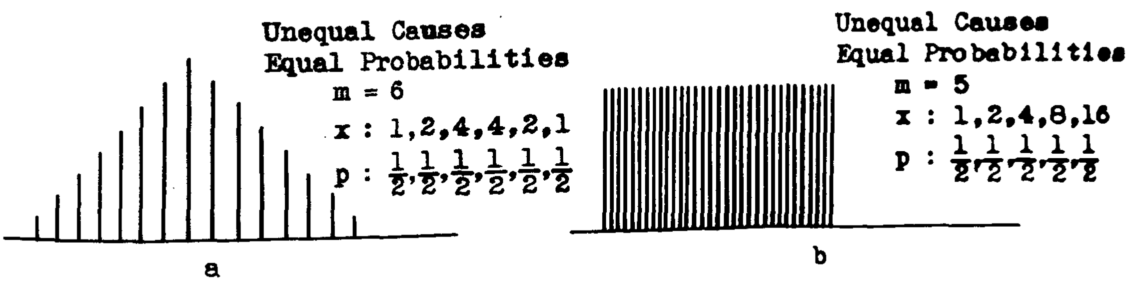


FIG. 23

producing this overwhelming effect could be found. It seems, then, that in imposing the above added restriction on this type of cause system, we have been able to state more precisely the necessary and sufficient conditions for maximum control in terms of causes.

Furthermore, the added restriction as to the causes makes normality a more general necessary condition for maximum control in terms of effects. For otherwise we would have to say that at least in some cases the necessary condition for maximum control in terms of effects is that the distribution be rectangular. Such a state of affairs is obviously undesirable in the light of results already obtained.

Passing now to the case where any cause C_i gives rise to a chance variable, we find by placing rather mild restrictions on the causes that the distribution of effects approaches normality as the number of causes becomes infinite. The nature of this restriction is that one of the causes produces an effect which is large compared to that of any other cause, but is small in comparison with the resultant of all the others.

In other words, taking the aforesaid properly restricted necessary and sufficient conditions for maximum control in terms of causes, we are able to show for this case also that the necessary condition in terms of the distribution of effects is that it approach normality, at least in the sense that β_1 approaches 0 and β_2 approaches 3.

Thus, let

$$(\mu_1)_i,\ (\mu_2)_i,\ (\mu_3)_i \text{ and } (\mu_4)_i$$

be the first four moments about the means for the ith cause

$$i = 1, 2, \ldots, m.$$

Then it can easily be shown[1] that the skewness $\beta_1 = k^2$ and degree of flatness or kurtosis β_2 of the resultant effect of the m causes are given by the expressions

$$\beta_1 = \frac{\left[\sum_{i=1}^{m} (\mu_3)_i\right]^2}{\left[\sum_{i=1}^{m} \sigma_i^2\right]^3},$$

1. Bowley A.L. "Elements of Statistics", Fourth Edition, Page 292.

and

$$\beta_2 = \frac{\sum_{i=1}^{m} \left[(\mu_4)_i - 3\sigma_i^4\right]}{\left[\sum_{i=1}^{m} \sigma_i^2\right]^2} + 3$$

respectively.

Assume now that there is one cause producing a distribution of effects much different from that of the others in the sense that its second, third and fourth moments are large in comparison with those of any other. For our present purpose we may assume that the second, third and fourth mc ents of the distributions of effects produced by the other (m-1) causes are identical one with another. If, then, we let μ_2, μ_3 and μ_4 be these moments of the distribution of effects of any one of the (m-1) causes, we may write the corresponding moments of the distribution of the remaining causes as $a\mu_2$, $b\mu_3$ and $c\mu_4$, where a, b and c are constants.

With these assumptions we get

$$\beta_1 = \frac{(m-1+b)^2}{(m-1+a)^3} \; \frac{\mu_3^2}{\sigma^6},$$

and

$$\beta_2 = \frac{(m-1+c)}{(m-1+a)^2} \; \frac{\mu_4}{\sigma^4} - \frac{3(m-1+a^2)}{(m-1+a)^2} + 3.$$

Clearly these expressions for β_1 and β_2 approach the values 0 and 3 respectively or the point of maximum control as m becomes infinite.

6. Sufficient Conditions

From a practical viewpoint, of course, what we really want is a set of sufficient conditions for maximum control. But here, as so often happens elsewhere in other fields of human endeavor, we must for the time being be satisfied with something a little short of the exact thing we want. We are to find that not any one or even all of the three necessary conditions previously considered are sufficient in themselves. But happily we shall find that this situation does not impose so serious a handicap as might at first be expected.

In Fig. 23 above, two distributions of effects are given which satisfy the conditions of smoothness and unimodality and yet in each case one cause produces a relatively large effect compared with the resultant effect of all the others and, since there are only a few causes, it is very likely that the cause producing the largest effect could be found, so that the system does not satisfy the condition of maximum control even though the distribution of effects is smooth and unimodal. Of course, however, these distributions, even when the number m of causes is very large, would be excluded on the score that they did not approach normality.

Now we shall consider one case where all three conditions, viz., smoothness, unimodality, and normality are satisfied in a practical sense and yet the cause system is not in a state of maximum control. Suppose that product is manufactured by two different processes where both show maximum control in respect to their own average qualities $\overline{X}_1$ and $\overline{X}_2$. For simplicity, let us assume that the standard deviation of each of the controlled processes is equal to σ. The accompanying Fig. 24 shows how closely the resultant distribution of product of the two machines (dots) approaches normality (solid line) even when the normal distributions of the two processes are separated by a comparatively large amount, in this case 1.5σ.

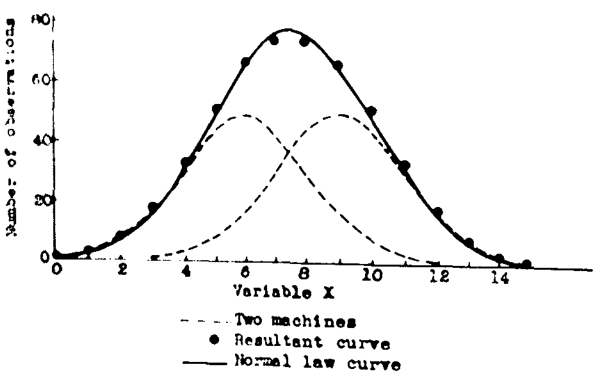


FIG. 24 - UNIMODALITY AND NORMALITY NOT SUFFICIENT CONDITIONS FOR LACK OF CONTROL

Here, however, the cause system obviously shows lack of constancy (in other words, lack of control within limits), which condition can usually be detected by sampling methods already referred to. Hence this cause for lack of sufficiency is not serious. Can we say then that for all constant systems of chance causes, the above three conditions in respect to distribution of effects are sufficient to indicate maximum control? From what has been said, it seems quite likely from a practical viewpoint that the answer to this question is yes, although it must be appreciated that this answer rests upon grounds which are far from rigorous in the mathematical sense.

7. Conditions for Control - Two or More Quality Characteristics

Now, let us consider very briefly the problem of specifying the necessary and sufficient conditions for control in terms of the distribution of effects when more than one quality characteristic is involved. A specific example would be that of the specification of the required characteristics for the distribution of resistance, capacity and inductance of cable, or that of the hardness, tensile strength and density of nickel silver sheet.

What we have said in respect to one quality characteristic holds for each of the several characteristics considered separately. In other words, smoothness, unimodality and normality are, in the sense above outlined, necessary conditions for maximum control for each of the characteristics. It follows that in general the frequency distribution in two or more dimensions will also possess these properties.

Under the special condition of maximum control, it follows, therefore, that the correlation coefficient is a measure of the commonness of causation. This is a very important result because it gives a possible physical interpretation to this correlation coefficient under the special condition of maximum control. For example, it is shown in Appendix 2 that, if ℓ of the causes are common to each of two variables where there are $\ell + s$ causes of variability in one and $\ell + n$ causes of variability in the other

$$r = \frac{\ell}{\sqrt{\ell + s}\,\sqrt{\ell + n}} .$$

8. Summary

We are now in a position to set down certain conditions which, at least from a practical viewpoint, may be taken as the necessary and sufficient conditions expressed in terms of effects for the two types of control, namely, control within limits and maximum control. These conditions are:

A. For control within limits the cause system must be constant. This condition is both necessary and sufficient.

B. For maximum control the necessary conditions are:-

1. Constant cause system.
2. Smooth distribution of effects.
3. Unimodal distribution of effects.
4. Approximately normal distribution of effects.

These four conditions taken together are sufficient for maximum control, subject to the limitations herein set forth.

CHAPTER V

THE LAW OF LARGE NUMBERS BASIC TO CONTROL

"When numbers are large, chance is the best warrant for certainty."

A. S. Eddington
The Nature of the Physical World.

1. Why a Law is Needed

In the first two chapters we got an occasional glimpse of the existing evidence showing that things sometimes happen in nature as though they were controlled in the sense of this bulletin. At that time, however, we did not have any definite picture in mind as to what we meant by a controlled state in terms of the causes themselves. We obtained such a picture in Chapter III. Then in Chapter IV we undertook to provide a tool whereby we could actually test whether or not controlled states exist in nature. This we did by finding the necessary and sufficient conditions for control expressed in terms of the distribution of possible effects of the causes.

However, as you will recall, the statement of the necessary and sufficient conditions was in terms of the effects of the cause system obtained by the process of letting the system operate once in each of all possible ways[1]. Now we must find out how to make use of these conditions in practice.

More explicitly, let us assume that a constant cause system operates, say n times, at random to give us as many values of the magnitude of some physical quality. No one would expect in general that the observed distribution of effects thus obtained would be identical with the distribution of possible effects. Nor would we expect two or more distributions of n observed effects to be the same. We say that such distributions exhibit sampling fluctuations.

In this chapter we shall show that certain kinds of natural chance cause systems, when allowed to operate at random for an indefinitely large number of times, appear to approach as a limit distributions of possible effects of constant cause systems. As in the case of all applied science, the evidence must be empirical. This phenomenon is, however, so thoroughly accepted by the

---

1. The meaning of this statement is clearly illustrated by the details presented in Appendix 1.

scientific world at large as to become one of the best established laws of nature. Through it - the so-called Law of Large Numbers - we can relate the probabilities in our assumed cause systems to hypothetically observable probabilities in nature.

Let us then give our attention to this very important law which relates actual observed distributions to those of the possible effects of the constant cause systems previously studied, thereby giving practical significance to our necessary and sufficient conditions for control expressed in terms of the distribution of possible effects.

2. The Law of Large Numbers

Flip a coin. Either the head or the tail must come up. Repeat the experiment again and again. There is a certain constancy in the nature of the results obtained in such an experiment which appears to be independent of whether you flip the coin or whether I flip it; whether the coin is flipped in some far off country or right at home. From every corner of the world, we get evidence of a certain constancy in the experimental results; i.e., it appears that the observed ratio of the number of times that a head comes up to the total number of throws approaches in a certain sense a constant value for a given coin.

But this kind of experience is not limited to coin throwing. In fact <u>wherever an event may happen in only one of two ways and the event is observed to happen under the same essential conditions for a large number of times, the ratio of the number of times that it happens in one way to the total number of trials appears to approach a definite limit as the number of trials increases indefinitely</u>. This statement is accepted as a law of nature and is called the Law of Large Numbers. Symbolically we may state this law in the form

$$\underset{n \to \infty}{L_s} \quad p = C$$

where L_s stands for what is termed an apparent or stochastic limit and C is a constant.

We should perhaps emphasize here the fact that this stochastic limit is different from the customary formal mathematical limit. For example, in the case of the coin, we do not reach a number n_0 of throws such that for all values of n greater than n_0, the ratio of the number of heads to the number of throws differs from some fixed value by less than some previously assigned small quantity ϵ.

Now we go one step further and call the limit of the relative frequency p, as the number of trials approaches infinity, an observed probability. Furthermore, it appears that we may use this observed probability as we did the probabilities in the constant cause systems to predict the distribution of effects once the observed probability has been determined. This we shall now see.

3. How to Use the Law of Large Numbers

We shall consider two simple illustrations in this chapter.

A. Let us go a little further with the experiment of flipping the coin. Suppose that we had made a large number N of throws in a way which, as before, we accept as random or as having been made under the same essential conditions. Let us break up the observed sequence of heads and tails into groups of m throws. Assuming a priori a probability of 1/2 in the mathematical sense used in our study of the cause system, it becomes merely a matter of elementary probability theory to calculate the relative frequencies of occurrence of 0, 1, 2, ..., m heads in m throws, for these frequencies are given by the terms of the point binomial $(1/2 + 1/2)^m$.

Suppose now that we let m = 12 and calculate the successive terms of this point binomial. We get the results shown in Column 2 of Table 3, where we have designated the mathematical probability by p' and the observed relative frequency by p. Now suppose that in the above mentioned experiment m is taken as 12. Column 3 of Table 3 shows the observed relative frequencies for 4096 sets of 12 throws[1]. Note the remarkable agreement between the observed frequencies p and p'.

Now let us go a little further. In this particular case the observed ratio of successes to total number of throws was .512. If we were to accept this ratio as the stochastic limit of the ratio or the objective probability, and if we used this probability in the point binomial, we would get again approximately the distribution we actually observed. In other words, we come to

1. These data are given by Yule in his book "An introduction to the Theory of Statistics", Eighth Edition, Page 258. He tabulates the results of throwing 12 dice 4096 times, a success being the turning up of the 1, 2, or 3 on a die. Clearly under the conditions described above for the coin tossing, the probability of getting a success on each die is the same as the probability of each coin coming up head.

see that this objective probability has in it much that is common with the a priori probability which we used in the point binomial. Without going further into this comparison at the present time, let us consider another very simple experiment.

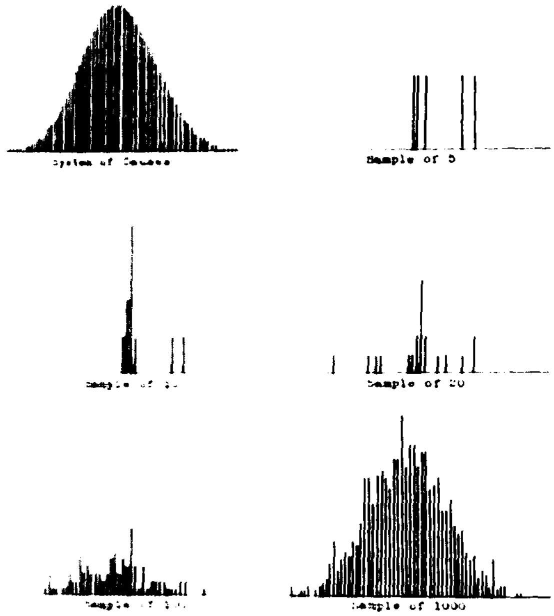
No. of Heads	Probability p' $(\frac{1}{2} + \frac{1}{2})^{12}$	Observed Relative Frequency p
0	.0002	.0000
1	.0029	.0017
2	.0161	.0146
3	.0537	.0483
4	.1208	.1050
5	.1934	.1785
6	.2256	.2315
7	.1934	.2068
8	.1208	.1309
9	.0537	.0627
10	.0161	.0173
11	.0029	.0027
12	.0002	.0000

TABLE 3

B. Suppose we take a bowl containing a large number of chips, each chip bearing a certain number. Let us assume that the relative frequency of a given number X_i is p'_i. Suppose we draw a series of n chips with replacement from this bowl, then a series of say 2n chips in the same way and then 3n chips and so on indefinitely. Now suppose that we plot the observed relative frequency distributions. It is common experience to find that the observed relative frequency p_i of the occurrence of a chip marked X_i fluctuates on the whole less and less about some fixed value as the number in the sample is increased. In other words, the distribution of observed relative frequencies appears to approach some limiting form of distribution. If the objective experimental probability of drawing a chip marked X_i is equal to the corresponding a priori probability, then the distribution of observed relative frequencies approaches that of the relative frequencies of the chips in the bowl. In the experimental results which we are about to give, it appears that this condition is fulfilled.

The following experiment of the above type was made. The distribution of chips in the bowl was normal in form as indicated in the upper left-hand corner of Fig. 25. Samples of 5, 10, 20, 100, and 1,000 were drawn with replacement. The observed relative frequency distributions of effects are shown in this same figure. In this series of distributions we note the evidence of the working of the Law of Large Numbers in the smoothing out of the distribution in the sense of the reduction, in general, of sampling irregularities. Incidentally, we note that the observed distribution of relative frequencies appears to approach in this case the distribution in the bowl.

In the consideration of these two simple experiments we have gone far enough to indicate the kind of experience generalized in the form of the Law of



[illegible] EXPERIMENTAL EVIDENCE FOR LAW OF LARGE NUMBERS

Large Numbers. Upon the basis of results such as these, we shall make the following assumption:

<u>There exist in nature systems of chance causes which operate in such a way that the probabilities reached through the stochastic limiting process are the same as those introduced in our study of constant cause systems. Stated in another way, we assume that there are discoverable[1] constant cause systems which could be used to predict within limits results to be expected in the future.</u> For the sake of clarity we shall from now on speak of the probabilities as introduced in our discussion of the cause systems as a priori and those approached in the stochastic sense as empirical. We must now examine a little further the relationship between these two.

4. <u>Empirical Versus A Priori Probability</u>

Practically we have the difficulty of finding the probabilities in the cause systems which upon the basis of evidence just presented we assume to exist in nature. It goes without saying that we can never <u>know</u> what these are any more than we can <u>know</u> anything about nature. Two courses of action are, of course, open:

A. One is to assume certain values for these probabilities in a given case.

B. The other is to take into consideration by some method or other observed experimental results in arriving at estimates of these probabilities.

Now, in the case of the coin, we are perhaps in a position to make a fairly shrewd guess as to what should be the constant value approached by the observed ratio. Here we can picture some of the causes which influence the

---

1. This means discoverable as a limiting process.

turning of the coin, and among them we do not distinguish any which could reasonably be expected to bias the result. If further the coin appears to be fairly symmetrical and homogeneous, we have no valid a priori reason for saying that it should come up head rather than tail and such considerations lead us to assume a priori probability of 1/2, which value, as we have already seen, serves to give quite a reasonable prediction of the distribution of observed frequencies. Similar arguments can be advanced in respect to the probabilities concerned in the drawing of the chips from the bowl. In the majority of cases, however, we cannot so easily arrive at what would perhaps be accepted as a reasonable guess at the a priori probabilities.

It is beyond the scope of the present bulletin to consider the generalized methods available under (B) above for finding the probabilities to insert in our postulated cause systems. However, we have gone far enough for our present purpose in laying a basis for control when we have shown the reasonableness of the assumption of the existence of these probabilities and have indicated that once these probabilities are known or assumed to be known, we can test our assumption upon the basis of the Law of Large Numbers.

5. Conclusions

A. The distribution of effects of every controlled system of causes approaches as a limit a definite functional form.

B. If the system of causes is one in which a given event may or may not happen, then the distribution of the frequencies of occurrence of this event in successive series of say n observations is of the point binomial type. It can be shown that the distribution is of the same type, if each cause may or may not produce an effect ΔX and if the limiting relative frequency of producing an effect is the same for each cause.

C. If the cause system is one in which the resultant effect X is a variable that may take on n different values $X_1, X_2, \ldots, X_n$, then the relative frequency distribution of effects of the cause system will be p_1 at X_1, p_2 at X_2, ..., p_i at X_i, ..., p_n at X_n where p_i is the limit of the observed relative frequency of occurrence of X_i as the number of observations increases indefinitely.

Now we begin to see the relationship between what we observe in nature and the postulated cause system of the two previous chapters. It is this: For every cause system obeying the Law of Large Numbers, the distribution of its observed effects approaches as a limit a definite distribution which can be produced by one or the other of the two types of constant systems of chance causes already discussed in this bulletin.

CHAPTER VI

UNIVERSAL NEED FOR CONTROL

"In short if it were not for the development of statistics, much of modern research would be impossible."

D. R. Buckingham
The Philosophy and Organization of Research, June, 1929.

"A situation like this merely means that those details which determine the future in terms of the past may be so deep in the structure that at present we have no immediate experimental knowledge of them and we may for the present be compelled to give a treatment from a statistical point of view based on considerations of probability."

P. W. Bridgman
The Logic of Modern Physics

"The future of theoretical chemistry is dependent on its (statistical theory) application and there will be a mutual and advantageous interplay in the development of these two sciences."

R. C. Tolman
Statistical Mechanics

"Independently of whether one regards the matter waves merely as probability waves or ascribes to them a physical reality, one must regard the physical happenings, when treated from a corpuscular standpoint, as undetermined. The earlier deterministic mechanics knew only the probabilities 1 and 0. It concerned itself with the study of those events whose probability is equal to 1, by means of integration of certain equations of motion with regard to exact initial conditions. The new mechanics knows no exact initial conditions. The classical alternative - either necessary or impossible - it does not know. For it anything is to be regarded as possible..... Correspondingly the new mechanics does not concern itself with the discovery of certain events having a given probability, but with the study of the probability of all possible events."

J. Frenkel
Einführung in die Wellenmechanik

1. Object of Science

For our present purpose we shall consider that the object of science is the prediction of the future in terms of the past - as it were, a search for constancy in what appears to be a world of change. The scientific procedure consists in the establishment of laws of nature. Through a knowledge of these laws, we hope to explain physical phenomena in a way to make possible the prediction of what is going to happen in the future, just as we now predict the occurrence of an eclipse of the sun or the rise and fall of the tide. In this sense, we are striving for the ideal or goal which would relate cause to effect. During the last few years we have come more and more to the feeling that for a long time to come we shall fall far short of this goal as is evidenced by the quotations at the beginning of this chapter.

In fact we are surrounded on every side with phenomena so complicated that it is almost beyond the hope of the scientists of the present day to be able to tell what is going to happen. We have, of course, the often cited example of predicting the future state of an isolated system of a gas confined within a fixed volume in terms of position and velocity of the constituent molecules at any future time. If there were only two molecules of gas within the enclosure, this would be possible, but if the volume contains the same number of molecules as a cubic centimeter of air under ordinary conditions, the problem of writing down the sets of equations which would determine this motion of the future would in itself be stupendous for it would have to take into account the positions and velocities of something like 10^{20} molecules.

Under such conditions we have, as it were, but one hope, namely that we can discover certain laws which will enable us to predict, at least within limits, something about what may be expected in the future. Without such a law, it is obvious that we cannot make any predictions whatsoever. The one and only law of this nature known to man is the one discussed in the previous chapter under the title "Law of Large Numbers". In the light of our own discussion, we can think of this as the law of control.

2. Control Basic to all Scientific Development

When one turns about in the field of science, he is astounded at the number of results which rest upon this practical principle of control expressed in terms of the Law of Large Numbers.

To start with, scientific progress depends upon the results of measurement, for what would there be to science without measurements? Can you imagine a science of physics, chemistry or engineering void of all experimental results? But, of course, measurements of a thing assumed to be constant differ among themselves because of chance causes. For example the charge on an electron is assumed to be constant but the observed values of this charge are not constant. More than this, most <u>things</u> we measure are in themselves variable under the influence of unknown or chance causes; witness, for example, pressure of a gas, rate of emission of electrons from a hot filament, physical properties of materials and so on indefinitely. In fact our assumption that anything such as the charge on an electron is a constant may be but a fantastic dream of the present generation. Be that as it may, however, it cannot be denied that all measurements are subject to the influence of chance causes; chance we have with us always.

In the last analysis, therefore, the things that go into all of our formulas of science are measurements subject to chance. Our predictions of the future in terms of the past through the use of these formulas are subject to chance. Now we have seen in previous chapters that when a chance system of causes is not controlled we can predict nothing about the resultant effect of the system. When it is controlled we can make predictions within certain limits, provided the Law of Large Numbers applies, as it usually appears to do.

<u>Hence we conclude that measurements of any kind to be of value in scientific prediction must have been taken under controlled conditions</u>.

It is one thing to talk about controlled measurements and quite another to feel it in our bones that they actually do exist. Let us therefore consider a notable illustration.

3. <u>Notable Example of Controlled Measurements</u>

It is natural that we should cast about in our search for such a series of measurements to obtain one which if possible is almost universally accepted as having been made under what were supposed to be conditions of maximum control. Perhaps the outstanding example of this character is the series of observed values of the charge on an electron obtained by Prof. Millikan. Because it was almost universally believed that he succeeded in eliminating all factors which should not be left to chance (thereby securing the conditions of maximum control), he was made the recipient of one of the highest scientific rewards - the Nobel prize.

Let us see if, in the light of our present discussion, this set of observed measurements exhibited the characteristics of maximum control. From what precedes, it is obvious that to do this we must have some criterion by which to judge whether or not the observed deviations within this set of measurements are greater than they should be provided everything had been done to eliminate those causes which should not be left to chance. One such powerful test applicable under these conditions is Modified Criterion 1 previously described in I.E.B. 1.[1] To apply this test we divide the series of observations into 14 groups of four each in the order in which they were taken, calculate the average of each group, obtain correct standard deviation based upon the standard deviations of the small samples and from this information construct a control chart as shown below in Fig. 26.

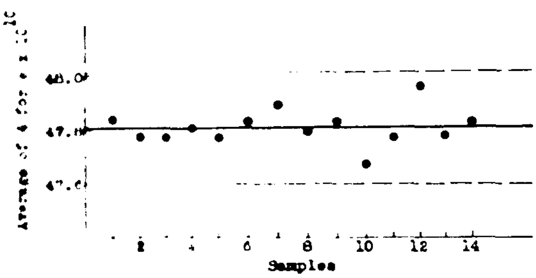


FIG. 26 - EVIDENCE THAT MEASUREMENTS CAN BE TAKEN UNDER CONDITIONS OF MAXIMUM CONTROL
Millikan's Measurements of Charge on an Electron

The fact that the observed averages, represented by the black dots in this figure, fall within the dotted control limits is consistent with the hypothesis that this series of observations was taken under the conditions of maximum control.

Naturally we may look at this result in either of two ways. First we may be willing to accept, as almost every one is, that these measurements were made under the conditions of maximum control. Then we have the conclusion that at least in this particular case our test for the conditions of maximum control gives a result consistent with our assumption of control. The other is that, if we accept the test and the philosophy underlying the test, it is reasonable to believe that the measurements in this case were made under the conditions of maximum control.

Enough has been said to show that all measurements to be of value in forecasting the future in terms of the past must be measurements representing controlled states which may be expected to continue to exist in the future. It is obvious, therefore, that in a very broad sense indeed, an engineer must be forever concerned with checking to see whether or not measurements which he uses from day to day give indication that they represent a state of control, but this is not the only broad generalization which makes it necessary for the engineer to be interested in the state of control. For example, often the very properties of matter itself with which he deals are, as we are now to see,

---

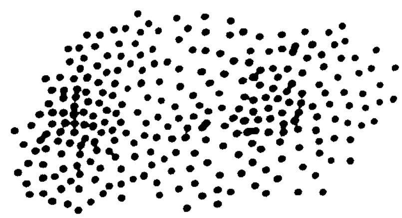
1. The philosophical basis justifying the use of this criterion will be considered in I.E.B. 5.

certain statistical averages arrived at under controlled conditions. Furthermore, we find that some of the most far reaching principles are but the description of controlled states. Such is true of the Second Law of Thermodynamics.

4. Control in Physics and Chemistry

It is far beyond the scope of the present very elementary discussion of this subject to do more than indicate the way in which we may justify some of the statements made in the previous paragraph[1]. First we shall show in what sense certain macroscopic properties of matter are the direct result of the existence of control.

Let us think of the properties of a gas. In the first place the gas itself, we believe, is made up of a large number of molecules dancing about in a manner characterized by the Brownian motion discussed in the first chapter of this bulletin. The properties of the single molecule of greatest importance are perhaps those of position, velocity, and mass. In engineering, however, we do not, in general, interest ourselves in the microscopic structure of the gas and therefore we are not for the most part interested directly in the above mentioned three properties of the molecules constituting that gas. On the other hand, the engineer deals with properties of a swarm or ensemble of molecules. In this case the most important properties are pressure, viscosity, temperature and entropy. These four properties are, however, average effects obtained under a state of control. They are, in other words, statistical properties which only have significance so long as the state of control exists.



THE VERY PROPERTIES OF MATTER OFTEN THE RESULT OF CONTROLLED STATES
A Swarm of Molecules

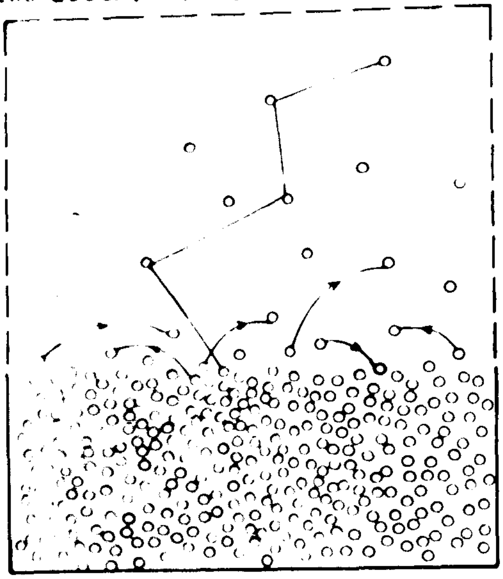
In a similar way, we can show that many of the properties of matter are but the properties of collections or swarms of molecules in conditions representing a certain state of control.

Furthermore, in engineering we often have to do with other phenomena of a statistical nature existing under controlled conditions. One such is the rate of evaporation of molecules from a liquid or solid surface. The simplest kind of case perhaps would be that of the evaporation of water molecules in an

---

1. The reader who is interested in extending his knowledge along this line is referred to any one of several important treatises on statistical mechanics.

enclosed space after it has reached equilibrium, schematically represented in the accompanying figure. This phenomenon represents a state of control in the sense that on the average the number of molecules leaving unit surface is equal, in a given interval of time, to the number of molecules condensing thereon, although at any instant these two numbers may not be identical. All that we mean is that the difference between these two numbers is on the average exceedingly small.



And so it is that we find this concept of control playing an important role in the discussion of such things as the Brownian motion, the fluctuation and density of a fluid, the distribution and velocities of electrons from hot filament, the distribution of thermal-radiation among its different frequencies, rates of diffusion and evaporation, rates of thermal and electric conduction, rate of momentum transfer, rate of thermal and photochemical reaction and so on indefinitely. Thus we get just a glimpse of the important part which the concept of control plays in physics and chemistry. We should not, however, leave this phase of the subject until we have at least one illustration showing how accurately certain physical phenomena do act as though they were controlled.

For this purpose, we shall consider the shooting off of alpha particles by polonium[1]. From the very nature of this phenomenon, it is reasonable to expect that the Law of Large Numbers will hold. In other words, it is reasonable to expect that the number of alpha particles striking the screen subtending a constant solid angle should fluctuate at random about some mean value. In other words, the observed distribution of frequencies with which 0, 1, 2,..., m particles strike the screen in a given interval of time should be given by the terms of the point binomial $N(q + p)^m$ where N represents the number of observations or intervals of time, p represents the stochastic limit or empirical probability of a particle hitting the screen and m is the maximum number of alpha particles that may strike the screen during the given interval.

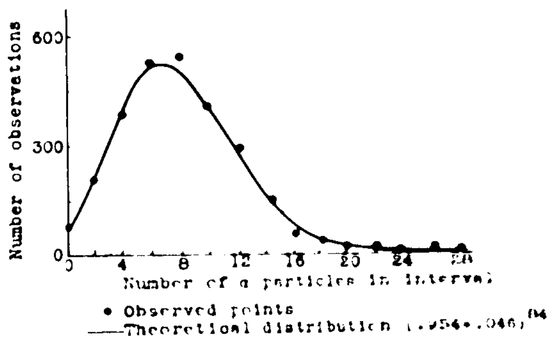
- -

1. Rutherford and Geiger, Philosophical Magazine, October, 1910.

Here, of course, we do not know either p or m except as these factors are revealed through the data. What we shall do here is to substitute the values of p and m determined from the first two moments of the observed distribution. In this way, we get from the two equations pm = 3.87 and the $\sqrt{pqm}$ = 1.92, the resulting values p = .046, q = .954, and m = 84. The observed frequency distribution is represented by the dots in Fig. 27 and the theoretical one derived upon the assumption that the phenomenon is controlled is represented by the solid curve.

In this case, we have a striking instance of the way in which the principle of control enables us to make use of data in forecasting what may be expected to happen in the future. Incidentally, these results may also be considered as an interesting illustration of the existence of control. It must be borne in mind, of course, that the factors p and m were obtained from the experimental data and hence our conclusion is limited to the statement that the observed distribution of data does not seem to differ materially from that which it would be expected to have upon the assumption that the shooting off of the alpha particles was a controlled phenomenon and that the values of p and m were equal to those observed in this experiment.

27-a - α particles being emitted under a constant system of causes



27-b - Evidence that the α particles are distributed the way we expected them to be

FIG. 27

5. Conclusions

A. All measurements to be of value in science must be taken under a state of control.

B. In general, the prediction of the future in terms of the past is possible either when the laws controlling that event in the future are known or it is known that the event is controlled by a constant system of causes.

C. Many of the properties of materials and many of the physical principles are merely statements of averages expected under controlled conditions.

D. Evidence in the field of physics and chemistry alone is sufficient to justify our firm belief in the utilitarian nature of the concept of control.

CHAPTER VII

MEASUREMENT OF LACK OF CONTROL

"When you can measure what you are speaking about, and express it in numbers, you know something about it, but when you cannot measure it, when your cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind."

Lord Kelvin

1. The Problem

All of the theory of control thus far presented would be of little value if it did not serve as a basis for the establishment of certain criteria by which to detect lack of control. That is to say, we need a yardstick to apply to observed variation in quality for the purpose of detecting the presence of assignable causes of variability. From this viewpoint, it is well to divide the assignable causes of variability into the following two types:

Type 1 - If the system of chance causes is not constant, it is said to contain an assignable cause of Type 1.

Type 2 - If the system of chance causes is constant but contains one predominating cause or group of causes, this cause or group of causes is an assignable cause of Type 2.

In terms of these definitions a controlled product is one free from the first type of assignable cause whereas a product under maximum control is free from both types of assignable causes. The problem, therefore, is to establish ways and means of detecting the presence of these causes through a kind of measurement of the variability of quality.

As already noted, the observed distribution of n effects of a controlled system differs in general from the distribution of possible effects of the cause system because of chance or sampling fluctuations. Assuming then that we know the distribution of possible effects of a cause system, how are we to tell whether an observed distribution supposedly coming from this constant system of causes differs from that of the possible effects to an extent sufficient to indicate that the cause system has changed? Perhaps the more common form in which this kind of problem arises is that where we do not know, to begin with, the distribution of possible effects and, therefore, we must determine whether differences in the distributions of observed effects in groups of measurements taken under supposedly the same conditions are sufficient to warrant the assumption that the cause system is not constant. These questions, of course, arise in connection with the detection of the presence of assignable causes of Type 1 whereas the detection of the presence of assignable causes of Type 2 must depend upon some measure of the difference between the observed distribution and that of an approximately normal one assumed to be representative of the state of maximum control in the particular case at hand.

Specific Illustration

A specific illustration of the problem of detecting the existence of assignable causes of Type 1 will now be considered. Table 4 gives the observed fractions defective found for a period of 12 months for two kinds of products designated here as Type A and Type B.

Apparatus Type A				Apparatus Type B			
Month	n No. Insp.	n_1 No. Def.	$p = \frac{n_1}{n}$ Fraction Def.	Month	n No. Insp.	n_1 No. Def.	$p = \frac{n_1}{n}$ Fraction Def.
Jan.	527	4	.0076	Jan.	169	1	.0059
Feb.	610	5	.0082	Feb.	99	3	.0303
Mar.	428	5	.0017	Mar.	208	1	.0048
Apr.	400	2	.0050	Apr.	196	1	.0051
May	498	15	.0301	May	132	1	.0076
June	500	3	.0060	June	89	1	.0112
July	395	3	.0076	July	167	1	.0060
Aug.	393	2	.0051	Aug.	200	1	.0050
Sept.	625	3	.0058	Sept.	171	2	.0117
Oct.	465	13	.0280	Oct.	122	1	.0082
Nov.	446	5	.0112	Nov.	107	3	.0280
Dec.	510	3	.0059	Dec.	132	1	.0076
Average	483.08	5.25	.0109		149.33	1.42	.0095

TABLE 4

We have for each month the sample size n, the number defective n_1 and the fraction $p = \frac{n_1}{n}$. The average fractions are $\bar{p}_A = .0109$ and $\bar{p}_B = .0095$.

Is there any indication that the product is not controlled, or in other words that the cause system is not constant? Possibly, you can answer this question better after visualizing the fluctuations in the fraction defective as shown in Fig. 28-a and Fig. 28-b. In each case, we see that there are two fluctuations out of the twelve which are large compared with the others. But are they large enough to indicate lack of control?

2. The Basis of Solution

There are two things of basic importance in our solution of this problem. They are:

A. The Law of Large Numbers showing that there is an objective or empirical probability P' that a piece of product will have a quality lying within any two fixed limits.

B. Empirical evidence justified by extensive application of the theory of control showing that it is feasible to select a value of P' which can be used economically in practice.

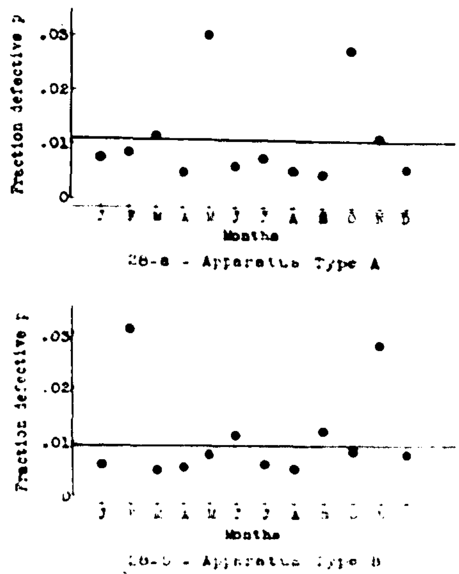


FIG. 28

Let us therefore consider the above mentioned simple problem to see wherein these two elements enter.

If the product is free from assignable causes of Type 1, then the Law of Large Numbers states that there is some value p' representing the limiting value of the observed fraction defective as the number n approaches infinity. It follows, as seen in Chapter V, that the fraction defective observed in successive samples of size n will be clustered or distributed about this objective limit p' in accord, in the long run, with the terms of the point binomial $(q' + p')^n$.

Graphically, this means that, if we take as ordinates the observed values of p and as abscissae a series of numbers corresponding to a sequence of samples of size n, then it follows from the application of the Law of Large Numbers that the observed fractions p will be distributed about the ordinate p' somewhat after the manner indicated schematically in Fig. 29 and such that when n approaches infinity, the distribution of values of p observed in the infinite sequence of samples will be some frequency curve such as indicated at the right of the figure. In other words, it follows that, if we could draw on this chart any two dotted lines parallel to the line p = p' and determine the relative number of times a value of p fell within these two limits, it would be some definite number foreordained as a consequence of the Law of Large Numbers. In other words, the fraction P' of the points falling within any pair of limits

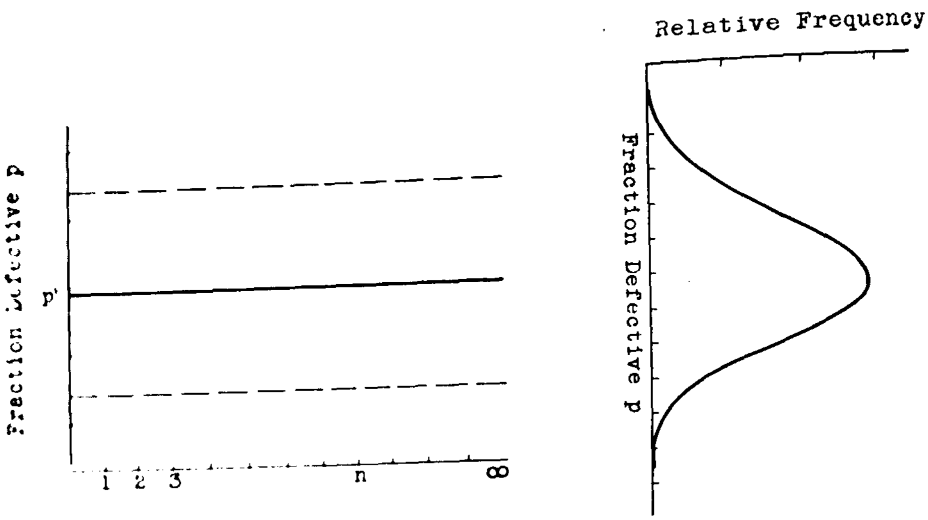


FIG. 29 - A CONTROL CHART - Representing What May Be Expected Under Controlled Conditions

based on such an infinite sequence of samples is itself an empirical probability or stochastic limit reached as a consequence of the Law of Large Numbers. Since Fig. 29 represents as it were what may be expected under controlled conditions, it has been termed a control chart.

In practice, it obviously would not be feasible to observe such a sequence even if the system of causes were constant. As mentioned in the previous chapter, we have first of all the problem of specifying p' and P' either on an a priori basis or by some method taking into account in a formal way the evidence furnished by the observed data. This point, however, we shall give no further consideration at the present moment, but pass on to another important phase of the problem.

Assume for the sake of argument that in some manner we have arrived at a control chart such as indicated in Fig. 29. In any sequence of observed values of p in this chart, since an infinite sequence is impossible, it follows from what has already been said that the observed fraction P will not in general be equal to the limiting fraction P'. Now, it is also true that any change in the constancy of the cause system also tends to change the fraction P found between any two limits. We must, therefore, set up some method of procedure whereby we shall look for assignable causes of Type 1 when the difference P - P' is sufficiently large or we must set the limits in such a way that we look for trouble every time an observed value of p is found outside these limits.

It is more or less obvious that the second method of procedure is the logical one. Choosing this method, however, we are faced with the difficulty of fixing upon the width of limits to be used in a given case. As we have already seen, an observed deviation outside any such limits does not necessarily mean the existence of an assignable cause of Type 1. Therefore, there is always a chance of looking for trouble when trouble does not exist. Similarly, the fact that the observed values of p are within these limits does not mean that there is no assignable cause of variability present. It follows that, if the limits are too narrow, one will be looking for trouble too many times when trouble does not exist. If, on the other hand, the limits are too wide, trouble may often exist without being detected.

So far as we know, there is no a priori method of establishing the width of these limits. Hence, one of the first problems which had to be considered in connection with the application of control was an empirical study of this problem of setting the limits. If it had turned out that no general methods of establishing such limits could have been justified upon an empirical or experimental basis, practically everything that we have said up to the present time would have been primarily of academic interest only. Luckily, however, experience seems to show that in most commercial problems coming to our attention it is easy to choose limits which have been quite satisfactory. In general, we have found that most problems come under one of the three cases now to be discussed.

In the laboratory where it is comparatively easy to make a search for assignable causes, it is believed that the limits corresponding to $P' = .50$ can be taken where, in general, it is understood that the limits are equally spaced on either side of the expected value. This means, however, that 50 times out of 100 one would be looking for trouble even though trouble does not exist and there are many instances where this amount of effort is not justified even in purely laboratory work. In searching for causes of variability in raw materials and production processes, it has sometimes been found feasible to use the limits corresponding to $P' \doteq .95$ or in other words, limits set a distance of $2\sigma'$ on either side of the expected value, where σ' is the expected standard deviation. For most commercial work, however, experience has shown that limits corresponding to $P' \doteq .99$ are the most satisfactory. These limits are equally spaced at a distance of $3\sigma'$ on either side of the expected value. In fact, almost all of our work has been conducted upon the basis of the last mentioned limits, and since in a majority of cases we have been able to discover a cause which when eliminated improved the quality of product, it is believed that we can consider this at least a good first step in the application of control.

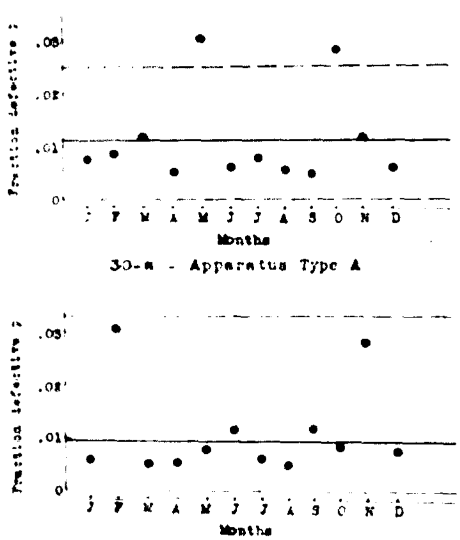
3. Application to Specific Problem

Returning to our simple problem, let us assume that

$$p'_A = \overline{p}_A ,$$

$$p'_B = \overline{p}_B ,$$

and then draw the limits corresponding to P' = .99. The results[1] are shown in Fig. 30.



30-a - Apparatus Type A

30-b - Apparatus Type B

FIG. 30

The fact that some of the points fell outside the control limits for Type A was taken as indicating lack of control, or in other words, the presence of assignable causes of Type 1. Therefore a search for such causes was instituted and some were found. On the other hand, since no points fell outside the control limits for Type B product, this product was assumed to be controlled and no such search was made for an explanation of the variability. Hence, instead of looking for lack of control in both cases, we looked for it only in one and there we found it. Now, in general, this illustrates the type of experience which leads to some of the generalizations already discussed in Chapter II.

We are ready to consider an outline of the basis for detecting lack of control in general.

4. General Method for Detecting Lack of Control

It follows from the Law of Large Numbers, along lines previously set forth, that states of control exist where the objective probability dy' of the

1. Here, of course, the sample sizes from month to month are not exactly the same and we should have variable limits. When we take this additional step however, our conclusions are not changed.

chance cause system producing a piece of product with a quality X lying within the range X to X + dX is expressable as a function φ' of the quality X and certain m' parameters represented by λ''s as follows:

$$dy'_{\lambda'} = \varphi'(X, \lambda'_1, \lambda'_2, \ldots, \lambda'_1, \ldots, \lambda'_{m'})\, dX. \qquad (6)$$

Obviously, if we knew φ' and the values of each of the m' parameters in Equation (6), it would be a simple matter to determine the probability P' that a quality X would lie within any specified limits. This would simply involve the integration of the function φ' within the corresponding limits. Where this could not be done analytically, it could be done with the necessary degree of precision for practical work through the use of a planimeter. Now we are in a position to generalize certain different phases of the problem already discussed. First we have:

A. <u>The Problem of Specification</u>

We must specify or determine the functional relationship in Equation (6). As already stated, however, the exact functional relationship in most practical problems must forever remain an unfathomable mystery. Otherwise, it would be necessary for us either to establish this specification on an a priori basis or to establish it through the analysis of an infinite series of observations. But who knows a priori exactly how any product is to be distributed or who, if he could, would be willing to take an infinite number of measurements?

These facts, however, do not in any way deter us from using the methods of control any more than the knowledge that Newton's Laws of Motion are not <u>the</u> laws of motion deters us from using Newton's Laws to solve certain problems in dynamics.

What we usually do under these conditions is to assume that Equation (6) may be represented by

$$dy'_{\theta'} = f'(X, \theta'_1, \theta'_2, \ldots, \theta'_1, \ldots, \theta'_m)\, dX \qquad (7)$$

where $dy_{\theta'}$ is assumed to be the objective probability of a unit of product having a quality X within the interval X to X + dX; f' is the assumed relationship and the θ''s represent the m parameters.

We should again emphasize the fact that φ' is likely not the same as f' in the sense that the law of motion is not Newton's Law of Motion. It follows, therefore, that θ'_i need not be the same as λ'_i and m need not be the same as m'.

In practice, what we do therefore is to specify some distribution such as represented by Equation (7) and make the assumption that it is the distribution corresponding to the given constant system of causes under study. For the case of maximum control, we shall assume as previously stated in this bulletin that the function f' is the Second Approximation Equation (3).

B. The Problem of Estimation

Obviously, the knowledge of the functional relationship now observed is not sufficient. We need to know the values of the m parameters in Equation (7). In other words, we must arrive at estimates of the parameters in terms of the data furnished by a sample or arrive at them upon some a priori basis. Such estimates we shall refer to as statistics. Hence, if we let θ_i represent the statistic for the parameter θ'_i, we may write Equation (7) in the form

$$dy_\theta = f(X_1, \theta_1, \theta_2, \ldots, \theta_i, \ldots, \theta_m)dX. \qquad (8)$$

In general, for every parameter θ'_i there may be numerous statistics $\theta_{i1}, \theta_{i2}, \ldots, \theta_{ij}$.

As an illustration, we might assume that the functional relationship in Equation (8) is normal and consider estimates of the standard deviation. Two very familiar ones are:

$$\theta_{21} = \sqrt{\frac{\pi}{2}}\,\Sigma\,\frac{|X - \overline{X}|}{n}$$

and

$$\theta_{22} = \sqrt{\frac{\Sigma(X - \overline{X})^2}{n}}$$

where the summation extends over all the X's in the sample of size n and $\overline{X}$ is the arithmetic mean. Now, it is well known that not all of the estimates of a parameter are equally efficient. Furthermore, as we shall see in Bulletin V, it is necessary that the Law of Large Numbers should lead to equality between any estimate derived from an infinite sample and the parameter itself.

C. The Problem of Distribution

It follows from what has previously been said that assignable causes of Type 1 may be detected by variability in any one of the parameters outside of certain limits. That is to say, we may establish control charts for any one of the parameters provided we know the distribution of the corresponding statistic in samples of size n. In other words, for each parameter θ'_i, we may choose two limits

$$\theta'_i \pm t\,\sigma'_{\theta_i}$$

such that the objective probability that the observed values of θ_i in samples of size n shall fall within these limits is some fixed value P'. For the reason given previously in this chapter, we generally take t = 3.

Conclusions

It follows from what has already been said that, in general, the detection of assignable causes of Type 1 may come about through the use of any one of the m parameters as soon as we have taken steps A, B, and C mentioned above. In fact, as we shall see in Bulletin V, we might set up other statistics by which to test this kind of lack of control, assuming in each case that the distribution of this statistic in samples of size n can be determined upon the basis of the knowledge of the specification as given by Equation (7) and assuming that we have evidence that the required statistics obey the Law of Large Numbers.

Now let us go one step further. Suppose that we have chosen in a given case a series of m different parameters $\theta'_1, \theta'_2, \ldots, \theta'_m$ which are to be used as a basis for detecting lack of control and suppose that we set up limits on each one of them corresponding to some probability P'. Obviously, if $m > 1$, the number of times in which we would look for trouble even though trouble did not exist would on the long run, be much larger than (1-P') of the number of observed values of the parameters. In other words, we must establish a satisfactory value of P' along lines indicated for the case m = 1.

5. Detection of Assignable Cause Type 2

Our study of the nature of the distribution of effects under the conditions of maximum control led in Chapter IV to the adoption of four conditions which for most practical problems are taken to be necessary and sufficient. The first of these was that the distribution of effects should indicate constancy of the system of chance causes. Therefore, the first step in the detection of assignable causes of Type 2 is to assure oneself through the procedure previously outlined in the present chapter that the system of causes is controlled. Then the second step is to apply some criterion to detect whether the specified distribution, in this case the Second Approximation Equation (3), differs sufficiently from the observed distribution to indicate that the condition of maximum control had not been satisfied. Criterion III of I.E.B. 1 is such a test.

6. Measure of Lack of Control - Conclusions

Enough has been said to indicate that it is feasible to set up certain criteria for detecting the presence of both types of assignable causes in the sense that it appears to be common experience that when a set of data do not satisfy these criteria, it is usually possible to find the causes of variability.

The relative frequency of occurrence of the conditions showing control is a measure of the extent to which we have eliminated all causes of variability that should not be left to chance.

CHAPTER VIII

ECONOMIC CONTROL

"It is therefore important to every technician who is dealing with problems of manufacturing control, to know the laws of statistics and to be able to apply them correctly to his problems. There is no need to fear that statements based only on probability would be too uncertain to base weighty conclusions upon them."

Becker, Plaut and Runge
Anwendungen der Mathematischen Statistik auf Probleme der Massenfabrikation.

"Statistical research is the logical method for the control of operations, for the research engineer, the plant superintendent, and the production executive."

K. H. Daeves
Author of Grundlagen und Anwendungen eines neuen Arbeitsverfahrens für die Industrieforschung mit zahlreichen praktischen Beispielen.

1. Why Control Quality

Chapter II was devoted to a consideration of five important reasons for controlling the quality of manufactured product. If only to insure that the quality of product which cannot be given a 100% inspection should lie within certain well-defined limits, the need for control would most likely be admitted by most producers, - particularly when we consider that the quality of a thing cannot be given a life test without destroying the thing itself, thus making it impossible to predict anything about the life of the quality of product unless we can be practically sure that it is in a controlled state.

The object of the present chapter, however, is to consider the need for control upon a much broader basis; viz., that it is an integral part of any program of industrial research.

In general, the object of industrial research is to establish ways and means of making better and better use of past experience. Insofar as research continues to reveal certain rules or laws which assist in the production of finished product whose quality characteristics satisfy some human need, we may expect industry to be interested in research. That industry today has such interest in this form of human endeavor, seems to be a well established fact. At least, during the year 1927, it is estimated that upwards of $200,000,000 was spent in industrial research alone in approximately 1,000 laboratories in the United States[1]. This gives us the order of magnitude of the large sum of money that is being spent annually in the effort to find out how to do something tomorrow that we do not know how to do today. All effort, however, in this direction is obviously not covered in the formal research programs of the laboratories included in the above mentioned survey.

Who has not in some way or other made use of past experience? It is rather startling indeed to see how much progress has been made in this way by that part of the human race which never had any knowledge of applied science as such. Long before any one worried over the physical principles which govern the use of the lever and that of the wedge, use had been made of both of these mechanical devices. Long before any one had arrived at the generalization known as the Law of the Conservation of Energy, our forefathers had transformed mechanical energy into heat energy to start their fires. These two illustrations are sufficient to illustrate the fact that progress in the use of past experience does not depend upon the knowledge of scientific laws as we know them today. The rate of progress on the other hand does depend upon this knowledge. In a similar way, we do not have to know the theory of control to make progress in the improvement of quality of product. But, as the physical sciences have led to useful generalizations which increase the rate of progress, so also does the knowledge of the principles of control.

To indicate the relationship which the theory of control bears to exact science, it is interesting to consider six stages in the development of better ways and means of making use of past experience. They are:

1. Belief that the future can not be predicted in terms of the past.
2. Belief that the future is pre-ordained.
3. Inefficient use of past experience in the sense that experiences are not systematized into laws.
4. Control within limits.
5. Maximum control.
6. Knowledge of all laws of nature - exact science.

---

1. Grondahl, L. O., "The Role of Physics in Modern Industry", Science, August 23, 1929, pp. 175-183.

It is a long way between this first stage where man listens to the oracles of the Gods to forecast the future to the last or sixth stage of belief admirably stated in either the words of Pope:

> "All nature is but art unknown to thee,
> All chance direction which thou canst not see",

or expressed in the form attributed to the great mathematician Laplace:

> "A perfect knowledge of the universe as had existed at any given moment, would give a perfect knowledge of what was to happen thenceforth and forever after"[1].

It is conceivable upon the philosophy of either Pope or Laplace that some time in the future man will have a knowledge of practically all of the laws of nature so that it would be possible for him to predict the future quality of product with absolute certainty. This, in fact, might be considered a goal for applied science. However, as we have already noted the indications today are that it would not be a practical one. We are a long way from such a goal, and for years to come the engineer must be content with the knowledge of only a comparatively few of the many conceivable laws of nature where we think of the term law in the sense of Newton's Laws of Motion. Furthermore, the engineer is fully aware of the fact that, whereas it might conceivably be possible with the knowledge of these laws to predict the future quality of product with absolute certainty, it would not in general be feasible to do so, any more than it would be feasible to write down the equations of motion (were it possible to do so) for a thimble full of molecules of air under normal conditions. Furthermore, the engineer is fully aware of the fact that, whereas in the laboratory one may often be able to hold conditions sufficiently constant that the action of a single law may be observed with high precision, this same degree of constancy cannot in general be maintained under what appear today to be necessary conditions of commercial production. Still further, if we are to believe, as do many of the leaders of scientific thought, that possibly the only kind of objective constancy in this world is of a statistical nature, then it follows that the complete realization of the sixth stage is not merely a long way off but impossible.

---

1. Jevons, W. S., "The Principles of Science", Second Edition, Page 738.

Enough has been said to indicate how the concept of control fits into the general scheme of scientific inference in respect to the future in terms of the past. In fact, this kind of inference that a thing will happen within certain limits may be the only kind that can be drawn. This leads us to the consideration of our second topic.

2. Control a Part of any Industrial Research Program

A. In the first place, we have already seen that the principle of control played an important role in laboratory research in what is ordinarily termed pure science. We have seen that it is necessary, in general, in all such work to attain as nearly as possible to certainty in the assurance that the observations supposed to have been taken under the same essential conditions have actually been taken in this way. As an efficient tool in testing whether or not this condition has been satisfied, we have the tests for lack of control previously referred to and illustrated throughout this bulletin. Going still further, we have seen that the criteria for maximum control give a test which indicates the limit to which it is reasonable that research may go in revealing causes of variability in the set of observations presumably taken under a constant system of chance causes. We have also seen that many of the quantities with which we actually deal in the so-called exact sciences are but averages of statistical distributions assumed to be given by what we have chosen to term a constant system of chance causes. Further discussion of this subject, however, is beyond the scope of the present bulletin.

B. Next let us consider the importance of control in the development and production of product. In most cases we can distinguish five more or less distinct steps in any modern scientific development program. They are:

1. A study of the results of research to provide principles and numerical data upon which to base a design.
2. The application of such information in the construction of an ideal piece of apparatus designed to satisfy some human want, where no attention is given to the cost.
3. Production of tool-made samples under supposedly commercial conditions.
4. Test of tool-made samples and specification of quality requirements that can presumably be met under commercial conditions.

5. Development of production methods.

It will be seen at a glance that from this viewpoint the results of design, development and production are grounded on the initial results of research What is more important, however, in our present study is the fact that often causes of variability enter in the last four steps above mentioned which by the very nature of the problem are not experienced in the research laboratory. That this is true is almost obvious because we have the possibility of assignable causes entering through different sources of material, the human element and variable conditions which affect the production process. One possible method of obtaining satisfactory quality under such procedure is through a 100% inspection of the product at the time it is ready for delivery, except for those qualities which cannot be tested because of the destructive nature of the tests. Then there is the cost of inspection to be considered. Furthermore, if indications of the presence of assignable causes of variability are discovered in the quality of final product, it is not easy to locate the causes because the data of final tests may have been taken long after the causes have ceased to function. What is even more important, as we have seen in previous chapters of this bulletin, the quality may appear controlled in the end and yet there may be assignable causes of variability at one or more steps in production. In the light of this information, therefore, it seems highly desirable that the measurements made in <u>each</u> of the last four of the above mentioned steps be tested to determine whether or not there is any indication of lack of control. If there is, it may be necessary that a further study be made in the laboratory to assist in finding the assignable causes of variability. An illustration of such a procedure was already referred to in Chapter II in connection with the development of an insulating material.

In this same connection, it is perhaps of interest to emphasize the importance of control in setting the standards for the raw materials that enter into the production process. As we have already noted, and as is well known, most physical properties are subject to the influence of presumably large numbersof chance causes. Therefore, it we are to make efficient use of data representing these properties, the data themselves must have been taken under controlled conditions. We may take as an illustrative example the problem of setting standards on the tensile strength of a particular aluminum die-casting as discussed elsewhere[1]. Before we can use the experimental results with any assurance of their giving a controlled product, it is highly desirable that we make use of tests to determine whether or not the data have been secured under controlled conditions[2].

Furthermore, in the development of processes of production, it should be of advantage to apply tests to detect lack of control and then to weed out the assignable causes of variability as they occur, with the assurance of the kind already indicated in this bulletin, that after this process of weeding out has once led to a product which appears to be controlled, future product will remain, in general, in the same state unless obvious assignable causes of variability are entering.

Here we see, therefore, how the theory of control plays an important part in the various stages of applied science. Furthermore, we see why it is desirable that the departments of design, development and production keep the laboratory research department informed as to evidence of existence of assignable causes wherever they arise up to the time that product goes to the consumer.

Let us next consider the application of theory in the study of the life history of product. Obviously, when equipment goes into the field, it meets many and varied conditions, the influence of which on the quality of product is not in general known. Such an example would be the varied conditions under which telephone poles are placed throughout the United States. A priori, it is reasonable to believe that the life of the pole depends in a large way upon the service conditions. Among the exceedingly large number of variables which may influence the life of the pole, little information is available to indicate the importance of any one. Obviously, the value of laboratory research in improving the quality of a pole through life must take into account ways and means of preservation suited to each of the various conditions. Naturally, therefore, it is of interest to know when the variability in the

---

1. Report presented at the meeting of the A.S.T.M. June 24-28, 1929.
2. Such a test, for example, is Modified Criterion 1 of Bulletin I.E.B. 1.

quality of the material at any stage in life is such as to indicate the existence of an assignable cause so that further research may be instituted to find ways and means of effectively removing this cause[1]. Field engineers, therefore, find need for analytical methods of detecting evidence of lack of control in the quality of product at any time as revealed by life data so that they can call this fact to the attention of the laboratory staff.

3. How Shall Control be Exercised?

It is perhaps well to summarize here a little more explicitly how control is to be exercised in the light of what we have considered in the present bulletin. We see that just two things are required.

A. The intuitive ability to make the right kind of hunches or hypotheses as to the cause of lack of control.

B. Ways and means of testing such hypotheses as already given in I.E.B. 1 and to be further developed in I.E.B. 5.

In the first step we see the play of scientific judgment and in the second the application of simple statistical methods, which have only become available within the past few years, to assist in checking that judgment.

We have also seen that a successful program of control calls for a close cooperation between laboratory research and the particular department where lack of control is discovered in the course of the process of fabrication from raw material to finished product and in the life of the material.

4. Economic Control

For obvious reasons, other things being equal, any industrial research program usually starts with the study of some particular problem, the solution of which most likely will give the largest improvement in quality for the least cost. Naturally, the same thing is true in the application of the principles of control. It is necessary to look about first to discover where the exercise of control will do the most good from an economic viewpoint.

When, however, we begin to exercise control we find several instances where evidence of lack of control is present and yet it is argued that the kind of assignable causes present (such as variability in source of raw material and even in finished product), are such that they cannot be regulated so that the

---

1. This point is already referred to in Chapter II.

resultant quality of all the product can be considered as coming from a single constant system of causes. In this connection, of course, it should be noted that the control of the quality of a given kind of product does not necessarily mean that all of that product should come from the same constant system of causes. It is sufficient for us to know enough about the production processes to be able to forecast within chance limits the variability in product at any future interval of time. Thus we may have situations where the product is controlled in the sense that certain known portions of it come from well defined constant systems of causes or it may be that the product comes from one constant system of causes superposed upon some trend in the expected value of this system which is a well defined function of time. In any case, however, it is necessary that the assignable causes of variability in the product be known. Under these conditions it then becomes feasible to predict the future quality in terms of the past upon the assurance given us by the apparently universal Law of Large Numbers.

Through applications of this law after the manner described in this bulletin, control within limits becomes a reality and by means of the criteria for detecting the necessary and sufficient conditions of maximum control, we have as it were an apparently practical limit beyond which research may not be expected to pay unless the manufacturing process is almost completely changed. The measure of control given in the previous chapter indicates the extent to which everything feasible is being done to maintain economic standard quality.

APPENDIX 1

HOW THE SIMPLE CAUSE SYSTEM OPERATES

Let us take a system consisting of only five causes,

$$C_1, C_2, C_3, C_4, C_5.$$

The effects of these causes are

$$1, 2, 3, 4, 5$$

respectively and the probabilities that the causes produce these effects are

$$\frac{5}{6}, \frac{4}{6}, \frac{3}{6}, \frac{2}{6}, \frac{1}{6}.$$

The causes act independently one of another and the resultant effect X of the system is made up of the sum of the effects of the individual causes. How are the possible values of X distributed?

Obviously to answer this question we must first find all the possible values that the cause system can produce and then determine the probability of the system producing each one of them.

A little study will suffice to show that resultant effects 0, 1, 2, 3,...., 15, are possible depending on the number of causes that produce effects or which "operate" as we shall say.

Clearly the effect $X = 0$ can be obtained in only one way - no one of the causes operates. Now if the probability that a cause operates is p, then the probability of the cause not operating is $(1 - p)$. For example, the probability that C_1 does not produce its effect is $1 - \frac{5}{6} = \frac{1}{6}$. Hence, the probability that no one of the causes operates is

$$\frac{1}{6}\cdot\frac{2}{6}\cdot\frac{3}{6}\cdot\frac{4}{6}\cdot\frac{5}{6} = \frac{120}{6^5}$$

and this product is therefore the probability that the system produces the value 0.

Next consider the resultant effect $X = 1$. This can be produced only if C_1 operates and the remaining causes do not. Hence, the probability that the system produces this value is

$$\frac{5}{6}\cdot\frac{2}{6}\cdot\frac{3}{6}\cdot\frac{4}{6}\cdot\frac{5}{6} = \frac{600}{6^5}$$

Similarly the effect $X = 2$ can be obtained in only one way.

However, the effect X = 3 can be obtained in two ways. That is, C_1 and C_2 may operate and the remaining three causes fail to operate or C_3 alone may operate. The probability for the first alternative is

$$\frac{5}{6}\cdot\frac{4}{6}\cdot\frac{3}{6}\cdot\frac{4}{6}\cdot\frac{5}{6} ,$$

and for the second is

$$\frac{1}{6}\cdot\frac{2}{6}\cdot\frac{3}{6}\cdot\frac{4}{6}\cdot\frac{5}{6}.$$

Hence, the probability that the cause system will produce the value X = 3 is the sum of these two probabilities or $1320/6^5$.

In this way the probabilities corresponding to all the possible values of the cause system are obtained. Fig. 1 shows for each effect the combinations of causes that produce it and the probability that these combinations occur.

The graphical picture of effects of the cause system is then obtained by erecting at the points 0, 1, 2, ..., 15 of the horizontal scale ordinates proportional to the probabilities of the different effects.

Possible Effect X of Cause System	Combination of Causes Acting to Give X	Probability of getting all the Combinations that give X, that is, prob. of getting the effect X	
0	....		$= \frac{120}{6^5}$
1	C_1		$= \frac{600}{6^5}$
2	C_2		$= \frac{240}{6^5}$
3	(C_1C_2), C_3.	$\frac{1200}{6^5} + \frac{120}{6^5}$	$= \frac{1320}{6^5}$
4	(C_1C_3), C_4.	$\frac{600}{6^5} + \frac{60}{6^5}$	$= \frac{660}{6^5}$
5	(C_1C_4), (C_2C_3), C_5.	$\frac{300}{6^5} + \frac{240}{6^5} + \frac{24}{6^5}$	$= \frac{564}{6^5}$
6	(C_1C_5), (C_2C_4), $(C_1C_2C_3)$.	$\frac{120}{6^5} + \frac{120}{6^5} + \frac{1200}{6^5}$	$= \frac{1440}{6^5}$
7	$(C_1C_2C_4)$, (C_2C_5), (C_3C_4).	$\frac{600}{6^5} + \frac{48}{6^5} + \frac{60}{6^5}$	$= \frac{708}{6^5}$
8	$(C_1C_2C_5)$, $(C_1C_3C_4)$, (C_3C_5).	$\frac{240}{6^5} + \frac{300}{6^5} + \frac{24}{6^5}$	$= \frac{564}{6^5}$
9	$(C_1C_3C_5)$, $(C_2C_3C_4)$, (C_4C_5).	$\frac{120}{6^5} + \frac{120}{6^5} + \frac{12}{6^5}$	$= \frac{252}{6^5}$
10	$(C_1C_2C_3C_4)$, $(C_2C_3C_5)$, $(C_1C_4C_5)$.	$\frac{600}{6^5} + \frac{48}{6^5} + \frac{60}{6^5}$	$= \frac{708}{6^5}$
11	$(C_1C_2C_3C_5)$, $(C_2C_4C_5)$.	$\frac{240}{6^5} + \frac{24}{6^5}$	$= \frac{264}{6^5}$
12	$(C_3C_4C_5)$, $(C_1C_2C_4C_5)$.	$\frac{12}{6^5} + \frac{120}{6^5}$	$= \frac{132}{6^5}$
13	$(C_1C_3C_4C_5)$.		$= \frac{60}{6^5}$
14	$(C_2C_3C_4C_5)$.		$= \frac{24}{6^5}$
15	$(C_1C_2C_3C_4C_5)$.		$= \frac{120}{6^5}$

FIG. 1 - ILLUSTRATIVE FORM FOR OBTAINING THE FREQUENCY DISTRIBUTION OF EFFECTS OF A CAUSE SYSTEM

APPENDIX 2

1. Commonness of Causation Measured by r

Suppose we have any two physical quantities X_1 and X_2. Variations in the first are produced by $(\ell+s)$ causes which we shall designate by

$$U_1, U_2, \ldots, U_\ell, V_1, V_2, \ldots, V_s,$$

while variations in the second are produced by $(\ell+n)$ causes

$$U_1, U_2, \ldots, U_\ell, W_1, W_2, \ldots, W_n,$$

so that ℓ of the causes are common to the two variables.

Let us consider first the following very simple hypothesis concerning the causes.

(1) Each cause produces a single effect and this effect is unity for all of the causes.

(2) The probability that any one of the causes produces its effect is constant and equal to p.

(3) The resultant effect X_1 or X_2 is made up of the sum of the effects of the individual causes.

These conditions of course would lead to a binomial distribution of effects for each of the variables X_1 and X_2.

Denote by z the contribution to X_1 and X_2 of the ℓ common causes. Denote the contribution of the s V's by x and of the n W's by y. Then for any particular operation of the cause systems

$$X_1 = x+z$$

and

$$X_2 = y+z.$$

We shall show that the correlation coefficient $r_{X_1X_2}$ between X_1 and X_2 is given by

$$\frac{\ell}{\sqrt{(\ell+s)(\ell+n)}}$$

Now by definition

$$r_{X_1X_2} = \frac{E(x+z)(y+z) - (\overline{x+z})(\overline{y+z})}{\sigma_{x+z}\ \sigma_{y+z}}$$

where E is a symbol for mathematical expectation and $\overline{x+z}$ and $\overline{y+z}$ denote the mean values of these two quantities.

Making use of the properties of the binomial distribution we have

$$E(x+z)(y+z) = E(xy+yz+xz+z^2)$$

$$= (ps)(pn) + (pn)(\ell p) + (ps)(\ell p) + (\ell p)^2 + \ell pq$$

$$(\overline{x+z})(\overline{y+z}) = p(\ell+s)p(\ell+n) = p^2(\ell^2+\ell s+\ell n+ns)$$

Also

$$\sigma_{x+z} = \sqrt{pq(\ell+s)}$$

$$\sigma_{y+z} = \sqrt{pq(\ell+n)}\ , \qquad q = 1 - p.$$

Hence

$$r_{X_1X_2} = \frac{\ell pq}{pq\sqrt{(\ell+s)(\ell+n)}} = \frac{\ell}{\sqrt{(\ell+s)(\ell+n)}}\ ,$$

or if $n = s = m$, in which case there are the same number $(\ell+m)$ of causes for each of the variables X_1 and X_2, we have

$$r = \frac{\ell}{\ell+m}\ ,$$

the ratio of the number of common causes to the total number of causes in either variable.

Consider now the more general case in which X_1 and X_2 are related their respective causes by some unknown functional relationship. Thus

$$X_1 = F_1\ (U_1, U_2, \ldots, U_\ell, V_1, V_2, \ldots, V_s)$$

and

$$X_2 = F_2\ (U_1, U_2, \ldots, U_\ell, W_1, W_2, \ldots, W_n)$$

Now, we shall think of the U's, V's and W's as symbols for groups of causes, each group producing a discontinuous distribution of effects.

Making certain definite assumptions we shall show that even in this case, $r_{X_1X_2}$ to a first approximation is given by

$$\frac{\ell}{\sqrt{(\ell+s)(\ell+n)}} .$$

Denote the mean values of

$$U_1, U_2, \ldots, V_1, V_2, \ldots, W_1, W_2, \ldots$$

by

$$\bar{U}_1, \bar{U}_2, \ldots, \bar{V}_1, \bar{V}_2, \ldots, \bar{W}_1, \bar{W}_2, \ldots$$

Then assuming that X_1 and X_2 can be expanded in a Taylor's Series about these mean values, we have

$$X_1 - X_{1_0} = (U_1-\bar{U}_1)\frac{\partial F_1}{\partial U_1} + (U_2-\bar{U}_2)\frac{\partial F_1}{\partial U_2} + \ldots + (V_1-\bar{V}_1)\frac{\partial F_1}{\partial V_1} + (V_2-\bar{V}_2)\frac{\partial F_1}{\partial V_2} + \ldots . \quad (1)$$

and

$$X_2 - X_{2_0} = (U_1-\bar{U}_1)\frac{\partial F_2}{\partial U_1} + (U_2-\bar{U}_2)\frac{\partial F_2}{\partial U_2} + \ldots + (W_1-\bar{W}_1)\frac{\partial F_2}{\partial W_1} + (W_2-\bar{W}_2)\frac{\partial F_2}{\partial W_2} + \ldots , \quad (2)$$

where

X_{1_0} and X_{2_0} are to a first approximation the mean values of X_1 and X_2 respectively and the partial derivatives are formed for the mean values of the variables.

Before proceeding further, we shall make the following further assumptions:

(a) Terms beyond the first powers in (1) and (2) may be neglected.

(b) $\frac{\partial F_1}{\partial U_1} = \frac{\partial F_1}{\partial U_2} = \ldots = \frac{\partial F_1}{\partial V_1} = \frac{\partial F_1}{\partial V_2} = \ldots = \frac{\partial F_2}{\partial U_1} = \frac{\partial F_2}{\partial U_2} = \ldots = \frac{\partial F_2}{\partial W_1} = \frac{\partial F_2}{\partial W_2} = \ldots$

The physical significance of this condition is that equal deviations in the U's, V's and W's produce deviations in X_1 and X_2 proportional to the number of variables contained in the X's.

(c) The standard deviation σ of each variable is the same as that of any other variable.

(d) The effect of any one cause is independent of that of any other.

Denoting the deviations in each variable by the corresponding small letter, we then have

$$x_1 = a(u_1+u_2+\cdots+u_\ell+v_1+v_2+\cdots+v_s)$$

$$x_2 = a(u_1+u_2+\cdots+u_\ell+w_1+w_2+\cdots+w_n),$$

where a stands for the common value of the partial derivatives.

The correlation coefficient $r_{x_1x_2}$ between x_1 and x_2 is then

$$r_{x_1x_2} = \frac{E(x_1x_2)}{\sigma_{x_1}\ \sigma_{x_2}}$$

where as before E is a symbol for mathematical expectation.

Now

$$E(x_1x_2) = a^2\ [E(u_1{}^2)+E(u_2{}^2)+\cdots+E(u_\ell{}^2)],$$

since the expected value of any product term is zero.

Hence

$$E(x_1x_2) = a^2\ell\sigma^2 .$$

$$\sigma^2_{x_1} = a^2\ E[u_1+u_2+\cdots+u_\ell+v_1+v_2+\cdots+v_s]^2$$

$$= a^2\ [\ell\sigma^2+s\sigma^2]\ ,$$

for the reason already stated.

Similarly

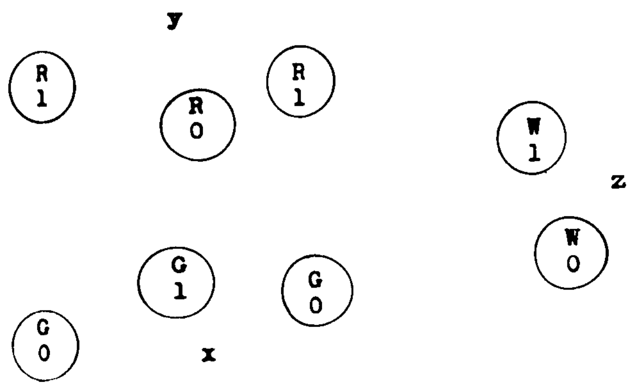
$$\sigma^2_{x_2} = a^2 E[u_1+u_2+\cdots+u_\ell+w_1+w_2+\cdots+w_n]^2$$

$$= a^2\ [\ell\sigma^2+n\sigma^2].$$

Therefore

$$r_{x_1x_2} = \frac{a^2\ell\sigma^2}{\sqrt{a^2\sigma^2(\ell+s)}\ \sqrt{a^2\sigma^2(\ell+n)}} = \frac{\ell}{\sqrt{(\ell+s)(\ell+n)}}$$

2. Simple Example Showing How r Measures Commonness of Causation



Two Systems Having Two Causes in Common

Consider as below the two chance systems of causes put into operation by the tossing of eight chips each of which has one side marked with a 1 and the other side with zero.

Each chip represents a cause. The effect produced by each cause is 1.

The probability that each cause produces its effect is 1/2.

Each cause operates independently of every other cause.

The first system of causes consists of:

3 green chips and 2 white chips.

The second system of causes consists of:

3 red chips and 2 white chips.

Hence there are two causes (2 white chips) which are common to both systems.

Denote by z the combined effect produced by the common causes: by x the combined effect produced by the remaining three causes of the first system and by y the combined effect produced by the remaining three causes of the second system. For example, the figure shows a typical result of the operation of the two systems in which

$$x = 1, \; y = 2, \; z = 1,$$

so that the effect produced by the first system is 2 and the effect produced by the second is 3. Obviously z can take on values 0, 1, or 2 and x and y may have the values 0, 1, 2, or 3.

In general the resultant effect of the first system is

$$X_1 = x + z,$$

and the resultant effect of the second system is

$$X_2 = y + z.$$

Inasmuch as each observed value of X_1 and X_2 has a common component, i.e., the effect of the common causes, we would naturally expect a certain relationship between the values of X_1 and X_2 in successive operations of the two systems.

Now the correlation coefficient $r_{X_1X_2}$ between X_1 and X_2 is a measure of this relationship and since these two systems of causes obey all the laws laid down for the general case above, we have merely to set $\ell = 2$, $m = 3$ and we have

$$r_{X_1X_2} = \frac{2}{2+3} = .400 \ .$$

By actual experiment we obtained 500 series of values of X_1 and X_2 and the observed correlation coefficient between the two sets was found to be .422 giving a rather close check on the expected value .400.

Fig. 1 gives the scatter diagram and lines of regression for these 500 values of X_1 and X_2.

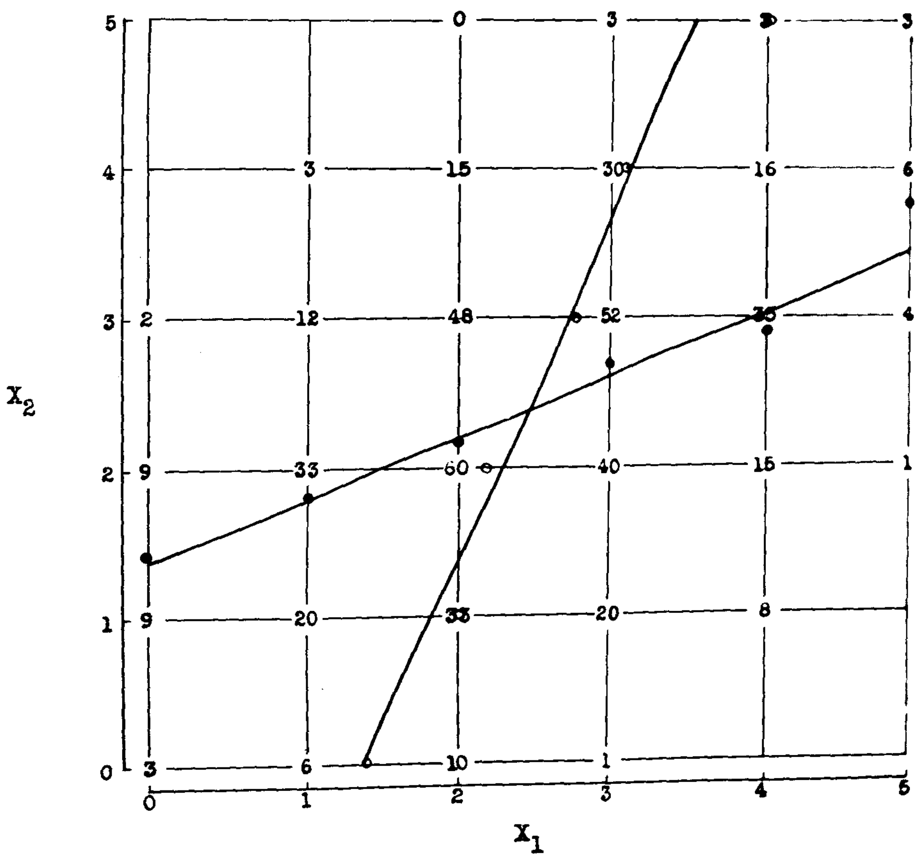


FIG. 1

APPENDIX 3

SIMPLE ILLUSTRATIONS OF KINDS OF CONTROL

In the discussion of economic control, we had occasion to refer to different kinds of control. For example, we pointed out that the product may come from different sources in respect to either materials, machines or shops, or, in fact, a combination of these. As an illustration, Fig. 1 shows the quality of two kinds of product from 24 suppliers. In such a case, it may not be economically feasible to so modify the systems of causes producing the product from the different sources so as to obtain a resultant product whose quality could be considered as given by a constant system of chance causes.

The theory discussed in the present bulletin forms a basis for setting control limits under any of four well defined conditions and thus securing the advantages of control. For the sake of simplicity, we shall take extremely simple illustrations.

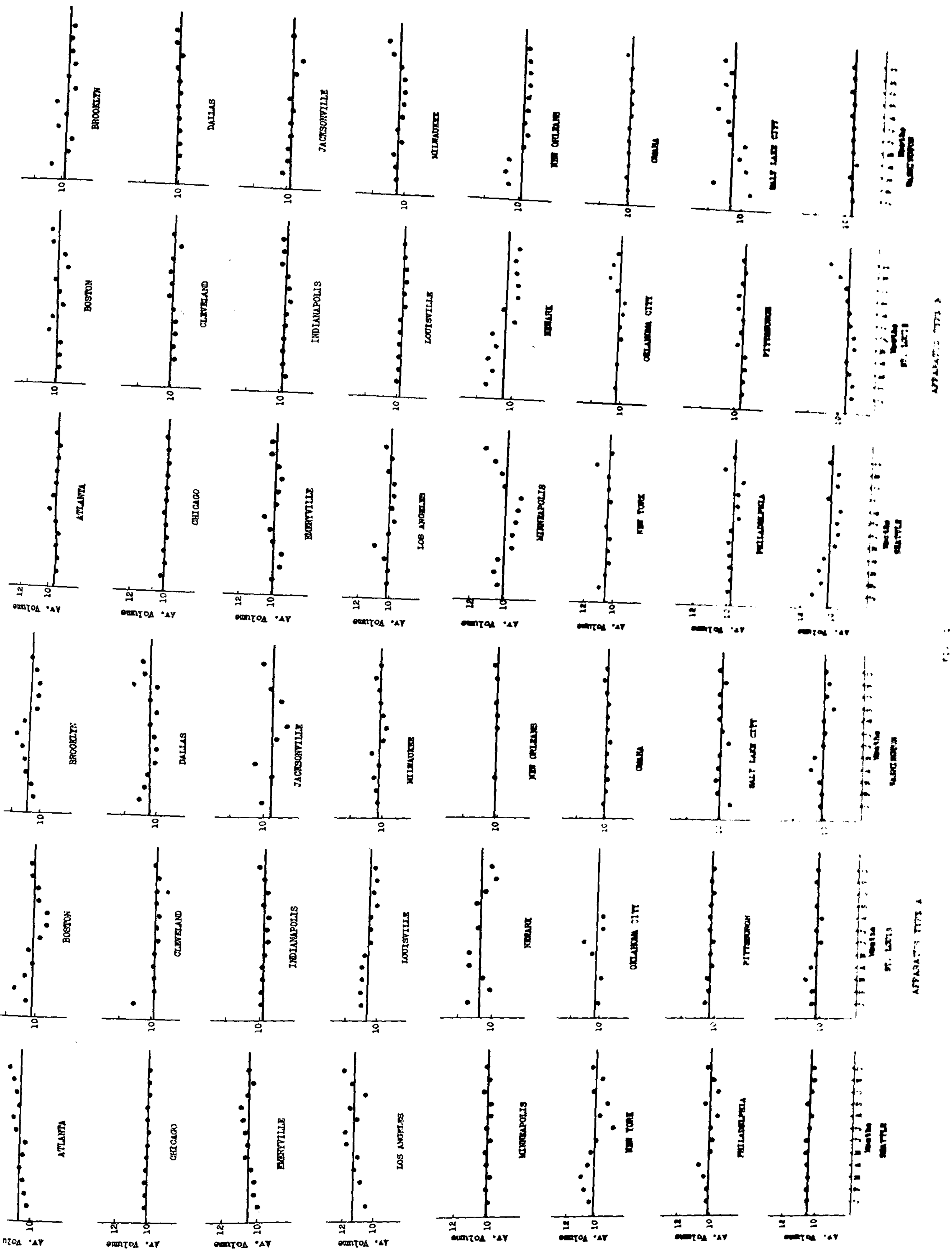
Bernouilli Control

Consider again the a's already introduced in Chapter 1 of this bulletin. It would indeed be a difficult task to set down all the causes of variability in your a's, but we nevertheless would probably be quite ready to grant that there are a finite number of such causes. Furthermore, even though we cannot see the individual causes operate we may be willing to assume that they work together to produce their effects without any regard to time. This done we have all the essential elements that characterize a constant system of chance causes.

Suppose now that the quality characteristic of your product (the a's) that you are interested in controlling is the shaded area within the a as just indicated. If this area were measured with a planimeter and its value X laid off on a horizontal scale, then the assumption of a constant system of chance causes, or Bernouilli Control, means that for any one of the M possible values that this system can produce the relative frequency or probability of its occurrence is constant.

Lexian Control

Now suppose that I try my hand at making a's. The human element in


BROOKLYN
DALLAS
JACKSONVILLE
MILWAUKEE
NEW ORLEANS
OMAHA
SALT LAKE CITY
BOSTON
CLEVELAND
INDIANAPOLIS
LOUISVILLE
NEWARK
OKLAHOMA CITY
PITTSBURGH
ATLANTA
CHICAGO
EMERYVILLE
LOS ANGELES
MINNEAPOLIS
NEW YORK
PHILADELPHIA
SEATTLE
Av. Volume
10
12

the process is now changed and therefore all the causes of variation associated with this element have in general different effects. In other words the cause system underlying the manufacture of the a's has changed, but it is natural that we should make the same assumptions concerning the cause system now to be put in operation as we made before.

Experience alone, without any theory of causation whatsoever would tell us that my a's are different from yours and hence that my areas or X's are different from yours. Moreover I would not expect to get the same value of frequency of possible X's as you did.

Now let a third, a fourth, a fifth, and in general an n^{th} person try the same experiment. Then for each individual there will correspond a constant system of causes different in general from all the rest and hence a distribution of resulting effects X different from all the others. Physically we may picture[1] the effects of these n systems of causes indicated in Fig. 2.

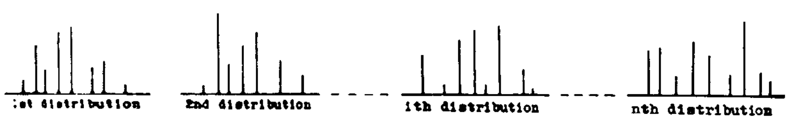


FIG. 2 - DISTRIBUTION OF PRODUCT UNDER LEXIAN CONTROL
Each System Exhibits Bernoulli Control

Each of the n persons who made a's put into operation a constant system of causes under Bernouilli Control, but if we think of the entire product of a's turned out by the n systems, this product does not exhibit Bernouilli Control because the probability of getting a given value of the quality characteristics changes from time to time. If, however, the same n systems of causes continue to operate, then the resulting product will be said to be Lexian controlled in the sense that the cause system always changes the same way in the future that it has in the past and no system of causes other than the n already put in operation enter.

Poisson Control

Assume now that the first system of causes produces the first a, the second system the second a, the third system the third a and so on, the n^{th} system the n^{th} a. This being done, the same n systems produce another

---

1. Actually of course the different types of a's do not differ greatly in size so that the various distributions of areas would overlap. However, for the purpose of illustration, we have pictured them as entirely separate one from another.

set of α's in the same fashion and then another set, and so on. The distribution of effects X resulting from such procedure would exhibit Poisson Control. The cause system changes after each piece of product is manufactured but each series of n pieces of product is produced by the same n system of causes.

General Form of Control

One general form of control would be that in which the expected $\bar{X}'$ of a constant cause system changes with time, which fact we may express symbolically by the equation

$$\bar{X}' = f(t)$$

where f represents the functional relationship and t is the time.

INDIAN STATISTICAL INSTITUTE LIBRARY
No. R 331
Date
CALCUTTA.

SHEWHART'S COLLECTION