TABLE OF CONTENTS

	Page
Part I - Role of Measurement in Inspection Engineering.	
1. What Measurement Involves	1
2. Object of this Bulletin	2
3. Determination of the True Quality of a Thing . .	2
4. Determination of True Quality of Product	5
5. How to Indicate if Quality is Economic Standard .	6
6. Presentation of Information of Greatest Service in the Improvement of Quality	7
Part II - Direct Measurement of Quality.	
1. Direct Measurement of Quality Defined	9
2. Formal Statement of Problem - Large Samples . . .	17
3. Formal Solution of Problem - Large Samples . . .	17
4. Illustration of Formal Method of Solution	19
Part III - Indirect Measurement of the Quality of a Thing.	
1. Measurement when Form of Function is Known a priori	20
2. Measurement when Form of Function is Unknown a priori	27
3. Measurement through Stochastic Relationship . . .	34
Part IV - Machine Measurement of Quality.	
1. Criteria for Machine	46
2. Assumptions Involved	47
3. Calibration of Machine	48
4. Maintenance of Machine	50
5. Errors of Measurement	51
Part V - Measurement of Quality of a Number of Things.	
1. Basis for Measurement	52
2. Practical Applications - Setting Standards . . .	57
A. Physical Properties of Die-Castings	57
B. Modulus of Rupture of Telephone Poles	61
C. Depth of Sapwood	66
3. Applications in the Control of Quality	69
A. Control of Treatment of Telephone Poles . . .	69
B. Control of Quality of Aluminum Die-Castings .	80
Part VI - Correction of Data for Errors of Measurement.	
1. How to Correct Data for Errors of Measurement . .	83
2. Special Applications	86

MEASUREMENT OF QUALITY

PART I

Rôle of Measurement in Inspection Engineering

1. What Measurement Involves

Naturally inspection engineering involves measurement and the sorting out of the good from the bad but that is not all. It involves the measurement of the right thing in the right way and the right number of times. It does not stop here. It involves the use of methods of analysis of data which will yield all the essential information contained in the data in a form to be of greatest service to the research, development, design and purchasing organizations in the better control of quality of product through the weeding out of those causes of variation which should not be left to chance.

There is another function of inspection engineering, namely, that of detecting whether or not quality of product differs significantly from economic standard. It is one thing for a consumer to know simply that a lot of product passes certain inspection requirements; it is quite another thing for him to know that the organization behind that product has applied all modern research methods in detecting and eliminating causes of quality variation which need not be left to chance and has thus arrived at an economic standard.

The rôle of measurement in inspection engineering must be considered in terms of the following four objects of inspection engineering:

a. Determination of the true quality of a thing by the most economical method. This inherently involves such things as the choice of the best method of measurement and the detection and elimination of errors of measurement.

b. Determination of the true quality of product in the most economical way.

c. Determination of the necessary steps to obtain economic standard quality.

d. Presentation of inspection information in a form that will indicate to design, development and research engineers what changes should be made in manufacturing methods and indicate to the purchasing engineer lack of control of the quality of raw material.

2. Object of This Bulletin

The object of the present bulletin is to present a formal solution of the problem of measurement involved in the attainment of the four objects of inspection engineering. In almost every instance a practical illustration is given in sufficient detail to indicate all of the necessary numerical steps in the solution of the problem.

In this sense the present bulletin is to serve as a manual of the theory underlying the establishment of specific methods of measuring the quality of product.

3. Determination of The True Quality of a Thing

Before we proceed with the detailed technical discussion of the problems which arise in measurement, we shall briefly outline some of them in a non-technical manner. It is natural that we start with the measurement of a single quality on a single thing, a simple example of which would be measurement of a length by means of a meter stick.

If we make a series of n measurements of this length, we will get n values, let us say,

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

where, in general, these values will differ one from another. The first problem which confronts us is to determine from this series of n observed values of the length some particular function of these values which will give us the most likely estimate of the true length, let us say, X'.

It is not sufficient, however, to know merely the numerical value of this estimate; we need to know something about the reliance which we are justified in placing upon this value. In other words, we must, in some way, indicate the magnitude of the error of measurement. This constitutes our second problem in the simplest form of measurement.

We shall find that the method of determining the error of measurement depends upon the size n of the sample. Furthermore, we shall find that three criteria govern our choice of the estimate of this error, if we are to assure ourselves that we have made use of all of the information contained in the original data and that we have made use of it in the most efficient known way.

This problem of measurement of length however is extremely simple compared with the majority of those which we meet in engineering work. As an illustration, assume that we wish to determine the modulus of rupture of a telephone pole, the charge on an electron, the coefficient of expansion of a rod or any one of a number of such quality characteristics. In such cases the quality is determined indirectly through measurement of the properties of the thing for which the quality is being determined.

The simplest problem of this nature arises when we know the functional relationship f between the quality Y under measurement and certain other characteristics, let us say, $X_1, X_2, \ldots X_m$. Such a relationship is indicated formally by the equation,

$$Y = f_1 \; (X_1, X_2 \ldots X_i \ldots X_m) \qquad (1.1)$$

Our problem, of course, is to determine the quality Y with the greatest precision obtainable at a given cost of measurement. Naturally the error in Y depends upon the errors in each of the m measured characteristics. Therefore we must find the functional relationship between the error in Y and those in the measured characteristics and then set up certain criteria which will indicate the minimum error that may be made by this method of measurement.

In many cases, however, it is possible to measure the quality Y through some other known functional relationship between this quality and certain other measurable characteristics of the things represented below by the Z's. For example we might have some functional relationship

$$Y = f_2 \; (Z_1, Z_2 \ldots Z_i \ldots Z_m) \qquad (1.2)$$

Proceeding as in the previous case, we can set down the conditions which will make the error in Y a minimum when measured in terms of the Z's and then we must choose between the two methods of measuring Y.

Of course there may be even more than two methods of measurement, a typical illustration of which is the measurement of the modulus of rupture of telephone poles by the one point, two point and cantilever loading methods.

So far, however, our problem is comparatively simple. We know the functional relation between certain variables and it is a very simple task to choose that method of measurement which will be most satisfactory in a given case, assuming of course, that we make use of recently developed methods for treating small samples.

So far we have assumed that the functions f_1 and f_2 in Equations 1.1 and 1.2 are known. In the majority of cases, the functional forms such as f_1 and f_2 are not known. In fact a large number of the so-called laws of nature are merely empirically determined functional relations. To cite a case which we shall treat in greater detail later, we may consider the resistance of a contact in a carbon microphone. The resistance of such a contact is a function of the voltage applied across the contact under conditions where, so far as we know, we have controlled all other variables. To begin with we do not know this functional relationship. If we are to use this relationship as a basis for measuring the resistance of carbon, as is being done in semi-commercial tests, it is necessary as a first step to determine empirically the most likely relationship between voltage and resistance. This introduces us to another fundamental problem, namely, having given an observed set of data, we are to justify our choice of hypothesis to explain this set of data.

In general we shall introduce criteria for determining the probability of the occurrence of the given set of data or one less likely (or of some value of a function of this set of data or one less likely) upon the assumption of a given hypothesis. Other things being equal, we shall then choose that hypothesis for which the probability of the observed set of data or of the given function of the observed set is a maximum.

The next difficulty in the way of measurement comes when we attempt to measure some quality characteristic, such as tensile strength of a material, in terms of some other characteristic, such as hardness or density of the material. The first thing that confronts us here is the fact that we know that there is no functional relationship between the quality Y such as tensile strength and such qualities as hardness and density. That is to say we know that, given the numerical values of hardness and density, the values of tensile strength are not uniquely determined. All that we have to go on in such a case is the fact that there appears to be some kind of a relationship and, in the literature of the subject, such an apparent relationship is defined as being <u>stochastic</u>.

Without further discussion at this point we shall rewrite Equations 1.1 and 1.2 in the form:

$$Y = f_{s1}(X_1, X_2 \ldots X_i \ldots X_m) \qquad (1.1')$$

$$\text{and} \quad Y = f_{s2}(Z_1, Z_2 \ldots Z_i \ldots Z_{m1}) \qquad (1.2')$$

where in general the subscript s on the function merely indicates that there is only an apparent or stochastic relationship. The first step in considering the use of stochastic relationships in measurement is that of reducing these to one or more functional relationships which express for the purpose in hand all of the essential information contained in the original sets of data. Having done this, we must then proceed with steps quite analogous to those considered above in the case of indirect measurement through a known functional relationship.

4. Determination of True Quality of Product

What has been said of course is just paving the way for the slightly more complicated problem of measuring the quality of product which, in general, involves the measurement of one or more quality characteristics on each of a number of things. For example, if we have a thousand instruments of a given kind, such as relays, and we wish to express the quality of these relays in terms of such factors as resistance, capacity and inductance, we would have, all told, a group of three thousand measurements, assuming that we made only one measurement of a given quality characteristic on each relay.

In the general case where we measure let us say m characteristics on each of n different things or pieces of apparatus we attain n sets of observations of m each, such as:

$$\left.\begin{array}{l} X_{11}, X_{12}, \ldots X_{1i}, \ldots X_{1m} \\ X_{21}, X_{22}, \ldots X_{2i}, \ldots X_{2m} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ X_{n1}, X_{n2}, \ldots X_{ni}, \ldots X_{nm} \end{array}\right\} \qquad (1.3)$$

We may take as our first problem the correction of this group of data for errors of measurement which arise in measuring any one of the m different characteristics. We wish to obtain our best estimate of the true quality characteristics of each of these n instruments. In other words, we must set up methods for correcting the observed distributions to allow for these errors of measurement.

Having performed this simple task, we then must consider ways and means of expressing the information contained in even a very large number of observations in some simple form such that it will contain all of the essential information required to answer the practical problems, for the solution of which data were taken. This means that we must substitute for the original set of data a number of functions of these data. Of course there is nothing strange in this for it is what we do every day when we use averages and standard deviations in the comparison of one set of data with another. We shall consider, however, very important engineering problems which cannot be solved efficiently, if at all, by the simpler methods which are customarily used in reducing a series of observations to a few statistics.

Formally we shall state our problem in the following way: Given a series of observations such as indicated above, we are to find the minimum number of functions $\theta_1, \theta_2 \ldots \theta_i \ldots \theta_k$ such that these functions contain all of the essential information given in the original data.

5. How To Indicate Whether or not Quality is Economic Standard

If an engineer could specify raw material and production processes so as to insure the quality of one piece of apparatus being identical with that of every other piece, the problem of the manufacture of apparatus in large quantities could be considerably simplified. However, it is of course, realized that we have little or no hope of obtaining such an ideal; instead it is to be expected that causes of variation in the quality of product will enter at every step in the fabrication process from raw material to the finished product. Appreciating that such causes of variation must of necessity enter in each and every one of the steps of production, we are confronted with the problem of determining if there be any criterion to guide us in our judgment of the economic significance of observed variations in quality.

Obviously it would be of little use to consider all of the ways and means of measuring the quality of product as we have just outlined above if it were not possible for us to set down certain conditions which should govern the variations in the observed measurements provided the quality of product is to be considered enonomically controlled. This means formally that, having chosen in a given case k different statistics $\theta_1, \theta_2 \ldots \theta_i \ldots \theta_k$ with which to measure the quality of product from month to month, we must naturally expect the observed values of these statistics to differ from period to period. We must therefore introduce some criteria by which to indicate the magnitude of the variations which may be expected in any one of these statistics, say θ_i, without these variations indicating need for concern about the quality of product. It is this general philosophy which lies back of the use of control charts in all of our inspection work. To make our present discussion of the measurement of quality of any practical value, we must therefore consider briefly the method of setting standards of quality and then present criteria by which to judge whether or not observed measurements of quality at a given period differ significantly from the chosen standard.

6. Presentation of Information of Greatest Service In The Improvement of Quality

After an inspection engineer has taken the above steps to determine the true quality of product in such a way that he can determine whether or not it is meeting economic standard, he is confronted with the even more important problem of showing why the quality is not up to economic standard provided he finds that it is not. In other words, the function of inspection engineering is far broader than that of merely measuring the quality of product, for after all if such measurements are to be of greatest use they must lead to improvements in quality or at least better and more economical methods of obtaining quality of a given standard. If quality is not satisfactory, it is by no means sufficient to have merely this indication; for such information to be effective it must come to the attention of the purchasing, design, development or research engineers so that one or more of these groups may introduce modifications of such a nature as to lead to the desired control of quality. But, before such information can become thus effective, it must indicate, wherever possible, the cause of the lack of control, which cause may be in the raw material, design, or some step in the development process.

Looked at in a certain way, quality production under present industrial conditions is but a gigantic laboratory experiment. When we appreciate how hard it is to control this experiment, we see that perhaps the greatest improvement in the service rendered by inspection engineering will come in the interpretation of the results of this great experiment in production in a way to give the most information to the research end of the organization including development, design and all other departments charged with the responsibility of making product of economic quality.

Of course the need for such information has long been realized as is evidenced by the extensive field tests which have been introduced in the study of many major items in a telephone plant. Perhaps too often disappointment has been felt at the results obtained by such tests. On the other hand it is generally appreciated that in our present state of scientific knowledge there remains a practical necessity of employing such extensive tests.

Quite naturally the problem of interpreting the measurements taken during any such field test is a very complicated one. Almost always the quality which we are interested in is related to certain other factors not in a functional but rather in a stochastic sense already referred to. In a large part, the methods of interpreting data obtained under controlled conditions in a laboratory cannot be used satisfactorily in the interpretation of the results of field tests. However, there is some hope that in the future we shall be able to obtain far more information than we have in the past. This naturally follows because remarkable strides have been made within the past few years in other than engineering fields toward the interpretation of results obtained under conditions where the contributing factors cannot be controlled, just as they cannot be controlled in the customary field test, and there appears to be no inherent reason why these same methods cannot be advantageously transferred to the solution of our engineering problems.

We shall now proceed with a detailed discussion of our subject.

PART II

Direct Measurement of Quality

Summary of Part II

Given a series of measurements, it is shown in Part II that customary error theory is not a satisfactory basis for the interpretation of data when the sample size is small; it is shown that in such cases we should tabulate two functions of these data, namely, the average and the root mean square deviation about this average.

The uses of these two functions in the interpretation of inspection data are then considered. It is shown that by using certain correction factors for the observed standard deviation we may estimate the expected probability associated with a given range about the average $\bar{X}$. This is of great interest in our inspection engineering work in the setting of standards.

It is also briefly indicated how these two functions of an observed set of data may be used in an estimate of the probability a posteriori that the average of the universe from which the sample was taken lies within a certain range subject to certain assumptions, although, in general, we have not found many inspection problems in the measurement of quality where such information can be readily applied.

For small samples we are limited to the use of the average and standard deviation as indicated above. When the sample size is large we can make use of the method of maximum likelihood to find the functions of the observed data which contain the essential information presented by the original set of data.

1. Direct Measurement of Quality Defined

Perhaps the simplest illustration of a direct measurement of quality is that of a length in terms of a unit of length. Suppose, for example, that you want to know the length X' of the line AB.

A ———————— B

The direct method of measuring this length is to take a unit of length and see how many times it is contained within the length AB. If you make a series of n measurements, let us say, $X_1, X_2, \ldots X_i, \ldots X_n$, in this way, it is not likely that all of these will be identical one with another. Now having made such a series of measurements, what will you give as the length of the line?

Being an engineer and familiar with the standard theory of errors, you will most likely answer that your best estimate $\bar{X}$ of the true length X' of the line is the arithmetic mean of the n observations, that is to say,

$$\bar{X} = \frac{\Sigma X}{n} .$$

You will go further and say that it is very likely that your estimate $\bar{X}$ is not identical with the true length X'. You will try to get around this difficulty by giving some estimate of the error of measurement, as you say, involved in your determination. If you follow the most prevalent practice you will likely say that the probable error of your estimate is $.6745\ \sigma_{1\bar{X}}$ where

$$\sigma_{1\bar{X}} = .8453 \frac{\sum_{i=1}^{n} |X_i - \bar{X}|}{n(n-1)} .$$

Now let us look at this matter a little more closely. Let us ask ourselves two questions:

a. Are these estimates of the true length and the error of our observation the most efficient estimates that we can use, that is to say, do they contain all of the relevant information presented in the set of data?

b. What do we really mean by probable error as calculated above?

Subject to the assumption that the errors are distributed normally, the answer to the first question is that the average is the most efficient estimate of X' but the estimate of the probable error given above contains only about 80% of the essential information contained in the data in respect to the error of an observation. As inspection engineers we cannot afford to throw away any information contained within our data because the cost of taking the data is a very important part of the cost of inspection engineering. Therefore, instead of the error of the estimate given above, we should use

$$\sigma_{\bar{X}} = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n(n-1)}}$$

subject to the limitations hereinafter set forth. Now let us consider the answer to the second question.

If σ' represents the true but unknown standard deviation of the normal universe of errors, then, as is well-known, 50% of the observed errors may be expected to fall within the range $X' \pm .6745\sigma'$. Now σ' is unknown, and we sometimes find the statement that 50% of the errors should fall within the range $X' \pm .6745\sigma$. Again we sometimes find the statement made that 50% of the observations should fall within the range $\bar{X} \pm .6745\sigma$ where

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}} .$$

Of course the statement made in terms of X' and σ' is true. At least one of the other statements must be false. It turns out that both are false as is shown by recent published work of these Laboratories.

Let us go one step further. We sometimes find the implication that by taking enough observations, that is, by making the number n large at will, we can get as close as we please to the true length of the line AB. More specifically, it is sometimes implied in the treatment of practical problems that the probability P that the error of the average will be less than any preassigned quantity ε may be made to differ from unity by less than any preassigned positive quantity δ by making n sufficiently large. This implication leads to ridiculous conclusions, because none of us believes that the length of the line pictured above is a real constant length. That very line we believe to be made up of a swarm of molecules jumping around in a random manner in such a way that the length of the line ceases to have meaning in the sense implied above.

Now, if we do not know exactly what we are talking about when we discuss the results of the measurement of the simplest kind, we cannot hope to know what we are talking about in the far more complicated problems which we must consider in the present bulletin. If, on the other hand, we think this problem through clearly and rigorously, it will keep us out of pitfalls in the consideration of those to follow. What we have in mind is admirably summed up in the words of the old song:

"If you would take an epsilon,
And I should take a delter,
Our dear old mathematics
Wouldn't be quite so helter skelter.
But if you make it easy,
By making it quite wrong,
You'll have to learn it all again -
'Twill take you twice as long."

Let us start with the least difficult problem mentioned above. Let us admit that the line AB does not have a true length X'. What then do we mean when we say that we measure the length of this line? In most instances we simply mean that X' is the expected value of X postulated for a given method of measurement, the errors of measurement being assumed to be distributed according to the normal law about the expected value X'. Under these assumptions, it can then be rigorously stated that the probability P that the following inequality will hold,

$$|X' - \overline{X}| < \varepsilon ,$$

can be made to differ from unity by less than any preassigned positive quantity δ by making the number of observations large at will. However, after all, this is really the only thing in which we are usually interested practically. In other words our conclusions are expressed in terms of measurements of a thing and not necessarily in terms of the magnitude of the thing itself. When we come to interpret what we mean by the error of measurement we are right up against a real problem. The interpretation of the significance of an observed error has been and still is the battleground of contending forces ever since the birth of probability theory.

Let us avoid as much trouble as possible and, to begin with, get down as close as possible to bed rock. In the first place, as we have said, the line AB does not have a true length in the sense often implied. No matter how many measurements we choose to make of the length AB, even though the errors are distributed normally, we can never come, through the use of probability theory, to the state where we are certain that our estimate $\overline{X}$ of the expected value X' is identical with X'. In other words, probability theory can tell us absolutely nothing unless we first make some postulate and probability theory can never be used backwards to prove that a postulate is correct. Enough for what we do not know and what we cannot say. Let us see what we can say. Let us take a practical problem and propose a practical answer and then discuss one by one the limitations of this answer.

1.1 Illustrative Problem

Five measurements of the length of a small piece-part gave the following:

3.242 cms.
3.243 "
3.240 "
3.241 "
3.238 "

1. What is the true expected value X'?

2. What is the error of measurement σ' assuming the error to be distributed normally?

We propose to answer these two questions by giving certain estimates and then show how to interpret the significance of these estimates. In tabular form we propose to give the following information:

No. of Measurements n	Estimate of X'	Observed Standard Deviation	Estimate of σ' (A) For use with normal law integral.	Estimate of σ' (B) For use with Student's integral.
n	$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$	$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n}}$	$\sqrt{\frac{n}{n-2}}\,\sigma = c_3\sigma$	$c_4\sigma$

The correction factor c_4 is derived in a manner to be explained after we have considered a little further the significance of this estimate of σ.

The use of the estimate $c_3\sigma$ is based upon the information originally presented in the paper, "Significance of an Observed Range", in the Journal of Forestry, November, 1928. It is obvious from theoretical considerations that the probability of an observation falling within the range $\bar{X} \pm t\sigma$ where $\bar{X}$ and σ are estimated from a sample should be less than that obtained through the use of the normal law integral. It is logical that no single constant value c_3 can be found such that the probability of an observation falling within the range $\bar{X} \pm tc_3\sigma$ should be given by the normal law integral.

Stated in another way, if we were to seek for that value of c_3 such that the expected probability associated with the range $\bar{X} \pm tc_3\sigma$ should be, say .5, we should expect to be able to find a value of c_3 to satisfy this condition. In the paper mentioned above, the following two values of c_3 were tried:

$$c_3 = \sqrt{\frac{n}{n-1}} \text{ and } c_3 = \sqrt{\frac{n}{n-2}}$$

for the particular case of n = 4. It was found there that for an expected probability of .5, the second value of c_3 gave .49 experimentally. This same value of c_3 for an expected probability corresponding approximately to the normal law 3σ gave .95 instead of .99+.

The experiment reported upon in the above paper, therefore, indicated that, if we use the first estimate (A) of σ', we can make use of this estimate to give us the expected probability associated with a given range subject to the limitations discussed in that paper. Furthermore, it was shown there that the same interpretation was roughly allowable for rectangular and triangular universes.

Now let us assume that we want to set down a range $\bar{X} \pm tc_4\sigma$ such that the expected probability associated with this range is a given value P. It is theoretically reasonable to expect that the c_4 corresponding to a given value of P can be determined from Student's integral in the following way. Enter Student's table[1] for a number of degrees of freedom two less than the size of the sample and find the value of t corresponding to the given value P. Then, for this particular value of probability P

$$c_4 = t\sqrt{\frac{n}{n-1}} \qquad (2.1)$$

Experimental results obtained in this laboratory and discussed in more detail in a book now in preparation seem to justify the use of this method.

Stated in another way, "Student's" integral can be interpreted as giving the expected probability that an average of a sample of size n will fall within the range

$$\bar{X} \pm t\sqrt{\frac{1}{n-1}}\,\sigma$$

where $\bar{X}$ and σ are the observed average and standard deviation in a sample of size n drawn from a normal universe.

We must pause, however, to note a very definite limitation to this use

1. Tables of Student's integral, Metron, Vol. V., No. 3, 1925.

of "Student's" integral imposed by the assumption that the universe is normal. For example, when we assume that it is normal, we assume that it is possible to observe magnitudes larger than any preassigned value. In other words, we assume that the universe extends from $-\infty$ to $+\infty$. In practical problems, however, we are confident that the range of possible values of X is limited. Hence as the probability P approaches unity, the value c_4 calculated from Equation 2.1 becomes unreasonably large.

However, the expected standard deviation for a small sample is, in general, less than the true standard deviation of the universe of errors of measurement, even though this universe is not a normal one. For this reason, some correction factor is always required. As pointed out in a recent paper (Reprint B-327) no integral of the "Student" type has been provided for other than a normal universe. We are, therefore, faced with a situation in which we have no theory to guide us in the choice of a possible factor similar to that presented in Equation 2.1 where the universe is not normal. It has been shown experimentally, however, in the paper already referred to (Reprint B-330), that the factor c_4 does not lead us very far astray even when applied to either a triangular or rectangular universe, provided at least we do not take values of P in excess of about .97.

Of course, recent work referred to in I.E.B. 1 makes it possible to express the moments of the distribution of variance (the standard deviation squared) of a sample of size n drawn from any universe in terms of the parameters of this universe. It is difficult, however, to obtain the exact frequency distribution of the variance from these assumed known moments such that it could be used in a manner analogous to that in which the normal law is used. Even if we were to use variance, we would be led into certain theoretical as well as practical difficulties which would throw grave doubt upon the significance of the small difference that would likely arise between the results obtained by using variance and those obtained by using the correction factor c_3 already discussed.

For the above reasons, it is felt that the factor c_3 used in connection with the normal law integral for values of probability P not in excess of .99 is to be recommended where, of course, the interpretation and limitations of this method as outlined in Reprint B-330 and in the above paragraphs are taken into account.

In what has gone before we have stressed particularly the fact that although we can never hope to know the expected value X' and the standard deviation σ' of the universe from which we are sampling, nevertheless we know quite accurately through extensive sampling experiments the expected probability associated with a given observed range, even though the universe is radically different from normal.

However, if we are interested in knowing the probability that the true mean of the sampled universe lies within a given range we may still use probability theory to answer this question, but the various assumptions that are made before a numerical result can be obtained must be kept clearly in mind in comparing this use of the theory of probability with the way we have just been using it.

More specifically we may sum up the subject matter contained in the last problem as follows:

A sample of n items has been drawn from a normal universe having an unknown mean X' and standard deviation σ'. Having looked at the sample, what is the a posteriori probability that the true mean X' lies within a given range say X_o to $X_o + dX_o$.

Now, the estimation of the a posteriori probability in question involves the calculation of two separate probabilities. (a) The a priori probability of getting from all possible normal universes those having means lying within X_o and $X_o + dX_o$, (b) For such normal universes and for any other possible normal universes, the probability of getting the observed sample.

Of course a knowledge of (a) and (b) involves knowing the a priori distributions of possible means and standard deviations, any pair of which would define a normal universe. However, knowing this and making certain definite assumptions, the a posteriori probability in question can be calculated and amounts essentially to this:

The ratio of the number of ways in which we could have obtained the observed sample from normal universes having means lying within X_o and $X_o + dX_o$ to the total number of ways the same sample could have been obtained from normal universes having means within all possible ranges.

Naturally the numerical answer depends entirely upon the nature of the general a priori assumptions made at the beginning and also upon the more detailed assumptions that we must make before a numerical result can be obtained.

Sometimes these assumptions can be rather easily justified from the nature of the practical problem but at other times they may reduce to something not far from pure guess work. In any case there is no absolute criterion for selecting from among the great number of possible answers the one that is the best.

2. Formal Statement of Problem - Large Samples

Let us assume that the errors are distributed in such a way that the probability dy' of an error x' falling within the range x' to $x' + dx'$ is a function f' of the error itself and m parameters. That is,

$$dy' = f'(x', \lambda_1', \ldots \lambda_i', \ldots \lambda_m')dx'. \tag{2.2}$$

We find ourselves confronted with two problems:

a. What is the law of error f'?

b. Knowing the law f', what are the best estimates of the parameters contained in Equation 2.2?

Having found a way of answering questions a and b, we may substitute the answers in Equation 2.2, and determine immediately the probability of a single observation lying within any specified range. However, to find what the error of the average is, we must express the law of error of the average in terms of the parameters in Equation 2.2. Now for large samples the error of the average can be shown to approach the distribution given by normal law irrespective of the parameters in Equation 2.2 provided none of these parameters are infinite, which condition we shall assume to be fulfilled in our practical problems.

3. Formal Solution of the Problem - Large Samples

Even when the samples are large we must first satisfy ourselves that the original set of data do not indicate the presence of any assignable causes of variation of the first type as can be done through the use of the four criteria presented in the first bulletin of this series, I.E.B. 1. A functional form f of f' is then assumed. Specifically, we assume that

$$dy = f(X, \theta_1, \ldots \theta_i, \ldots \theta_m)dX \tag{2.3}$$

is the probability of an observed value of X falling within the range X to X + dX. Assuming this law to be correct, the probability of getting the observed series of observations is obviously proportional to the following expression:

$$P = \prod_{i=1}^{n} f(X_i, \theta_1, \ldots \theta_i, \ldots \theta_m)dX \qquad i=1,2,\ldots m$$

One comparatively reasonable method of choosing the estimates θ's of the λ's is to take that set of estimates which make the probability P a maximum; in other words, to take that set of values for which

$$\frac{\partial P}{\partial \theta_i} = 0. \qquad i = 1,2, \ldots m$$

This formal method of obtaining the parameters is sometimes referred to as the method of maximum likelihood.

Obviously, however, this set of estimates of the parameters gives a maximum probability of occurrence of the observed set of data only provided our function f is correct. We need some criterion, in other words, to determine whether or not this assumption is justified.

This criterion has been provided by the work of Pearson in the so-called χ^2 test. If we break up the original range of observations into u different cells and let $n_1, n_2, \ldots, n_u$ be the observed frequencies in the u cells and then, if we calculate from Equation 2.3, the theoretical frequencies, $n_1', n_2', \ldots n_u'$, in these cells, the differences

$$n_i - n_i' \qquad i = 1,2, \ldots u$$

constitute a group of u correlated variables such that

$$\chi^2 = \sum_{i=1}^{u} \frac{(n_i - n_i')^2}{n_i'} .$$

From Elderton's original tables we can calculate the probability of obtaining value of χ^2 as large or larger than that observed, where, of course, we enter his table with the proper number of degrees of freedom. If this probability is very small, for example, less than .001, it is necessary to consider carefully whether or not we should look for some other function f. Thus briefly we have before us an outline of the machinery which is available for the solution of our problem in the case of large samples.

The work of Fisher in 1922 showed that the estimates of the parameters derived by this method of maximum likelihood for the case of large samples lead to sufficient and efficient estimates. The meaning of these two terms is as follows. A sufficient statistic is one such that it contains all of the essential information given by the original set of data for the estimate of a given parameter in the true law of error. The most efficient estimate of a parameter is one such that in large samples its error of estimate is less than that of any other estimate.

4. Illustration of Formal Method of Solution

Let us assume that the law of error is normal and that it is therefore given by the expression

$$P = dy' = \frac{1}{\sigma'\sqrt{2\pi}} e^{-\frac{(X - X')^2}{2\sigma'^2}} dX$$

where X' is the true expected magnitude and σ' is the true standard deviation of the error of measurement. Let us apply the method of maximum likelihood to determine the best estimates of X' and σ'. If this is the true law of error the probability of the observed set of observations is

$$P = \prod_{i=1}^{n} \frac{1}{\sigma'\sqrt{2\pi}} e^{-\frac{(X_i - X')^2}{2\sigma'^2}} dX_i \qquad i = 1, \dots n$$

Taking the log of both sides we have

$$\log P = -n \log \sqrt{2\pi} - n \log \sigma' - \frac{\Sigma(X_i - X')^2}{2\sigma'^2} \qquad i = 1, 2, \dots n$$

The relationship

$$\frac{\partial P}{\partial X'} = -\frac{1}{\sigma'}\left[(X_1 - X') + \dots + (X_n - X')\right] = 0$$

gives us the following estimate of X', namely,

$$X' = \frac{\sum_{i=1}^{n} X_i}{n} \qquad i = 1, 2, \dots n$$

or, in other words, the arithmetic mean.

Similarly, the relationship

$$\frac{\partial P}{\partial \sigma'} = 0$$

can be shown to lead to the following estimate of σ',

$$\sigma' = \sqrt{\frac{\Sigma(X_i - X')^2}{n}}$$

where we substitute for X' its estimate, the mean of the sample, given above.

Of course in the case of the normal law, the knowledge of X' and σ' gives us all of the desired information.

PART III

Indirect Measurement of the Quality of a Thing

1. Measurement when Form of Function is Known a Priori

Summary of Section 1

In the first section of Part 3 we start with the assumption that the quality Y to be measured is a known function f of a given set of measurable characteristics. We then show how to determine, wherever possible, those magnitudes of the measurable characteristics which make the error of measurement of Y a minimum. This provides a method of comparing different methods of measurement of a given quality to see which one inherently gives the smallest error of measurement. Naturally, other things being equal, that method having the minimum error of measurement will require the minimum amount of inspection to maintain quality within a given range.

This discussion shows very definitely what factors an inspection engineer needs to consider in respect to the method of measurement aside from those items which are customarily taken into account.

Let us assume that the quality characteristic Y to be measured is a known function f of $X_1, X_2, \ldots X_m$ such that

$$Y = f(X_1, X_2, \ldots X_m). \tag{3.1}$$

For example, in the measurement of modulus of rupture of a rectangular beam by the one-point loading method, we have the relationship

$$Y = \text{Modulus of rupture} = \frac{cFL}{bh^2}$$

where c is a constant, F is a force, L is the length between supports, b is the breadth and h is the height of the rectangular beam. In all such cases Y, here the modulus of rupture, is a quantity which we must measure indirectly through measurements of certain other physical characteristics, in this particular example F, L, b and h.

Theoretically speaking, of course, we might make F, L, b and h any value whatsoever. Practically, we do not have such a broad choice, first, because there is only a limited range of possible values of rectangular beams of a given species of timber and second, there are, in general, limitations involved because of the excessive cost of the machines required to measure extremely large beams. Assuming for the time being, however, that modulus of rupture is a constant for a given species of timber and that there are no errors of measurement in any one of the characteristics F, L, b and h involved in the above functional relationship, it would be possible to obtain the modulus of rupture of a given species of timber accurately by using one size of beam just as well as by using any other size.

In practice, however, none of the factors F, L, b or h can be measured with absolute precision. In other words, measurement of any one of these factors is subject to certain errors and the resultant error of measurement in modulus of rupture is a function of the errors of measurement of the characteristics involved in the functional relationship.

In any practical problem, therefore, we must express the quality Y as a function of the expected or average values of the X's. Suppose, for example, that we make a single measurement on each of the m different characteristics represented by the X's. Making use of Equation 3.1, we would obtain a value for the quality Y. However, since X is subject to error, Y will also be subject to error. In other words, if we let

$$Y = Y' + y$$

and

$$X_i = X'_i + x_i \qquad i = 1, 2, \ldots m$$

we have as a result of a given set of observations

$$Y' + y = f(X'_1 + x_1, \ldots, X'_i + x_i, \ldots, X'_m + x_m)$$

where y and x_i are errors of measurement. In any given experiment, naturally, we do not know the true values represented by the primes and hence we customarily take the relationship

$$\overline{Y} = f(\overline{X}_1, \overline{X}_2, \ldots \overline{X}_i, \ldots \overline{X}_m) \qquad (3.2)$$

as forming the basis of measurement, where the $\overline{X}$'s are the average values found from measurement, and $\overline{Y}$ is the quality resulting from these average values of the characteristics.

Our problem, therefore, is to express the error of measurement of Y in terms of the standard deviations $\sigma_1, \sigma_2, \ldots \sigma_i, \ldots \sigma_m$, of the errors of measurement in $X_1, X_2, \ldots X_i, \ldots X_m$. To do this let us assume that Equation 3.2 can be expanded in a Taylor's series about the point $\bar{X}_1, \bar{X}_2, \ldots \bar{X}_i, \ldots \bar{X}_m$. As a first approximation we have

$$y = \sum_{i=1}^{m} a_i x_i$$

where

$$a_i = \left(\frac{\partial f}{\partial X_i}\right)_{\bar{X}_1, \bar{X}_2, \ldots \bar{X}_m} \qquad i=1,2,\ldots m,$$

$$y = Y - \bar{Y},$$

and

$$x_i = X - \bar{X}_i.$$

Under these conditions it is well-known that the standard deviation σ of the measured quality characteristic Y is given by the following expression

$$\sigma = \sqrt{\sum_{i=1}^{m} a_i^2 \sigma_i^2}\ , \qquad (3.3)$$

σ_i being, as stated above, the standard deviation of X_i.

Thus σ can be expressed as a function of the m mean values and our problem is then the purely mathematical one of making σ a minimum with respect to the expected values $\bar{X}_1, \bar{X}_2, \ldots \bar{X}_i, \ldots \bar{X}_m$. Before proceeding further, however, we must note that only m-1 mean values can be considered as independent variables because the remaining one must be determined in such a way as to satisfy Equation 3.2.

For purposes of illustration, then, we may solve for $\bar{X}_m$ from Equation 3.2 and substitute its value in Equation 3.3, thus obtaining an expression for σ in terms of $\bar{X}_1, \bar{X}_2, \ldots \bar{X}_i, \ldots \bar{X}_{m-1}$ and the constant value $\bar{Y}$.

The necessary condition that σ be a minimum now becomes

$$\frac{\partial\sigma}{\partial\bar{X}_i} = 0. \qquad i = 1,2,\ldots(m-1) \qquad (3.4)$$

These m-1 equations[1] together with Equation 3.2 are sufficient to determine the m unknowns $\bar{X}_1, \bar{X}_2, \ldots \bar{X}_i, \ldots \bar{X}_m$.

Naturally, it may turn out that no practical set of values $\bar{X}_1, \ldots \bar{X}_m$ can be found to make σ a minimum. Thus, some of them may turn out to be zero, infinite or imaginary, all of which is of little value in practice unless σ is known to be a monotonic decreasing function of the m different characteristics and continuous in the neighborhood of the $\bar{X}$'s determined above. In this case, the choice of a practical set of $\bar{X}$'s which do not differ greatly from the determined set may yield a value of σ not greatly in excess of its minimum value.

The set of equations 3.4, however, form the first criterion for judging whether or not the method of measurement by means of a known functional relationship has been set up so as to provide a minimum error of measurement.

In the general case, we may know of other functional relationships between the quality Y to be measured and certain other quantities. Thus, we might have

$$Y = f_1(Z_1, Z_2, \ldots Z_m)$$

and similar expressions. A practical example is given by the three formulas expressing the modulus of rupture as already noted. Our problem then becomes that of minimizing, if possible, the overall errors by each method and then determining which of the formulas gives the smallest error at a minimum of cost.

<u>Application of Theory to Practical Problem</u>

<u>Problem 1</u> - In determining the modulus of rupture of pole timber we have the following three formulas for calculating the modulus of rupture depending upon the type of mechanical test used.

a. One point loading: $M = \dfrac{69.62\ L_1F_1}{C_1^3}$

b. Two point loading: $M = \dfrac{52.64\ L_2F_2}{C_2^3}$

c. Cantilever method: $M = \dfrac{315.8\ L_3F_3}{C_3^3}$

---

1. Obviously we might have solved Equation 3.2 for any one of the m variables in terms of the other m-1 and have substituted its value in Equation 3.3. Equation 3.4 could then have been modified to correspond to the particular choice of independent variables assumed. However, the final solution of the m expected values $\bar{X}_1, \bar{X}_2, \ldots \bar{X}_i, \ldots \bar{X}_m$ would in all cases be the same no matter [illegible] independent.

As before F is the force required to break the pole, L is the distance between supports in a and b and in c the distance from point of application of load to the point of failure, C is the circumference of break section and M is an abbreviation for modulus of rupture. The subscripts have been attached to the force, length and circumference because they are not in general the same in all three cases.

If we are going to break poles to determine the modulus of rupture of a certain species of pole timber, which formula will yield the smallest overall error in the determination of this modulus?

We may of course treat this case according to the above general mode of procedure considering F, L, and C as variables subject to the condition that

$$M = \frac{KLF}{C^3}$$

and arrive at a best value of F, L, and C corresponding to the minimum error in M. However, we may affect a simplification of our problem by taking note of the following practical consideration. In most cases the taper of the pole is small. Hence, we may without serious error assume that C is constant for the three methods and proceed to find the most efficient way of determining the modulus of rupture as already outlined in general.

This condition being fulfilled we may represent all three formulas for determining modulus of rupture by a single equation,

$$M = K_i L_i F_i, \qquad (i = 1, 2, 3)$$

where

$$K_i = \frac{a_i}{C^3} .$$

Proceeding as in the general theory we find

$$(\sigma_M)_i = \sqrt{K_i^2 F_i^2 \sigma_{L_i}^2 + K_i^2 L_i^2 \sigma_{F_i}^2}$$

$$= \sqrt{\frac{M^2}{L_i^2}\sigma_{L_i}^2 + K_i^2 L_i^2 \sigma_{F_i}^2} \qquad (i=1,2,3) \qquad (3.5)$$

Differentiating with respect to L_i and setting this derivative equal to 0, we find on solving

$$L_i = \sqrt{\frac{M\sigma_{L_i}}{K_i\sigma_{F_i}}}$$

from which

$$F_i = \sqrt{\frac{M\,\sigma_{F_i}}{K_i\sigma_{L_i}}} \qquad (i = 1, 2, 3).$$

Substituting these values of F_i and L_i in equation 3.5 we find

$$(\sigma_M)_i \text{ (min.)} = \sqrt{2K_iM\sigma_{F_i}\sigma_{L_i}} \ .$$

Now, if

$$\begin{cases} \sigma_{F_1} = \sigma_{F_2} = \sigma_{F_3}, \\ \sigma_{L_1} = \sigma_{L_2} = \sigma_{L_3}. \end{cases}$$

then the method giving the least error in the determination of M is the two point loading method, i.e., the one having the smallest K.

If, however, σ_{F_1}, σ_{F_2}, σ_{F_3} are not all the same, as is probably the more correct assumption, then we should choose that method which makes σ_M a minimum when these three σ's are known.

Problem 2 - Now consider the modulus of rupture as determined from small pieces. Here again we have three formulas depending upon the type of mechanical test used and these three may be combined into a single formula as follows:

$$m = \frac{a_if_il_i}{b_ih_i^2} \qquad (i = 1, 2, 3)$$

where of course a_i is different from the a_i used in case of poles. As before m is the modulus of rupture of the species, f and l have the same meaning as before, b is the breadth and h the height of the beam tested.

If we try to apply the general method of minimizing σ_m with respect to all four variables, we find that the problem does not have a real solution. In other words this means that for no positive set of values of f, l, b and h does σ_m have a minimum value.

In such a case we may use cut and try methods to decrease σ_m but having arrived by this process at certain limiting values for the four variables, all that we can say is that σ_m for this set is less than it is for certain other sets that we have tried. However, by making as before certain assumptions about some of the variables, we can obtain a quite definite result.

In the practical case the size of the machine may be an important economic factor in the sense that, if large pieces are to be used, the machine itself must be made sufficiently large. Hence, let us assume that a machine is chosen which makes it necessary to keep the force f and length l the same for each method of test. Under these conditions, what dimensions should b and h have in each of the three methods of test to give minimum error in measurement of modulus? Which method is capable of measuring the modulus with least error?

Under these conditions we may set

$$m = \frac{k_i}{b_i\ h_i^{\ 2}}, \qquad (i = 1,\ 2,\ 3).$$

where

$$k_i = a_i f_i l_i.$$

Proceeding in exactly the same way, we find

$$(\sigma_m)_i = k_i \sqrt{\frac{\sigma_{b_i}^{\ 2}}{b_i^{\ 4}\ h_i^{\ 4}} + \frac{4\ \sigma_{h_i}^{\ 2}}{b_i^{\ 2}\ h_i^{\ 6}}},$$

$$= k_i \sigma \sqrt{\frac{1}{b_i^{\ 4}\ h_i^{\ 4}} + \frac{4}{b_i^{\ 2}\ h_i^{\ 6}}} \qquad (3.6)$$

provided we assume, as it is reasonable to do, that $\sigma_{b_i} = \sigma_{h_i}$, and both have a value σ.

Now substituting in Equation 3.6 the value of b_i in terms of h_i and setting the derivative with respect to h_i equal to zero, we find

$$h_i = \sqrt[3]{\sqrt{2}\ \frac{k_i}{m}} \text{ and } b_i = \frac{k_i}{m\left(\frac{2k_i^{\ 2}}{m^2}\right)^{2/6}} = \sqrt[3]{\frac{k_i}{2m}}$$

Hence the minimum value of σ_m for the i^{th} method is

$$k_i\ \sigma \sqrt{\frac{m^4}{k_i^{\ 4}}\left(\frac{2\ k_i^{\ 2}}{m^2}\right)^{4/6} + \frac{4\ m^2}{k_i^{\ 2}\left(\frac{2\ k_i^{\ 2}}{m^2}\right)^{2/6}}} = \sigma\ \frac{m^{4/3}}{k_i^{\ 1/3}} \sqrt{2^{2/3} + 2^{5/3}}$$

This tells us at once that, subject to the assumptions previously set down, that method will give minimum error for which k_i is the greatest and furthermore that the cross sectional area for minimum standard deviation for a particular method is $\frac{(k_i)}{(m)}$. Of course, even in this case, it may not be practical to give b and h the values which make the standard deviation of measurement a minimum. Be that as it may, however, the above discussion illustrates the details of the methods available in comparing indirect methods of measurement in respect to their inherent errors of measurements

2. Measurement When Form of Function is Unknown a Priori

Summary of Section 2

This section of Part 3 outlines the method of procedure to be followed in minimizing the error of measurement of quality Y which quality is determined indirectly by measurement of certain characteristics assumed to be functionally related to the quality in an unknown manner. The problem divides itself into two parts: (a) the assumption of a functional relationship to be used as a basis of measurement; (b) the testing of this functional relationship.

The results are applied to the problem of minimizing the error of measurement of tensile strength of nickel silver sheet in terms of a measurement of hardness and gauge number. In this particular instance it turns out that the reduction in error of measurement over that of the customary method of measurement is very appreciable indeed and in this way indicates the nature of investigations to be pursued in the consideration of such things as specifications on raw materials.

In many practical cases we are not fortunate enough to know the relationship connecting the variables, and even though we postulate that there is a functional relationship between a set of X's and the quality Y which we wish to measure, we must find empirically the nature of this functional relationship through a study of observed values.

To obtain a functional relationship f between the two variables, we must make a guess or hypothesis as to what the relationship actually is, and then test this hypothesis. Before proceeding further, we shall outline the formal method of attack.

Let us assume that we have a series of n observed values of X and Y.

Assume that the function f is expressible in terms of a Taylor's series so that it can be written to a first approximation as a polynomial

$$Y = \sum_{i=0}^{i=c} a_i X^i \qquad (3.7)$$

where c is the number of terms which it is necessary to retain in order to make our polynomial a sufficiently close approximation to f. Under these conditions, if we substitute in Equation 3.7 each of the n pairs of values of X and Y, we obtain n different equations. If as in the usual case, n is greater than c+1, we cannot in general solve the n equations for the c+1 coefficients represented by the a's. What we must do, therefore, is to reduce this series of n equations to another of c+1 equations, and one way of doing this is to determine the a's in such a way that the sum of the squares of the errors in Y is made a minimum.

To this end, let x_j be the error in Y for $X = X_j$ i.e.,

$$x_j = Y_j - \sum_{i=0}^{i=c} a_i X_j^i$$

and let

$$u = \sum_{j=1}^{j=n} x_j^2$$

Then the condition that u, the sum of the squares of the errors be a minimum with respect to a_0, a_1, ...a_c is

$$\frac{\partial u}{\partial a_i} = 0 \qquad i = 0,1,2.....c \qquad (3.8)$$

This gives us the desired number of equations, the solution of which will, in general, lead to unique determinations of the coefficients represented by the a's.

It will be recalled that the solution which we have thus obtained depends upon the assumption that the deviations, the sum of whose squares are to be minimized, are those in the dependent variable Y. Of course, we might choose to minimize the sum of the squares of the perpendicular distances of the points from the assumed curve. This process could be readily carried out provided the functional relationship is a linear one, but this leads into a discussion which will be given elsewhere.

Practical Applications

Figure 3.1 shows tensile strength of nickel silver sheet plotted against hardness for gauges 14 and 36. It appears that the curve for gauge 14 could be represented by a parabola or second degree polynomial

$$Y = a_0 + a_1X + a_2X^2 \qquad (3.9)$$

and for gauge 36, by a linear relationship

$$Y = a_0 + a_1X \qquad (3.10)$$

Having once made this choice of function, the formal process of fitting these two curves by the least square method outlined above can be easily carried through and we will come out with the equation

$$Y = 290251.5471 - 7000.0807X + 52.6559X^2$$

for gauge 14 and

$$Y = -505844.0315 + 7307.9006X$$

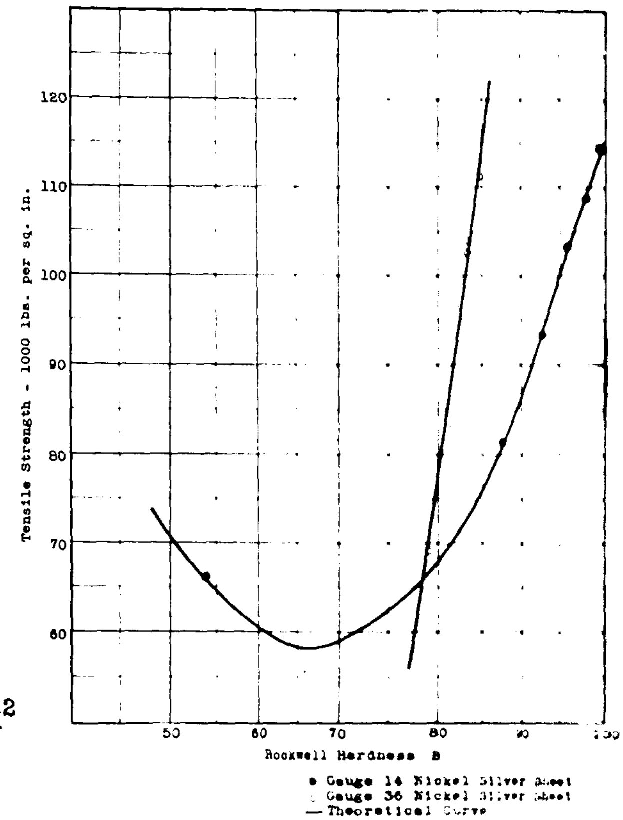


FIG. 3.1

for gauge 36. Computing the values of tensile strength for the six given values of hardness we get the smooth curve and straight line shown in Figure 3.1.

To the eye, these two curves seem to fit their respective points very closely, particularly the parabola. In fact, computing the standard deviation of the points from each of these curves, we find

$$\sigma_{(parabola)} = 583 \text{ lbs. per sq. in. and}$$

$$\sigma_{(line)} = 4315 \text{ lbs. per sq. in.}$$

A mathematical explanation for the large standard deviation from the straight line can be offered in the fact that even though the points lie very close to the fitted line, their vertical deviations from this line are large, particularly since the slope of the line is large. Physically, however, it may mean one of two things: (a) for a given hardness, the standard deviation of tensile strength increases with increasing gauge or (b) the relationship between tensile strength and hardness for gauge 36 is not linear but some kind of curve. Present available data are not sufficient to tell whether or not the former statement is true. Concerning the latter, the method for judging how well a particular functional relationship fits the data is taken up later. At present, we are primarily concerned with a discussion of the resulting determination of tensile strength on the assumption that the curves we have chosen to use are the correct ones.

Let us see then what degree of accuracy may be obtained in using these curves to measure tensile strength by making use of the curve for 14 gauge, Fig. 3.1

Assuming that tensile strength is homoscedastic with respect to hardness for a given gauge, i.e., the standard deviation of tensile strength from the curve is the same for all values of hardness, we obtain a very accurate measure of tensile strength. Thus, if we select a value of observed hardness, say X_1, we may associate with it a value of tensile strength given by the curve and the range of variation in tensile strength corresponding to a probability of about .99 is about 3600 lbs. per sq. in. this range being equally divided on either side of the curve value. A similar interpretation holds for the straight line corresponding to 36 gauge, only here the measure of tensile strength is must less accurate, the 99% error range being about 13000 lbs. per sq. in. on either side of the line.

These error ranges are much less than those previously obtained and hence it is now possible to set closer quality limits by means of the Rockwell Hardness Test than was previously possible without fear of throwing out satisfactory material. .This result is typical of what it is reasonable to believe can be done in many other cases of setting quality limits for raw materials.

Thus we see that, by applying the method of least squares in the previous paragraph, we are able to establish an empirical relationship between the quality Y, in this case tensile strength, and a certain measurable characteristic X, in this case hardness. What we need now is some method for indicating how good this empirically determined relationship really is.

First Criterion for Judging Method of Measurement

We shall first show that assuming our guess at the function f is correct then the coefficients determined as above by the method of least squares are the best in the sense that they make the likelihood of the occurrence of the given set of values a maximum, subject to the condition that the errors are distributed normally.

Assume that

$$Y = f(X)$$

gives us the smooth curve in Figure 3.2 and that the observed deviations are as indicated in this figure.

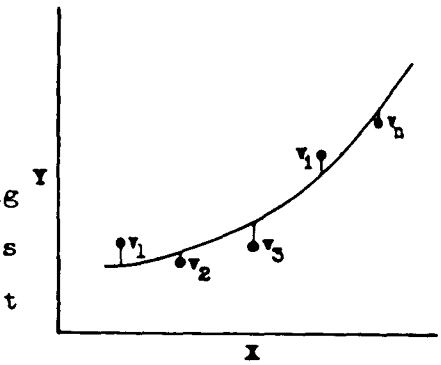


FIG. 3.2

where the i^{th} method deviation v_i is given by

$$v_i = Y_i - f(X_i)$$

Let us consider the probability of obtaining the set of deviations $v_1, v_2, \ldots. v_i, \ldots. v_n$ in a series of n observations. In the first place, we assume that there is no correlation between the deviation in one observation and that in another or in other words, that the v's constitute a non-correlated system of variables. Let us assume further that the deviations are distributed normally. Under these conditions, the probability of getting the observed set of values is

$$P = C e^{-h^2 \sum_{i=1}^{n} v_i^2} \qquad (3.11)$$

where C and h are constants.

The most likely set of values of the v's is obviously that which makes $\sum_{i=1}^{i=n} v_i^2$ a minimum, but the very process of obtaining the empirical relationship by the above method involves making the sum of the squares a minimum. Hence, under the assumptions which we have made concerning the v's, the coefficients in the equations already determined by the method of least squares are the same as those which would have been obtained by the above method of making the probability of the observed set of deviations a maximum.

This, however, does not necessarily mean that the chosen function f is the best. It is only best in the sense that, having chosen the form of the function, the coefficients entering into the first approximation of this function and derived from the data are such as to make the probability of the observed set of deviations a maximum.

Of course, one might at first suggest that we should try to find that function f which would make the sum of the squares of the deviations a minimum. Obviously, to follow such a suggestion would be a false move, because we could make this function of the deviations zero by retaining sufficient terms in our approximation, i.e., as many (constants in our equation) as we have pairs of observations. However, it is ridiculous to assume that the relationship thus determined is the true one. In other words, we need some method for allowing for the fact that another similar set of observations would, in general, not be identical with the observed set. Stated otherwise, we must allow for the fact that the closeness of fit as measured by Σv^2 must be corrected for the effect introduced by changing the number of terms retained in the function f(X).

By making enough observations of each value of X to determine the standard deviation of Y at each point, we can make use of Pearson's Chi Square Test provided we allow for the loss of the correct number of degrees of freedom taken up in fitting the function F(X). Suppose, for example, that Y is measured at u values of X and that $\sigma_1, \sigma_2, \ldots \sigma_i, \ldots \sigma_u$ are the respective standard deviations determined by measurement at these points. Since there is no correlation between the deviations in the observed values of Y for one value of X and those for another, we have

$$\chi^2 = \sum_{i=1}^{u} \left(\frac{\left[Y_i - f(X_i) \right]^2}{\sigma_i^2} \right) \qquad (3.12)$$

This value of χ^2 has u-m degrees of freedom where m is the number of parameters involved in the function F(X). When we calculate the probability $P(\chi^2)$ of the occurrence of as large or larger values of χ^2 from Elderton's tables allowing for the right number of degrees of freedom, we are supposed automatically to take account of the effect of increasing the number of coefficients used in fitting the function F(X).

The probability of fit $P(\chi^2)$ determined in this way from one function F(X) may be compared with that obtained in a similar way from another function subject to limitations as to interpretation discussed later.

Second Criterion for Judging Method of Measurement

The approximate empirical relationship between Y and X may, in general, be put into the form

$$Y = a_0 F_0(X) + a_1 F_1(X) + \dots + a_i F_i(X) + \dots + a_m F_m(X)$$

where $F_0(X) = 1$, $F_i(X)$ is a polynomial of degree i in X, i taking on integral values between 1 and m and the functions of X are orthogonal in the sense that

$$\sum_{k=1}^{k=n} F_i(X_k) F_j(X_k) = 0, \qquad (i \neq j).$$

Under these conditions, R. A. Fisher has shown[1] that the probability of getting a value of t as large or larger than that observed is given by "Student's" integral previously referred to, where t expresses the error of a coefficient a_i and is calculated in the particular way indicated in the cited reference.

If, in general, the probability of getting a value of t as large or larger than that observed is less than .001, we have reason to believe that there may be a better choice of regression coefficients. Of course, in such a case we must decide whether it is more likely that the regression coefficient is in error than that the functional form is not properly defined. This, however, is merely a technical point which must be considered by the analyst and need not concern us at the present time. Here, as in part II, extreme caution must be exercised in interpreting the significance of the test just given particularly when a_i is not known.

It is sufficient to note that we have in this case a method of testing the coefficients in an empirically determined formula subject to the conditions outlined above even though it is not feasible to group the observations into certain classes as we shall do later where we do not assume the existence of a mathematical relationship but instead assume the existence of a stochastic relationship between the quality Y and the characteristic X in terms of which it is measured.

We shall not go into the details of applying this test to a practical

---

1. Fisher, R.A. Application of "Student's" Distribution Metron Volume V, No. 3, Pages 3-18, 19-25.

example but before we leave the subject we should consider a little further the importance of what might be termed a third method of testing a functional relationship. In the practical example already cited where we try to find the most satisfactory relationship for measuring tensile strength in terms of hardness for #14 gauge nickel silver sheet, we present a parabola which gives quite a pleasing appearance of fit to say the least. Now even though the two tests just discussed indicated this parabola to be a good fit, it is dubious whether we should consider it so looked at from a physical point of view. Common sense suggests that the results of the Rockwell hardness test should be influenced by the roughness of the surface of the material. Now the lowest point on the 14 gauge curve corresponds to measurements on stock of zero per cent reduction. Hence it is reasonable to believe that this point should not be included with the others. This leads to a simplification in the functional form required to fit the remaining points. In fact it appears that a straight line may be used if this one point is eliminated and information given later seems to indicate that instead of a parabola, two straight lines should be used to cover the range for 14 gauge, if we are to include material of zero reduction by rolling.

In what we have just said however, we have brought into the picture the factor of reduction by rolling which information, of course, is not taken account of when we consider only the tensile strength values corresponding to certain hardness numbers for 14 gauge.

In other words, we should never rely upon the two tests involving the use of either χ^2 in fitting a mathematical curve to observed points or of t in determining the error of a coefficient in an assumed functional relationship unless we have first made use of all available a priori information.

3. Measurement Through Stochastic Relationship

Summary of Section 3

In the third section of Part 3 we introduce the concept of stochastic relationship as a basis of measurement and give an outline of the steps to be taken in making use of it in the measurement of quality with minimum error. Specific applications of the principles are made to the testing of tensile strength of nickel silver sheet in terms of hardness and other characteristics presented in one of the recent LRM Bulletins and to the measurement of the physical properties of aluminum die-castings. Similar applications are found in practically every phase of our work.

In general, it is shown that a very appreciable reduction in the error of measurement can be obtained by making the best use of information at hand. Hence this part of the bulletin furnishes a background for the development of improved methods of measuring quality, particularly of raw materials, so as to make possible a much closer quality control without increasing the cost of inspection.

<u>Stochastic Relationship Defined</u>

Let us start with a very simple illustration, namely that of measuring the tensile strength in terms of hardness where all that we know about this relationship is the set of observed data presented graphically in Fig. 3.3.

We see that, in general, the harder this kind of material is, the stronger it appears to be. For a given observed hardness, however, the strength is not always the same. In other words, hardness does not determine one and only one value of tensile strength. Naturally this state of affairs is somewhat attributable to the fact that neither hardness nor strength can be measured without error. In this particular case, however, we have plenty of a priori evidence to indicate that tensile strength is not uniquely determined by hardness even when both the tensile strength and hardness are measured without error. We say that Y, the tensile strength, is stochastically related to hardness X in the sense that hardness tells us something about tensile strength. Let us amplify the distinction between stochastic and functional relationship.

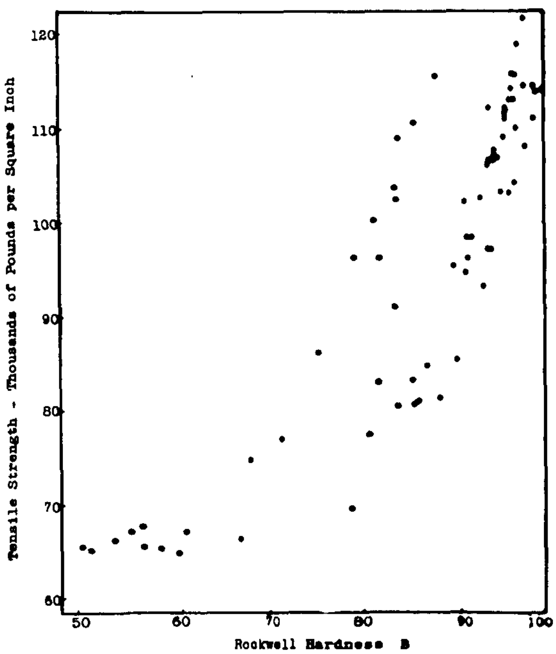


FIG. 3.3

To do this, let us consider that we have a series of n pairs of values

X_1Y_1, X_2Y_2, ... X_nY_n. We say that Y is a function of X when for a given value of X the possible values of Y are uniquely determined. The simplest kind of functional relationship is that for which there is a one to one relationship so that for any value of X there is one and only one value of Y, this value being thus uniquely determined. In general, of course, the functional relationship may be a multiple valued one, i.e., for a given value X_i, Y may take on say k different values $Y_{1i}, \ldots Y_{ki}$. However, instead of having a functional relationship, we may have the situation represented by the ordinary scatter diagram where certain pairs of numbers occur more than once. In this case, all that we can say is that for a given value of X, let us say X_i, there is some function of Y, let us say $\varphi_i(Y)dY$, which represents approximately the number of Y's falling within the range of Y to Y + dY when X takes on a particular value X_i. The quality Y is said to be stochastically related to the quality X when the frequency distribution $\varphi(Y)$ corresponding to a given value of X is not the same for all values of X. In the ordinary scatter diagram, this merely means that the frequency distributions in the columns are not all the same.

In a later bulletin we shall have occasion to consider more in detail the physical significance of the difference between functional and stochastic relationship. Here we shall see that Y can be thought of as being functionally related to X or, in fact, any number of variables when these variables are the independent causes of Y, whereas Y can be thought of as being stochastically related to X or any number of variables when this set of variables cannot be considered as the independent causes of Y. For the present, however, we shall be interested in the use of stochastic relationship as a basis of measurement and not in the physical explanation thereof.

Solution of the Problem

Without further discussion let us consider the steps which must be taken in the use of a stochastic relationship in the measurement of quality. In our discussion of this problem it will be necessary to make free use of statistical terms used in the theory of correlation.

We start as already noted above with the formal expression of the stochastic relationship representable in the following way

$$Y = f_s\ (X_1,\ X_2\ \ldots\ X_n).$$

Knowing that such a stochastic relationship exists we must in general take three steps:

(a) Postulate a curve or surface of regression.

(b) Apply certain tests to determine whether or not the values of the quality Y are distributed in a homo-scedastic fashion about the curve or surface of regression.

(c) Apply certain tests for determining whether or not the observed set of data is a likely sample based upon the assumed stochastic relationship and the assumed form of regression.

After these three steps have been taken we must carefully consider whether or not there are any reasons a priori for discounting the use of the curve or surface of regression in determining the expected value of Y associated with the given set of n characteristics represented by $X_1,\ X_2\ \ldots\ X_n$. Furthermore, we must make sure that there are no practical difficulties in the way of using the curve or surface of regression. Assuming that the three steps have been taken and that there is no reason for rejecting the results so obtained, we must proceed to calculate the standard deviation of the quality Y about the curve or surface of regression obtained as a result of the first three steps.

Although it is beyond the scope of the present discussion to enter into a consideration of all of the details involved in taking the first three steps outlined above, it is nevertheless advisable to consider a little further the nature of the procedure which must be followed in practice. For example, because of practical difficulties involved in the calculation we are generally limited to the use of curves of regression representable by a series of orthogonal functions as discussed in a previous paragraph. For several reasons we are customarily limited to the use of planes of regression.

Customary methods may be used for testing the homo-scedasticity of the regression provided we allow for the effect of small samples upon the expected value of the standard deviations. The principles underlying such corrections have already been sufficiently well considered in Bulletin IEB-1 and Part 2 of the present bulletin.

Now we come to a little more careful consideration of step 3. We choose a form of regression curve or surface and by one method or another (in most cases the method of least squares) we determine the coefficients in this assumed functional curve or surface of regression based upon the observed data. We may use either of the methods discussed under paragraph 2 above for determining whether or not certain functions of the observed data are reasonable upon the basis of the assumed functional form of the curve or surface of regression. We may go even further and consider the use of certain other tests or criteria. We shall consider one of these which may be used in the case of linear regression[1]. In this case χ^2 becomes

$$\chi^2 = (n - u)\frac{\eta_{YX}^2 - r^2}{1 - \eta_{YX}^2} .$$

It is generally assumed that χ^2 has u - 1 degrees of freedom. Fisher has also given other methods[2] of estimating χ^2.

Choosing that method of estimating χ^2 most suitable to the problem in hand and following the scheme outlined by Fisher, we could calculate the probability of obtaining as large or a larger value of χ^2 upon the basis of the assumed form of curve of regression. This method could naturally be extended to the plane of regression. If the probability of obtaining as large or larger value of χ^2 than that observed is less than .001 we might assume that the use of t is curve of regression was not justified and in that case, continue our search for another. In the interpretation of the results of this criterion however, we must proceed with extreme caution in its use.

As above indicated, we are treading upon the kind of shifting logical sand that we have already met in Part II. In this interpretation we must remember that as a limiting case corresponding to one degree of freedom the χ^2 test breaks down to the normal law and we have already seen how carefully we must hedge our conclusions in respect to the probability associated with any specified range about the observed average measured in terms of the observed standard

---

1. Fisher, R.A. "The Goodness of Fit of Regression Formula", Journal of Royal Statistical Society, Volume LXXV, July, 1922, Page 603.

2. These will be discussed in detail in Part III of the book on Quality Control now under preparation.

deviation. Not only must we be cautious on this ground but also we must take account of the fact that a hypothesis which is not true might give a smaller value of χ^2 with the observed data than a hypothesis which is true. Furthermore, there are other limitations to the use of χ^2 which must be carefully considered.

With these facts in mind we might be tempted to say that the test is of little value for the purpose in hand. But, before we jump to this conclusion, let us remember that if we throw out the χ^2 test, we also throw out the normal law.[1] Let us not therefore throw away the test completely unless we have something better to put in its place.

Practical Applications

Obviously the field of application of the results of this section of Part III as well as of the previous two sections, is very extensive indeed. This method of attack has already been applied in the testing of physical properties of timber and those of certain other raw materials such, for example, as are discussed in some of the LRM bulletins. The initial results which have been obtained to date indicate that in certain instances the above methods when applied to available data reduce the standard deviation of measurement in some instances by more than 50% of that obtained by the customary method and thus make possible without additional labor the control of quality within limits which may be less than 50% of those now in use. Naturally, these results suggest a very important field of investigation in the development of methods for specifying the quality of raw materials and product in general within narrower limits without increasing the cost of inspection. In what follows we shall not detail all of the steps.

We take as a first illustration the measurement of quality, tensile strength of aluminum die-castings, in terms of density and hardness. The data for this study were furnished by Mr. H. A. Anderson, Chairman of the Committee on aluminum die-castings of the American Society for Testing Materials and are presented in the two scatter diagrams of Fig. 3.4.

One of these shows the stochastic relationship between tensile strength and hardness in Rockwells of sixty die-castings. The other scatter diagram shows the corresponding relationships between tensile strength and density for these

- -

1. See Part III of the book on Quality Control.

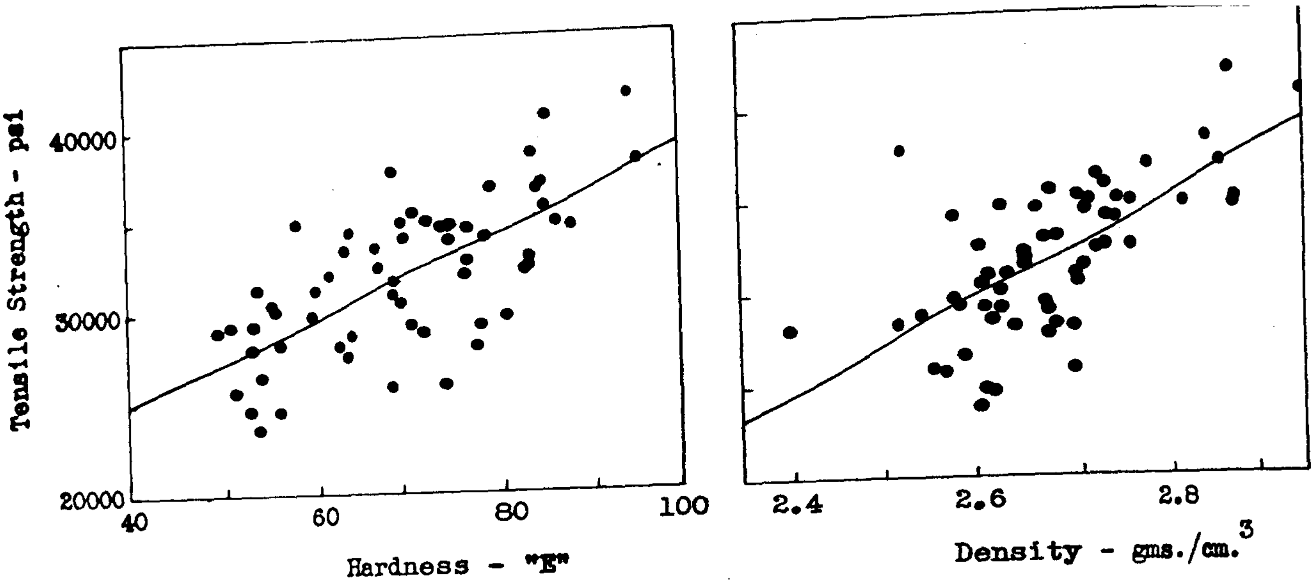


FIG. 3.4 - ALUMINUM DIE CASTING SPECIMENS

same specimens. The lines of regression of tensile strength upon hardness in one case and of tensile strength upon density in the other are shown in these figures. The distributions appear to be almost homo-scedastic and the results of the application of step 3 seems to justify the assumption that it is reasonable to assume linear regression. Under these conditions we might use either hardness in Rockwells or density as an indication or measure of the tensile strength of the specimen. Through the use of such a test it would be possible to measure the tensile strength of the specimen without breaking it. The standard deviation of such a measurement, however, depends upon the scatter of the observed points about the line of regression and for the Rockwell test the standard deviation is 2,894 pounds per square inch whereas for the density test it is 2,987 pounds per square inch. Roughly speaking, one of these tests would be as good as the other.

If, instead of using one or the other of these lines of regression, we make use of the plane of regression of tensile strength upon the density and hardness and calculate the standard deviation from this plane, we find that it is only 1,384 pounds per square inch. In other words, a knowledge of density and hardness enables one to specify the tensile strength of a die-casting within a range approximately one half of that within which it may be specified if only one of these measurements is known. We see at once from an inspection viewpoint how much more accurately the quality of product may be controlled through the use of these two variables than it can be through the use of only one. A somewhat similar result has been obtained in the study of the physical properties of certain timbers.

We must not conclude, however, that the use of a regression surface involving more than one stochastically dependent variable always leads to the best method of measurement. To illustrate this we shall return to a consideration of the measurement of tensile strength of nickel sheet already discussed at some length in this bulletin. Making use of the data[1] presented in Table 3.1, we have

TABLE 3.1

Tensile Strength psi X_1	Rockwell Hardness 100 Kg. 1/16" ball X_2	Per Cent Red. by Rolling X_3	Nos. Hard	Gauge Number X_4
66,100	53.6	0	0	14
81,300	87.6	22.6	2	14
93,400	92.8	37.2	4	14
103,400	96.0	47.8	6	14
108,600	97.9	58.8	8	14
114,300	99.1	68.3	10	14
65,100	51.1	0	0	16
85,600	89.6	21.6	2	16
97,400	93.8	40.1	4	16
104,500	96.6	51.3	6	16
111,400	98.9	60.8	8	16
114,500	99.7	67.6	10	16
65,300	50.6	0	0	18
81,100	85.3	19.4	2	18
97,400	93.4	37.6	4	18
103,600	94.9	50.9	6	18
110,400	96.8	62.1	8	18
115,000	98.8	66.8	10	18
67,600	56.2	0	0	20
80,700	85.0	19.9	2	20
96,300	90.9	35.4	4	20
107,200	94.6	51.4	6	20
109,500	95.3	61.4	8	20
113,600	96.7	67.8	10	20
74,600	67.6	0	0	22
80,600	83.4	16.5	2	22
98,700	91.3	37.8	4	22
107,000	94.0	49.6	6	22
112,200	95.5	60.6	8	22
113,500	96.0	68.7	10	22

1. See General Apparatus Development Bulletin No. 8, Issue 1, Nov. 22, 1927.

Tensile Strength psi X_1	Rockwell Hardness 100 Kg. 1/16" ball X_2	Per Cent Red. by Rolling X_3	Nos. Hard	Gauge Number X_4
66,900	60.6	0	0	24
94,900	90.6	19.4	2	24
98,700	90.9	34.9	4	24
107,000	93.4	49.1	6	24
112,400	95.5	58.8	8	24
116,200	96.6	68.9	10	24
65,200	57.9	0	0	26
85,000	86.3	21.9	2	26
108,200	94.2	35.8	4	26
106,600	93.3	46.3	6	26
112,600	95.4	60.2	8	26
114,700	96.2	67.4	10	26
64,500	59.8	0	0	28
83,300	84.9	20.3	2	28
102,800	92.4	38.5	4	28
116,300	96.3	51.4	6	28
111,500	95.4	58.4	8	28
116,400	96.3	68.0	10	28
65,300	56.2	0	0	30
77,500	80.6	18.4	2	30
95,600	89.2	35.9	4	30
107,500	94.0	46.4	6	30
119,500	97.1	59.1	8	30
115,100	96.7	66.7	10	30
67,000	55.1	0	0	32
76,800	71.5	17.0	2	32
91,200	83.0	34.6	4	32
102,400	90.4	49.1	6	32
112,600	93.4	60.5	8	32
122,200	97.6	66.9	10	32
66,200	66.4	0	0	34
86,200	75.7	23.5	2	34
96,300	79.1	40.0	4	34
100,500	81.0	50.0	6	34
109,200	83.3	61.1	8	34
115,900	87.2	68.3	10	34
69,400	78.9	0	0	36
83,200	81.5	23.0	2	36
96,500	81.6	35.3	4	36
102,700	83.1	45.0	6	36
104,100	82.9	57.7	8	36
111,100	84.9	67.1	10	36

calculated the equations of certain curves and surfaces of regression as indicated in Table 3.2 below and in each case found the standard deviation of the quality Y from the given curve or surface of regression. These standard deviations in pounds per square inch are given in the third column of the table. We see that they vary all the way from 3,129 pounds per square inch to 166,000

TABLE 3.2

SUMMARY OF STUDIES OF RELATIONSHIPS BETWEEN TENSILE STRENGTH AND CERTAIN OTHER PHYSICAL CHARACTERISTICS OF NICKEL-SILVER

Relationships	Curves or Surfaces of Regression	Standard Deviation From Curve or Surface of Regression
T.S.(X_1) vs. Hardness (X_2)	$x_1 = -2927.8347 + 1410.7024 x_2 + 16.0687 x_2^2.$	7,627
T.S. vs. Nos.Hard (X_3)	$X_1 = 46,732 + 6354 X_3 - 303 X_3^2.$	166,535
T.S. vs. Gauge (X_4)	$X_1 = 96585 + 1.7 X_4$	17,252
T.S. vs. (Hardness (X_2) (Nos.Hard (X_3)	$x_1 = \frac{\sigma_1}{\sigma_2}\left(\frac{r_{12} - r_{13}r_{23}}{1 - r_{23}^2}\right) x_2 + \frac{\sigma_1}{\sigma_3}\left(\frac{r_{13} - r_{12}r_{23}}{1 - r_{23}^2}\right) x_3$	3,908 3,129 (Hardness & Reduction)
T.S. vs. (Hardness (X_2) (Gauge (X_4)	$x_1 = \frac{\sigma_1}{\sigma_2}\left(\frac{r_{12} - r_{14}r_{24}}{1 - r_{24}^2}\right) x_2 + \frac{\sigma_1}{\sigma_4}\left(\frac{r_{14} - r_{12}r_{24}}{1 - r_{24}^2}\right) x_4$	8,131
T.S. vs. (Nos.Hard (X_3) (Gauge (X_4)	$x_1 = \frac{\sigma_1}{\sigma_3}\left(\frac{r_{13} - r_{14}r_{34}}{1 - r_{34}^2}\right) x_3 + \frac{\sigma_1}{\sigma_4}\left(\frac{r_{14} - r_{13}r_{34}}{1 - r_{34}^2}\right) x_4$	5,683
T.S. vs. (Hardness (X_2) (Nos.Hard (X_3) (Gauge (X_4)	$\Delta x_1 = \frac{\sigma_1}{\sigma_2}\left\{r_{21}(1-r_{34}^2) - r_{23}(r_{31} - r_{34}r_{41}) + r_{24}(r_{31}r_{43} - r_{41})\right\} x_2 - \frac{\sigma_1}{\sigma_3}\left\{r_{21}(r_{32} - r_{34}r_{42}) - (r_{31} - r_{34}r_{41}) + r_{24}(r_{31}r_{42} - r_{32}r_{41})\right\} x_3 + \frac{\sigma_1}{\sigma_4}\left\{r_{21}(r_{32}r_{43} - r_{42}) - (r_{31}r_{43} - r_{41}) + r_{23}(r_{31}r_{42} - r_{32}r_{41})\right\} x_4$ $\Delta = \begin{vmatrix} 1 & r_{23} & r_{24} \\ r_{32} & 1 & r_{34} \\ r_{42} & r_{43} & 1 \end{vmatrix}$	3,842

pounds per square inch. Based upon the standard deviation alone it would appear that the best method of measurement was in terms of the plane of regression of tensile strength upon hardness and per cent production. We find, however, that the distribution is not homo-scedastic, hence we cannot make direct use of this plane of regression. It is obvious therefore, that extreme caution must be exercised here in the use of this plane. Possibly we would not meet this difficulty, if it were possible to use a plane of regression of tensile strength upon hardness and density but the data for such a determination are not available. What we must conclude in this case is, therefore, that perhaps the best method of measuring the tensile strength of nickel silver sheet is that already discussed in paragraph 2 above.

Enough has been said to show that when we are called upon to measure some physical property Y in terms of certain other variables stochastically related thereto we must carefully consider the results obtained by carrying out the steps of analysis outlined above and in this way arrive at the most satisfactory and practical measure. Before leaving this subject however, it may be of interest to consider the various lines of regression between tensile strength and hardness for nickel silver sheet when the original data are grouped according to numbers hard. These lines of regression are shown in Fig. 3.5 and the standard deviations of tensile strength determined from these lines are on the average less than 3000 lbs. per square inch.

For the case of zero numbers hard the standard deviation from the line of regression is about 2100 pounds per square inch, but in general the standard deviation determined in this way for numbers hard greater than zero is considerably larger than the figure just stated. Moreover numbers hard are in general not accurately known.

It seems, therefore, as a result of our studies on nickel silver

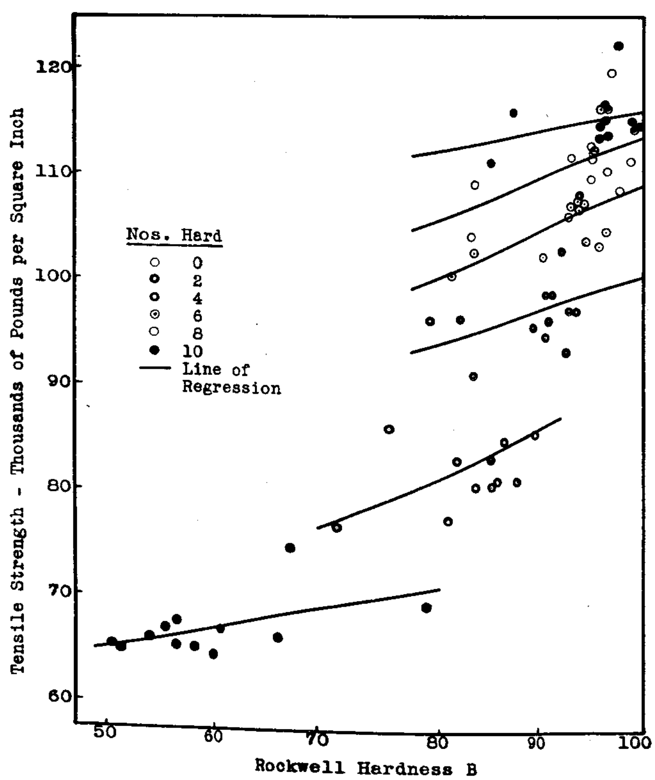


FIG. 3.5

sheet that, subject to further study, we should measure tensile strength as follows. For zero numbers hard use the line of regression (the lowest line in Fig. 3.5) of tensile strength on hardness obtained by allowing the gauge to vary from 12 to 36.

In all other cases use the line of regression of tensile strength on hardness for a given gauge, the numbers hard varying from 2 to 10.

The set of calibration lines which would be obtained under these conditions would appear schematically as in Fig. 3.6.

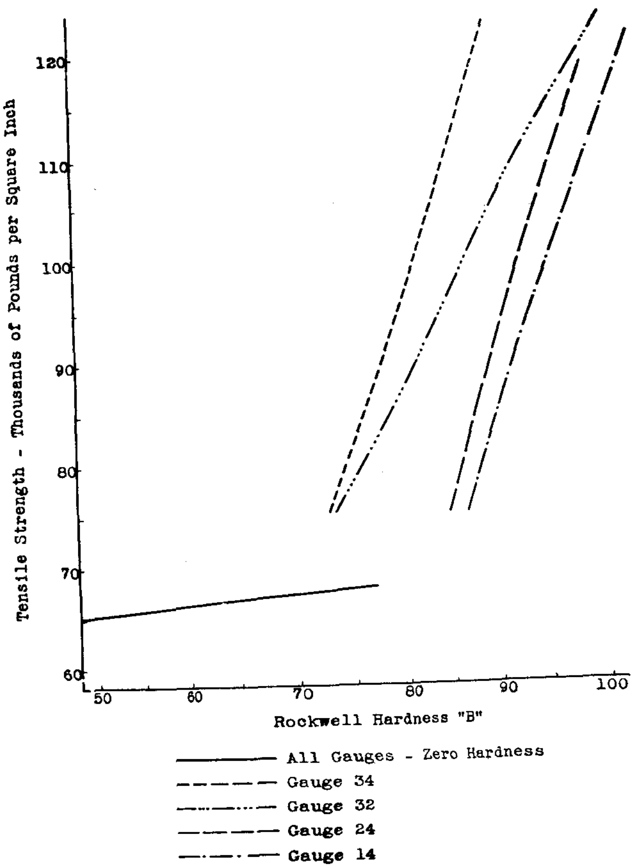


FIG. 3.6

Of course further studies may reveal the necessity for using higher order regression, in which case the lines in Fig. 3.6 should be replaced by curves. At present it appears that the best way of measuring tensile strength of nickel silver sheet indirectly through measurement of other physical characteristics is as stated above.

PART IV

Machine Measurement of Quality

Summary

Evidence is presented to show that certain methods now followed in calibrating some of the machines used in measuring quality of product may introduce errors so as to give an entirely misleading picture of the quality. A definite method is outlined for calibrating and maintaining these machines so as to remove this difficulty and thus secure adequate quality control.

A memorandum giving detailed steps in the application of the method outlined above to transmitter and receiver testing machine is now under preparation by Mr. P. S. Olmstead.

1. Criteria for Machine

Many important machines used in the measurement of quality of telephone equipment depend upon the use of such a stochastic relationship. Examples of such machines are those for measuring burning and efficiency of transmitters and the efficiency of receivers. The same type of machine problem would arise if we were to try to calibrate a hardness testing machine in terms of tensile strength.

Let us start by considering a practical illustration. Fig. 4.1 presents the scatter diagram of a machine measure, in this case agitated resistance in ohms, of transmitters whose efficiencies are represented by the ordinates of the scatter diagram. The practical problem in this case is to consider the use of the agitated resistance as a measure of the efficiency. This particular set of data was taken at a very early stage in the development of machine measurement of efficiency and is used here only to illustrate the methods involved

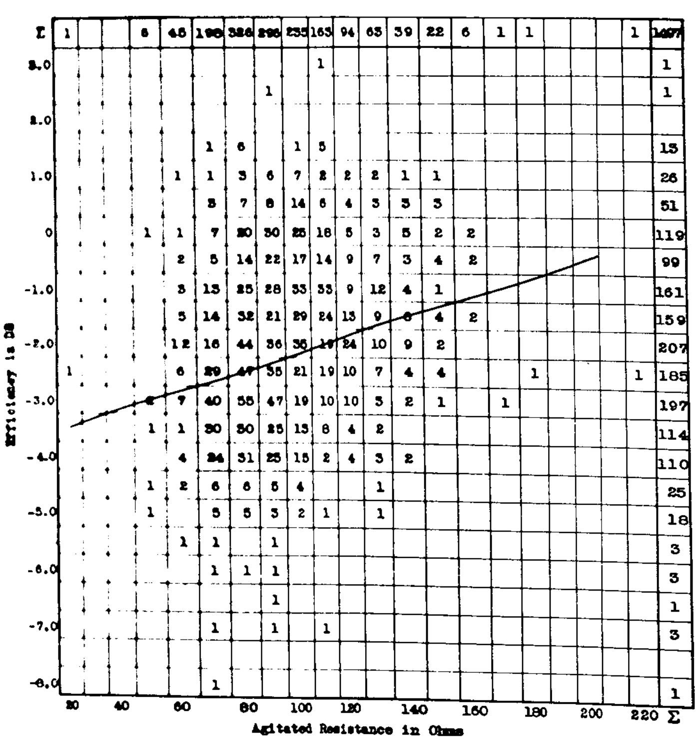


FIG. 4.1

in determining whether or not a proposed machine measurement is desirable.

In this particular instance a group of 1497 transmitters were selected at random from product, measured for efficiency Y and agitated resistance X. The observed distributions in respect to efficiency Y and resistance X are given at the borders of the scatter diagram and are shown graphically in Fig. 4.2.

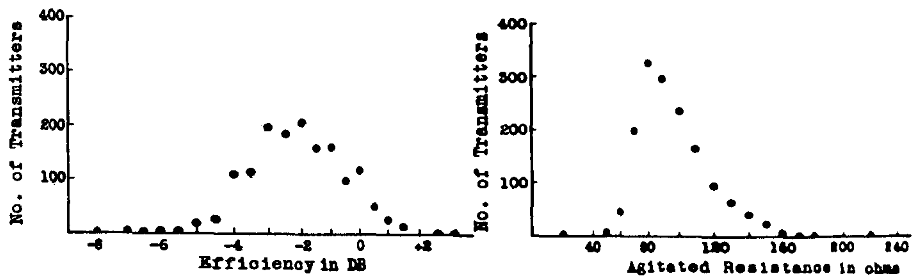


FIG. 4.2

Two requirements are sometimes placed on the machine. They are:

(1) That the machine throw out bad product, that is to say, in this case, instruments having an efficiency below some limiting value Y_L,

(2) That it detect lack of control of product.

Obviously a perfect machine would be one such that the meter deflection is functionally instead of stochastically related to the measured quality.

2. Assumptions Involved

It will be assumed in what follows that there are errors of measurement in both the quality Y and the machine measure X. In other words it is assumed that, if the same instrument is tested for efficiency several times, the measure of this efficiency, will have a certain standard deviation which we may represent by σ_{Ye}. Similarly, if this instrument is tested several times by machine, the readings given by the machine will be found to have a certain standard deviation σ_{Xe}. Furthermore, it is assumed that the expected efficiency of a transmitter is stochastically related to the expected machine measure of this efficiency. Under these conditions it is obvious that were it not for the fact that there were errors of measurement in both Y and X, the scatter diagram of Fig. 4.1 would have a slightly different appearance but on the other hand it would still remain a scatter diagram and not a functional relationship between the two variables considered. This means, in other words, that all instruments having an efficiency Y will not have the same machine measure X.

3. Calibration of Machine

Let us assume that the group of transmitters, in this case composed of 1497 instruments, is a random sample of controlled product.[1] It has been the practice to adopt some functional relationship between the meter deflection X and the efficiency in d.b. so as to calibrate the meter deflection directly in terms of efficiency. One customary method of doing this has been to adopt the line of regression of a quality Y on the measure X. If we were to adopt this line for the case of the data given in Fig. 4.1 we would get the distribution shown in Fig. 4.3. In other words, all of the transmitters would appear to have efficiencies within the range -3.3 to -.25 whereas the actual efficiencies were observed to extend over the range -8 to 3.0.

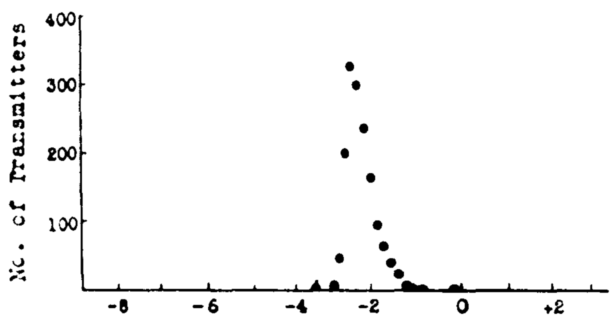


FIG. 4.3

It has also been suggested that instead of using this line of regression we should use a line which bisects the angle between the two lines of regression. If we were to use this line we might get a distribution with an increased range as shown in Fig. 4.4.

In the first place we should note that both of these suggested methods of calibration introduce difficulties because they indicate that the range of variation in the instruments is different from what it really is. The first line does have this to its credit, namely, that the meter reading obtained in this way does give the expected or average value of efficiency associated with a given value of X. Furthermore, the standard deviation of the error of measurement introduced by using this line is $\sigma_Y(1-r^2)^{1/2}$. Obviously if r is small there is little gained by using this line. The second line, however, even may give an efficiency associated with a given value of X which does not

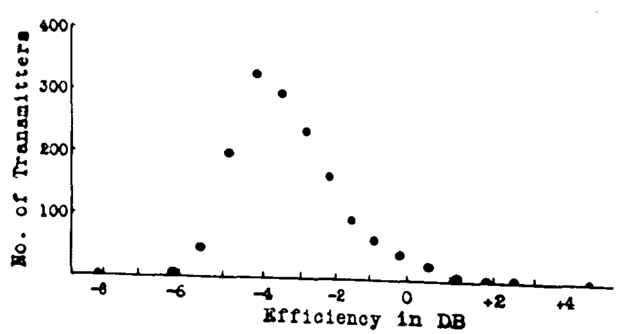


FIG. 4.4

---

1. Control is used here in the technical sense discussed in Reprint B-277, Quality Control.

occur, as is readily seen from the figure. Of course, if it were true that there was a direct functional relationship between the quality Y and measurement X when the errors of measurement were excluded, we would have some justification for the use of a line such as the second one indicated above, because under the conditions which actually exist the errors of measurement of both the efficiency and machine reading are such that a line drawn as indicated above would be approximately that obtained by the method of least squares where we minimize the distance of all the points in the scatter diagram to this line.

Enough has been said already to show that a machine calibrated by either one of the two lines indicated above would not perform the first function of the machine, namely, to throw out instruments below a certain level. In fact it is easily seen from the scatter diagram itself that there is no method available for using a meter deflection X to accomplish this end. In other words, it is not possible for a machine based upon a stochastic restraint to achieve this object of measurement. Let us turn, therefore, to a consideration of the extent to which such a machine can be used in controlling product.

The first step to be taken is to insure by methods presented in the first bulletin of this series that the assignable causes of variation in quality of product have been removed and then to choose a representative sample from which we may set up, as per the method suggested in the reprint on quality control[1], the specification of the distribution of product as measured in terms of the meter deflection X. In the present example, assuming that we had not additional information, we would take the distribution of the 1497 values of X given in the scatter diagram of Fig. 4.1 as representing standard product. We would then establish a quality control chart in the standard way using the parameters of this distribution as a basis for the chart. Formally this amounts to accepting a distribution, let us say $\Phi'(X)$ to be taken as standard. Now so long as the distribution of product does not change sufficiently from that given by $\Phi'(X)$ so that it may be detected by the control chart method, we may assume that the product is being controlled satisfactorily in respect to the quality Y in which we are really interested. If, however, the observed distribution from time to

- -

1. Loc. Cit.

time differs from $\Phi'(X)$ by more than may be attributed to chance, this fact indicates the possibility of lack of control of the quality of the product in respect to the quality Y. Such variations in X are necessary to indicate lack of control in Y but they are not sufficient, because it is possible for the distribution of X to vary by more than can be attributed to chance without at the same time implying that the quality Y has changed.

Before leaving the subject of calibration we must consider briefly some of the errors introduced by not choosing a random sample by which to calibrate the machine. Such a method would be justified provided there was a functional relationship between the quality Y and the meter deflection X when both had been corrected for all errors of measurement. This condition, however, does not exist as we have already noted. Hence in general, if we choose other than a random sample, we shall obtain a correlation coefficient different from the true correlation which exists. In several instances coming to our attention, this particular practice has been employed in such a way as to give a much higher correlation than that which really existed between the quality Y and the meter deflection X in a random sample. When such a correlation coefficient is used for establishing the line of regression to be used as a calibration curve, this line no longer necessarily measures the expected value of the quality Y in terms of a given value of meter deflection X. In other words, this particular calibration curve loses some of the advantages that it would have if the true correlation existing between them were used as a basis for determining this line. For this reason, it seems highly desirable that careful consideration be given to the methods of calibrating the machines which are used as a basis for the measurement of quality. Of, course, this difficulty is removed if the suggestion made above is followed, namely, to use the meter deflection X as the basis for detecting quality changes.

4. Maintenance of Machine

There are obviously several means of checking the maintenance of a machine. One of these customarily used, of course, is to select a group of transmitters which are kept as standards. The machine reading on any one of these transmitters should differ from the original value accepted as a standard for the machine by not more than an amount which must be left to chance (I.E.B. 1).

5. Errors of Measurement

It is well-known of course that the correlation r_o between the observed Y and the observed X is less than the correlation r between the values of the quality and the meter reading when corrected for errors of measurement. These two functions are related one to the other as follows:

$$r = \frac{\sigma_{X_o} \sigma_{Y_o}}{\sigma_X \sigma_Y} r_o.$$

In general, therefore, the observed correlation r_o is always smaller than the true correlation. By increasing the number of observations of Y and X on a single instrument we can increase the observed correlation. The fact that r is in general larger than r_o does not help us, however, in using the correlation between single pairs of observations for controlling the quality of product. If we are to improve our method of measurement by taking more than one observation on each instrument, the amount of improvement can be calculated by means of the above relationship, between r and r_o.

In the case of transmitters the standard deviation of measurement σ_{Y_e} (by means of the ear) of the efficiency Y is practically the same as the standard deviation σ_Y of the efficiency of transmitters and hence in that particular test

$$r_o = \frac{1}{\sqrt{2}} \frac{\sigma_X}{\sigma_{X_o}} r.$$

Hence r_o becomes a maximum when the error of machine measure σ_{X_e} is zero and is then equal to .7071 r .

PART V

Measurement of Quality of a Number of Things

Summary of Part V

In Part V we point out that, having taken a series of n observations on each of m different characteristics $X_1, X_2, \ldots X_i, \ldots X_m$, the essential information contained in the observed results is expressible for the most part in terms of the averages and standard deviations of the characteristics together with the correlation coefficients between sets of observations on different characteristics. Starting at this point we show how information expressed in terms of averages, standard deviations and correlation coefficients can be used in setting standards of quality and detecting lack of control of quality of manufactured product.

1. Basis for Measurement

In the general case we assume that the quality of a single thing is determined by, let us say, m different characteristics

$$X_1, X_2 \ldots X_i, \ldots X_m$$

The quality of a group of n things in terms of these m characteristics gives the following information:

$$\begin{array}{cccc} X_{11}, & X_{21}, & \ldots X_{i1}, & \ldots X_{m1} \\ X_{12}, & X_{22}, & \ldots X_{i2}, & \ldots X_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ X_{1n}, & X_{2n}, & \ldots X_{in}, & \ldots X_{mn} \end{array}$$

In the simplest case where only one characteristic X is measured on each of n things, the quality of the group of n things may be thought of as a frequency distribution of values of X along the axis of X. In a similar way, the quality of a group of n things, expressed in terms of m different characteristics, may be thought of as a frequency distribution in m dimensions.[1]

---

1. This concept of quality is amplified in Part I of the book on "Quality Control".

Theoretically speaking, the quality of a group of n things in respect to single characteristic X can be represented by n points along the straight line, i.e., in one dimension. In a similar way the quality of n things in respect to m different characteristics can be represented by n points in an m-dimensional space. In any practical case, of course, if the number n is large, it is very difficult indeed to appreciate the significance of such a set of n points. Particularly is this true when we come to the problem of considering the significance of the difference in the qualities of two groups of similar things. What we need in this case is obviously certain functions of these quality characteristics which will summarize, as it were, the essential information contained in the original data. Hence the measurement of the quality of a number of things reduces down to the problem of finding certain simple functions of an observed set of quality characteristics which may be used for the purpose of specifying the quality of a group of n things.

Considered from an abstract point of view, it would be a difficult task indeed to find functions which would always contain the essential information in any given set of data. Most of us perhaps would be inclined to give up the search for such a function, because a little study soon reveals that a particular function can satisfy this requirement only under very limited conditions.

It turns out, however, in the study of control of quality of manufactured product that the problem is somewhat simplified because, in general, as shown elsewhere,[1] there are two equations of condition which make it possible to proceed with some confidence in the setting up of ways and means of measuring the quality of a number of things for the purpose of Quality Control. These conditions are:

For a controlled product in respect to a single characteristic X, the probability dy that a single thing will have a quality X within the range X to X + dX can be approximated closely by

$$dy = \frac{1}{\sigma\sqrt{2\pi}}\left[1 - k\left\{\frac{x}{\sigma} - \frac{1}{3}\frac{x^3}{\sigma^3}\right\}\right]e^{-\frac{x^2}{2\sigma^2}}\,dx \qquad (5.1)$$

---

1. Parts 2 and 3 in the Book on Quality Control and in a published article on Quality Control.

where $x = (X - \bar{X}')$; $\bar{X}'$ is the expected quality X for the controlled process, σ being the standard deviation, and k being the skewness of the distribution of this controlled product. Similarly we assume that for a controlled product in respect to m different characteristics, the probability dz that a single thing will have a quality $X_1, X_2, \ldots X_i, \ldots X_m$ within the range X_1 to $X_1 + dX_1$, X_2 to $X_2 + dX_2 \ldots X_m$ to $X_m + dX_m$ is given by the following expression:

$$dz = Ce^{-\frac{1}{2}\left[\Sigma_1 \left\{\frac{R_{ii}\ x_i^2}{R\ \sigma_i^2}\right\} + 2\ \Sigma_2 \left\{\frac{R_{ij}\ x_i\ x_j}{R\sigma_i\ \sigma_j}\right\}\right]} dx_1, dx_2, \ldots dx_m$$

$$= Ce^{-\frac{1}{2}\chi^2}\ dx_1,\ dx_2,\ \ldots\ dx_m \qquad (5.2)$$

where C is a constant, and

$$R = \begin{vmatrix} 1 & r_{12} & \cdots & r_{1m} \\ r_{21} & 1 & \cdots & r_{2m} \\ r_{31} & r_{32} & \cdots & r_{3m} \\ \vdots & \vdots & \cdots & \vdots \\ r_{m1} & r_{m2} & \cdots & 1 \end{vmatrix}$$

and R_{ii} and R_{ij} are the minors obtained by striking out the ith row and ith column and the ith row and jth column, respectively, of the determinate R, Σ_1 is the sum for every value of i, Σ_2 is the sum for every pair of values of i and j.

It will be recognized at once that χ^2 in equation 5.2 is the generalized form of the function which we have already met so often in particular guises in this bulletin. Equations 5.1 and 5.2 give us a clue to the nature of the functions which must be used in the measurement of a quality of a number of things:[1]

---

1. For a description of χ^2 function as given above, see Pearson, Karl, Philosophical Magazine S.5, Vol. I, 1900, page 157.

σ_i = standard deviation of X_i and

r_{ij} = correlation between the ith and jth variable characteristics X_i and X_j.

Broadly speaking, measurement of quality of a number of things is used for one or the other of two purposes:

(a) We may start out with n sets of m characteristics derived from measurement of n things and attempt to specify standards for each of the m characteristics, or in other words from the observed sets of data we may try to establish the set of parameters required in one or the other of Equations 5.1 and 5.2.

(b) Having specified the standard of quality of a given product in terms of either Equation 5.1 or 5.2, we measure the quality of product to see if it differs significantly from that of standard quality[1].

In either case the following requirements should be met. The functions of an observed set of data must be efficient[2] and satisfactory estimates of the following functions of quality characteristics of controlled product:

The expected value $\bar{X}'_i$ for the ith characteristic X_i,

Standard Deviation σ'_i for the ith characteristic X_i,

Skewness k'_i for the ith characteristic X_i,

Correlation Coefficient r'_{ij} between the ith characteristic X_i and the jth characteristic X_j.

In this we have introduced the prime notation since our use of Equations 5.1 and 5.2 in quality control work is, for the most part, on the assumption that the parameters are known. We shall refer to our estimates of these parameters by the corresponding symbols $\bar{X}_i$, σ_i, k_i and r_{ij}.

If the sample size n is large, we have little or no trouble in estimating these parameters. In all cases with which we deal in practice, perhaps the most satisfactory estimate of $\bar{X}'_i$ is the arithmetic mean of the n values of X_i. When the sample size n is small (certainly when it is not over 25) we must

---

1. Reprint B-277 "Quality Control".
2. Book on Quality Control, Part III.

make correction for the fact that the expected standard deviation in sample size n will, in general, be less than the true standard deviation of the population from which the sample was taken. This general subject has already been discussed in considerable detail in Part II of this bulletin and all we need to do at the present time is to translate that discussion into terms of our present problem by merely remembering that we may consider a value of X in the discussion of Part II from a mathematical viewpoint in identically the same way as we consider a value of X in our present discussion. We may use an estimate of skewness derived from the sample for the purpose of specification provided the sample size from which it is determined is not less than 500. In general the coefficient of correlation r_{ij} derived from a small sample must be corrected for size of sample[1]. The type of corrections that we apply to the functions derived from the observed set of data depend entirely upon whether we are using these functions in the specification of standards or for the other purpose of determining whether or not the quality of a given set of n things differs significantly from standard.

In any case, we need the following information contained in a set of data:

1. The number n of things measured for the quality characteristic X_i.
2. The average $\overline{X}_i$.
3. The root mean square deviation σ_i.
4. The correlation coefficient r_{ij}.
5. The skewness k, if $n \geq 500$.

In what we have just said, we do not in any way mean to limit our analysis of data simply to the use of the stated simple functions. Rather, it is to be understood that the above functions constitute the minimum number required to express the information contained in a set of data, at least for most purposes of quality control. More specifically, it should be made a definite practice in most investigations to arrange to have sets of data obtained in a way that these simple functions can be readily calculated. Starting with these results we can then make use of the data with perhaps a minimum amount of

---

1. Ways and means of doing this are presented in Part III of the book on Quality Control.

effort and at the same time secure the maximum amount of information.

In what has already been said we have arrived at what may appear to be a somewhat empirical method of determining what functions of an observed set of data shall be used to express the quality of a number of things. It is true, of course, that anyone who stops to follow through the references cited will come away with a feeling that the method is not so empirical as it might appear before due consideration has been given to the two fundamental problems involved in quality control; that is, the setting of standards of quality and the determination of whether quality variations from this standard are significant.

Under these conditions it is believed to be worth while for us to consider briefly some of the ways in which we have made use of the above simple functions of the data for the purpose of solving some very important problems which have come to our attention within the past few years.

2. Practical Applications - Setting Standards

A. Physical Properties of Aluminum Die Castings

We have already had occasion to call attention to the work now in progress in one of the Committees of the American Society for Testing Materials in the determination of the physical properties of aluminum die castings. This problem is not peculiar to the work of this particular Committee. Quite the contrary, it is the problem which confronts every one of these Committees having to do with the specification of standards for the quality characteristics of raw materials. In fact, similar problems have been brought to our attention by two other Committees of the American Society for Testing Materials and by different members of our own organization who have the problem of setting standards on quality characteristics. Naturally, this problem falls in the domain of Inspection Engineering, because without knowledge of the standards of quality for a given kind of material, it obviously would be impossible to inspect the quality of product to determine whether or not it was significantly different from the standard quality.

Company	Tensile Strength X_1 in lbs. per sq. in.	Hardness X_2 in Rockwells	Density X_3 in grams per cc.
A	30120	55.3	2.627
B	25680	51.1	2.575
C	29248	50.7	2.585
D	30496	70.2	2.700
E	28900	49.4	2.669

TABLE 5.1 - Data: Alloy #4

In Table 5.1 we give the original observations in respect to three quality characteristics, tensile strength X_1, hardness X_2 and density X_3 of aluminum alloy #4. This particular alloy happens to be the one which is used in the present desk stand. We may consider Table 5.1 as being typical of the ways in which the original results should be presented. From these data it is necessary for us to arrive at certain values which we are willing to accept as standards for these three characteristics.

As soon as the analyst receives information such as that presented in Table 5.1 he should proceed in the following way.

First, he should make sure to find out from the engineers who have had to do with the taking of the data and the setting of the specifications on the material and all others concerned if they have any reason whatsoever to believe that the observed values of the quality characteristics have not been taken under a controlled Constant System of Chance Causes as defined, for example, in IEB-1. If the engineers in charge answer that the data have not been accumulated under such a condition, it remains for them to define further how they would group the data taken under one set of constant causes so as to distinguish it from the data taken under another set of constant causes. In other words, the analyst must demand to know all of the apriori information before he ventures to add much to what the engineer can do. This point cannot be emphasized too much.

Let us assume now that we have received the answer of the engineers in charge that the data of Table 5.1, to the very best of their knowledge, has been accumulated under a constant system of chance causes. It is then the duty of the analyst to test this postulate. Ways and means for doing this are discussed at length in I.E.B. 1.

In the present case samples of the alloy were furnished by each of the five companies to each of eight different testing laboratories, thus making a total of 40 tests on each alloy. We need not go through all the details of applying the criteria of I.E.B. 1 to the set of data given in Table 5.1 because the method of doing this has been given adequately in Bulletin I.E.B. 1. Suffice it to say that it is necessary to determine whether or not there is any indication of assignable causes of variation between the testing laboratories and also between the companies furnishing samples. In practically every case which was investigated we did come out with positive indications of assignable causes of variation and in certain meetings of this Committee the nature of the most likely causes of such variation were suggested by members of the Committee. If we were to follow through all of the steps, therefore, in taking account of our findings in respect to assignable causes in this particular instance, it would be necessary for us to exclude certain of the data presented in Table 5.1. For our present purpose, however, since we are merely interested in illustrating the method of setting the standards of quality after having taken the two steps previously mentioned for finding whether or not assignable causes are present, we shall assume that the data of Table 5.1 failed to indicate in any way whatsoever the presence of assignable causes of variation.

Physical Property	Average	Observed Standard Deviation σ	Corrected Standard Deviation $c_3\sigma$	Corrected Standard Deviation $c_4\sigma$
Tensile St. X_1 lbs. per sq. in.	28888.8	1704.3763	2200.3403	1059.3721
Hardness X_2 in Rockwells	55.34	7.6891	9.9266	4.7792
Density X_3 grams per cc.	2.6312	.04790	.06184	.02977
Sample Size n	5			
$r_{12} = .5449$ $r_{13} = .6604$ $r_{23} = .6791$	$\check{r}_{12} = .20$ $\check{r}_{13} = .27$ $\check{r}_{23} = .27$			

TABLE 5.2 - Analysis of Data: Alloy #4

Starting at this point the analyst prepares the information given in Table 5.2 where c_3 and c_4 are the correction factors already discussed in Part II of the present Bulletin. This table also includes the correlation coefficients r_{12}, r_{13}, and r_{23} and the values of these coefficients corrected for the size of sample[1].

Now, we come to the problem of interpreting the results in Table 5.2 as finally given by the analyst. Having no information other than that presented in this table and subject to the limitations already set forth, we accept the averages, $\bar{X}_1$, $\bar{X}_2$ and $\bar{X}_3$, as representing our best estimates of the expected values $\bar{X}'_1$, $\bar{X}'_2$ and $\bar{X}'_3$ of the corresponding physical characteristics. In a similar way, subject to the conditions already stated, we assume that the corrected standard deviations $c_3\,\sigma_1$, $c_3\,\sigma_2$ and $c_3\,\sigma_3$ represent estimates of standard deviations of the properties, tensile strength, hardness and density, respectively. In other words they may be taken as a measure of the variability of these three characteristics. More specifically we make the following assertion in the light of information already presented in Part II of this Bulletin: The a priori expected probability associated with the range $\bar{X}_i \pm t\, c_3\, \sigma_i$ is given by the normal law integral with no greater error than that indicated in reprint B-330.

Similarly, we may state, upon the basis of both theory and experiment, that subject to the assumption that the distributions of the physical properties are produced under controlled conditions such that the distributions themselves are normal, the expected probability associated with the range $\bar{X}_i \pm t\, c_4\, \sigma_i$ is given by "Student's" integral so long as t is not made larger than that corresponding to a probability of about .97.

Similarly, we may state, upon the basis of both theory and experiment, subject to the conditions of the previous paragraph, that on the average the fraction P of the expected values, $\bar{X}'_1$, $\bar{X}'_2$ and $\bar{X}'_3$, will not differ in absolute value from the averages $\bar{X}_1$, $\bar{X}_2$ and $\bar{X}_3$ of the respective samples by more than $c_5 t$ times the respective standard deviations σ_1, σ_2 and σ_3, where

---

1. Ways and means of doing this are discussed in Part III of the book on Quality Control.

$$c_5 = \frac{1}{\sqrt{n-1}}$$

and P is obtained from "Student's" integral for a given value of t.

The corrected values of r, that is, $\check{r}_{12}$, $\check{r}_{13}$ and $\check{r}_{23}$, can be used in a way similar to that in which the corrected sigmas, $c_3\,\sigma_1$, $c_3\,\sigma_2$ and $c_3\,\sigma_3$, have been used[1].

B. Standard Modulus of Rupture of Telephone Poles

Practical Situation - Let us consider another important problem in the setting of quality standards. For several years the accepted moduli of rupture for the four important classes of telephone poles have been those shown in Column 2 of Table 5.3. These figures were probably set upon the basis of tests made on sawn timbers. Within more recent years the question has arisen as to whether or not these previously accepted figures are satisfactory standards. There are some who feel that the figures are too low and others who feel that they are too high.

	Previously Accepted Standard	More Recently Proposed Standard	Possible Standard		
			$\bar{\bar{X}}$	σ	$c_3\sigma$
Southern Yellow Pine	5000-5100	6800	8164	1453	1468
Western Cedar	5000-5100	5600	5787	1052	1057
Eastern Cedar	3600	3600	3538	880	889
Chestnut	5000-5100	6000	6761	1204	1226

TABLE 5.3 - Modulus of Rupture in psi.

These previously accepted values for the modulus of rupture were early brought into question because it was argued that tests on sawn timbers might give results different from those on full size poles. Hence within the last few years several pole tests have been made to provide additional data as a basis for reconsidering the previously accepted standards. A preliminary review of these new data by some of the engineers in charge of the work led to proposed standards given in the third column of Table 5.3.

---

1. This statement is subject to limitations which will be set forth in Part III of the book on Quality Control.

It was shown by those interested in the design of pole lines that the acceptance of the proposed new standards would give an annual saving running into hundreds of thousands of dollars.

At this point in the study the original pole test data were submitted for statistical analysis to determine whether or not the more recently proposed values were completely justified in the light of available experimental information.

Formulation of Practical Problem - Let us look at the figures in Column 2 of Table 5.3. Are we supposed to interpret the previously accepted standards to mean that every southern yellow pine, western cedar or chestnut pole will have a modulus of rupture between 5000 and 5100 psi and that every eastern cedar pole will have a modulus of rupture exactly equal to 3600 psi? If these figures cannot be interpreted in this way, what is their interpretation? This is the kind of question that immediately comes to the mind of the analyst. Obviously the method of giving a standard in the form presented in Column 2 of Table 5.3 is open to the very serious criticism that it does not make possible an answer to this question and, without knowing the answer to this question, we cannot make the most satisfactory use of available information. What we should do in this case, as we have already pointed out, is to give our best estimate of the expected value together with an adequate measure of the dispersion about this expected value which can be used as a basis for an estimate of the expected number of poles having a modulus of rupture within any given range.

The criticism levelled at the previously accepted standards may also be levelled at the more recently proposed standards given in Column 3 of Table 5.3.

Results Obtained to Date - In our previous discussion of the method of setting standards, we have been careful to point out that in every instance the first step that should be taken is to obtain all available a priori information which would make possible the division of the original data into rational subgroups in the sense that it is reasonable to believe that the sub-groups could have come from different constant systems of chance causes. The next step is to record the number of observations, average, and standard deviation for each of the rational sub-groups of data. We may then use the correction factor c_3 for the observed standard deviation and be able to say, subject to the limitations previously set forth in this bulletin, that not less than 95% of the poles may be expected to have moduli of rupture within the range of the observed average $\bar{X}_i$ plus or minus $3c_3\sigma_i$ for the ith rational sub-group.

At the initial stage of the analytical investigation no a priori information could be obtained to make possible sub-division of the original data into rational sub-groups. Initially, then, we were forced to assume that all of the tests on a given species of pole came from a constant normal system of chance causes. Upon the basis of this assumption, the results presented graphically in Fig. 5.1 were obtained. The two groups of standards are also shown in this figure. We see at once that neither group of standards are located in the same way in respect to the distribution of modulus of rupture.

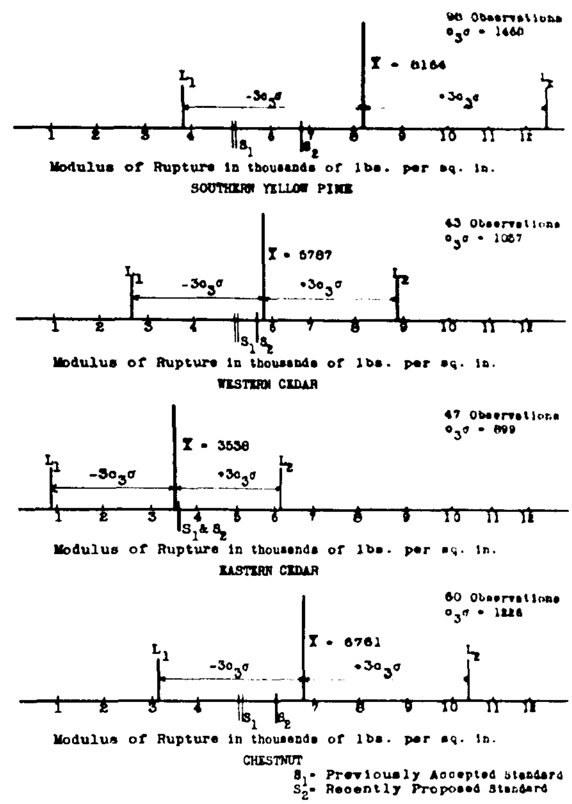


FIG. 5.1

Later investigation gave an a priori basis for the division of the original data into certain rational sub-groups. One such sub-group was that containing the saturated untreated poles. For this particular group the results obtained are presented in Table 5.4 and are shown graphically in Fig. 5.2. Here again we find that neither the previously accepted or more recently proposed standards offer a satisfactory basis for comparison of the different species.

Naturally, an engineer is interested in the corresponding standards for various other rational sub-groups and some of this information has already been secured. For example, it is found that certain of the poles which must be treated before being placed in line lose upwards of one thousand pounds in expected modulus of rupture. Sufficient has been said, however, to indicate that for every rational sub-group of the data we should give at least the average $\bar{X}$, standard deviation σ and corrected standard deviation $c_3\sigma$ and not merely some single figure such as given in Columns 2 and 3 of Table 5.3 since such figures do not bear any previously known definite relationship to the distribution of the quality characteristic.

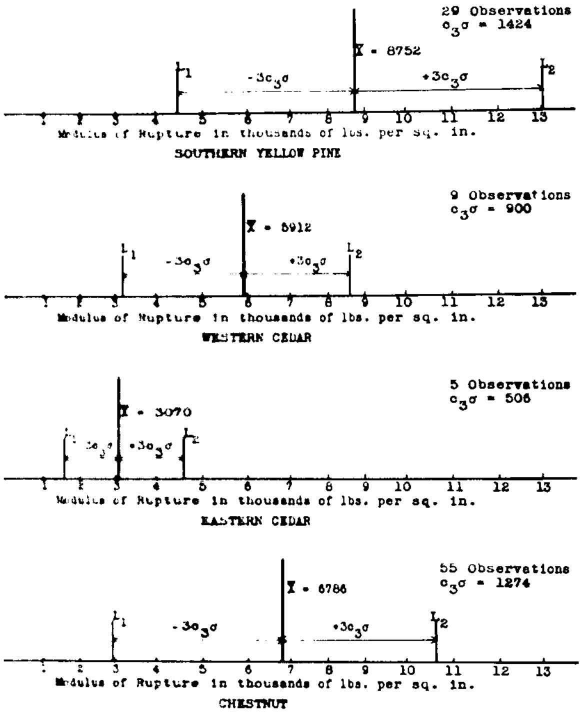


FIG. 5.2

Now, let us go a little further to point out what should be an even better way of setting a standard in this case. For example, moisture content is an assignable cause of variation. Little can be done, however, in the study of the original data such as that considered in connection with Table 5.4 because the method of measuring moisture content is subject to considerable error.

	n	$\bar{X}$	σ	$c_3\sigma$
Southern Yellow Pine	29	8752	1374	1424
Western Cedar	9	5912	793	900
Eastern Cedar	5	3070	392	506
Chestnut	55	6786	1251	1274

TABLE 5.4 - Saturated Untreated Poles - Possible Standards.

In a more recent work on the study of modulus of rupture of lodgepole pine, the error of measurement of moisture content was somewhat reduced and an attempt was made to relate the two factors, modulus of rupture and moisture content. The scatter diagram showing the observed relationship between these two factors is presented in Fig. 5.3. We see at once that this relationship is not functional but instead is stochastic. Further study revealed that the regression in this particular case could not adequately be represented by a straight line. However, the data definitelv indicated that the expected modulus of rupture depends upon the moisture content. Now an engineer in charge of building a line in a dry area could afford to use a much higher modulus of rupture than an engineer designing a line in which the same kind of poles were to be used in a wet area. It is proposed, therefore, in a problem such as this, that we try to establish, by means previously discussed in this bulletin, curves of regression between modulus of rupture and moisture content and then determine the standard deviation of the observed data about this accepted curve of regression. Quality standard then would be presented in terms of the equation of this curve involving the averages, correlation coefficients, and possibly higher moments, and the standard deviation about this curve.

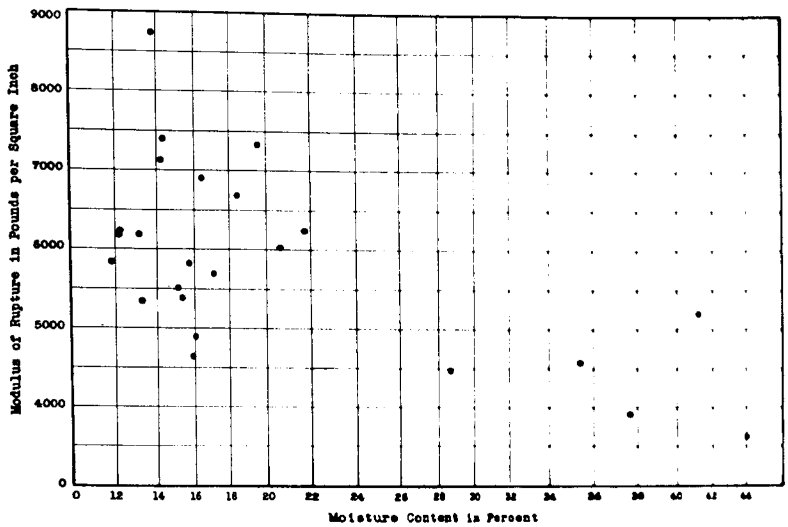


Fig. 5.3

We should, perhaps, call attention in more detail to one particular type of question which arises in setting certain standards, particularly in so far as the preceding discussion bears upon it. For example, we may be called upon to set a standard for modulus of rupture for Southern Yellow Pine Poles. Now in general, four different classes of Southern Yellow Pine Poles are distinguished one from another; namely, Loblolly, Shortleaf, Longleaf and Slash. Evidence is at hand to indicate that each of these subdivisions of Southern Yellow Pine Poles may be significantly different, one from the other. It is well recognized, however, that very serious practical difficulties arise when one tries to draw a close division line between the four subgroups. In fact, certain experienced timber men claim that it cannot be done. The fact remains, however, that observations on pole strengths of timbers have been divided into four groups bearing the names indicated above and that some of these groups are significantly different from others. This fact indicates the presence of assignable causes which should not be left to chance.

Assuming then, for the sake of argument, that the four subgroups are admitted to be different, how shall we set a standard for Southern Yellow Pine Poles so as to include all four of the subgroups. Obviously there are a very large number of answers to this question in the sense now to be explained. If we were to examine a group of n poles and all that we knew was that they were Southern Yellow Pine, it is obvious that there are a very large number of ways in which the poles might have been divided among the four subgroups. It is obvious that the standard deviation of modulus of rupture for a sample of size n of Southern Yellow Pine Poles will depend upon the way in which the sample is supposed to be divided among the subgroups. Therefore, we can not, in general, set standards for Southern Yellow Pine as a group unless we mean by that standard that a certain previously assigned percentage of the poles shall be obtained from each of the four subgroups.

In practice it may not be possible always to make sure that in the future a certain fixed percentage of poles will be drawn from each of the subgroups but it may be possible to estimate roughly what this percentage should be so that a standard may be set subject to the limiting assumption as to the way the poles are supposed to be divided among the four subgroups. Obviously it is necessary in proposing such a standard that we make sure that the individual who is to make use of the standard fully appreciates that the proposed standard only holds under the definite assumption that a certain fixed percentage of poles is to be taken from each of the subgroups.

C. Standard for Depth of Sapwood

We choose this third problem because it indicates certain features not previously considered in either of the other two. Some time ago, 1370 Southern Yellow Pine Poles were shipped to one treating plant and, before treatment, the depth of sapwood on each pole was determined by one measurement on each pole. This experiment gave us a series of 1370 observations of sapwood of Southern Yellow Pine Poles which we might use in setting the standard for depth of sapwood of Southern Yellow Pine Poles.

The frequency distribution of these results is presented in Table 5.5

$f(x)$ = Second Approximation

Cell No. (0)	Mid-cell values (1)	Cell boundaries (2)	[illegible] at δ, X (3)	Obs. freq. y (4)	yX (5)	yX² (6)	yX³ (7)	yX⁴ (8)	Theor. freq. y_0 (9)	$y-y_0$ (10)	$(y-y_0)^2$ (11)	$\frac{(y-y_0)^2}{y_0}$ (12)
0	1.0	.850	0	2	0	0	0	0	9	-7	1	5.444
1	1.3	1.150	1	29	29	29	29	29	25	4	16	.640
2	1.6	1.450	2	62	124	248	496	992	55	7	49	.891
3	1.9	1.750	3	106	318	954	2862	8586	99	7	49	.495
4	2.2	2.050	4	153	612	2448	9792	39168	149	4	16	.107
5	2.5	2.350	5	186	930	4650	23250	116250	189	-3	9	.048
6	2.8	2.650	6	193	1158	6948	41688	250128	207	-14	196	.947
7	3.1	2.950	7	188	1316	9212	64484	451388	193	-5	25	.130
8	3.4	3.250	8	151	1208	9664	77312	618496	159	-8	64	.403
9	3.7	3.550	9	123	1107	9963	89667	807003	116	7	49	.422
10	4.0	3.850	10	82	820	8200	82000	820000	77	5	25	.325
11	4.3	4.150	11	48	528	5808	63888	702768	46	2	4	.087
12	4.7	4.450	12	27	324	3888	46656	559872	25	2	4	.160
13	5.0	4.750	13	14	182	2366	30758	399854	13	1	1	.077
14	5.3	5.050	14	5	70	980	13720	192080	6	-1	1	.167
15	5.6	5.350	15	1	15	225	3375	50625	2	-1	1	.500
	Σ			1370	8741	65583	549977	5017239				X^2=10.843

m = UNITS(b) PER CELL = .3   DEGREES OF FREEDOM 12   P = .545

MOMENTS ABOUT ORIGIN δ

NOTE: THE ORIGIN δ IS THE MID-CELL VALUE OF CELL #0

$$\mu_1 = \frac{\Sigma yX}{\Sigma y} = \frac{8741}{1370} = 6.380292 \qquad \bar{X} = \delta + m\mu_1 = 1.0+.3(6.380292)=2.914088 \qquad \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}} = \frac{.798211}{37.013511} = .021565$$

$$\mu_2 = \frac{\Sigma yX^2}{\Sigma y} = \frac{65583}{1370} = 47.870803 \qquad \sigma = m\mu_2^{1/2} = .3(2.660704) = .798211 \qquad \sigma_\sigma = \frac{\sigma}{\sqrt{2N}} = \frac{.798211}{52.345009} = .015249$$

$$\mu_3 = \frac{\Sigma yX^3}{\Sigma y} = \frac{549977}{1370} = 401.443066 \qquad k = \frac{\mu_3}{\mu_2^{3/2}} = \frac{4.613423}{18.836059} = .244925 \qquad \sigma_k = \sqrt{\frac{6}{N}} = \sqrt{.0043796} = .066178$$

$$\mu_4 = \frac{\Sigma yX^4}{\Sigma y} = \frac{5017239}{1370} = 3662.218248 \qquad \beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{134.299660}{50.117111} = 2.679717 \qquad \sigma_{\beta_2} = \sqrt{\frac{24}{N}} = \sqrt{.0175183} = .132357$$

(CORRECTED MOMENTS)

UNCORRECTED MOMENTS ABOUT ARITH. MEAN $\bar{X}$

$$\mu_2 = \mu_2 - \mu_1^2 = 47.870803 - 40.708126 = 7.162677$$

$$\mu_3 = \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3 = 401.443066 - 916.289104 + 519.459461 = 4.613423$$

$$\mu_4 = \mu_4 - 4\mu_1\mu_3 + 6\mu_1^2\mu_2 - 3\mu_1^4 = 3662.218248 - 10245.295930 + 11692.384081 - 4971.454567 = 137.851852$$

$3\sigma_{\bar{X}}$ = .064695

$3\sigma_\sigma$ = .045747

$3\sigma_k$ = .198534

$3\sigma_{\beta_2}$ = .397071

CORRECTED MOMENTS ABOUT $\bar{X}$ (SHEPPARD'S CORRECTIONS)

$$\mu_2(cor) = \mu_2 - 0.083333 = 7.079344$$

$$\mu_4(cor) = \mu_4 - 0.5\mu_2 + 0.0291667 = 137.851852 - 3.581339 + .029167 = 134.299660$$

Cell No. (0)	Sapwood in inches (1)	Cell bound. (2)	Dev. from $\bar{X}$, x (3)	z $(\frac{x}{\sigma})$ (4)	F(z) (5)	±f(z) (6)	±kf(z) (7)	F(z)±kf(z) (8)	Diff. (9)	Freq. (10)
		.85	2.0641	2.5859	.4952	.0799	.0196	.5148		
0	1.0								.0065	9
		1.15	1.7641	2.2101	.4865	.0889	.0218	.5083		
1	1.3								.0181	25
		1.45	1.4641	1.8342	.4667	.0958	.0235	.4902		
2	1.6								.0399	55
		1.75	1.1641	1.4584	.4277	.0924	.0226	.4503		
3	1.9								.0719	99
		2.05	.8641	1.0825	.3605	.0729	.0179	.3784		
4	2.2								.1084	149
		2.35	.5641	.7067	.2601	.0406	.0099	.2700		
5	2.5								.1378	189
		2.65	.2641	.3309	.1296	.0105	.0026	.1322		
6	2.8								.1501	207
		2.95	.0359	.0450	.0180	.0005	.0001	.0179		
7	3.1								.1412	193
		3.25	.3359	.4208	.1631	.0164	.0040	.1591		
8	3.4								.1160	159
		3.55	.6359	.7967	.2871	.0488	.0120	.2751		
9	3.7								.0850	116
		3.85	.9359	1.1725	.3795	.0791	.0194	.3601		
10	4.0								.0560	77
		4.15	1.2359	1.5483	.4392	.0945	.0231	.4161		
11	4.3								.0336	46
		4.45	1.5359	1.9242	.4729	.0942	.0232	.4497		
12	4.6								.0184	25
		4.75	1.8359	2.3000	.4893	.0867	.0212	.4681		
13	4.9								.0091	13
		5.05	2.1359	2.6759	.4963	.0779	.0191	.4772		
14	5.2								.0040	6
		5.35	2.4359	3.0517	.4988	.0718	.0176	.4812		
15	5.5								.0017	2
		5.65	2.7359	3.4275	.4997	.0684	.0168	.4829		
	Σ								.9977	1370

TABLE 5.5

together with the detailed steps of the analysis involved in the calculation of moments and the application of Criterion 3, I.E.B. 1, for determining whether or not there are any indications of the presence of the assignable causes of variation. The observed and theoretical distributions are presented graphically in Figure 5.4. As seen from the data sheet, the value of X^2 was 10.84 corresponding to a probability of fit P = .54. It will be recalled that a probability of fit greater than .001 is to be taken as not indicating the presence of assignable

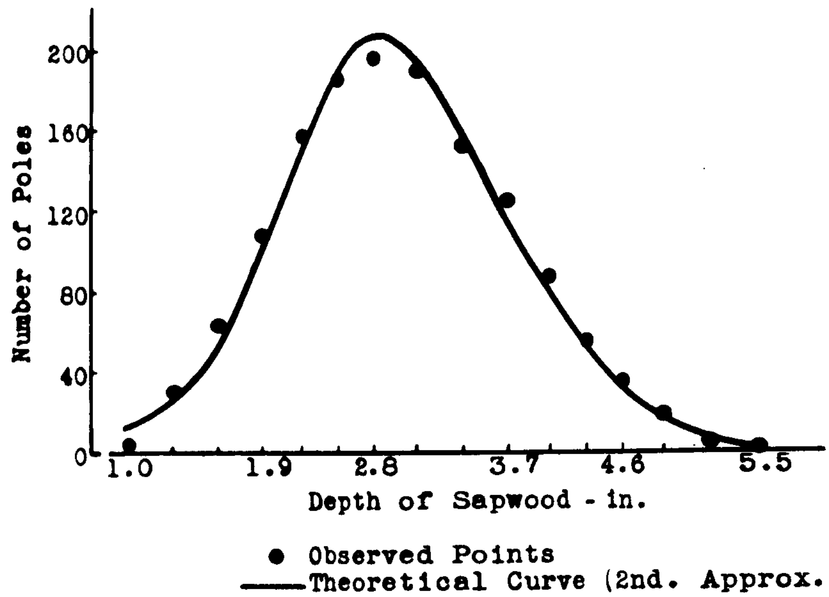


Fig. 5.4

causes of variation although it will be remembered also that it is quite possible that this test may fail to detect assignable causes although the conditions under which such failure might take place are not very likely in most practical problems. Since in the present case we had no way of subdividing the 1370 observations into rational subgroups we must rely upon the use of Criterion 3 and conclude that we have no indication of the existence of assignable causes of variation. Hence we take the following information as representing the standard for depth of sapwood of Southern Yellow Pine Pole.

Estimate of expected value $\bar{\bar{X}}'$ is average $\bar{X}$ = 2.91"
Estimate of standard deviations σ' is $c_3\sigma$ = .798"
Estimate of skewness k' is k = .24

Now, we should call attention to two points in connection with the acceptance of these standards. In the first place we should consider the bearing of the previous section having to do with the setting of standards for Southern Yellow Pine Poles as a group when it is assumed that this group of poles may be rationally subdivided into at least four subgroups. In the first place the standard as presented above could be accepted as a standard for the depth of sapwood of Southern Yellow Pine Poles made up of four subgroups in the proportion that these subgroups existed in the original series of 1370 observations. It is of course true that this standard might even have a broader significance. For example, it is possible that the subgroups do not differ significantly in respect to depth of sapwood although it is doubtful that this is actually the situation.

In the second place the contention may be made generally that we cannot measure accurately the depth of sapwood. One reason for this is that it is very difficult to establish any sharp line of demarcation between sapwood and heartwood. In other words, there is an element of personal judgment in deciding on the depth of sapwood. In addition to this the depth of sapwood is determined for each pole on the basis of a single observation although it is recognized that the sapwood in general is not the same over the butt section of the pole where it is customarily taken. Hence, each of the 1370 observations were subject to error. This leads us to a consideration of the effects of error of measurement which will be taken up in detail in Part 6 of this bulletin. It is sufficient to say here, however, that we may use the standard suggested above as the basis for still further correction which will take into account the standard deviation of the error of measurement. In general, it is reasonable to assume that this error of measurement is distributed symmetrically and hence that it does not influence either our estimate $\bar{X}$ of the expected value $\bar{X}'$ or our estimate k of the skewness k'. On the other hand, the standard deviation of the true depth of sapwood is most likely less than $c_3\sigma$ given above. In fact, as we shall see in Part 6 our best estimate of the true standard deviation σ_T' is given by the following relation

$$\sigma_T' = \sqrt{(c_3\sigma)^2 - \sigma_E^2}$$

where σ_E is the standard deviation of the error of measurement.

3. Applications in the Control of Quality

A. Control of Treatment of Telephone Poles

The quality of a number of things, as we have already seen, may be used in two ways for control purposes. For example, we may compare the quality of a group of things with that of another group of the same kind to see if they are significantly different or we may compare the quality of a group of things with standard quality for that kind of thing to see if the difference between them is significant. In the treatment of telephone poles it is necessary to have several plants located in different sections of the country so as to keep the freight charges as low as possible. Naturally, one of the functions of inspection is to determine whether or not the treatments given at one plant are assignably different from those given at other plants.

In the past, two characteristics have been used to indicate quality of treatment,- depth of penetration and the amount of creosote or oil contained in a cubic foot of sapwood. Some time ago the following problem was called to our attention. Each of seven plants had been sent a number of telephone poles of a given kind for treatment, and the depth of penetration for each of the poles had been measured at each of the plants. Making use of this information we were supposed to determine whether or not the plants were significantly different in respect to the quality of treatment as measured by depth of penetration.

Naturally the method of procedure in such a case is to apply Modified Criterion I, I.E.B. 1, to determine whether or not there are any indications of the presence of assignable causes of variation. The application of this test gave a positive indication of the existence of assignable causes. A little consideration of the physical problems involved indicated that we should expect to find that the depth of penetration depended very definitely upon the kind of poles submitted to a given plant in addition to the inherent quality of treatment which that plant was capable of giving.

For example, one of these factors which we should expect to influence the depth of penetration is the depth of sapwood. If the depth of penetration does depend upon the depth of sapwood, it is obvious that in any comparison of different plants it is necessary that we consider not only depth of penetration but also depth of sapwood of the poles delivered to a given plant. In general, we can say that, if there are several quality characteristics, say $X_1, X_2, \ldots X_m$, which influence the depth of penetration, then it is necessary to take all of these factors into consideration in comparing the quality of treatment given by one plant with that of another. Following the general method of attack outlined in the first section of Part V of this bulletin, the comparison of the quality of treatment given by the different plants requires a knowledge of the average and standard deviation of each of the m quality characteristics, and of the coefficients of correlation between them. This general problem of adequately comparing the qualities of treatment given by different plants is far from being solved up to the present time although we may outline some of the steps which have been taken to date so as to indicate the method of procedure to be followed in studies of this nature.

X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
4.60	1.90	3.00	1.20	3.20	1.50	3.70	2.00	2.20	2.15	3.00	1.80	3.25	2.20	2.60	1.50	2.20	1.10	1.35	0.70	4.40	1.60	4.50	2.40	3.90	1.40	3.10	2.15
2.15	1.30	2.90	1.80	2.30	1.75	2.70	0.80	1.90	1.50	1.60	1.00	3.20	1.85	2.10	1.20	2.20	1.20	3.30	1.05	3.40	2.40	2.80	1.30	2.70	1.60	4.10	3.40
3.70	1.60	3.55	1.50	4.10	2.00	4.40	1.90	4.10	2.90	4.20	1.05	2.75	2.20	3.50	2.15	2.70	1.10	3.10	1.40	2.05	1.60	2.10	1.70	2.05	1.00	1.30	0.90
2.70	2.20	3.30	2.00	3.50	1.45	2.50	1.20	2.80	2.45	2.90	2.30	3.50	1.60	3.60	1.70	3.15	0.80	3.80	1.20	2.90	1.10	1.15	0.90	3.30	1.00	4.00	1.80
3.20	1.50	3.40	1.85	1.60	0.90	3.60	2.05	2.15	1.90	3.00	0.95	2.35	1.10	3.60	1.45	2.25	0.90	2.70	0.80	3.30	2.20	3.30	1.60	2.80	1.60	2.40	1.80
2.40	1.70	3.80	2.40	3.70	1.90	3.70	1.80	1.90	1.30	2.85	0.85	2.50	1.20	3.20	0.85	2.40	1.10	1.55	1.00	2.80	1.80	1.75	0.95	1.15	0.80	2.60	1.35
2.15	1.15	1.70	1.00	1.75	1.10	1.50	0.85	1.80	1.15	1.95	0.80	2.70	2.50	2.90	2.10	1.60	1.45	2.30	1.00	3.60	2.30	2.10	1.50	1.65	1.40	2.60	1.10
1.90	1.50	3.10	1.80	2.50	1.50	2.60	1.60	3.60	2.20	2.90	1.40	2.00	1.05	2.35	1.00	3.05	1.40	2.60	1.90	3.40	1.70	3.20	2.00	2.80	1.80	2.55	1.40
3.50	1.50	2.50	1.90	2.30	1.55	2.00	1.20	1.75	0.90	2.60	1.10	2.45	1.35	3.30	1.15	3.65	2.00	3.35	1.85	1.65	1.15	3.70	1.00	2.00	1.35	2.10	1.00
3.50	2.70	2.95	1.35	3.40	1.50	3.60	2.15	3.90	2.55	1.90	1.50	3.05	1.25	3.20	2.65	1.95	1.80	2.65	1.80	2.60	2.05	2.25	1.65	1.95	1.15	3.55	2.05
2.10	1.05	3.00	2.00	2.85	1.25	1.90	1.20	1.90	1.60	2.85	1.25	3.00	2.30	2.35	1.80	4.40	0.90	2.65	1.10	2.45	1.15	3.80	1.45	3.70	1.80	3.40	1.35
2.45	1.60	2.75	1.20	2.60	1.50	3.65	2.00	2.25	1.60	2.80	1.40	2.05	1.50	3.70	1.50	2.40	1.40	3.05	0.70	3.55	1.65	4.00	3.60	1.45	1.10	3.50	1.50
3.30	1.45	2.65	1.65	3.50	1.40	3.00	1.10	2.30	0.90	2.20	0.90	2.60	1.00	2.60	1.15	2.50	0.75	1.50	0.80	1.40	0.75	1.90	1.30	2.70	2.40	3.40	1.10
2.90	1.65	2.25	1.35	3.70	1.35	3.05	1.50	2.10	0.95	3.10	1.50	1.90	0.90	1.60	1.05	3.40	1.20	1.60	1.45	3.05	1.20	3.60	1.65	3.70	1.60	3.00	2.80
3.10	2.10	3.90	3.30	3.90	1.30	4.00	2.85	2.30	1.75	2.60	0.80	4.20	2.10	3.85	1.85	3.75	1.95	2.10	1.15	3.80	2.75	3.00	1.05	4.00	2.65	3.30	2.50
3.50	1.60	1.70	1.55	2.05	1.60	2.70	2.00	3.35	2.85	3.70	1.10	3.25	1.95	1.95	1.50	2.65	1.00	3.25	2.75	1.85	1.05	1.75	1.10	3.40	1.60	2.90	1.40
1.50	0.85	2.70	1.25	3.15	1.75	3.10	1.50	3.25	2.30	2.90	0.90	3.25	1.25	2.45	1.55	2.20	1.40	2.85	1.60	1.55	0.85	2.30	1.70	2.70	1.40	2.45	1.25
2.95	1.55	3.10	1.35	3.35	1.90	3.20	1.75	3.30	1.65	2.80	2.55	2.80	0.90	2.45	1.00	2.90	1.45	3.20	2.00	3.50	2.40	2.75	1.30	3.25	2.20	2.35	1.35
2.90	0.90	4.10	1.10	4.05	1.30	2.40	1.50	2.15	1.20	3.05	1.90	2.60	0.90	2.25	1.60	2.70	1.00	2.70	2.40	4.30	1.30	4.80	1.05	3.30	1.70	2.50	1.50
1.70	0.85	2.55	1.40	2.10	1.40	2.80	1.40	2.30	1.50	2.40	1.45	3.40	0.95	1.95	1.35	2.20	1.10	3.40	1.90	2.85	1.90	2.80	1.90	2.70	1.35	2.85	2.00
1.10	1.00	2.40	1.30	3.05	1.05	3.40	2.10	1.65	0.95	3.20	1.70	3.00	2.50	2.50	1.80	2.50	1.00	2.90	1.35	2.70	1.80	2.60	1.70	2.50	0.75	2.70	1.60
2.70	1.20	2.80	1.60	2.20	0.90	1.65	0.75	3.05	1.30	1.85	1.60	2.50	1.80	2.90	1.40	1.95	1.20	3.00	2.00	3.70	2.10	2.10	1.80	3.00	0.65	4.40	1.55
3.10	0.85	4.70	1.50	2.40	0.90	4.50	1.90	3.90	0.70	5.00	2.60	4.90	1.70	4.00	2.50	3.10	0.85	3.40	2.90	2.70	2.10	3.05	1.95	3.15	1.60	3.10	1.50
3.30	1.80	4.15	3.00	4.25	2.65	3.40	1.40	3.60	1.95	2.70	1.55	3.05	1.35	4.20	3.10	2.25	1.10	3.00	0.70	2.40	0.40	4.00	1.60	4.30	2.55	3.50	1.30
2.80	1.15	2.75	0.80	3.20	2.10	1.30	0.60	2.85	0.90	4.00	2.40	2.85	1.30	3.60	1.20	3.20	2.20	2.30	1.65	2.25	0.85	3.50	2.30	2.55	1.55	3.00	1.15
3.20	2.55	2.80	1.90	2.65	1.15	2.10	1.10	2.05	1.50	2.15	1.65	2.65	1.20	3.65	2.40	2.05	1.25	2.90	2.60	2.70	1.80	3.10	0.80	1.90	1.25	5.20	4.70
3.50	1.80	3.70	1.70	4.80	2.70	4.80	2.90	4.00	1.80	3.70	1.70	2.70	2.05	3.20	2.05	2.60	0.95	2.70	1.75	1.15	0.55	1.75	1.60	1.40	1.05	3.60	2.65
1.70	1.00	1.90	1.50	1.95	1.00	3.75	3.35	2.85	2.30	3.20	1.25	3.00	1.60	2.70	2.00	1.90	1.00	3.00	2.85	2.05	1.30	2.40	0.90	2.70	2.60	2.55	1.40
2.50	0.80	3.45	2.90	2.15	1.50	2.05	1.10	2.20	1.40	2.40	1.40	3.45	2.20	1.20	0.80	2.00	1.05	3.45	0.85	1.60	0.80	3.10	2.15	4.00	2.70	4.40	1.50
4.10	2.60	3.10	1.50	3.30	0.80	3.05	2.00	4.55	3.15	4.05	2.50	3.10	1.00	3.75	2.50	3.30	0.80	2.35	1.05	2.00	1.60	1.25	0.75	2.40	1.10	4.15	1.75
2.70	1.05	3.90	1.25	4.00	1.25	3.90	3.20	3.70	1.80	3.40	2.30	3.40	2.40	2.50	1.10	2.30	1.40	2.90	1.70	1.30	1.00	2.00	1.70	2.50	1.10	3.90	2.30
2.70	1.20	1.60	0.90	3.60	0.90	2.50	1.00	1.60	0.90	2.70	1.50	3.10	1.20	1.90	0.70	3.30	1.40	2.00	1.20	1.80	1.00	2.30	1.30	3.30	1.10	1.70	1.40
2.60	1.90	4.70	1.20	3.50	1.70	1.60	0.90	2.80	2.10	2.10	0.70	2.10	0.90	2.80	1.00	1.90	0.80	4.00	1.20	4.30	2.10	4.20	1.50	2.60	1.40	3.90	1.80
3.40	1.80	2.60	1.50	4.10	2.70	3.10	1.50	2.40	0.70	1.95	0.70	1.90	1.25	2.20	1.25	2.50	1.30	2.35	1.25	2.00	0.75	3.20	1.80	2.15	1.10	4.15	2.80
4.50	3.55	5.00	4.15	3.35	1.05	2.20	1.25	2.50	0.60	2.55	1.00	1.60	0.60	2.40	1.20	1.70	1.15	3.15	1.40	2.60	1.10	4.10	2.60	3.60	1.00	3.65	2.60
2.95	1.20	2.95	2.40	2.40	1.50	3.40	1.40	3.30	1.55	3.80	1.90	3.80	2.30	3.40	0.90	4.00	2.65	3.40	1.45	3.80	1.05	4.80	3.50	2.05	1.35	2.50	2.30
3.15	1.25	4.20	2.10	4.20	2.20	3.40	1.25	2.15	0.70	2.95	2.10	3.00	2.10	3.50	1.80	3.25	1.15	4.30	2.00	2.30	0.95	3.60	2.10	3.30	1.80	4.60	2.75
2.30	0.80	2.30	1.80	3.50	1.55	3.30	1.25	2.30	1.20	1.60	1.05	3.50	1.80	2.90	1.90	2.85	1.55	3.60	2.60	1.65	1.30	1.25	0.90	2.95	1.40	3.10	1.50
1.80	1.20	2.35	1.35	1.65	0.75	1.40	0.80	1.55	0.75	3.15	2.90	3.25	1.35	2.70	1.10	1.65	1.05	3.25	1.10	2.85	1.50	2.35	1.15	2.15	1.25	2.65	0.90
1.80	0.95	3.00	1.80	1.70	0.70	3.25	2.30	2.80	1.95	3.45	1.60	2.00	1.00	2.75	2.20	1.75	1.25	2.85	1.00	2.25	1.00	2.15	1.30	2.25	1.60	3.00	1.10
1.70	1.05	2.80	2.00	3.50	1.70	2.90	1.50	3.70	2.00	3.60	1.00	3.60	1.60	2.45	1.50	3.60	2.20	3.95	2.15	2.90	1.50	1.95	1.60	3.25	1.35	2.90	1.35
1.55	1.40	3.50	1.85	2.50	1.45	4.10	2.15	2.80	1.65	3.50	1.30	3.20	1.30	2.85	2.10	2.75	2.25	3.10	1.20	4.00	2.40	3.60	1.10	3.80	1.60	3.70	2.60
2.05	0.80	3.40	1.60	2.55	0.90	2.55	1.60	2.90	1.85	3.80	2.80	1.75	1.35	2.10	1.00	2.35	1.05	3.65	2.25	3.95	0.85	2.05	1.80	2.50	1.85	2.40	1.50
2.80	1.15	2.35	1.70	3.10	1.40	4.00	2.10	2.10	1.30	2.20	1.30	4.10	2.30	2.00	0.90	4.40	1.20	3.20	1.60	3.80	0.80	3.75	3.40	2.40	1.95	4.10	1.50
4.20	2.65	3.50	2.90	3.50	1.50	2.25	0.90	3.30	1.65	2.50	1.70	3.60	1.75	2.55	2.20	3.80	1.30	3.40	2.80	3.15	2.95	2.70	1.30	2.30	1.70	2.60	1.70
2.05	0.85	2.50	2.05	2.90	1.10	3.05	1.55	3.20	1.05	3.20	1.50	2.05	1.30	1.75	1.10	2.85	1.40	2.00	0.90	1.80	1.05	2.65	2.35	3.70	1.50	2.50	0.80
3.20	1.20	1.70	0.60	1.70	0.50	3.10	1.60	2.10	0.80	3.70	1.65	3.10	2.60	3.60	2.00	2.60	1.05	2.00	1.15	3.30	2.30	3.70	2.10	2.00	0.75	2.65	1.90
2.90	1.50	2.50	1.10	3.70	2.05	3.30	1.10	3.00	1.60	3.00	1.20	4.00	1.20	3.90	1.50	3.50	2.20	2.35	0.70	3.80	1.40	3.90	2.40	4.60	1.50	3.90	1.60
2.30	1.25	3.85	1.80	3.90	1.60	2.80	1.00	3.30	1.50	3.10	1.80	3.00	1.30	2.85	1.20	2.90	1.40	2.90	1.60	2.70	1.50	1.80	1.20	1.65	1.00	2.20	1.70
1.50	1.50	2.30	1.50	1.90	1.05	2.30	1.10	3.80	1.00	3.60	3.00	2.80	1.10	2.90	1.60	2.70	1.55	4.00	0.80	2.30	1.40	2.60	1.30	2.00	1.30	3.05	1.80
3.05	2.30	2.25	1.45	3.30	2.20	2.70	1.60	2.00	0.85	1.95	0.90	2.90	0.75	1.20	0.30	2.55	1.20	4.80	1.50	4.70	1.40	2.45	1.15	1.95	1.00	1.90	1.00
2.15	1.30	3.10	1.15	2.80	1.65	2.90	1.25	1.85	1.05	2.20	0.80	2.35	1.15	2.25	1.20	2.10	0.70	2.45	1.20	2.70	0.90	1.55	1.10	2.05	0.70	2.90	1.30
2.50	1.75	2.80	1.85	3.00	0.80	2.70	1.30	1.70	1.20	3.60	1.60	3.80	1.40	3.90	2.90	3.70	2.20	4.60	2.90	3.80	2.20	2.50	0.70	2.50	1.10	2.45	1.50
2.05	1.65	2.65	0.95	2.60	1.75	3.00	1.20	2.45	0.90	4.50	2.15	2.55	2.45	3.30	1.90	2.60	0.95	2.85	1.50	3.10	2.00	4.00	2.10	3.30	1.90	1.40	1.15
2.20	1.45	3.00	1.70	1.50	0.85	1.60	1.35	2.75	1.85	2.50	1.30	3.70	2.10	2.30	1.90	2.35	1.20	1.90	1.35	1.95	0.95	2.00	1.10	3.50	3.80	5.00	1.70
3.50	1.10	1.80	1.05	3.00	1.60	3.70	1.70	3.50	1.40	3.40	1.75	2.00	1.30	2.00	0.60	3.10	1.50	1.80	1.50	2.60	1.80	2.65	1.70	1.60	1.10	1.25	0.80
2.50	1.25	1.80	1.20	3.30	1.30	2.30	1.70	2.95	1.40	2.70	2.40	3.10	1.30	1.65	0.75	1.20	0.95	4.35	2.50	2.50	1.00	3.35	1.75	3.65	1.05	1.80	1.00
1.85	1.05	1.50	0.80	4.35	3.00	2.00	1.15	2.35	1.00	3.80	1.40	3.60	2.80	2.85	2.65	2.90	2.70	3.30	1.30	3.70	1.30	3.60	1.80	1.45	1.00	2.55	1.35
2.15	1.00	2.40	1.10	2.80	1.00	3.50	1.00	5.00	3.70	2.80	1.80	2.30	1.10	2.40	2.20	3.20	2.00	3.40	2.10	3.10	1.00	2.10	1.30	3.70	2.10	2.60	1.30
2.00	0.80	4.00	1.15	2.30	1.70	2.10	1.50	3.90	2.30	2.70	1.30	2.55	1.10	2.60	1.70	1.55	0.95	5.20	4.00	2.65	1.00	1.75	0.95	5.10	3.00	3.60	2.45
2.00	1.25	2.75	1.80	3.20	1.55	2.20	0.80	3.95	3.25	2.60	2.05	2.95	0.70	2.85	1.45	3.60	3.10	3.05	1.50	2.85	2.00	3.65	2.30	3.80	2.55	2.95	1.70
2.75	0.80	3.15	2.15	3.15	1.90	2.55	1.70	3.20	1.90	2.90	1.00	2.35	1.25	2.00	0.95	1.95	0.90	2.85	1.40	2.60	0.90	2.25	1.60	2.45	1.10	2.00	1.15
3.40	1.00	3.60	1.20	2.90	1.05	3.70	1.55	4.10	2.00	3.70	2.20	3.40	3.25	2.50	0.80	3.80	1.65	2.90	0.95	1.50	0.85	3.70	2.00	3.10	1.65	3.00	1.80
4.65	3.90	3.90	2.85	4.40	2.60	3.60	2.50	4.60	4.00	4.25	3.30	4.20	3.00	3.10	1.10	3.40	1.80	2.50	1.05	2.05	1.80	4.40	2.00	1.80	0.50	2.40	1.00
4.80	3.80	2.00	1.30	3.20	2.10	3.15	1.70	4.40	2.40	2.50	1.80	3.05	1.20	3.20	1.50	2.70	2.60	2.80	0.80	2.35	1.10	2.60	1.35	2.60	1.80	3.05	2.20
3.20	1.80	3.50	2.25	4.40	3.50	3.70	3.10	3.75	3.50	1.90	1.60	3.70	2.30	2.40	1.20	3.15	1.60	2.25	1.80	4.80	2.40	2.65	1.50	4.30	2.10	3.95	1.50
2.55	1.05	4.15	1.60	2.85	1.20	2.80	1.10	3.20	2.10	2.40	1.15	3.70	2.20	2.10	1.00	4.10	1.15	3.15	0.90	2.15	0.90	3.00	1.40	2.80	1.30	1.90	0.80
2.30	0.90	1.45	0.90	2.35	1.20	1.40	0.80	4.00	1.50	2.00	1.05	1.95	1.25	2.55	1.20	4.10	2.10	4.30	3.60	3.50	2.20	3.00	1.50	3.00	1.30	4.00	2.70
2.60	1.90	1.90	1.80	2.50	1.05	2.70	1.30	1.70	0.90	2.50	1.60	3.35	1.25	2.15	1.10	1.80	1.60	2.30	2.20	2.65	1.10	3.15	1.00	3.50	1.80	3.60	1.05
3.90	2.10	3.70	2.05	3.65	1.80	3.00	0.95	3.30	1.55	2.80	1.50	3.40	2.35	2.80	1.60	3.00	1.45	3.10	1.20	3.50	1.60	3.05	1.80	2.85	1.10	3.00	1.45
3.05	1.80	3.60	1.40	1.85	1.15	2.50	1.30	1.50	1.30	3.25	2.45	2.10	0.50	1.90	1.30	1.85	0.90	3.60	3.25	3.20	1.20	2.50	1.60	3.30	0.90	2.50	1.25
4.40	1.90	2.50	1.20	3.30	1.30	2.70	0.95	1.95	1.35	1.95	1.30	2.50	1.00	2.90	1.00	4.30	1.10	2.00	1.00	3.50	2.80	2.55	1.30	3.15	1.40	2.75	1.45
2.75	1.35	4.50	1.90	3.60	1.75	3.20	1.10	3.05	0.95	4.80	1.90	2.65	1.15	2.35	1.55	3.95	2.30	3.60	1.60	2.00	1.80	3.50	2.55	1.45	1.10	3.90	2.90
1.35	0.60	2.20	1.70	2.30	1.35	1.95	0.60	3.95	2.20	3.00	2.10	2.85	1.05	3.50	1.40	3.30	1.75	2.90	1.50	2.95	2.70	3.20	1.90	3.30	1.40	2.50	2.20
3.30	1.40	4.00	2.35	2.05	0.85	2.20	1.30	3.60	1.85	2.30	1.30	2.00	0.85	1.70	0.90	2.50	1.15	2.70	1.80	3.05	0.65	2.10	1.15	2.50	0.80	1.30	1.10
2.10	0.85	2.95	2.30	2.65	1.60	2.25	1.50	1.70	0.85	2.65	0.90	3.20	2.10	1.40	0.75	2.55	1.10	2.75	1.00	3.35	1.50	4.20	1.95	2.10	1.10	3.75	1.80
2.60	0.80	2.20	0.95	1.90	1.10	3.00	2.00	2.00	1.00	2.90	2.30	2.00	0.75	3.95	2.95	2.50	0.90	3.30	2.30	2.70	1.00	1.35	0.85	2.55	2.00	2.80	1.00
2.30	1.20	2.60	1.00	2.30	1.40	2.45	0.90	3.10	1.00	1.95	1.00	2.40	1.20	2.90	2.60	3.50	3.10	3.65	1.50	4.10	1.70	3.65	2.80	3.30	1.50	3.40	0.70
4.05	2.15	4.70	0.80	2.55	1.80	2.35	1.30	4.20	2.10	3.50	2.10	3.70	0.85	3.70	2.40	4.70	3.00	4.90	3.15	2.60	1.25	4.40	3.00	4.50	3.25	4.40	2.40
4.10	1.70	4.40	2.00	4.00	1.35	4.50	2.65	2.55	0.85	1.65	1.20	4.35	1.20	4.00	2.00	3.50	2.45	3.30	2.20	2.00	0.55	4.60	1.15	4.60	2.30	2.40	1.30
4.30	1.40	4.50	3.00	2.40	0.80	2.80	1.70	3.05	1.90	3.60	2.30	2.60	2.15	2.75	1.50	2.40	2.00	2.20	1.30	2.70	1.40	3.50	1.70	3.45	1.40	4.05	2.80
2.05	1.10	2.40	1.40	2.15	1.20	2.15	1.10	3.40	2.50	3.90	2.50	2.50	1.30	3.00	1.40	3.40	1.90	2.65	1.35	2.50	1.20	3.00	1.90	1.25	0.90	3.30	1.40
2.20	0.65	3.10	0.85	1.80	1.15	1.55	1.25	3.15	1.70	2.00	1.25	3.00	2.00	3.15	2.20	3.15	2.10	4.55	1.65	1.85	1.35	1.75	1.00	4.15	1.70	3.15	1.25
3.40	1.70	2.20	1.50	1.60	1.00	2.55	0.70	3.10	1.60	2.55	1.55	3.80	1.20	3.45	2.45	3.35	1.05	3.15	2.15	3.20	1.90	2.15	0.95	2.70	0.90	2.65	1.70
2.70	1.30	4.35	2.80	3.00	1.40	2.15	1.40	2.30	0.70	3.60	1.85	4.25	2.50	4.00	3.30	1.20	1.00	3.20	2.50	1.65	1.50	4.50	2.20	1.70	1.45	3.10	1.50
4.10	2.70	2.10	2.00	2.70	1.60	3.00	1.50	4.30	2.35	1.85	1.40	2.40	1.70	2.30	1.40	1.15	0.80	2.50	1.50	2.50	1.55	3.90	2.25	3.80	2.45	4.80	2.25
2.80	2.20	3.80	2.15	4.50	3.50	2.70	1.40	2.90	1.30	1.20	0.95	2.45	1.85	2.55	1.95	2.25	1.35	3.30	2.80	3.50	2.90	3.15	2.80	2.60	1.30	2.80	2.10
3.50	3.25	2.70	2.00	3.80	3.00	2.60	1.10	2.60	0.90	2.70	1.55	2.15	1.50	3.00	0.95	2.80	1.45	2.95	2.05	2.65	1.20	3.90	2.20	3.20	2.10	2.85	1.50
3.40	2.00	2.95	1.60	2.05	1.70	2.25	1.45	2.90	1.65	2.80	2.55	3.30	2.20	3.05	1.90	2.40	1.35	2.75	1.30	3.35	2.10	3.05	2.00	3.00	1.95	2.35	1.10
5.10	3.80	3.35	1.60	2.05	1.60	2.45	1.30	3.70	1.90	3.60	2.00	3.40	2.75	0.90	0.65	2.65	1.20	3.30	1.25	3.60	0.80	3.90	2.90	4.00	2.15	2.30	1.00
1.95	0.90	3.40	2.45	3.50	2.10	3.00	2.40	1.40	1.00	2.35	1.00	2.15	1.10	3.00	1.50	3.10	1.15	3.15	1.90	3.80	2.40	5.20	1.00	3.00	1.60	3.40	2.05
2.45	2.00	2.05	1.60	3.80	1.70	3.30	1.30	2.10	1.40	5.00	0.85	4.65	1.45	4.35	1.60	1.65	0.80	2.80	1.20	2.70	1.35	3.80	2.70	1.65	1.05	2.80	2.80
4.00	2.65	1.70	1.00	2.40	1.70	3.45	2.20	2.20	1.50	2.45	1.50	2.45	1.30	3.80	2.65	3.50	0.70	2.05	1.20	2.30	1.35	3.60	1.75	3.00	2.25	3.90	1.80
3.00	1.05	2.05	0.85	2.95	1.80	3.60	1.95	2.70	1.65	3.00	1.30	2.25	1.10	1.35	0.90	2.50	1.75	2.25	1.30	3.90	2.60	3.15	1.15	1.75	0.75	2.20	1.80
3.05	1.20	3.00	2.00	3.05	1.10	2.50	1.40	2.50	1.40	4.00	1.50	3.40	2.10	3.10	1.60	3.80	2.10	2.70	1.20	2.85	1.05	2.05	1.00	3.30	1.30	1.80	1.10
1.30	0.75	2.50	0.90	2.50	0.85	2.40	1.10	3.05	2.70	2.25	1.90	1.70	1.35	3.00	2.20	3.00	2.65	2.20	1.60	2.30	1.70	2.90	1.50	2.95	1.30	3.50	2.85
3.70	2.30	2.35	1.40	3.40	2.00	2.70	2.30	2.35	1.50	3.00	1.50	3.00	1.50	2.30	1.70	2.65	1.50	3.50	2.30	2.75	2.10	3.70	2.30	3.50	2.40	2.15	1.50
2.00	1.30	2.20	1.15	2.55	1.90	3.50	2.10	2.80	1.50	2.70	1.70	3.60	2.10	2.10	1.40	2.30	1.80	2.90	1.30	2.90	2.30	3.65	2.20				

TABLE 5.6

In the first place it was of interest to determine whether or not there was any experimental evidence indicating correlation between the depth of sapwood and depth of penetration in a homogeneous group of poles of a given species. A short time later we obtained experimental information which made such a study possible. The data in this instance consisted of 1370 pairs of observations of depth of sapwood and depth of penetration in a group of as many telephone poles of the same species as had previously been investigated. So that the reader may not only follow the line of argument but understand some of the steps which may be unfamiliar, we shall outline in some detail the method of analyzing these data.

The original results are presented in Table 5.6. One of the first things to do, of course, is to make sure that the group of poles was homogeneous in respect to depth of sapwood. In other words, it was necessary to determine whether or not there was any indication of the presence of assignable causes of variation in the depth of sapwood within the group of 1370 poles. Now this particular group of poles is the one used above in establishing a standard depth of sapwood where it is shown that the data give no indication of the existence of assignable causes of variation.[1]

Table 5.7 gives the necessary details required to calculate the correlation coefficient between depth of sapwood and depth of penetration. By means of analysis we arrive at the following figures which are supposed to give the essential information contained in the original data in so far as we are immediately concerned:[2]

$$\bar{X} = 2.91 \text{ inches}$$

$$\sigma_X = .80 \text{ "}$$

$$\bar{Y} = 1.59 \text{ "}$$

$$\sigma_Y = .62 \text{ "}$$

$$r_{YX} = .60 \text{ "}$$

---

1. It will be shown in Part III of the book on Quality Control that, had there been any evidence of assignable causes of variation, there would have been some reason to believe that any observed correlation did not necessarily indicate that sapwood was a controlling factor.
2. It was necessary to make further investigation to determine whether or not the stochastic relationship between Y and X was linear. The details of this are presented in Part III of the book and need not be considered further here, since they indicated that what we are about to say is justified.

X = Depth of Sapwood   Y = Depth of Penetration

(1) X	(2) Y	(3) n	(4) nXY	(1) X	(2) Y	(3) n	(4) nXY	(1) X	(2) Y	(3) n	(4) nXY
1.0	.7	1	.70	3.1	.7	10	21.70	4.0	3.4	5	68.00
	1.0	1	1.00		1.0	22	68.20		3.7	1	14.80
1.3	.4	1	.52		1.3	40	161.20	4.3	1.0	4	17.20
	.7	15	13.65		1.6	42	208.32		1.3	4	22.36
	1.0	12	15.60		1.9	36	212.04		1.6	7	48.16
	1.3	1	1.69		2.2	24	163.68		1.9	7	57.19
1.6	.4	2	1.28		2.5	6	46.50		2.2	6	56.76
	.7	11	12.32		2.8	7	60.76		2.5	7	75.25
	1.0	33	52.80		3.1	1	9.61		2.6	4	48.16
	1.3	11	22.88	3.4	.7	3	7.14		3.1	5	66.65
	1.6	5	12.80		1.0	15	51.00		3.4	3	43.86
1.9	.7	13	17.29		1.3	29	128.18		3.7	1	15.91
	1.0	41	77.90		1.6	28	152.32	4.6	.7	1	3.22
	1.3	36	88.92		1.9	22	142.12		1.3	3	17.94
	1.6	14	42.56		2.2	27	201.96		1.6	5	36.80
	1.9	2	7.22		2.5	11	93.50		1.9	3	26.22
2.2	.4	1	.88		2.8	12	114.24		2.2	3	30.36
	.7	11	16.94		3.1	2	21.08		2.5	1	11.50
	1.0	42	92.40		3.4	2	23.12		2.8	3	38.64
	1.3	48	137.28	3.7	.7	1	2.59		3.1	3	42.78
	1.6	39	137.28		1.0	10	37.00		3.4	2	31.28
	1.9	10	41.80		1.3	13	62.53		3.7	1	17.02
	2.2	2	9.68		1.6	21	124.32		4.0	2	36.80
2.5	.4	1	1.00		1.9	24	168.72	4.9	1.0	1	4.90
	.7	14	24.50		2.2	28	227.92		1.6	3	23.52
	1.0	50	125.00		2.5	11	101.75		1.9	1	9.31
	1.3	59	191.75		2.8	7	72.52		2.2	1	10.78
	1.6	34	136.00		3.1	4	45.88		2.5	2	24.50
	1.9	19	90.25		3.4	4	50.32		2.8	2	27.44
	2.2	7	38.50	4.0	.7	2	5.60		3.1	1	15.19
	2.5	2	12.50		1.0	2	8.00		3.7	2	36.26
2.8	.7	6	11.76		1.3	10	52.00		4.3	1	21.07
	1.0	37	103.60		1.6	10	64.00	5.2	1.0	1	5.20
	1.3	51	185.64		1.9	9	68.40		3.1	1	16.12
	1.6	45	201.60		2.2	15	132.00		3.7	1	19.24
	1.9	22	117.04		2.5	12	120.00		4.0	1	20.80
	2.2	18	110.88		2.8	14	156.80		4.6	1	23.92
	2.5	12	84.00		3.1	2	24.80	5.5	2.5	1	13.75
	2.8	2	15.68								

$n = 1370$

$\Sigma XYn = 6765.77$ $\qquad \overline{X}\overline{Y} = 4.637654$

$\frac{\Sigma XY}{n} = 4.938518$ $\qquad \sigma_X \sigma_Y = .498779$

$$r = \frac{\frac{\Sigma XYn}{n} - \overline{X}\,\overline{Y}}{\sigma_X \sigma_Y} = \frac{4.938518 - 4.637654}{.498779} = .603201$$

TABLE 5.7 - CALCULATION OF CORRELATION COEFFICIENT

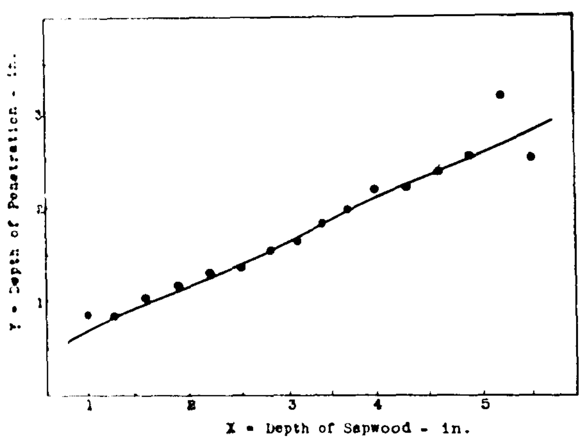
Since the correlation coefficient is as large as it is, we have reason to believe that there is a stochastic relationship between the depth of penetration and depth of sapwood. Furthermore, since, as we have pointed out in a previous footnote, further analysis of the data revealed that the regression was linear, we may make use of the correlation coefficient thus obtained to give us the line of regression,

$$Y - \overline{Y} = r_{YX} \frac{\sigma_Y}{\sigma_X} (X-\overline{X}),$$

and this equation becomes upon substitution of the values:

$$Y = .472209\, X + .215401$$

Fig. 5.5 shows how closely this line of regression fits the observed means of the columns.



Regression Y on X

Fig. 5.5

This concluded the first step in the investigation which showed quite definitely that one of the quality characteristics which must be considered in the comparison of the quality of penetration as given by different plants was the depth of sapwood.

Soon we had occasion to make use of this information in the further study of quality of treatment. In this case we started with the information presented in Fig. 5.6 which shows the scatter diagrams of depth of sapwood vs. depth of penetration for groups of the same kind of pole submitted to the seven different plants. In line with the results previously obtained it was necessary to reduce these data to the averages, standard deviations and correlation coefficients presented in Table 5.8. Immediately we note that there is quite a wide variation not only in the average depths of penetration but also in the average depths of sapwood. Naturally we would look to see if the depth of sapwood is the controlling factor in producing an assignable difference in depth of penetration. Even a little investigation of the information presented in

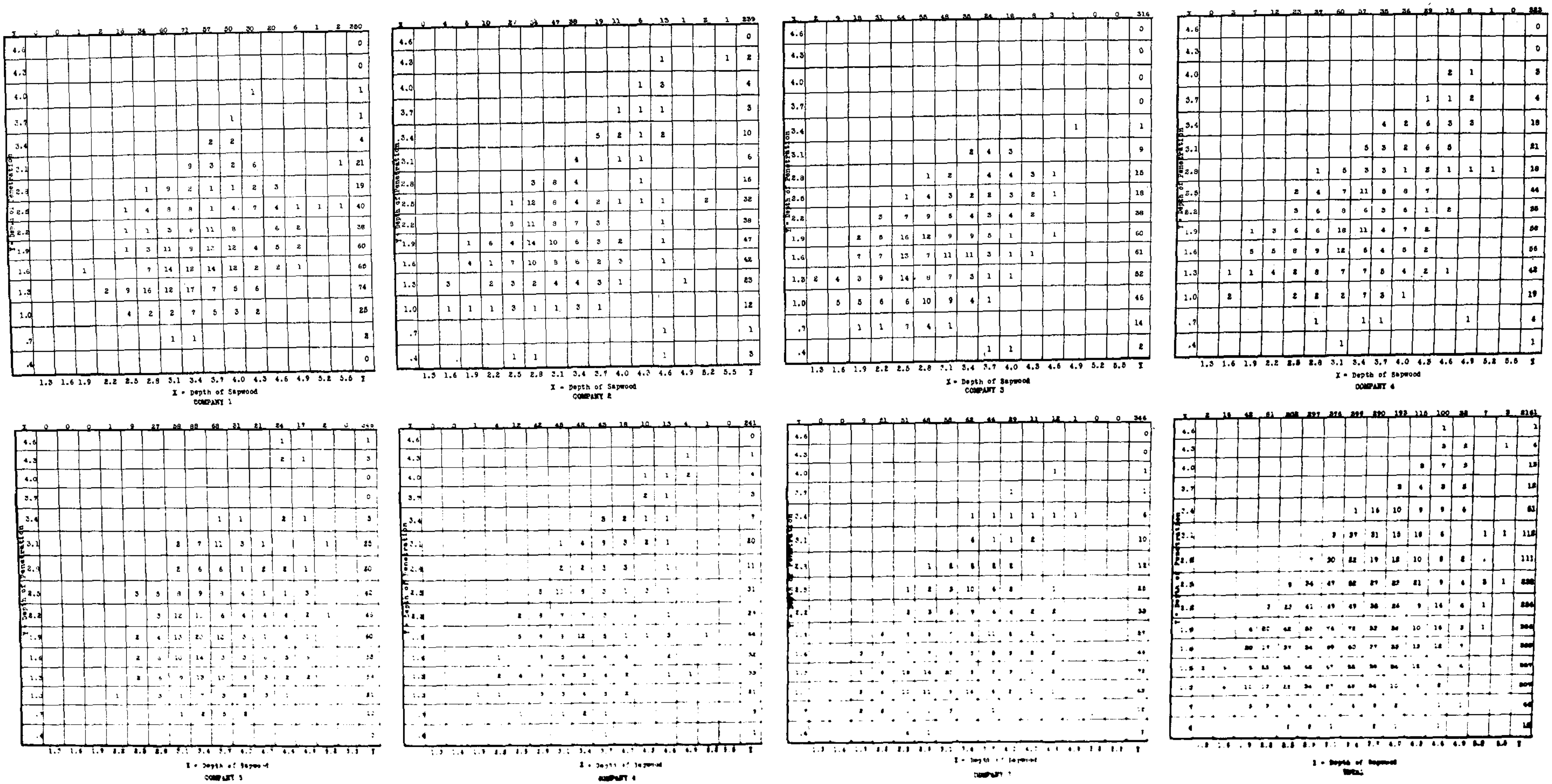

Y = Depth of Penetration
X = Depth of Sapwood
COMPANY 1
COMPANY 2
COMPANY 3
COMPANY 4
COMPANY 5
COMPANY 6

FIG. [illegible]

X = Depth of Sapwood in inches
Y = Depth of Penetration in inches

Company	No. of Poles	No. of Borings	$\bar{X}$	σ_X	$\bar{Y}$	σ_Y	r_{XY}
1	48	350	3.5611	.6060	1.8966	.6326	.2597
2	50	239	3.1552	.6922	2.0795	.7091	.4403
3	50	316	2.8959	.6667	1.7016	.5925	.4913
4	47	323	3.3963	.7093	2.0653	.7153	.1584
5	48	346	3.6107	.5935	1.9642	.6865	-.1815
6	50	241	3.4012	.5987	2.0320	.7546	.4181
7	50	346	3.1850	.6385	1.6832	.6563	.3855
Total	343	2161	3.3242	.6863	1.9053	.6911	.3926
			$\bar{\sigma}_X$ = .6436		$\bar{\sigma}_Y$ = .6767		$\bar{r}_{XY}$ = .2817
			*$\bar{\sigma}_X$ = .6422		*$\bar{\sigma}_Y$ = .6720		*$\bar{r}_{XY}$ = .2656

*Weighted Average

TABLE 5.8

Table 5.8 reveals that this is not the case. Then the next logical step is to consider the control charts of the type given by Criterion I modified, I.E.B. 1, for the averages, standard deviations and correlation coefficients. These are presented in Fig. 5.7. Immediately we get an indication that there is some assignable cause of variation in the correlation coefficient itself. With this information at hand, an attempt was made to find what other information the data yielded which might indicate the nature of some of the assignable causes of variation. Along with the data presented in Fig. 5.6, other data were given indicating the number of poles of a given depth of sapwood which were penetrated. An examination of these showed that the frequency distribution in a column array of a scatter diagram including not only saturated but also unsaturated poles was of the general nature indicated in Fig. 5.8.

Upon the basis of the kind of reasoning presented in I.E.B. 1, in connection with Criterion 3, we see at once that the existence of this type of distribution indicates quite clearly the presence of an assignable cause of difference in the poles other than depth of sapwood. In this connection, Fig. 5.9 shows the ratio of the number of saturated to the total number of poles having a given depth of sapwood for each of the seven different companies. A rather

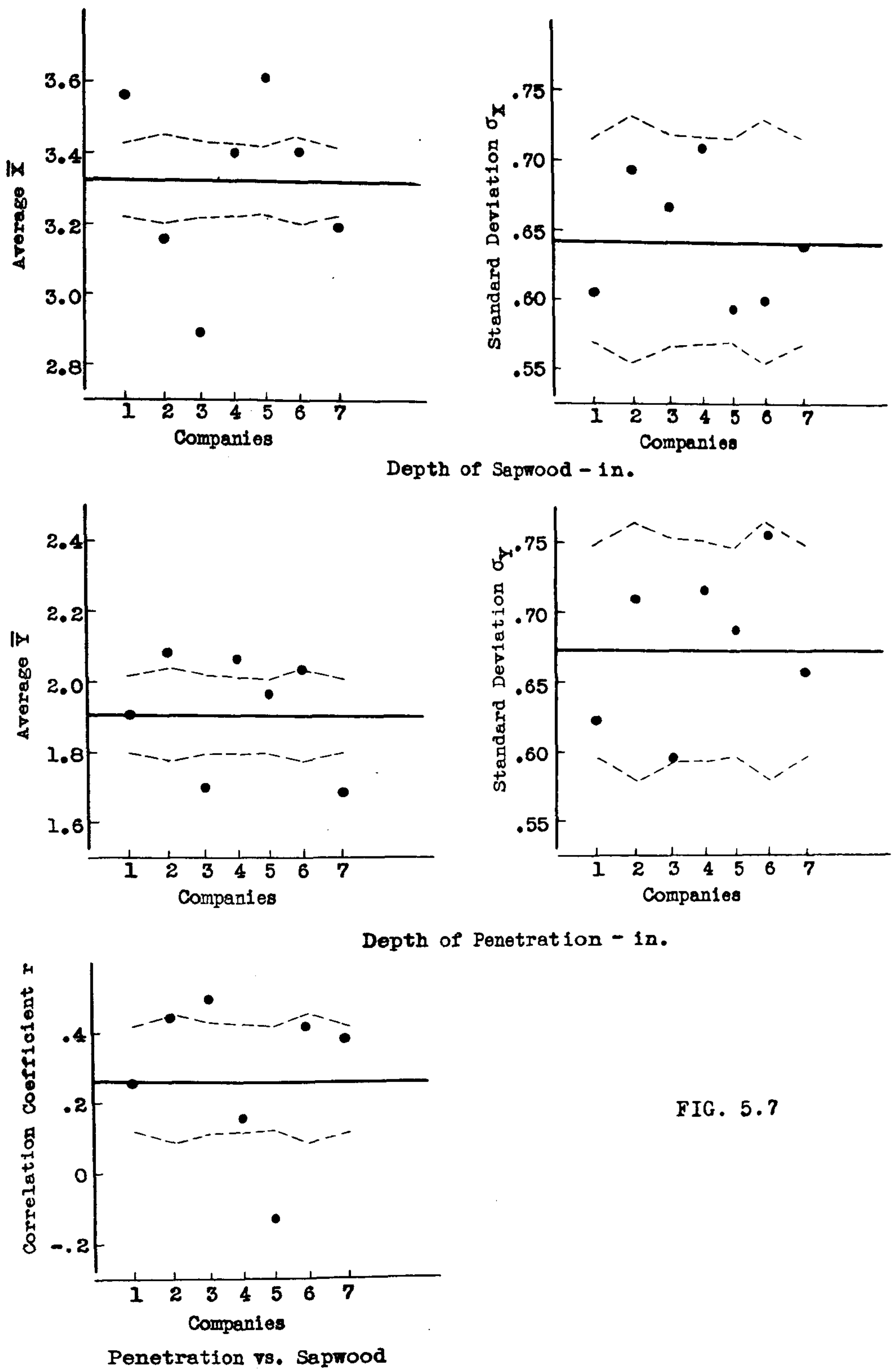


FIG. 5.7

interesting observation can be made that this ratio remains comparatively large even for poles of very deep sapwood, although, of course, we must allow for the

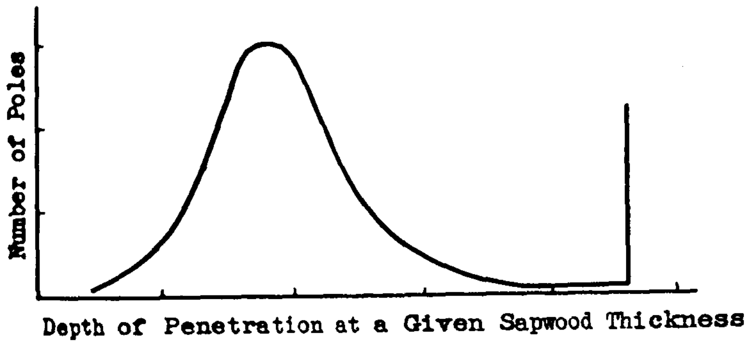


FIG. 5.8

fact that the ratios for very high or very low depths of sapwood are based upon only a very small number of poles and hence are subject to very large sampling fluctuations. The information presented in this figure, however, gives still further evidence of the existence of assignable causes of variation other than depth of sapwood.

As a result of this study, what then are we in a position to say in the light of the information presented above?

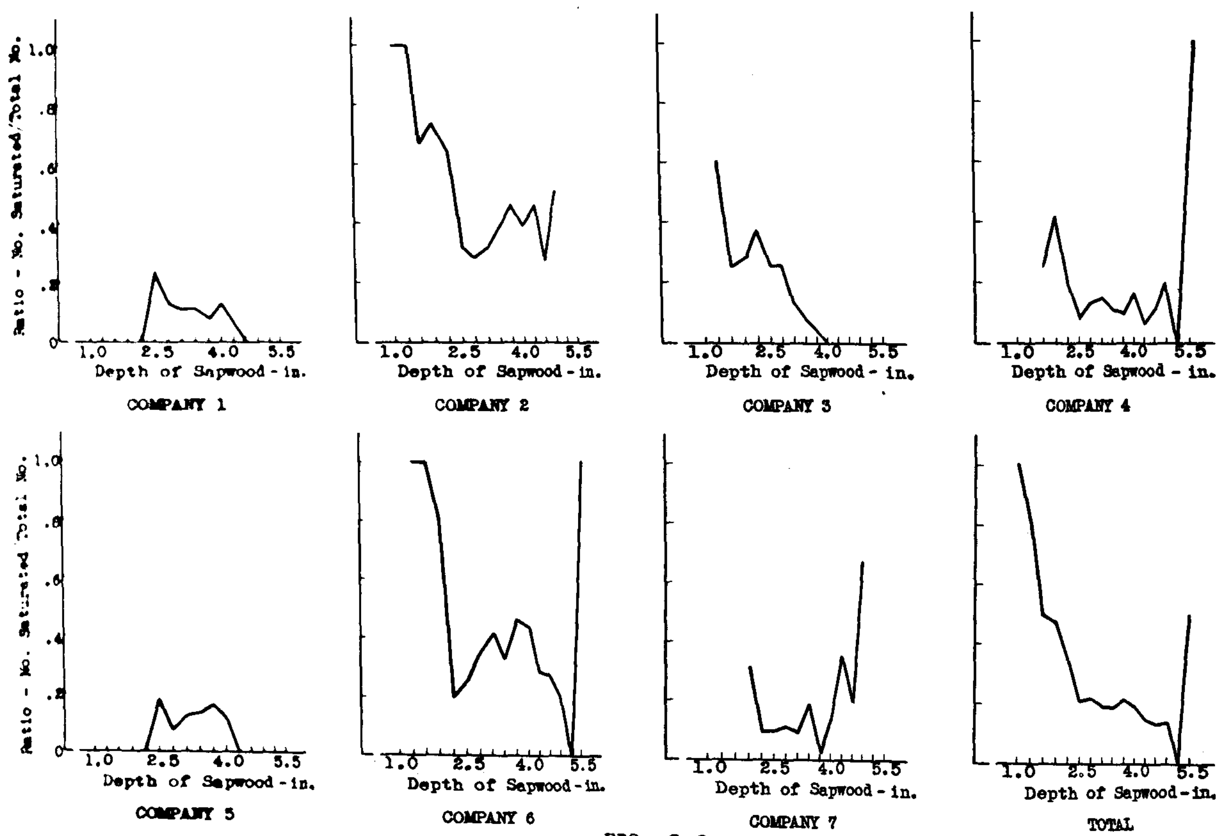


FIG. 5.9

We can say quite definitely that, at least for certain homogeneous groups of poles of a given kind, there is a very definite correlation between depth of sapwood and depth of penetration, making it necessary for us to consider depth of sapwood along with depth of penetration in comparing plants in respect to the quality of their treatments. Furthermore, the information presented in Fig. 5.7 indicates quite conclusively that there are other at present unknown assignable causes (or perhaps you might say assignable characteristics of telephone poles) which it is necessary to tabulate before an adequate comparison of the treatments at different plants can be made.

It so happened that during the time these analytical results were being obtained, one of the engineers interested in this work did visit several of the plants entering into the investigation presented above. He came back with several a priori suggestions as to the assignable characteristics of poles and assignable causes of variation in the method of treatment which serve to explain several of the significant fluctuations beyond the limits in the control charts of Fig. 5.7, although the investigation up to the present time is far from complete in the sense that not all of these significant variations have been accounted for. The point which we wish to make in this connection is that the method of analysis as presented above naturally led to the assumption of the existence of assignable causes of variation in treatment other than those previously suggested and the direct engineering work up to the present time has definitely justified the conclusions based upon the analysis. Quite naturally, even more significant results could have been obtained had we had to begin with the a priori information which is now available so that groups of data coming from each of the different plants could have been rationally sub-divided in a way that is not now possible since the records were not kept in a form to make such sub-divisions possible. This illustrates the point which should always be kept in mind by inspection engineers, namely, that far greater use can be made of data where to begin with, rational hypotheses are available so as to make possible the tabulation of the data in rational subgroups. It seems reasonable to believe that we can look forward to even more applications of the simple functions outlined above for examining the control of quality when engineers in charge of experimental investigation become more familiar with the importance of recording data so that they may be divided into rational subgroups.

As stated above, not only depth of penetration but also quantity of oil or creosote absorbed is used as a measure of the quality of treatment. Therefore it may be of interest to consider briefly here an outline of the study which has been made to determine whether or not the amount of oil or creosote absorbed is correlated with the depth of sapwood. The data representing the combined results for the seven companies presented in Fig. 5.10 make such a study possible. The analysis of these gives a correlation coefficient of .21.

1	3	6	14	33	49	64	61	51	29	22	7	5	343
[illegible]								1					1
[illegible]								2					2
[illegible]					1								1
[illegible]							1						1
[illegible]			1			1	2	1			1		6
[illegible]	2					3	4	2	1	1			13
[illegible]				1	5		2	3	1	2		1	15
[illegible]	[illegible]	[illegible]	1	6	6	6	10	2	2	2		1	39
[illegible]			5	9	12	14	14	14	4	4	1		76
[illegible]		2	3	2	10	22	14	12	11	8	4	2	90
[illegible]			2	10	9	11	11	9	9	5	1	1	68
[illegible]		2	1	4	5	7	3	5	1				28
[illegible]			1	1	1								3
[illegible]				1									1
	1.6	1.9	2.2	2.5	2.8	3.1	3.4	3.7	4.0	4.3	4.6	4.9	Σ

X = Depth of Sapwood in Inches

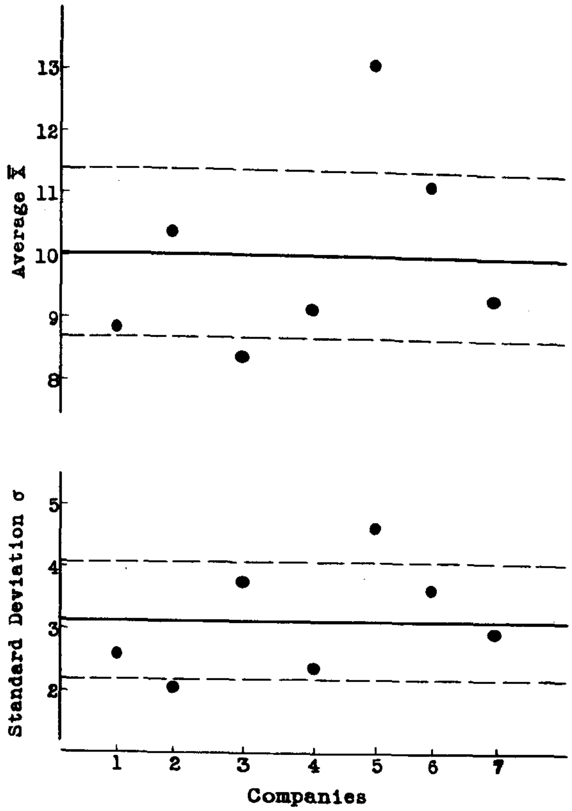
Fig. 5.10

If we interpret this small value of correlation coefficient as indicating lack of correlation, we would be justified in considering the results presented in the control charts of Fig. 5.11 as indicating the existence of assignable causes of variation in the treatment of poles, as measured by the extraction of oil, that cannot be explained in terms of the depth of sapwood. In other words, these results point to the fact that in the comparison of different plants in respect to quality of treatment there are assignable factors which should be known but which up to the present time are not known. In other words, these results definitely indicate that in the study of the treatment of telephone poles there are assignable causes of variation as yet unknown which should not be left to chance.

F. Control of Quality of Aluminum Die-Castings

We have already had occasion to refer to some of the results which have been obtained to date in one of the committees of the American Society for Testing Materials charged with the investigation of the physical properties of aluminum die-castings. One of the particular properties which is of interest is that of tensile strength. This in turn is stochastically related to at least two characteristics of the material, namely, hardness and density.

We have no intention of doing more than indicating how some of these



Extraction of Oil per cu. ft. of Sapwood

Fig. 5.11

results are being analyzed and how the results of the analysis are being used. For example, Table 5.9 presents the analytical results obtained in connection with a certain group of the data divided into rational subgroups as indicated in the table. Control charts for these various quantities, similar to those presented in Fig. 5.7, gave very definite indication of the presence of assignable causes of variation. These results when reviewed by some of the members of the committee were explained in terms of what appeared to be possible assignable causes.

The significance of these results is for the most part this, - that so long as the data analyzed in this way indicate the presence of assignable causes of variation, there is reason to be-

X_1 = Tensile Strength in pounds per square inch

X_2 = Hardness in Rockwells "E"

X_3 = Density in grams per cubic centimeter at 25 °C.

	Average			Standard Deviation			Correlation Coefficient		
Company	$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$	σ_1	σ_2	σ_3	r_{12}	r_{13}	r_{23}
C	33399.167	68.4917	2.66642	2564.958	10.19448	.08294	.683	.150	.421
D	28215.833	68.0250	2.65542	4317.552	14.48937	.08331	.876	.891	.819
G	30312.667	66.5667	2.63250	2187.992	10.17254	.11334	.714	.786	.801
W	33150.167	76.1167	2.68942	3954.093	11.08090	.07664	.715	.858	.622
S	34269.000	69.9250	2.74850	2715.022	9.88147	.09073	.805	.637	.552

TABLE 5.9

lieve that much more work should be done in specifying the method of making the alloys before we can expect to get a controlled product.

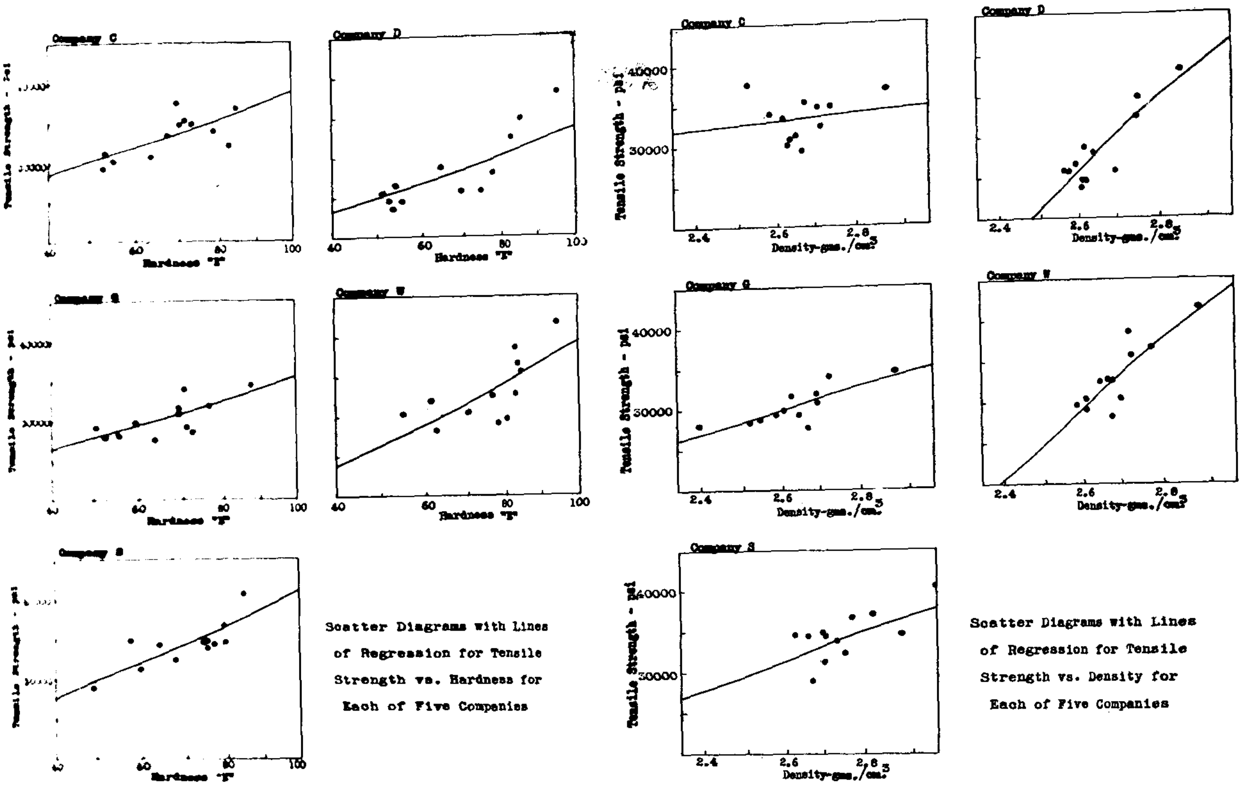


FIG. 5.12 - ROUND SPECIMENS ALUMINUM DIE-CASTING

Fig. 5.12 presents graphically some of the results which have been of particular interest to the members of this Committee. It shows the regression between some of the various factors presented in detail in Table 5.9. Even a casual observation indicates that there are large variations in the slopes of the regression lines and assignable differences in these slopes are easily shown to exist by means of Modified Criterion I of I.E.B. 1. Just to cite one particular instance, we note that the regression in the upper left hand corner between tensile strength and density indicates very low correlation. The likely assignable cause of this result, as proposed by members of the Committee, is the existence of blow holes in the castings. This is merely indicative of the way this information is being used by members of this Committee.

PART VI

Correction of Data for Errors of Measurement

1. How to Correct Data for Errors of Measurement

In a previous part we have considered the problem of measuring the quality of a number of things in terms of the observed values of quality characteristics. Obviously these measurements, like all others, are subject to errors of measurement. Now in reporting on the results of measurements of quality, it is very essential indeed that our quality reports reflect the true quality of product as nearly as we can attain to this, instead of merely the observed quality of product.

In Part I we started with the simplest concept of measurement, namely, that involved when we measure something in terms of its own yardstick. In that discussion we see imposed the necessity for always considering not only the estimate of the expected value $\bar{X}'$ of the measured characteristic X but also of the error or standard deviation of measurement σ'. We may state, therefore, as a simple principle which should be made general practice in all inspection engineering work that we should always obtain the best estimate σ of the error of measurement σ'. Numerous instances have been found within the last two or three years where it has been impossible to obtain any adequate estimates of this error of measurement primarily because no estimate had ever been made and the original data which would have made possible such an estimate had been destroyed. This is particularly true in connection with the study of the qualities of raw materials. It should be made a principle that, whenever measurements are taken for a definite purpose, every single observation and not merely averages of groups of observations be tabulated.

Passing on to Part II, we have considered the problem of estimating the error of measurement of the quality Y expressed in terms of m different characteristics $X_1, \ldots X_2, \ldots X_i, \ldots X_m$, by the relationship

$$Y = f(X_1, X_2, \ldots X_i, \ldots X_m).$$

We found there that to minimize the standard deviation of the quality characteristic Y, it is necessary to know the standard deviation of each of the m different characteristics. Here again in an important application of available theory we find it necessary to have recorded every single measurement on every characteristic. In this case it should be remembered that the standard deviation of the quality Y is given by the expression

$$\sigma_Y = \sqrt{\sum_{i=1}^{m} a_i^2 \sigma_i^2} \qquad a_i = \left(\frac{\partial f}{\partial X_i}\right)_{\bar{X}_1, \ldots \bar{X}_m} \qquad (6.1)$$

to a first approximation. Equation 6.1 is generally known as the law of propagation of error.

Also in Part III we considered the methods of minimizing the error of measurement of quality through the use of curves and planes of regression. In general, it was shown that the standard deviation s_Y of the quality Y measured from the line of regression is

$$s_Y = \sigma_Y (1 - r^2).$$

Similarly, for the plane of regression of Y on the m characteristics,

$$s_Y = \sqrt{\frac{\sum_{i=1}^{n} \left[y - (a_1 x_1 + a_2 x_2 + \ldots + a_m x_m)\right]^2}{n}}$$

where y and $x_1, x_2, \ldots x_m$ are deviations from their respective mean values and n is the number of sets of observations. The standard deviation from the general curve of regression is

$$s_Y = \sqrt{\frac{\sum_{i=1}^{n} \left[(Y - f(X)\right]^2}{n}}$$

where f(X) represents the regression curve fitted to the data and n is the number of pairs of values of X and Y.

Under the conditions of practice, it can be stated as a universal fact[1] that: The standard deviation of the error of measurement of an average of n observed values of quality Y derived in any way whatsoever is

$$\sigma_{\bar{Y}} = \frac{\sigma_Y}{\sqrt{n}}.$$

---

1. This is true assuming only finite universes as will be pointed out in Part III of the book on quality control.

Hence by taking the average of a number n of measurements we may decrease the error of measurement in any quality characteristic on a single thing by increasing the number n.

Now of course, when we pass over to the measurement of quality of a number of things we find, as already stated, that even in the cimplest case, we make use of three different functions of the observed set of data representing quality characteristics on n different things. These are the average $\bar{X}_i$, standard deviation σ_i and correlation coefficient r_{ij}, where the meaning of these is that given in Part V.

Now any estimates such as those just mentioned derived from the observed measurements will be subject to the errors of measurement of the respective variables. The general method of correcting for these errors of measurement has previously been discussed.[1] Here it was shown that, if a characteristic X is measured by a method subject to error, then the observed standard deviation σ_0 of a set of n observed values of X on as many different things of the same kind will be given by the following relationship

$$\sigma_0 = \sqrt{\sigma_T^2 + \sigma_E^2} \tag{6.2}$$

where σ_T is the standard deviation of the true quality X and σ_E is the standard deviation of the errors of measurement. In general, we must solve this expression for the standard deviation σ_T of the true quality characteristic X.

In a similar way the correlation coefficient r_0 between any two sets of observed values will be related to the true correlation r between the two qualities in the following way:

$$r = \frac{\sigma_{1_0}\,\sigma_{2_0}}{\sigma_{1_T}\,\sigma_{2_T}}\,r_0 \tag{6.3}$$

where σ_{1_0} and σ_{2_0} are the observed standard deviations of the quality characteristics X_1 and X_2 respectively and σ_{1_T} and σ_{2_T} are the standard deviations of the true values of these quality characteristics.

Therefore, any report on the quality of a number of things should always be corrected by means of Equations 6.2 and 6.3 as was pointed out in the publication referred to above.

---

1. Shewhart, W.A. "Correction of Data for Errors of Measurement" Reprint B-186.

2. Special Applications

Few relationships are required more often in the measurement of quality than those given above in Equations 6.1 and 6.2. Numerous instances have arisen where we have had to make use of these in correcting measurements of quality of both raw materials and finished product. In fact, as previously pointed out, these relationships are made the basis of corrections in the presentation of our inspection reports on the true quality of product. We shall now briefly review some of the applications of these formulas to the study of a typical special problem in the testing of pole timber for modulus of rupture.

Suppose we have a pole whose modulus of rupture we wish to determine. One way is to break it by one of the three methods described in Part II of this bulletin and thus measure the modulus Y directly. The other is to take a number n_1 of small clear specimens from the pole and break these using the average of these n_1 tests as a measure X of the modulus. There are several reasons why it is cheaper to test a reasonable number n_1 of small clears from a pole than it is to test the pole itself by breaking.

The problem, however, is not quite so simple as it looks, because we must first show that there is a definite relationship between Y and X and that the comparatively large variation in the modulus of rupture of small clears from the same pole does not so reduce the efficiency of the small clear method of measurement as to compensate for the extra cost of measurement by pole test method. Some of the considerations now to be reviewed have been influential in bringing the Laboratories to their present position of placing greater reliance in the results of the pole test method than in those of the small clear method as a basis for setting standard strength figures for telephone poles at least at the present stage of the development of the two tests.

Suppose that we were to break n poles and thus determine their moduli of rupture $Y_1, Y_2, \ldots Y_n$, and then that we were to break n_1 small clears from each of the n poles thus getting n sets of values as shown schematically in Table 6.1, where $X_1, X_2, \ldots X_n$ are the averages of n_1 tests on pole 1, n_1 tests on pole 2 and so on respectively. Assume for the time being that all of these tests could be made on poles and small clears at the same moisture content. We would get some stochastic relationships as shown schematically in Fig. 6.1.

Pole Number	Small Clear Specimens						Pole Test
	Test 1	Test 2	Test 3		Test n_1	Av.	
1	X_{11}	X_{12}	X_{13}	. . . . .	X_{1n_1}	X_1	Y_1
2	X_{21}	X_{22}	X_{23}	. . . . .	X_{2n_1}	X_2	Y_2
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
n	X_{n1}	X_{n2}	X_{n3}	. . . . .	X_{nn_1}	X_n	Y_n

TABLE 6.1 - Modulus of Rupture - psi.

Assuming the regression to be linear, we could use the line

$$Y - \overline{Y} = r_o \frac{\sigma_{Y_o}}{\sigma_{X_o}} (X-\overline{X})$$

as in the earlier parts of this bulletin where σ_{Y_o} is the observed standard deviation of Y and σ_{X_o} is the observed standard deviation of the averages X_1, X_2,...X_n of Table 6.1, and r_o is the observed correlation coefficient.

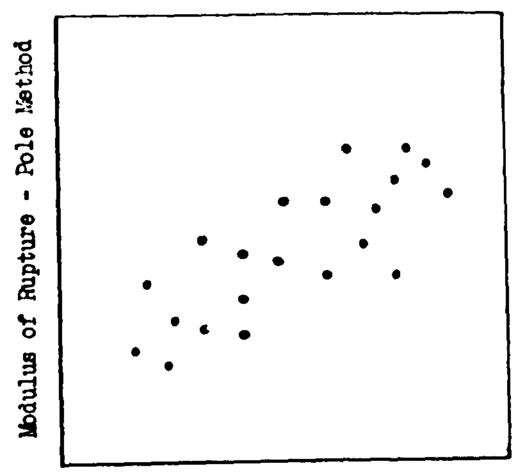


FIG. 6.1

Now, of course, the standard deviation σ_{X_o} is made up of two parts, such that

$$\sigma_{X_o} = \sqrt{\sigma_{X_T}^{2} + \frac{\sigma_{E_X}^{2}}{n_1}}$$

where σ_{X_T} is the estimate of the standard deviation σ'_{X_T} of the expected values obtained by the small clear method and σ_{E_X} is the estimate of the standard deviation σ'_{X_E} of the measurements on small clears from a given pole.

In some of our cooperative studies with members of the Forest Products Laboratory it has been possible to obtain estimates of the standard deviation

σ_{E_X} for certain species of timbers although this work has not progressed to such a state that these values can be given as standards. We have available, however, a series of recent tests on small clears taken from lodgepole pine poles which we can use to illustrate the method involved in calculating the above mentioned standard deviations.

In this series of measurements sixteen poles were used and a number of small clears from each pole were broken. The number of clears from each pole, the average modulus of rupture of these and the observed standard deviation of the measurements are presented in Table 6.2 We see at once that the average

Number of Pole	Number of Pieces from each Pole	Average Modulus of Rupture	Observed Standard Deviation of Tests on Small Clear Specimens from one Pole in psi
50	17	12139	1195
51	19	12034	1889
52	20	8972	1576
53	8	11263	1592
54	17	10063	1907
55	17	11089	1657
56	22	11385	1590
57	16	10518	1139
58	27	11982	1484
59	15	9553	1635
60	19	11586	1145
61	18	13388	1313
62	19	11389	2452
63	16	11894	1150
64	16	11604	934
67	22	13466	2109
		11395	
			Av. 1548

TABLE 6.2 - Data Showing Variability Among Observed Values of Modulus of Rupture of Small Clears from the Same Pole.

standard deviation, giving equal weight to that of each pole, is 1548 psi. This obviously may be taken as a reasonable estimate of the standard deviation σ_{E_X} of measurement by the small clear method of lodgepole pine poles.

Returning now to Equation 6.3 we see that for the case of lodgepole pine poles, making use of the above data.

$$r_o = \frac{\sigma_{Y_T}\ \sigma_{X_T}}{\sigma_{Y_o}\ \sigma_{X_o}}\ r < \frac{1}{1.4}\ r$$

where only one clear is tested for each pole and $\sigma_{Y_T} = \sigma_{Y_o}$ in this case. Obviously the observed correlation will be much less than the true correlation unless we make n_1 appreciable, that is unless we test several small clears from each pole. We see at once that the standard deviations of the values of Y about the line of regression given by the relationship

$$S = \sigma_{Y_O} \sqrt{1 - .7^2 r^2}$$

is almost as large as σ_{Y_O} unless the correlation coefficient r is very near unity.

Now the standard deviation of the averages in Table 6.2 is 1168 psi and from this we get a value for σ_{X_T} equal to 1097 psi. From a series of pole tests on the same species we get a value for σ_{Y_O} which is less than 1000 psi. In other words it appears that σ_{Y_O} is likely not greater than σ_{X_T} which is a very important fact indeed. Stated in another way the standard deviation by pole tests is possibly less than the standard deviation of averages of n_1 small clear tests per pole even when $n_1 \longrightarrow \infty$. This leads us to ask how many poles n_2 would have to be tested by the small clear method, n_1 clears to the pole, in order to give a standard deviation of the average of the $n_1 n_2$ tests equal to the standard deviation of n pole tests. Under the above conditions[1] it follows that the approximate relationship between n, n_1 and n_2 is

$$\frac{1}{n} = \frac{1}{n_2}\left(1 + \frac{1}{n_1}\right)$$

Plotting n_1 as abscissae we may represent as ordinates the number of poles n_2 tested by the small specimen method expressed as a ratio of the number n of poles tested by the pole method.

For example, we see that if it has been decided to test five small pieces from each pole we would have to test by the small test method 1.2 times the number of poles tested by the pole method. Now, if we are going to break 100 poles to estimate the modulus of rupture of the species, we would have to test five specimens from each of 120 poles to get the same precision as would be obtained from the 100 pole tests. In a similar way we may find corresponding values of n and n_2 for any desired value of n_1.

---

1. Specifically this assumes that $\sigma_{Y_T} \doteq \sigma_{X_T} \doteq \sigma_{E_X}$.

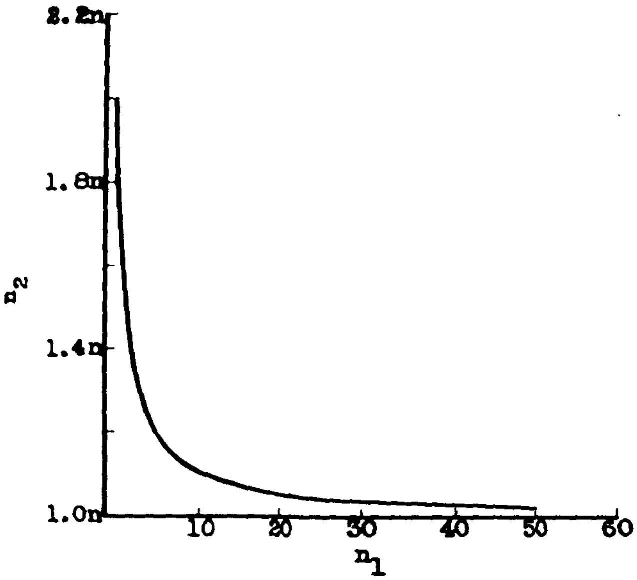


FIG. 6.2 - RELATIVE NUMBER OF POLES TESTED BY THE TWO METHODS CORRESPONDING TO A GIVEN PRECISION

The efficiency (usually expressed in per cent) of the small specimen method of test as compared to the pole method is defined by the ratio $\frac{n}{n_2}$. For example, if it is necessary to test twice as many poles by the small specimen method as by the pole test method to get the same precision we would infer that the small specimen method of test were only 50% efficient. In other words, the efficiency of the small test method as compared to the pole method will be expressed by the inverse of the ratio of the respective number of poles tested by the two methods.

In general then it will be enlightening to graph the function

$$\text{Efficiency} = \frac{n}{n_2} = \frac{n_1}{1 + n_1}$$

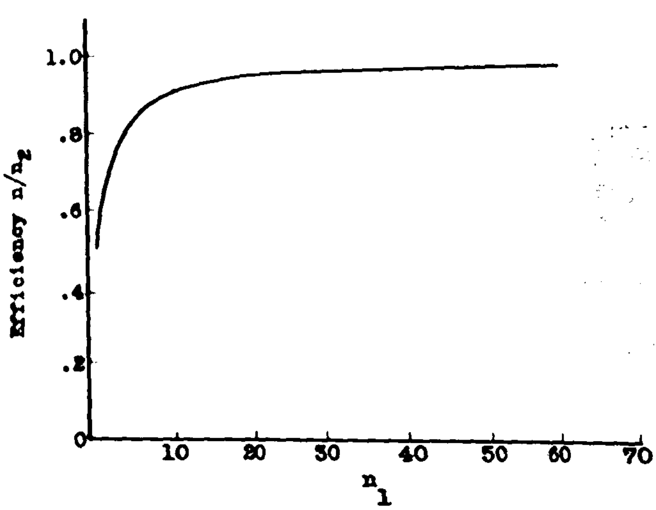


FIG. 6.3 - INHERENT INEFFICIENCY OF SMALL SPECIMEN METHOD OF TEST

Fig. 6.3 makes the efficiency situation very clear. Thus if $n_1 = 5$, the chart shows that the small specimen method is only about 83% efficient, and only as n_1 becomes large do we approach very closely the efficiency of the pole test method.

Of course this whole discussion applies only to lodgepole pine poles at a given moisture content assuming that the modulus of rupture of small clears is independent of the section of the pole from which they are taken. Even under these conditions we see that the small clear method is not very efficient even assuming that the correlation between the modulus of rupture of poles by pole test method and the expected moduli of rupture for small clears from the same poles is large.

Going further we find that it is not practically possible to measure poles and clears from the same pole at the same moisture content and that we do not have sufficient information to make possible a correction for moisture content. Furthermore we find that the modulus of rupture for small clears taken from a pole is not independent of the section of the pole from which they are taken. In this situation it is evident that error corrections would become extremely complicated and uncertain, at least until such time as we have obtained much more definite information than we now have even after cooperative studies with members of the Forest Products Laboratory about the effect of moisture on modulus of rupture of both poles and small clears and the way in which the modulus of rupture of small clears vary throughout the pole.

For these reasons it is believed that pole tests must be relied upon, at least at the present time, to furnish data as a basis for establishment of standard values for modulus of rupture of telephone poles.