Basic Concepts of Statistical Quality Control*

PAUL IRICK
Research Statistician
Highway Research Board

1. *The Purpose of Statistical Quality Control*

I would like to describe a few concepts that I think are basic to the understanding of statistical quality control. Perhaps we should start by distinguishing between quality control and statistical quality control. It seems to me that quality control of a manufactured product involves three activities that are repeated over and over:

a. the development of information about the process and product

b. engineering decisions based on this information

c. implementation of the decisions

As an example, suppose a base course has been finished and is up for acceptance. Density measurements may give information from which to decide whether or not the job is satisfactory, and this decision will be implemented in one way or another. I think we will agree that good quality control requires good information, good decisions, and good implementation.

Now statistical quality control is mainly concerned with the development of information that is objective, unbiased, and adequate for decisions that will be based on this information. In the example that we have mentioned, statistical quality control would be concerned with sampling techniques, measurement procedures, and with converting the measurements into quantities that are immediately useful for making engineering decisions. In a nutshell, then, statistical quality control is aimed at producing good information for quality control activities, either during the manufacturing process or when the manufactured product is submitted for acceptance.

2. *Variability in All Aspects of a Manufacturing Process*

The most basic concept in statistical quality control is that there is variability in every aspect of a manufacturing process. Materials vary

* Presented at the Purdue Road School, Purdue University, Lafayette, Indiana, March 23, 1964.

from place to place and from time to time. So do construction procedures. As a result, the properties of the manufactured product have what we will call inherent variability.

When we measure any property such as density, we use measurement systems that involve variable equipment and variable test procedures. Thus we expect to get testing variability even if the inherent variability is very small. Finally, when a variable property is evaluated with a variable measurement system, we are almost certain to get variable measurements[1] as a basis for our decisions.

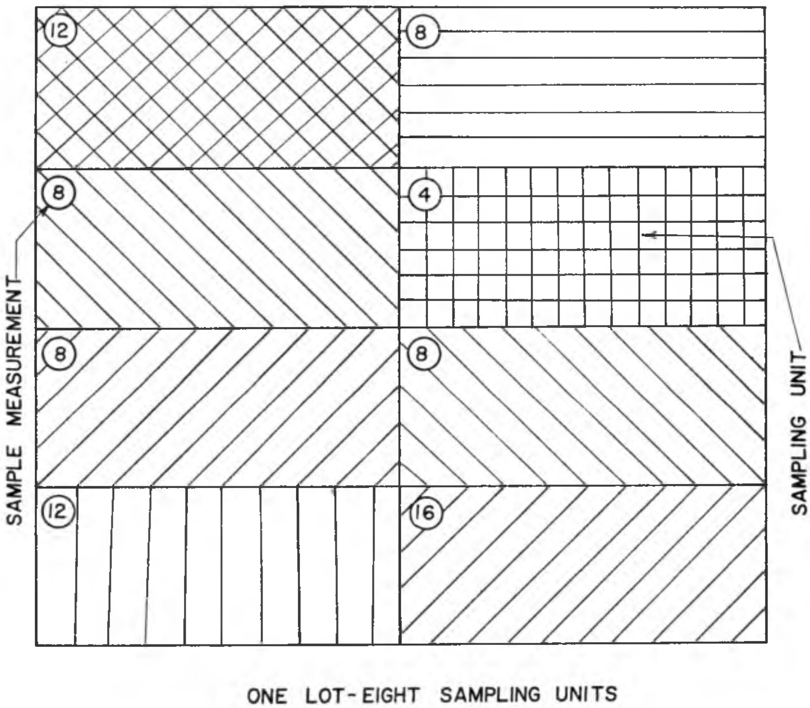

Fig. 1. Lot, sampling units, measurements.

3. *Sampling Units and Lot Distributions*

To talk intelligently about statistical quality control we must first define quantities of material, or completed product, to which quality control decisions apply. Let us call one of these quantities a *lot*. For

---

[1] Analysis of variance for measurements from designed experiments can be used to obtain separate estimates of inherent and testing variability for any controlled process. In general, testing variability can be made arbitrarily small by using the average of a sufficient number of repeated measurements on the same sampling unit.

example, a lot might be an aggregate stockpile, a batch of concrete, an hour's production of hot mix, one hundred yards of subgrade embankment, one-half mile of surfacing, etc. Next we assume that each lot consists of a relatively large number of *sampling units* such that any measurement we obtain applies to just one sampling unit. In Figure 1 we see a lot that consists of eight sampling units, and one measured value of some property is shown for each sampling unit. The sampling units of a lot could be time units, volume units, area units, etc., but the important concepts are: (a) an observed measurement applies to just one sampling unit, and (b) a quality control decision applies to the whole lot. We see that the illustrative measurements are variable. If we make a frequency distribution of the measurements from all sampling units we have a *lot distribution* which shows how many sampling units have each measured value. Figure 2 shows the lot distribution of measurements that were given in Figure 1.

Statistical quality control is mostly concerned with three characteristics of lot distributions:

(a) the mean of the distribution, $\overline{X}'$. For the example $\overline{X}' = 9.5$.

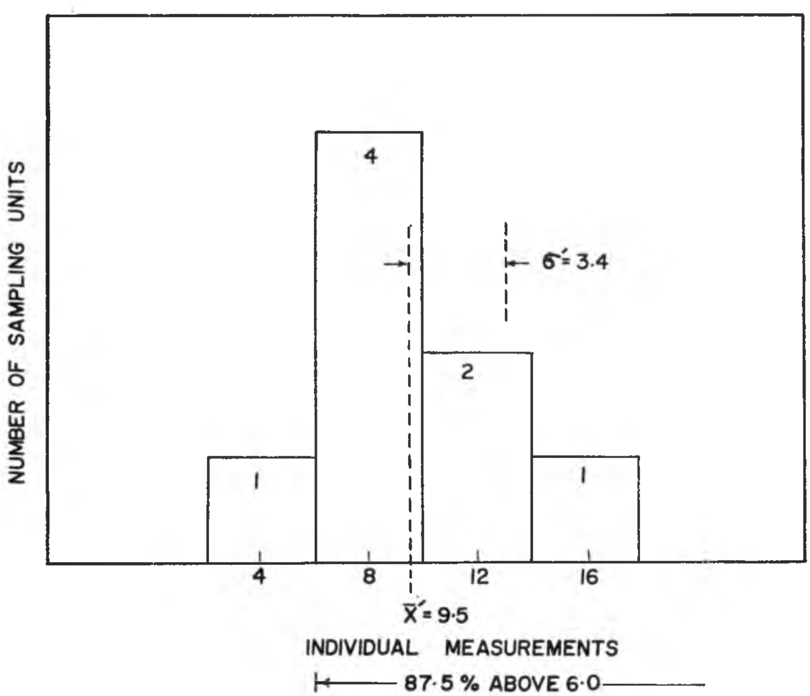

Fig. 2. Lot distribution of test measurements.

(b) the standard deviation of the distribution, σ'. The standard deviation of a set of measurements is an average difference between individual measurements and the distribution mean. Thus σ', which is 3.4 for the example, is a measure of variability in the lot distribution.

(c) the percent of the distribution that lies below or above a single limit, or the percent that lies between two limits. As an example, $7/8 = 87.5$ percent of the illustrative measurements are greater than 6.0.

Many distributions of measurements have the general shape of a normal curve. If we have values for $\overline{X}'$ and σ' we can use normal curve area tables to approximate the percent of individual measurements that fall between any two limits. In Figure 3 we see superimposed on the previous lot distribution a normal curve whose mean is $\overline{X}' = 9.5$ and whose standard deviation is $\sigma' = 3.4$. In a normal distribution, for example, about 96 percent of all measurements are between $\overline{X}' - 2\sigma'$ and $\overline{X}' + 2\sigma'$. As another example, the figure shows that about 85 percent of the normal curve area is above 6.0, very close to the 87.5 percent of the actual lot distribution which lies above 6.0.

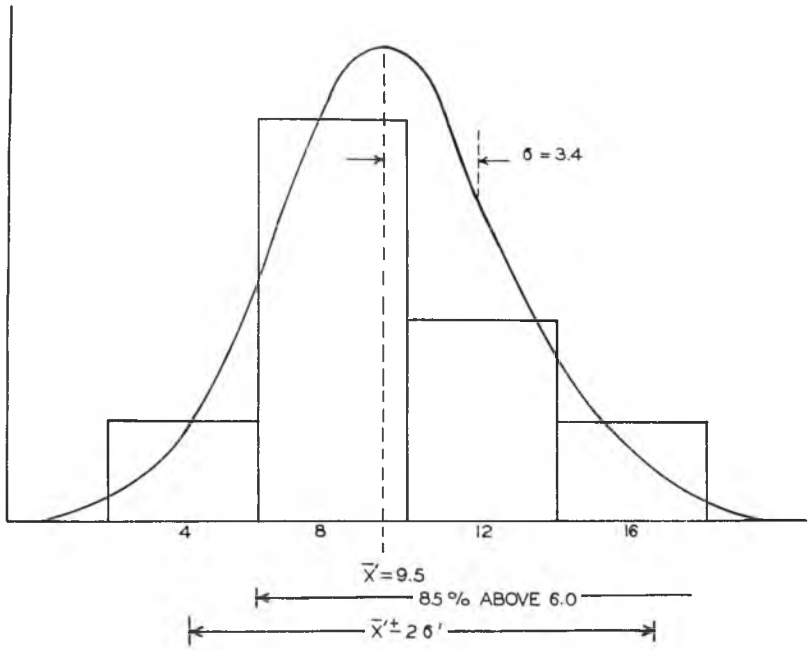

Fig. 3. Normal curve approximation to lot distribution.

We can see how information on these three distribution characteristics can be applied to quality control. For example, we may be able to decide whether the lot quality is acceptable—in terms of $\overline{X}'$, σ', and the percent of the distribution that falls between given limits. If the lot is not acceptable we may decide to alter the materials and/or construction procedures to change $\overline{X}'$ and/or σ', and thus to change the percent of measurements that fall between the given limits.

Another example of a lot distribution is shown in Figure 4 which gives a frequency distribution of thicknesses for over 1,000 asphaltic concrete cores. We will assume that each core represents a small amount of surface area, and the whole distribution represents about one-quarter mile of two-lane pavement, 24 feet wide. The mean of this lot distribution is $\overline{X}' = 3.2$ inches, and the standard deviation is .37 inches. Suppose the specified thickness for this lot is 3.0 inches and it is desired that nowhere should the thickness be less than 2.5 inches. The lot distribution shows, however, that two percent of the sampling units have thicknesses below this limit. We will not pass any judgment here, but we may suppose that this information could be used for accepting or rejecting the lot—as far as thickness is concerned. As a matter of fact, if this were an actual distribution we would have a quarter mile of pavement that had over 1,000 holes in it. This brings us to our next basic concept—that we never expect to see a complete lot distribution.

4. *Inferences About Lot Distributions*

In practice we know we must operate from measurements on only a few of the sampling units in any lot, perhaps on just two or three out of hundreds or thousands that actually exist. Thus quality control decisions for lots are based on uncertain information, that is, on inferences rather than facts about lot distributions. This is a basic concept in quality control. We must admit from the start that our sample measurements will not give sure information about $\overline{X}'$, σ', or the percent of the lot distribution between any two limits. But we can use statistical methods to produce objective and unbiased estimates of these quantities, and to regulate the degree of uncertainty that is involved in our inferences about a lot distribution. Offhand we might assume that the more samples we observe, the more certain we will be about the lot distribution. This assumption is valid only if the sampling units have been selected at random from all possible sampling units in the lot. In this illustration, for example, 500 cores all taken from the right side of the distribution will give us quite biased estimates of both $\overline{X}'$ and σ'. A random sample of five or ten measurements, however, could give

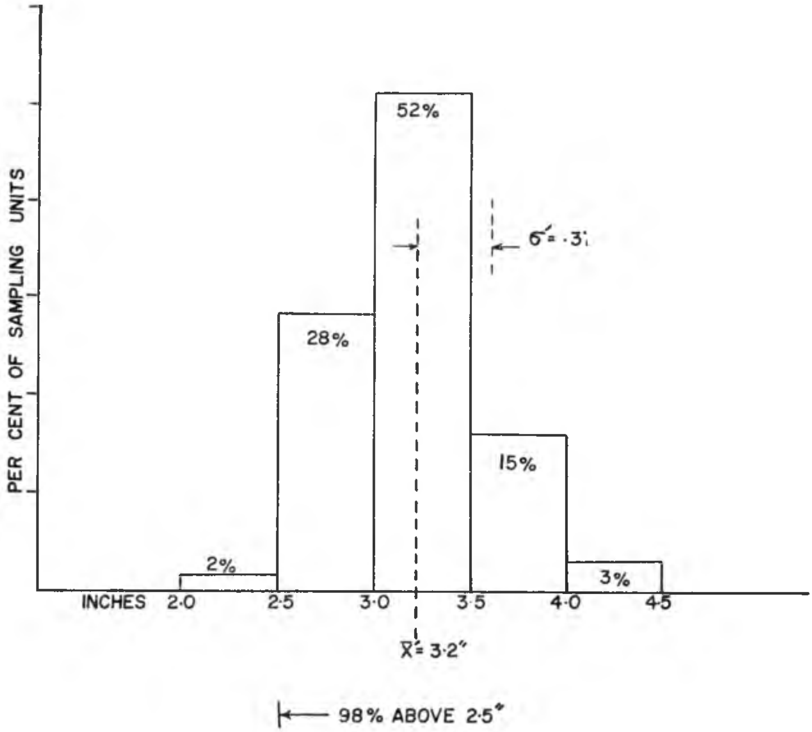

Fig. 4. Lot distribution of surfacing thickness—1030 cores.

a fairly accurate picture of the lot distribution, and in addition could be used to determine the degree of uncertainty of the picture they do give.[2]

If we agree that we must have random samples and that we can only draw inferences about lot distributions, then we are ready to look at two major techniques of statistical quality control—the use of control charts, and the use of acceptance sampling plans.

5. *Use of Control Charts in Manufacturing Processes*

We can't discuss control charts in any detail, but we will indicate the kind of information they provide. Suppose that we select n sampling units, perhaps only two or three, from each successive lot of material or constructed item that is being inspected and measured for quality. For each set of n measurements we can calculate a mean, $\overline{X}$, and a standard deviation, σ, then plot these values as shown in Figure 5. If we wish,

[2] Statistical methods can be used to determine confidence limits, with any degree of confidence, for lot characteristics that are obtained from a given number of sample measurements.

we can imagine that this figure is concerned with slump measurements on two sampling units from each successive batch (lot) of concrete.[3]

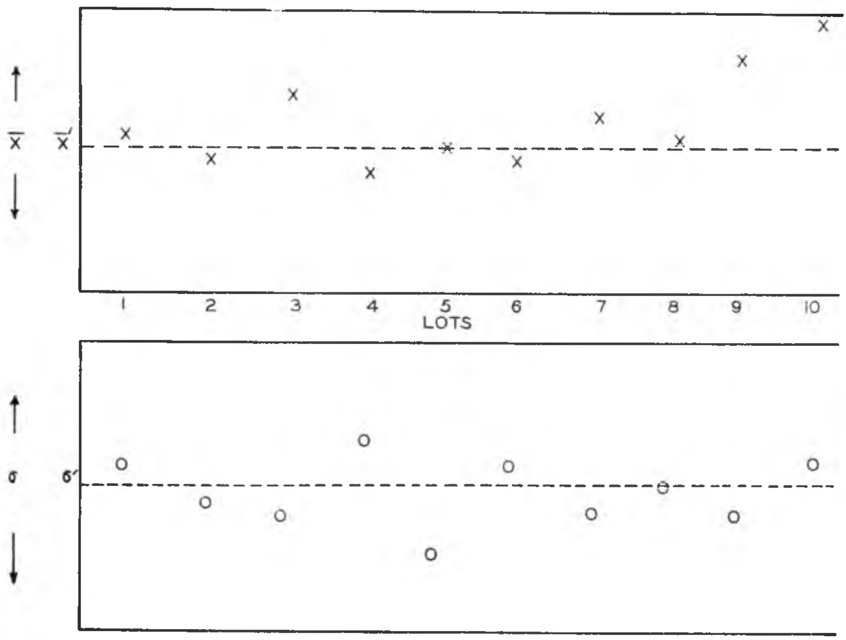

Fig. 5. Sample means and standard deviations.

We will suppose here that control charts are used to decide whether the materials or construction procedures should be altered during the manufacturing process. For our illustration we will suppose that the centerlines for the $\bar{X}$ and σ charts represent aimed-at values for the lot distributions mean, $\bar{X}'$, and standard deviation, σ', respectively.

Before the charts are complete we must add control limits in such a way that when we compare the plotted points with the control limits we will be able to infer whether the process is running along satisfactorily or whether it should be changed. To discover how to compute control limits let's go back to our first illustration whose lot distribution is now shown on the left in Figure 6. Here we had a lot consisting of eight sampling units whose measurements formed a distribution

---

[3] The standard deviation of a set of n measurements is found by summing the squared deviations from the mean, dividing this sum by n - 1, then taking the square root. Thus $\sigma = \sqrt{\Sigma(X-\bar{X})^2/(n-1)}$.

When n is small, variability is almost always expressed in terms of the range, R, from the largest X to the smallest X.

having $\overline{X}' = 9.5$ and $\sigma' = 3.4$. Suppose we can afford to observe $n = 4$ randomly selected sampling units from this lot. It turns out there are seventy possible sets of four that can be selected from the eight units, and any of these sets is equally likely to occur with random sampling. The figure shows that the means, $\overline{X}$, of the seventy possible selections form a frequency distribution whose mean is the same as the mean of the lot distribution, $\overline{X}' = 9.5$. The distribution of means, however, is much more normal than was the lot distribution and is not as variable, ranging from 7 through 12 instead of from 4 to 16. It is a mathematical fact that the distribution of means, for all possible sets of sampling units, will have a standard deviation of $\sigma'/\sqrt{n}$. Since the example is for $n = 4$, the distribution of $\overline{X}$ has only one-half the variability that existed in the lot distribution. Furthermore, we can find from normal curve tables that there is only a very small chance that any set of n selected units will have a mean, $\overline{X}$, that is any farther from $\overline{X}'$ than $2\sigma'/\sqrt{n}$, say. We won't attempt to say why, but there is also a very small chance that the standard deviation, σ, of n random measurements will be farther from σ' than $2\sigma'/\sqrt{2n}$.[4]

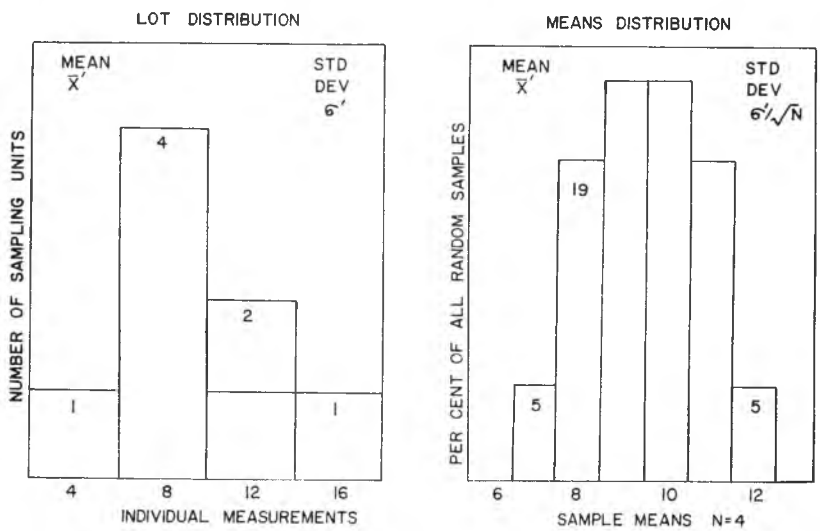

Fig. 6. Distribution of individual measurements and of means.

We can use these facts to construct control limits on our charts, as in Figure 7, at distances $2\sigma'/\sqrt{n}$ from $\overline{X}'$ on the $\overline{X}$ chart, and at

4 Limits set in this way would be "two-sigma" limits, and include about 96 percent of the sampling variation of X about X' and of σ about σ'. In many control chart applications, "three-sigma" limits are used to include about 99.7 percent of the sampling variation.

distances $2\sigma'/\sqrt{2n}$ from σ' on the σ chart. We are now ready to draw inferences about the sampled lots:

(a) If the sample means, $\overline{X}$, and standard deviations, σ, show no trends and if all fall inside the control limits, we infer that the lot distributions are being controlled at the levels indicated by $\overline{X}'$ and σ'. *The control limits allow for nearly all uncertainty* that is associated with our incomplete information about the lot distribution.

(b) If points on the σ chart show a trend or fall outside the control limits, we infer that the corresponding lots have a different variability than σ'. We could be wrong, since the control limits do not cover all uncertainty that is associated with sampling variations, but we prefer to believe that σ' has changed. If the inferred change is toward a higher σ', and we are dealing with slump tests, we may decide that the mixing time should be increased—to give more uniform batches. If the inference is toward a lower σ', we might decide to check the measurement system—to see if the slumps of two specimens from the same batch are really as alike as the measurements indicate.

(c) If the σ chart shows control, but $\overline{X}$ points show trends, or fall outside the control limits, we infer that such lots have means that are different from $\overline{X}'$—and decide to take action. If this $\overline{X}$ chart is for slump measurements, for example, we infer that the average slump for lots 9 and 10 is actually higher than $\overline{X}'$. In this case it might be decided to use less water in the mixing operation. We could be wrong, but it's a good bet that the process mean has shifted away from $\overline{X}'$ in lots 9 and 10, and that corrective action needs to be taken. Samples from succeeding lots will soon indicate on these charts whether or not the process has been brought back in control.

There are many types of control charts and many ways to use them, but we have the basic idea if we see that they can provide objective and useful information for decisions that must be made in the face of uncertain information about the process.

6. *Acceptance Sampling Plans*

I will conclude with the application of statistical quality control to decisions which result in the acceptance or rejection of submitted lots. An acceptance plan involves four considerations:

a. a procedure for selecting and measuring sampling units

b. one or more statistics computed from the measurements

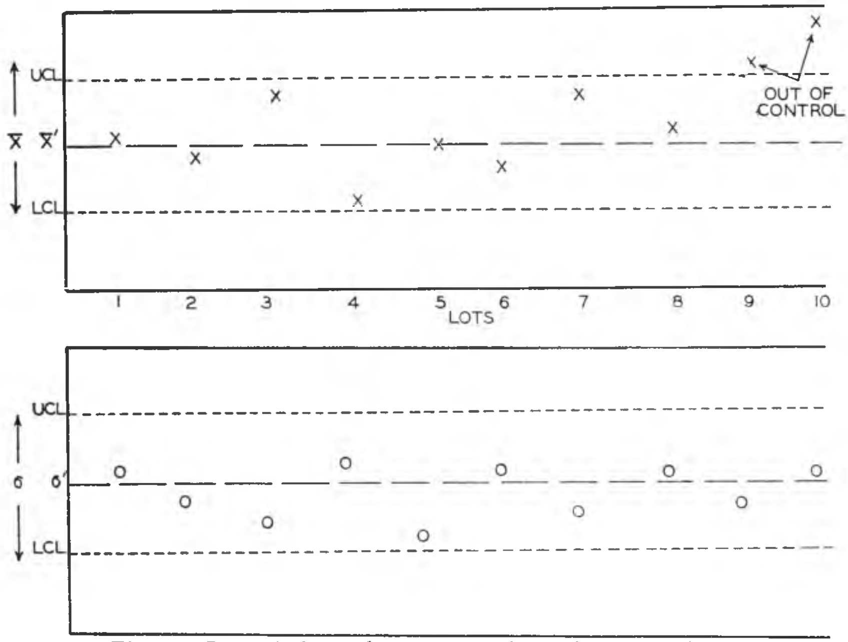

Fig. 7. Control charts for means and standard deviations.

c. a rule for accepting or rejecting the submitted lot—in terms of the measurement statistics

d. an operating characteristic for the acceptance plan

We have time to look at only one illustration. Suppose that one hundred (lineal) yards of compacted subgrade must be accepted before any further construction is permitted. This lot of material might cover, say, 15,000 square feet. Let's suppose that this area consists of sampling units each ten feet by ten feet, and that density measurements will be made for each of four randomly selected sampling units. To simplify our example we will assume that the standard deviation of the lot distribution is known to be $\sigma' = 6$ pcf from past experience, and that the subgrade will be accepted or rejected according to the mean value, $\bar{X}$, of our four sample densities. Thus $\bar{X}$ is the statistic whose value will be computed from the four sample measurements. Next we will suppose that our rule is simply to accept the subgrade (as far as density is concerned) if $\bar{X}$ is 100 pcf or more, and reject it if $\bar{X}$ is less than 100 pcf.[5] In effect we have just spelled out a specification for subgrade

[5] This example deals with only one specification limit. In other cases this rule might call for acceptance if $\bar{X}$ is between two limits and rejection if $\bar{X}$ is outside either limit.

density. But before we can tell whether this rule makes any sense we must look at its *operating characteristic,* that is, what actually happens when we start using this acceptance specification.

Two lot distributions are shown in Figure 8, both with $\sigma' = 6$ pcf, but one with $\overline{X}' = 95$ pcf and the other with $\overline{X}' = 110$ pcf. We will suppose that we want to reject the subgrade if its mean density is as low as 95 pcf and that we want to accept the subgrade if $\overline{X}'$ is as high as 110 pcf.

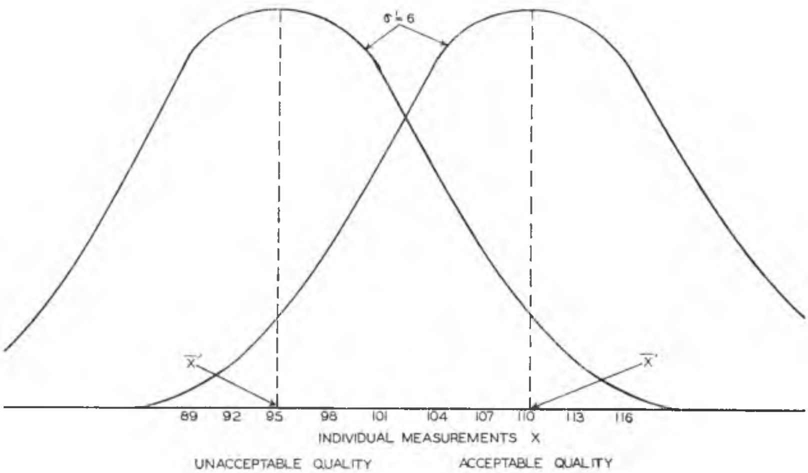

Fig. 8. Poor and good quality lot distributions.

We realize that we do not get to look at the complete lot distribution, but only at four measurements from any one lot. The two distributions in Figure 9 show all the possibilities for sample means, $\overline{X}$, when $n = 4$ sampling units are observed from either lot distribution of the previous figure.

Our acceptance rule is shown across the bottom of the figure. Now about 5 percent of all means from the undesirable distribution extend above $\overline{X} = 100$. Thus when undesirable lots with $\overline{X}' = 95$ pcf are completed, our acceptance rule will result in acceptance about one time in twenty. This is a *consumer risk*—that poor lots will be accepted.

On the other hand, we see that a very small percent of means from the desirable distribution extends below $\overline{X} = 100$. This tells us that our acceptance plan will result in the rejection of lots with $\overline{X}' = 110$ pcf perhaps once in 2,000 submissions. This is the *producer's risk*—that good lots will be rejected. A basic concept of quality control is that *any* acceptance plan or specification, statistical or not, involves the chance that poor material will be accepted or that good material will be rejected. The element that statistical quality control adds is simply a way to know what these risks are and thus to set specifications in terms of agreed-upon risks. We can see from the figure, for example, that moving our acceptance statistic, $\overline{X} = 100$ pcf, in either direction will change both risks. Thus if we accept when $\overline{X} = 105$ or more, we will greatly reduce the chance of accepting poor lots, but at the same time we will increase the risk of rejecting good lots. To use statistical acceptance plans in highway work, the producer (contractor) and the consumer (highway department) have to decide how to share the two risks.

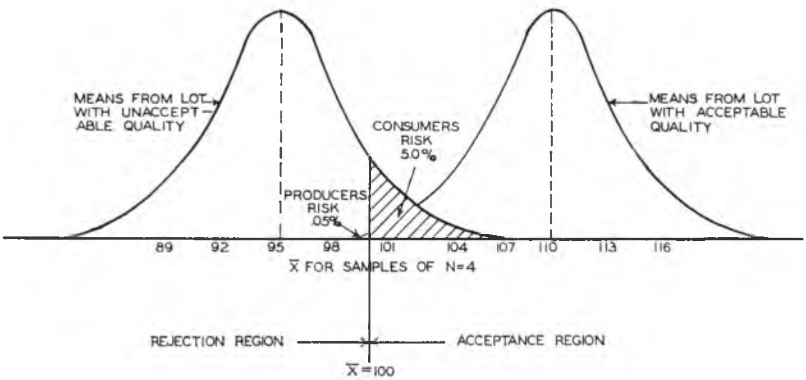

Fig. 9. Means distributions—acceptance rule.

Before we leave this figure we should note that these curves are for n = 4 sampling units selected randomly from a completed lot. If we were to decrease the size of n the curves would be wider, and *both* risks would increase. If we increase the sampling, however, the curves will become narrower and both risks will decrease. Thus we have another way to control producer and consumer risks.[6]

Finally we must note that we have talked about acceptance and rejection as though there were only two lots involved, one with $\overline{X'} = 95$ pcf and one with $\overline{X'} = 110$ pcf. We need to see how our plan will operate for any submitted lot quality. In the next figure (Figure 10), the horizontal scale gives the average quality characteristic, $\overline{X'}$, in submitted lots—ranging from 80 to 120. For the acceptance rule we have been discussing, two operating characteristics are given by the curves in the figure, one for n = 1 and one for n = 4.

[6] The risks would also be different for different values of the lot standard deviation, σ'.

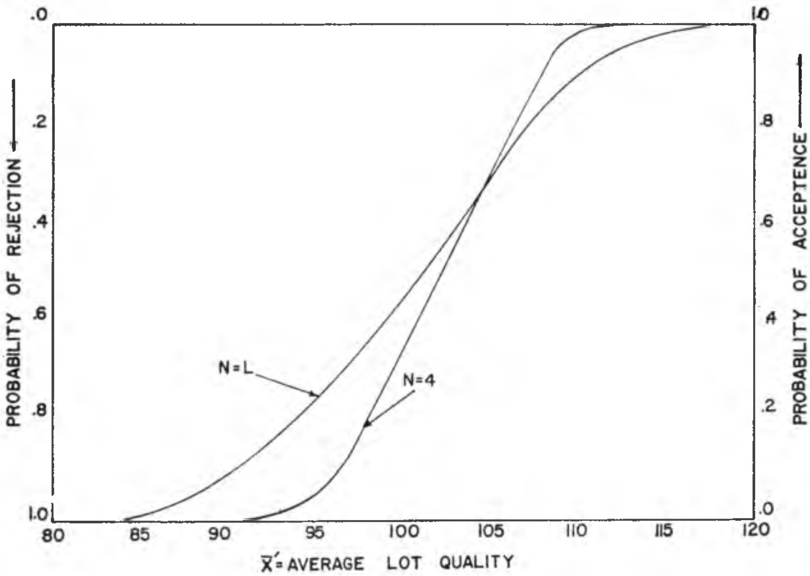

Fig. 10. Operating characteristic of the acceptance plan.

If we read from the bottom up to a curve and over to the left scale we find the probability that the submitted lot will be accepted. If we read from the curve to the right hand scale we find the probability that the lot will be rejected when the rule is used. We see that low quality lots, say when $\overline{X}'$ is less than 95 pcf, will be rejected 95 percent of the time if we use n = 4 samples, but only 80 percent of the time if we use only n = 1 sample. The curves also show that high quality lots, say when $\overline{X}'$ is 110 pcf or more, will be accepted almost always if n = 4, but only 95 percent of the time if n = 1. We can thus see that n = 4 gives a better operating characteristic than does n = 1. By now it should be rather clear that we can have almost any operating characteristic we want for an acceptance specification—by altering the number of samples, by altering the definitions of high and low quality, or by altering the risks we will assume for accepting lots with poor quality or rejecting lots with good quality. But the main point is that statistical acceptance plans or acceptance specifications, if you please, are objective and have known operating characteristics. If we cannot draw the operating characteristic of an acceptance specification, I am afraid we really don't know what is going on.

Before I conclude we should decide what to do with our illustrative subgrade. Suppose our four random density measurements are 94, 98, 103 and 109 pcf. These numbers look pretty variable but their standard deviation is 6.5 pcf, only slightly more than the assumed value of $\sigma' =$ 6.0 pcf. The sample mean is $\overline{X} = 101$ pcf, and since this exceeds $\overline{X} = 100$, our rule says to accept the embankment—at least as far as density is concerned. We realize this may be the one time in twenty that relatively poor material will be accepted by our plan, but we have already agreed to assume this risk. If the sample mean had been less than 100 we would reject the embankment in its present form—and perhaps decide that more compaction is necessary.

We have now discussed many of the basic concepts of statistical quality control. Through further study and through practical applications I think we will find that statistical methods can be quite useful in the quality control of highway materials and construction.

REFERENCE

Grant, E. L., *Statistical Quality Control,* McGraw-Hill, 1952.

(This is only one of many available texts on the subject of statistical quality control.)