rt No. FHWA-RD-73-77

ALITY ASSURANCE FOR PORTLAND
MENT CONCRETE

J. B. DiCocco

September 1973
Final Report

This document is available to the public
through the National Technical Information
Service, Springfield, Virginia 22151

3/4

Prepared for

FEDERAL HIGHWAY ADMINISTRATION
Offices of Research & Development
Washington, D.C. 20590

U.S. Federal Highway Administration.
Report no. FHWA-RD-73-77.

Technical Report Documentation Page

1. Report No.
FHWA-RD-73-77

2. Government Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle
QUALITY ASSURANCE FOR PORTLAND CEMENT CONCRETE

5. Report Date
September 1973

6. Performing Organization Code

7. Author(s)
John B. DiCocco

8. Performing Organization Report No.
NYSDOT-ERD-73-RR-10

9. Performing Organization Name and Address
Engineering Research and Development Bureau
NYS Department of Transportation
State Campus, Albany, New York 12226

10. Work Unit No. (TRAIS)
FCP 46C1-113

11. Contract or Grant No.
SSN 45-1

12. Sponsoring Agency Name and Address
Offices of Research and Development
Federal Highway Administration
US Department of Transportation
Washington, D.C. 20590

13. Type of Report and Period Covered
Final Report

14. Sponsoring Agency Code
P-0032

15. Supplementary Notes
Prepared in cooperation with the Federal Highway Administration through HPR Program. Implementation Manager E. C. Granley (HDV-22)

16. Abstract
Quality assurance has been successfully used in most industries, but the construction industry and State agencies are only now beginning to recognize its importance. The work reported here results from this recognition and is one attempt to show how the elements of quality assurance can be applied to concrete. Chapter I deals with concrete uniformity and compliance with specifications requirements, under the current degree of process control and inspection. Both uniformity and compliance are found to be poor. Chapter II deals with techniques that producers should use to achieve compliance with concrete requirements and includes three case studies to show that process control does indeed lead to compliance. It is also stressed that process control is the responsibility of producers, but recognizing that inspecting agencies (having done most of the testing in the past) may have more pertinent information than producers, guidelines are suggested for producers to use until they accumulate data of their own. The next three chapters deal with acceptance sampling (inspection). In Chapter IV, current inspection schemes are reviewed with particular attention to those used in New York. It is concluded these schemes afford little protection to concrete buyers and alternative statistically sound sampling plans are suggested for inspection of fresh and hardened concrete. The advantages and shortcomings of the suggested plans are discussed in Chapter V, along with the monetary value of such plans as applied to concrete. Finally, recognizing the trend toward the improper use of acceptance control charts for inspection of construction materials, these tools are reviewed in Chapter VI. It is concluded that concrete acceptance control charts are inappropriate and should not be used in concrete inspection.

17. Key Words
Quality Control
Acceptance Sampling

18. Distribution Statement
No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22151.

19. Security Classif. (of this report)
Unclassified

20. Security Classif. (of this page)
Unclassified

21. No. of Pages
157

22. Price

Form DOT F 1700.7 (8-72)

Reproduction of completed page authorized

PREFACE

Statistical tolerancing, quality control methods, and acceptance sampling have been successfully used in most industries, but the construction industry and state inspection agencies have seldom taken advantage of these powerful tools. In recent years, emphasis has been placed on statistical tolerancing. Many states have undertaken uniformity studies and have used the results to establish practical specification limits. Unfortunately, in most cases the effort has ended there. Statistical tolerancing is only one step in the establishment of quality assurance programs. The results of uniformity studies have little value unless properly used to design mathematically sound control procedures and acceptance sampling plans that can guarantee quality levels consistent with engineering requirements and costs.

Quality assurance has been defined in many ways. Miller and Freund (1) see it as consisting of three elements -- quality control, acceptance sampling, and the establishment of tolerance limits. Following this definition, this report has three parts, each dealing in turn with these elements of quality assurance. The first (Chapter I) concerns the uniformity of concrete properties in relation to specification limits, while the second and third deal respectively with concrete quality control (Chapter II) and acceptance sampling (Chapters III through VI).

The primary objectives of the research reported here were 1) to determine the current uniformity of portland cement concrete in New York in terms of measurable properties, and 2) to use the results a) to set specification limits of the properties measured in inspection at achievable levels, and b) to use the measured variations in designing statistically sound control procedures and acceptance sampling plans. Secondary objectives were to establish possible correlations between the various concrete properties measured, and to gather information on the concrete-making process so as to identify possible sources of variability.

Readers unfamiliar with quality assurance terminology will find a glossary of terms (with abbreviations used in this report) in Appendix A.

This study was performed under administrative and technical direction of William C. Burnett, P.E., Director, Engineering Research and Development Bureau. The author gratefully acknowledges the helpful suggestions and guidance of Peter J. Bellair, P.E., Associate Civil Engineer, during both the course of this research project and the preparation of the final report. The author also thanks Gerald L. Anania, Senior Engineering Technician, for numerous, time-consuming computations of probabilities used for the *OC* curves shown, and for preparation of the original figures and tables. William P. Chamberlin, P.E., Associate Civil Engineer, and Duane E. Amsler, Assistant Civil Engineer, were helpful in providing some of the data presented, and in reviewing parts of the manuscript. Finally, the author acknowledges assistance of the Materials Bureau, H. H. McLean, Director, and James J. Murphy, Research Liaison Representative, in conducting the case studies reported in Chapter II, and for providing record sampling and other data.

CONTENTS

TABLES

FIGURES

I. VARIABILITY OF CONCRETE PROPERTIES

A. Importance of Data Concerning Concrete Variability

Knowledge of product variability is essential to any mathematically based quality assurance program. Specifically, it is needed 1) to set specification limits properly, 2) to design statistical quality control procedures, and 3) to design efficient acceptance sampling plans. Knowledge of product variation is indispensable both in determining a product's ability to meet specification limits dictated by engineering requirements, and in re-evaluating specification limits for products in use when compliance becomes a problem.

Ideally, in choosing a product for engineering use, its quality determinants are identified, and the limits determined within which these properties can be allowed to vary without impairing performance. Having set these limits, variation in the product's essential properties is measured to see if it can meet the engineering constraints. If the product's variation is so large that only a small portion complies with the specifications, a new product is chosen, the production process is altered to reduce variation, or the design is revised.

Often, however, products and processed materials are put to engineering uses without following this ideal procedure. In such cases, the specification limits for their essential properties are likely to be set at desirable but unachievable levels, and enforcement of the specifications then causes high rejection rates. This often results in resetting specification limits at levels consistent with the product's inherent variability. As in the ideal case, product variation must be measured to make the proper adjustments.

The importance of variation does not end with setting proper specification limits. Once these limits have been set at achievable levels, the manufacturing process must be monitored to achieve compliance. Monitoring is necessary to guard against assignable causes forcing the process to operate at a different level of control than required to obtain the desired results. The process is usually monitored with control charts. Setting proper control limits for these charts requires knowledge of product variability.

In most industries, process control is exercised by the producer. In such cases, the buyer can choose to use the producer's process control data as a criterion to judge compliance, or to ascertain that compliance is actually attained by acceptance sampling. When the buyer resorts to acceptance sampling, knowledge of the product's achievable variation again becomes important, in designing efficient sampling plans.

Because of its importance, product variation is generally determined from data obtained by designed experiments. Historical data are usually shunned because their reliability is often questionable and usually do not include sufficient replicate testing and sampling to determine sampling and testing variations.

Until 1965, most historical data available on concrete had all these shortcomings. The Federal Highway Administration (then the Bureau of Public Roads), recognizing this problem, sponsored a large effort to collect reliable data on concrete properties. To this end, FHWA published The Statistical Approach to Quality Control in Highway Construction (2). This gave guidelines for experiments to measure variations, including an analysis of variance model and a computer program for analysis of the data. New York participated in this effort, and the data presented in this chapter of this report were obtained according to those guidelines.

B. Experimental Design

In 1967, provisions were made to obtain data to measure the uniformity of concrete produced in New York State under the level of process control normally exercised by inspection forces. No provisions were made further to control concrete-making processes, or to perform process capability studies. (See the Appendix A glossary of terms, and for further discussion see 3, pp. 1-25 through 1-27, 4-6 through 4-7, 11-17 through 11-41, 19-3 through 19-8.)

Ten projects were chosen to assure equal representation of the major concrete production methods (truck, central, and paving mixers), and project locations selected to cover various regions of the state. Properties chosen as measures of concrete uniformity were slump, air content, unit weight, and compressive strength. These are the properties usually measured in inspection and recognized as quality determinants.

To measure variations in these properties, the experimental design suggested in the FHWA guidelines was adopted. It consists of the analysis of variance model in Table 1, and a sampling scheme to obtain the necessary replicate sampling and testing. The latter provides for testing 50 randomly selected batches ("batch" here means production units, i.e., a truckload, a paver batch, etc.) from each concreting job, as shown in Figure 1. Thus, the number of batches b for the analysis of variance is 50, the number of independent samples n is 2, and the number of portions or replicates per sample k is also 2. In addition to the information necessary for the analysis of variance, provisions were made to recover a

TABLE 1
ANALYSIS OF VARIANCE TABLOID*

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square and Quantity Estimated
Between Batches (Material)	$b - 1$	$nk\sum(\bar{y}_b - \bar{y})^2$	$(MS)_2 \rightarrow \sigma_0^2 + k\sigma_1^2 + nk\sigma_2^2$
Within Batches (Sampling)	$b(n - 1)$	$k\sum(\bar{y}_s - \bar{y}_b)^2$	$(MS)_1 \rightarrow \sigma_0^2 + k\sigma_1^2$
Analytical Error (Testing)	$bn(k - 1)$	$\sum(y - \bar{y}_s)^2$	$(MS)_0 \rightarrow \sigma_0^2$
Totals	$bnk - 1$	$\sum(y - \bar{y})^2$	

*n = samples per batch, k = repeats per sample, b = batches per job, $\bar{y}_s$ = sample average, $\bar{y}_b$ = batch average, σ^2_0 = variance of analytical error, σ^2_1 = variance of sampling error, σ^2_2 = materials variance; $\sigma_0^2 + \sigma_1^2 +$ $+ \sigma_2^2 \approx$ overall variance.

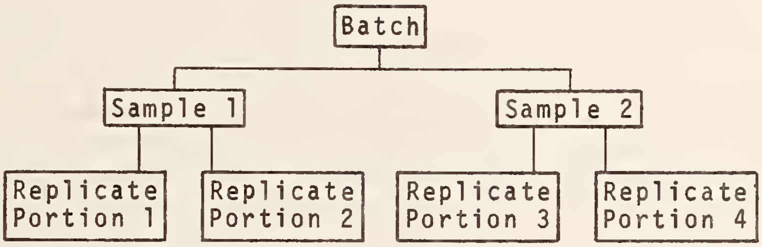


Figure 1. Sampling scheme.

random sample of cores from each concrete paving job included in the study.

C. Sampling and Testing

Most data for this study were collected during the 1967 construction season by a seven-man team. To assure that results reflected the quality of concrete routinely produced, project engineers and producers were urged to do business as usual: of the ten concreting jobs, three were supplied with central-mixed concrete, three with paving-mixer concrete, and four with truck-mixed concrete. Fifty production units (batches) were randomly selected from each project. Two independent samples were recovered from each batch, and each was split into two equal replicates or portions for testing as shown in Figure 1. This procedure was followed everywhere but on Job 9, where due to field conditions two independent samples could not be recovered, and instead one large sample was recovered and then split equally into two.

Each production unit was tested for air content, slump, unit weight, and compressive strength. Sampling and testing for slump, air content by the pressure method, and cylinder strength complied with the applicable ASTM standard procedures. Unit weight was measured by using 1/4 cu ft samples instead of the 1/2 cu ft sample recommended by ASTM. Air content was measured by the pressure method and with the Chace indicator, the manufacturer's recommendations being followed for the latter device.

Each replicate was tested once for slump, for air content by the presure method, and for unit weight. To determine air content with the Chace indicator, two readings were taken from each portion and averaged for one air content determination. Two cylinders were cast from the first portion of every other batch, and from each of the four test portions of every tenth batch (Nos. 10, 20, 30, 40, 50). This deviation from the FHWA guidelines was adopted to limit the number of cylinders to 80 per job. Had the FHWA recommendation been followed, 400 cylinders per job would have been cast, for a total of 4,000 -- which was felt to have been an unmanageable number. Compressive strengths for companion cylinders cast from the same portion were averaged for one strength determination. The results thus obtained for slump, air content and unit weight were subjected to analysis of variance. For strength, however, only cylinder strengths obtained from every tenth batch of each job were used for analysis of variance. The remaining cylinders were used only to estimate the overall standard deviation. In addition, a random sample of cores was recovered from each of the three concrete paving jobs. Core lengths and strengths were measured to determine variations in these properties.

General information about the jobs visited as well as concrete and air temperatures were recorded. A brief summary for each job visited is given in Appendix B. The information in that appendix is helpful to those interested in actual job conditions and the sampling and testing methods used.

D. Data Analysis

Data analysis for this part of the study concentrated on three subjects:

1. Determining variations of the concrete properties tested and the partitioning of these variations into component parts whenever possible,
2. Investigating possible effects of the various production methods on the uniformity of concrete properties, and
3. Comparing the variations obtained with those implied in the New York State specifications and with those obtained by other agencies in similar studies.

The components of variance were determined for all concrete properties for which replicate sampling and testing were performed, using the analysis of variance model in Table 1. But no attempt was made to include "job" as a factor in the analysis of variance, since specifications called for different levels of concrete properties for each job. Under these circumstances, testing the hypothesis of equality of means for the properties of concrete used on the ten jobs would be inappropriate, since it is already known that these means were different by design.

Using the analysis of variance in Table 1, sampling, testing, materials, and overall variances for each job were computed for slump, air content, and unit weight. Cylinder strengths obtained for every tenth batch of each job (except Job 9, which was excluded because of improper sampling) were combined and also subjected to analysis of variance, with the results given in Table 2.

Sampling variances were tested for significance with respect to testing variances, and material variances for significance with respect to testing and sampling variances combined. This was accomplished by computing F-ratios, using the appropriate mean squares, and comparing the computed F-values with tabular F-values for the corresponding degrees of freedom and confidence intervals desired. Referring to Table 1, the ratio to test the significance of sampling variance with respect to testing variance is $(MS)_1/(MS)_0$, and the degrees of freedom $b(n - 1)$ for the numerator and $bn(k - 1)$ for the denominator. Similarly, the appropriate ratio to test the significance of the materials variance, with respect to the combined testing and sampling variance is $(MS)_2/(MS)_1$, and the degrees of freedom are $b - 1$ for the numerator and $b(n - 1)$ for the denominator. F-test results for the 95-percent confidence interval are given in Table 3.

The results show that materials variances were significant for all jobs, and sampling variances for all jobs except for slump on Jobs 1 and 10, and unit weight on Job 4. This means that real material variations exist in all cases, and that variations due to sampling are almost always significant for all concrete properties tested -- i.e., real sampling variations exist.

Testing, sampling, and materials variances were tested for homogeneity from job to job with the Bartlett test (4). Non-homogeneity prevailed in all cases. Non-homogeneity of materials variance means that the magnitude of the variations for concrete properties differs from job to job, and that each concrete-making process is producing concrete of different uniformity. This is not too surprising in view of the differences in mixers, plants, operators, and materials but contradicts any intuitive assumption that under the present degree of control, concrete produced with similar processes should have comparable uniformity. It also makes it difficult to choose a "typical" standard deviation.

TABLE 2
SUMMARY OF ANALYSIS OF VARIANCE

Job	Mixer Type*	Mean	Variances				Overall Standard Deviation
			Material	Sampling	Testing	Overall	
A. AIR (PRESSURE)							
1	C	4.55	0.798	0.070	0.043	0.911	0.954
2	C	5.91	0.499	0.041	0.126	0.666	0.816
6	C	6.18	0.340	0.105	0.058	0.503	0.709
3	P	5.14	0.479	0.050	0.067	0.596	0.772
7	P	4.94	0.386	0.076	0.084	0.546	0.739
8	P	4.82	0.698	0.043	0.048	0.788	0.887
4	T	7.90	1.745	0.292	0.159	2.196	1.482
5	T	6.11	1.270	0.042	0.035	1.347	1.161
9	T	5.80	2.388	0.135	0.052	2.575	1.605
10	T	6.08	1.642	0.136	0.088	1.866	1.366
B. AIR (CHACE)							
1	C	6.39	0.652	0.545	0.201	1.398	1.182
2	C	7.42	0.698	0.134	0.221	1.053	1.026
6	C	7.39	0.400	0.137	0.256	0.793	0.890
3	P	7.38	0.326	0.099	0.147	0.572	0.756
7	P	6.41	-0.960	2.549	0.144	1.733	1.316
8	P	6.51	1.024	0.251	0.260	1.535	1.239
4	T	10.20	1.378	0.148	0.335	1.860	1.364
5	T	8.75	0.484	0.989	0.452	1.925	1.387
9	T	8.43	1.790	0.325	0.431	2.546	1.596
10	T	6.33	1.712	0.238	0.283	2.233	1.494
C. SLUMP							
1	C	2.04	0.155	-0.006	0.074	0.222	0.471
2	C	1.86	0.373	0.061	0.061	0.494	0.703
6	C	2.12	0.140	0.034	0.033	0.208	0.456
3	P	2.34	0.424	0.025	0.081	0.530	0.728
7	P	2.41	0.201	0.086	0.084	0.371	0.609
8	P	2.26	0.507	0.047	0.158	0.713	0.844
4	T	1.77	0.206	0.012	0.028	0.245	0.495
5	T	2.37	0.305	0.030	0.066	0.401	0.633
9	T	3.36	0.679	0.118	0.279	1.076	1.037
10	T	2.53	0.505	0.015	0.072	0.592	0.769
D. UNIT WEIGHT (CONCRETE)							
1	C	145.21	0.865	2.153	0.996	4.013	2.003
2	C	145.48	1.450	0.271	0.699	2.419	1.555
6	C	141.07	2.249	1.657	0.585	4.491	2.119
3	P	145.12	1.564	0.287	0.760	2.611	1.616
7	P	151.13	1.253	0.682	1.372	3.307	1.819
8	P	145.41	2.070	0.404	0.700	3.174	1.782
4	T	139.50	4.204	0.226	1.157	5.588	2.364
5	T	141.34	2.370	1.950	0.450	4.770	2.184
9	T	145.24	5.585	0.196	0.903	6.683	2.585
10	T	144.42	4.756	0.335	0.300	5.391	2.322
E. CYLINDER COMPRESSIVE STRENGTH							
1,2,3,4,5 6,7,8,10	C,P,T	4119.92	430222.87	8024.19	83822.27	522069.33	722.54

*C = central mixer, P = paver, T = truck mixer.

TABLE 3
SUMMARY OF SIGNIFICANCE TESTING AT THE 5-PERCENT LEVEL
S = Significant, NS = Not Significant

Source of Variation	Job 1	Job 2	Job 3	Job 4	Job 5	Job 6	Job 7	Job 8	Job 9*	Job 10	All Jobs*
Sampling											
Slump	NS	S	S	S	S	S	S	S	--	NS	--
Air (Chace)	S	S	S	S	S	S	S	S	--	S	--
Air (Pressure)	S	S	S	S	S	S	S	S	--	S	--
Unit Weight (Concrete)	S	S	S	NS	S	S	S	S	--	S	--
Cylinder Compressive Strength	--	--	--	--	--	--	--	--	--	--	NS
Material											
Slump	S	S	S	S	S	S	S	S	S	S	--
Air (Chace)	S	S	S	S	S	S	NS	S	S	S	--
Air (Pressure)	S	S	S	S	S	S	S	S	S	S	--
Unit Weight (Concrete)	S	S	S	S	S	S	S	S	S	S	--
Cylinder Compressive Strength	--	--	--	--	--	--	--	--	--	--	S

*Job 9 testing invalid because of improper sampling.

The non-homogeneity of sampling and testing variances from job to job is harder to explain. The former may be due to the different sampling methods and operators, and the latter to different operators. But although not homogeneous, sampling and testing variances are generally low.

1. Air Content

Results of the analysis of variance presented in Table 2 and illustrated in Figure 2 show that for air content, as measured by the pressure method, testing and sampling variances remained low for all jobs, but that the materials variance of truck-mixed concrete was higher than that produced by central mixers and pavers. The same may be said for air content measured by the Chace indicator, if Jobs 5 and 7 results are disregarded.

In the case of Job 7, as can be seen in Figure 2B, no materials variance was detected for air content measured by the Chace indicator, and most variance appears as sampling variance. Actually, the materials variance was computed as a negative quantity (this is possible because it was computed by subtracting the sum of sampling and testing variances from the overall variance), and was taken to be zero. In this case, the reason for no materials variance and exceedingly high sampling variance is known -- one tester contantly misread the indicator for part of the job, but this was not discovered until it was too late to identify the incorrect readings. It is also suspected that incorrect readings were taken on Job 5, but the results are presented for sake of completeness and because by accident they illustrate a very important point -- materials variance can be underestimated if sampling and testing variations are predominant.

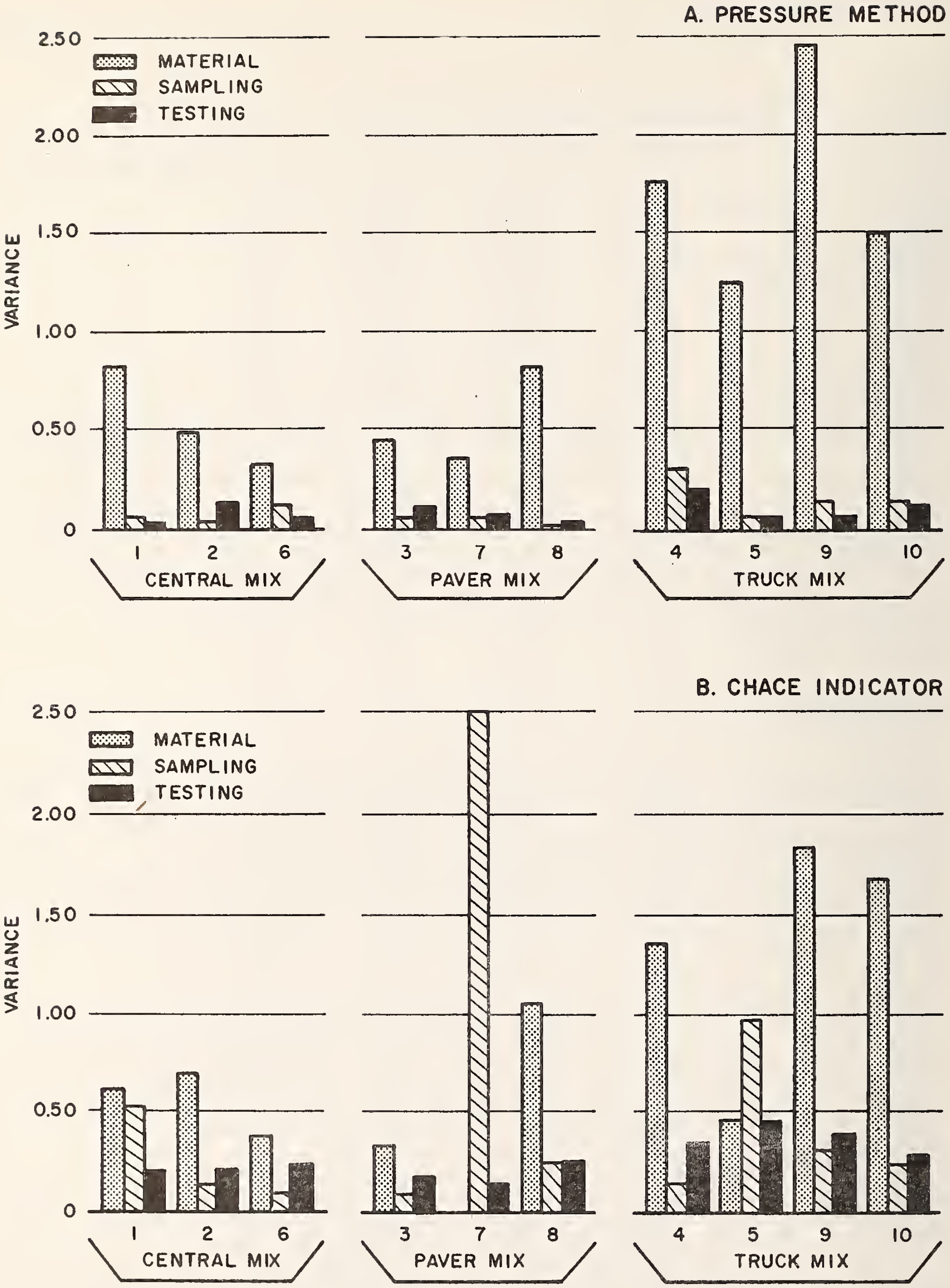


Figure 2. Magnitude of variances of air content measured by two methods; abscissa numbers are the same job numbers used in the tables.

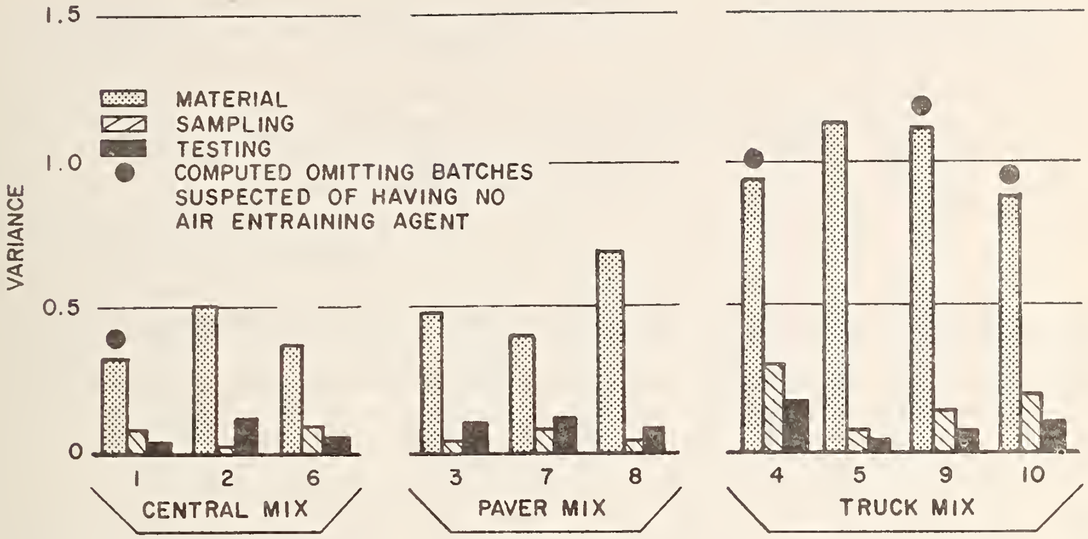


Figure 3. Magnitudes of variances of air content measured for three production methods, with non-air-entrained batches omitted.

Mean air contents obtained by the Chace indicator are constantly higher than those obtained by the pressure method, but the overall variances obtained by each method are comparable, as can be seen from Figure 2 if Jobs 5 and 7 are omitted.

The magnitudes of overall standard deviations for air content (Table 2) range from approximately 0.7 to 1.6 percent. This wide range indicates that the air content of concrete produced in New York under the present degree of process control is quite variable.

Air content variances for truck-mixed concrete were compared with those of concrete produced by the other two methods. Individual comparisons by the F-test showed that for this property the overall variances of truck-mixed concrete were significantly larger in all cases. This led to a search for an explanation. Review of data for each job revealed that the air content of relatively few batches was extremely low. These air contents were 3 percent or less -- values that can be obtained in non-air-entrained concrete. This occurred rarely for central- and paver-mixed concrete, but relatively often for truck-mixed. In the case of central-mixed concrete, it was verified that the low air content was due to lack of air-entraining agent. This led to speculation that low air content for truck-mixed concrete was also due to failure to introduce air-entraining agent into the batches. This of course would result in very high variations, which are independent of mixer

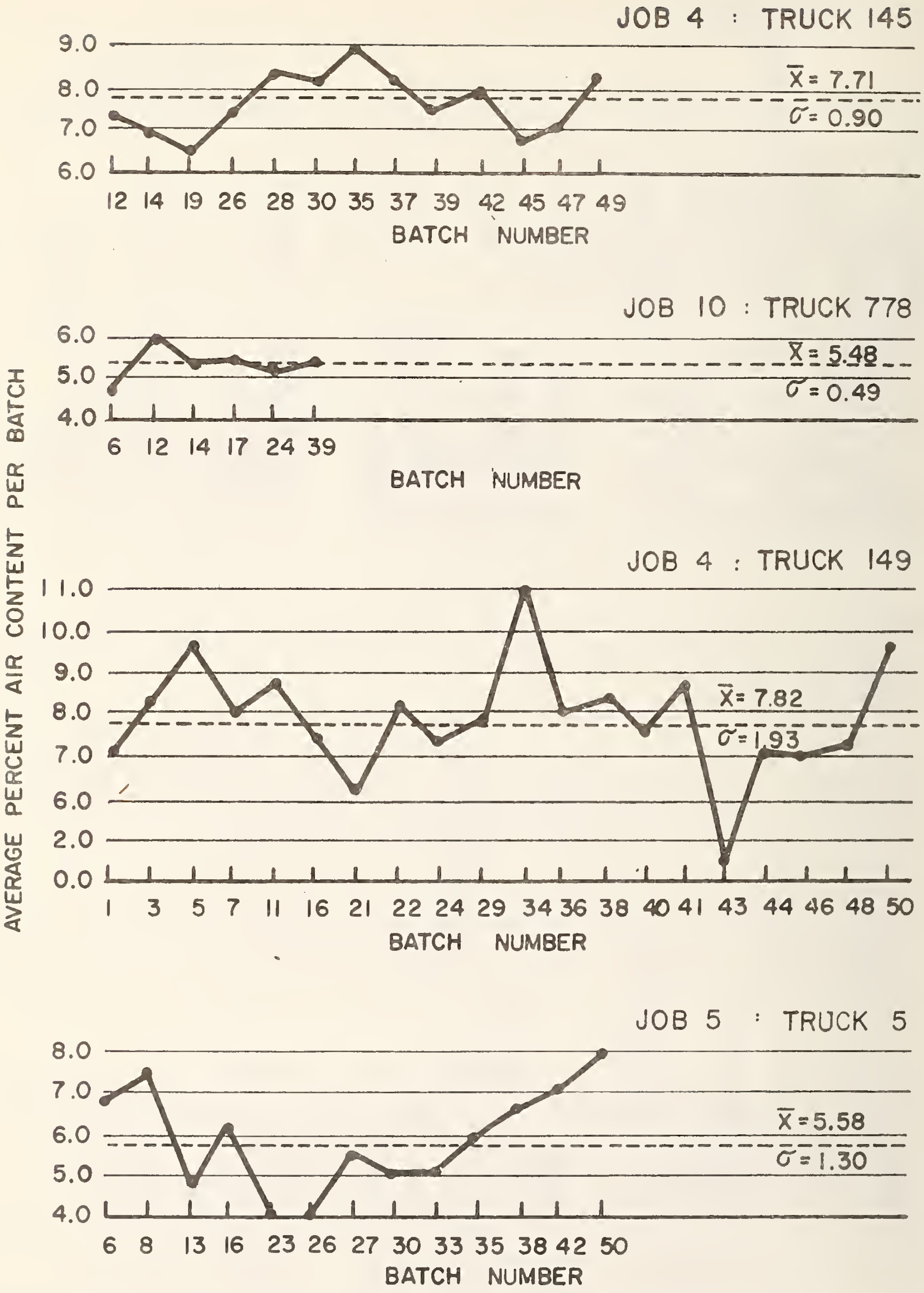


Figure 4. Air content uniformity achievable with truck mixers.

type. Thus, conservatively taking an air content of 2.5 percent as the cutoff point between air entrainment and non-air-entrainment, the values below 2.5 percent could be dismissed on the basis of assignable cause (see Appendix A glossary). This was done and variances re-computed omitting these "no-air" batches. The materials variances thus computed for the pressure air content were very similar, but statistically still significantly greater than those of central- and paver-mixed concrete. These results are shown in Figure 3.

With the truck-mixed concrete variances not completely explained by the "no-air" batches, the search for an explanation continued. It can be reasoned that each truck mixer constitutes a unique mixing process, and that grouping the output of different trucks is equivalent to grouping the output of many central mixers or pavers. Such groupings could result in large variations, although the output of each mixer could be very uniform. For trucks this is illustrated by the graphs in Figure 4. It can be seen that some trucks can produce concrete with relatively uniform air content, while others can produce concrete having quite variable air content. If the air content of all concrete produced in these four trucks were grouped together, the resulting variation would be large, despite the fact that air content for concrete produced in two of these trucks is quite uniform. This means that if the air content data were grouped according to the truck in which the concrete was produced, variations in air content should be lower.

Accordingly, all batches produced in the same truck were grouped together, and the analysis of variance repeated on a truck basis, both including and excluding the "no-air" batches. The results are presented in Table 4, where variances computed excluding the "no-air" batches are shown in italics. These results show that as expected, overall variances for air content produced in some trucks can be low, and that in the three cases involving "no-air" batches, exclusion of these batches resulted in considerable reduction in variances. However, for some trucks the overall variances remained large. Examples are Trucks 1 and 13 on Job 5, Truck 22 on Job 9, and Truck 788 on Job 10. But it is believed that in these cases, the trucks themselves are not responsible for the large variances. They happened to be used at different times, when the batching processes supplying them were themselves operating at different levels of control, and the batching processes are more likely responsible for the large variation than the trucks.

A mathematical analysis of this effect cannot be given without some knowledge of the process capability, but Figure 5 illustrates this point. If Truck 1 on Job 5 is followed in Figure 5A, it can be seen that this truck was used when the process was producing concrete having 5-percent air content, as well

TABLE 4
COMPONENTS OF VARIANCE FOR AIR CONTENT (PRESSURE METHOD)
FOR BATCHES MIXED IN INDIVIDUAL TRUCKS
Italicized Values Identify Batches
Suspected of Having Air-Entraining Agent Omitted

Truck	Total Batches	Mean, $\bar{X}$	Standard Deviation, S	Variances			
				Testing	Sampling	Materials	Overall
A. JOB 4							
149	20	7.82	1.93	0.1840	0.2875	3.2556	3.7271
149	*19*	*8.15*	*1.27*	*0.1936*	*0.3010*	*1.1205*	*1.6151*
152	11	8.61	1.15	0.2364	0.2391	0.8547	1.3301
4	6	7.24	0.92	0.0783	0.2817	0.4850	0.8450
145	13	7.71	0.90	0.0937	0.3490	0.3670	0.8097
B. JOB 5							
2	8	5.83	0.74	0.0422	-0.0013	0.5105	0.5514
3	7	6.29	0.96	0.0307	0.0364	0.8526	0.9197
5	13	5.85	1.30	0.0438	0.0915	1.5487	1.6840
6	5	5.60	0.78	0.0195	0.0240	0.5604	0.6039
1	4	6.37	2.30	0.0144	0.0169	5.2635	5.2948
4	4	5.58	0.56	0.0275	0.0369	0.2484	0.3128
13	4	6.79	1.20	0.0344	0.0025	1.3982	1.4351
C. JOB 9							
27	6	6.05	1.56	0.0442	0.1379	2.2415	2.4236
21	4	5.58	0.78	0.0275	0.1681	0.4103	0.6059
30	4	6.14	0.71	0.0731	0.0412	0.3854	0.4998
43	6	5.80	2.51	0.0525	0.1404	6.1172	6.3101
43	*5*	*6.77*	*0.90*	*0.0580*	*0.1550*	*0.6019*	*0.8149*
40	4	5.43	0.54	0.0894	0.1788	0.0226	0.2907
22	5	4.98	2.00	0.0445	0.0575	3.8828	3.9848
D. JOB 10							
786	6	5.40	2.51	0.1696	0.0296	6.1099	6.3091
786	*4*	*6.96*	*0.77*	*0.1512*	*0.0919*	*0.3469*	*0.5900*
798	6	5.48	0.49	0.0338	0.0788	0.1275	0.2400
800	6	6.37	0.81	0.0567	0.1617	0.4337	0.6520
820	6	6.69	0.83	0.0388	0.2083	0.4456	0.6927
788	5	6.72	1.17	0.0845	0.2000	1.0938	1.3782

as when the process was producing concrete with 9.5-percent air content. It seems that in this case the truck itself cannot be blamed for the non-uniform air content, because the whole process mean shifted from a relatively low to a relatively high level. The large variation is due mainly to the shift in process mean. This was the most obvious case, but the same can be said for the individual trucks in Figures 5B and 5C.

This discussion suggests that the air content of concrete mixed in a single truck is not necessarily more variable than that of concrete produced in a central mixer or paver. However, when concrete produced in different trucks is grouped together, the air content variability can be expected to be greater. This greater variability is likely to be due to the grouping of data from different trucks each representing a distinct production process, and to the drift of the overall process mean, rather than to the inherent variability of truck mixers. Similar results would be expected if data from different central mixers or pavers were grouped in the same manner.

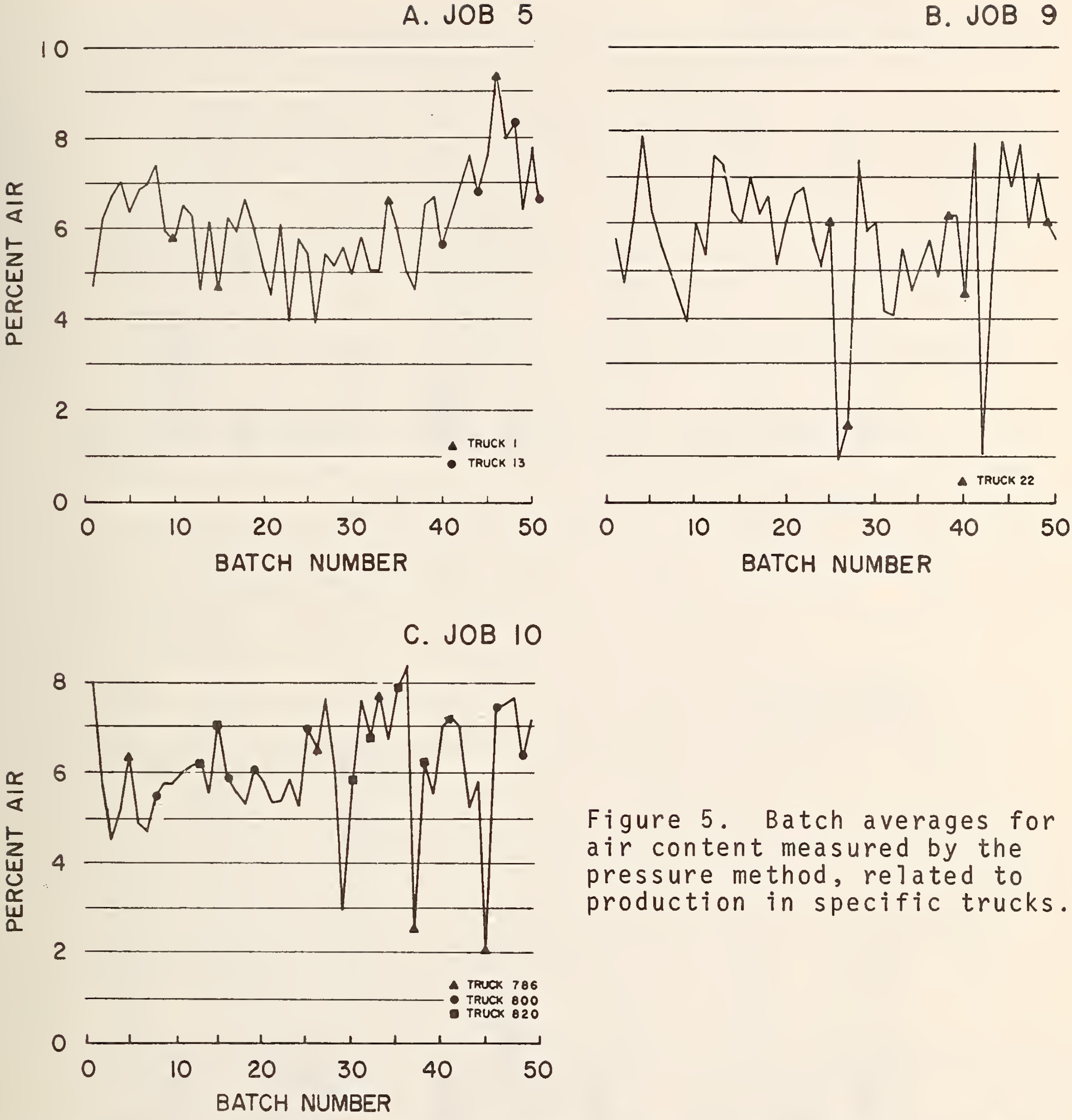


Figure 5. Batch averages for air content measured by the pressure method, related to production in specific trucks.

---

2. Other Properties

Unlike air content, overall slump variances remained approxmately constant, although not homogeneous for all mixer types. The sampling and testing variances were low and the overall variances about 1 sq in. or less. Results for each job are shown in Table 2, and the variances are grouped according to production method in Figure 6A.

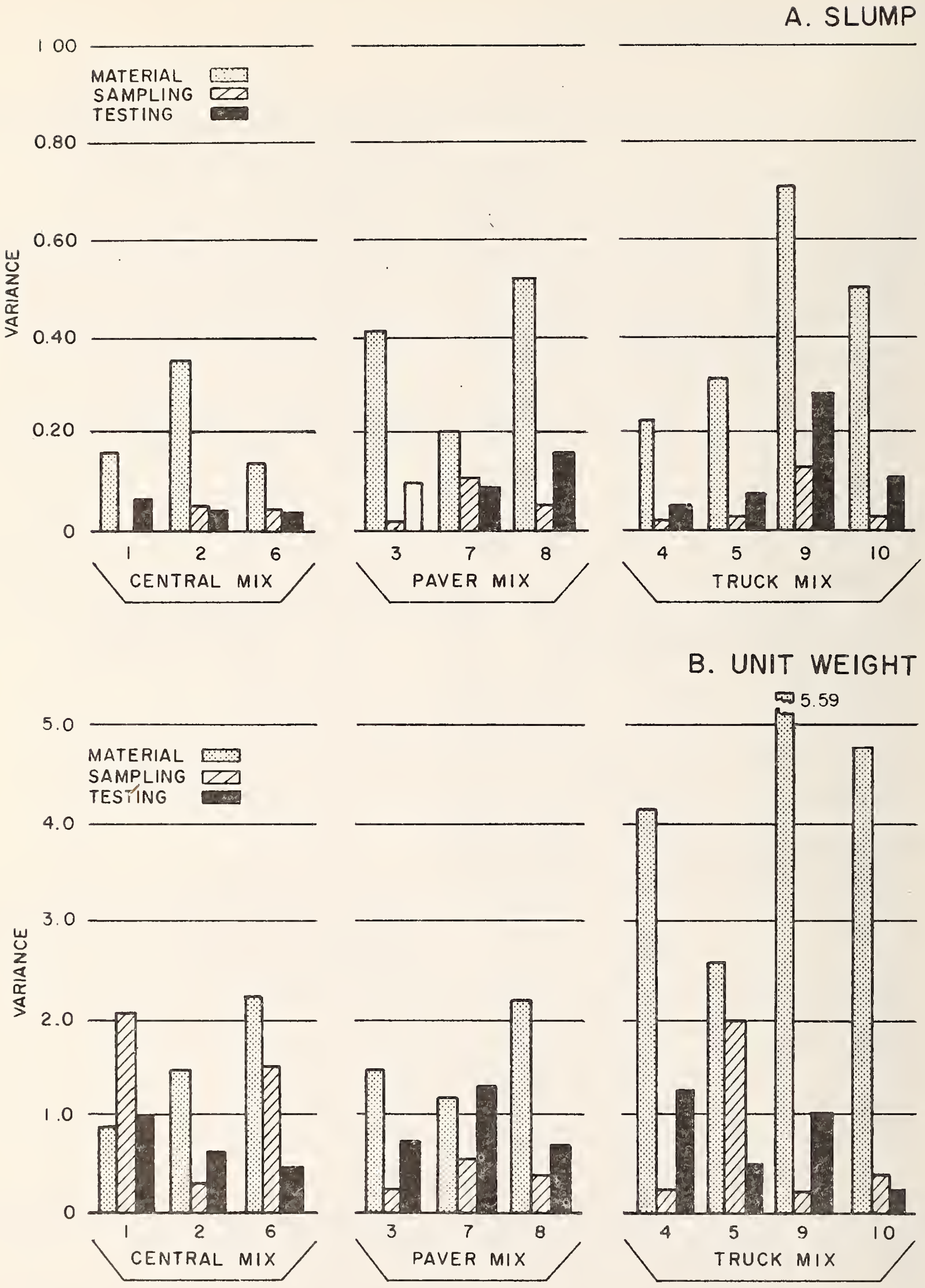


Figure 6. Magnitudes of variances for slump and unit weight.

TABLE 5
MEANS AND STANDARD DEVIATIONS OF CYLINDER AND CORE PROPERTIES

	Cylinder Compressive Strength, psi			Core Strength, psi			Core Length			
Job	Sample Size	Mean, $\bar{X}$	Standard Deviation, S	Sample Size	Mean, $\bar{X}$	Standard Deviation, S	Sample Size	Design Thickness, in.	Mean Length, in.	Standard Deviation, S
1	27	3833	669	25	3520	667	25	9	9.15	0.23
2	25	3755	449	24	2826	441	24	8	8.23	0.33
3	25	5031	569	25	4468	779	25	9	9.14	0.20
4	25	4460	569	*	*	*	*	*	*	*
5	26	3490	583	*	*	*	*	*	*	*
6	25	4450	530	25	3952	659	25	9	9.63	0.47
7	25	4301	481	23	4112	556	23	9	9.02	0.26
8	25	3779	512	25	3680	730	25	8	8.24	0.26
9	25	3740	631	25	3250	715	25	9	9.76	0.20
10	25	4454	474	*	*	*	*	*	*	*

*Structural concrete, no cores taken.

Results of the analysis of variance for unit weight are also summarized in Table 2. Overall variances are large, ranging from 2.42 for Job 2 to 6.68 for Job 9, and reflect variations in air content. Sampling and testing variances are also large, perhaps because of having used 1/4 cu ft samples instead of the 1/2 cu ft sample specified by ASTM. As can be seen in Figure 6B, truck-mixed concrete had the greatest unit weight variations.

As previously noted, replicate cylinders were cast from each of the four test portions of every tenth batch. Those cast from Jobs 1 through 8 and from Job 10 were grouped for the analysis of variance producing the results shown in Table 2. When interpreting these figures, it should be kept in mind that different classes of concrete were grouped, and that due to this indiscriminate grouping of the data, the overall and materials variances have no real meaning because they include job-to-job variations. In fact, this scheme was adopted only to arrive at estimates of sampling and testing variations. The sampling and testing standard deviations combined were found to be approximately 300 psi.

For some properties not enough replicate data were available to perform the analysis of variance. In those cases, the means and standard deviations of the properties were computed to estimate the overall variability. Specifically, this was done for cylinder strength, core strengths, and core lengths.

Besides the cylinders cast from each tenth batch for the analysis of variance, two cylinders were cast from all other even-numbered batches, thus casting cylinders from 25 batches of each job. Cylinder strengths were averaged to obtain one value per batch; using the resulting batch averages, means and standard deviations were computed for each job. Similarly, the means and standard deviations of core strengths were computed for all jobs, with the results summarized in Table 5. It can be seen that core and cylinder strengths show comparable variations, with the standard deviations of core strengths generally higher than those of the corresponding cylinders. Standard deviations of both cylinder and core strengths are large, ranging from 449 to 669 psi for cylinders and from 441 to 779 psi for cores.

The means and standard deviations of core lengths were computed for all paving jobs. The standard deviations ranged from 0.20 to 0.45 in., and no core lengths fell outside the specification tolerance of -1/2 in. These results are also presented in Table 5.

E. Comparison of Results with Those of Others And With Those Implied in the Specifications

Data obtained from the Federal Highway Administration and summarized in Figure 7 show how overall standard deviations for concrete properties measured in New York compare with those found by others in similar experiments. No attempt was made to analyze the data statistically. It was believed that any such analysis would be inappropriate because the variances were found to be non-homogeneous -- i.e., not from the same population -- and pooling these variances would have resulted in meaningless averages. Thus, the data were used only for general comparisons.

Generally, the standard deviations measured in New York are of the same order of magnitude as those measured by others. Exceptions are the variations in unit weight and cylinder strength. In most cases, standard deviations measured by others for these properties are smaller than obtained in New York, as may be seen from Figures 7D and 7E.

For air content and slump, the range of the magnitude of standard deviations obtained in New York and elsewhere is similar, but the grouping of standard deviations within this range is different. Figure 7A shows that for air content, those measured by others group in the lower part of the range, while those measured in New York tend to fall in the middle or upper part of the same range. For slump (Fig. 7C), standard deviations obtained in New York group in the middle of the common range.

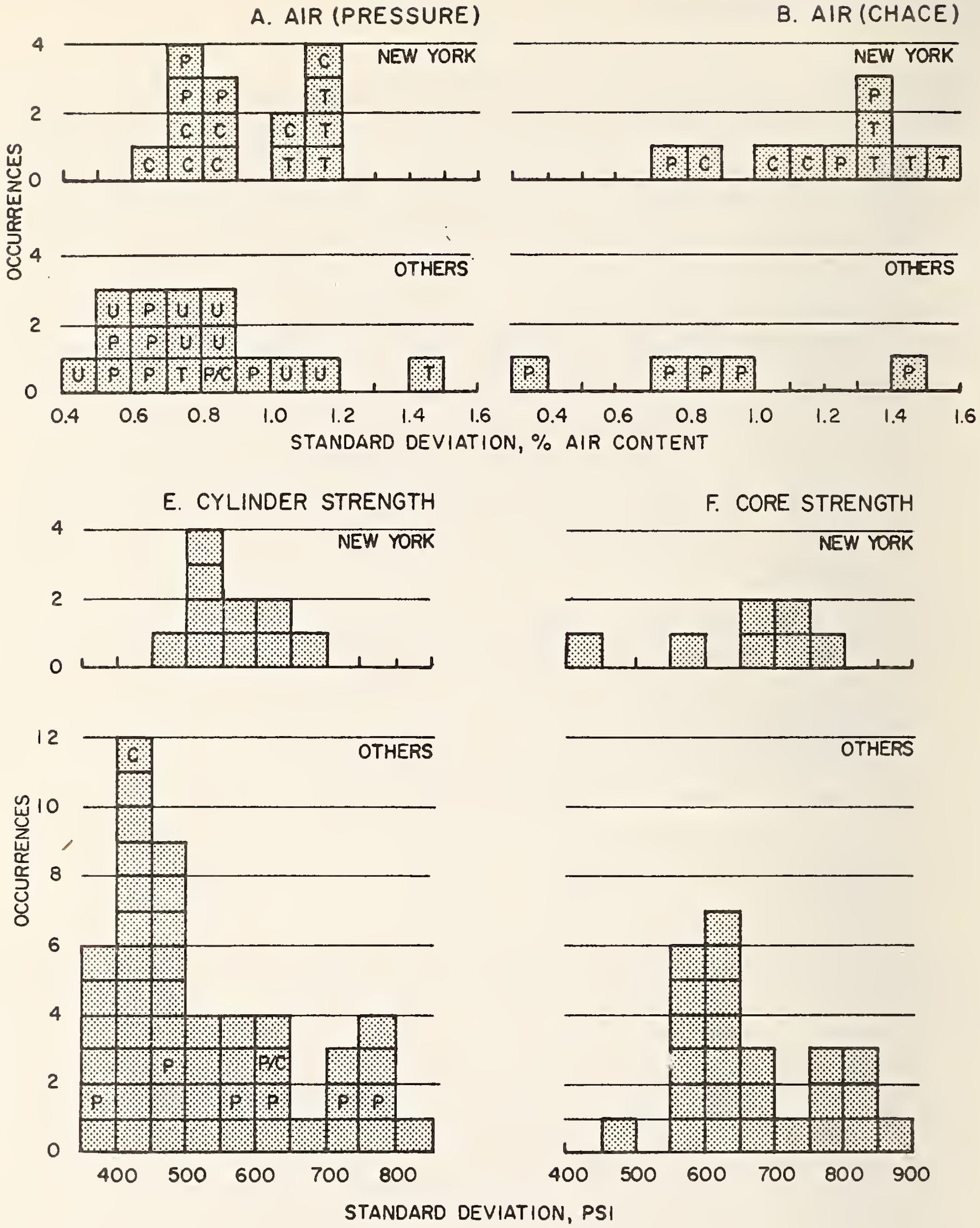
Some concrete properties measured in this study are controlled by specifications either formally or informally. Those for which formal specification limits are given are slump, air content, and core engths. No formal specification is given for strength, but cylinders are routinely tested for structural concrete and specification limits are informally applied. The variations measured were consistent with specification limits for strength and core lengths. But the variations in slump and air content were found to be much larger than those implied in the New York specification in force in 1967:

> All concrete produced shall have an air content as ordered by the Engineer within tolerance for any one test of ±1%. The column in the following table captioned "Suggested Air Content" is presented as a guide to the Engineer. The columns captioned "Min. % and Max. %" are minima and maxima for averages of five (5) Chace tests. Concrete which falls outside the minimum and maximum limits by average of five (5) Chace tests or one (1) pressure test shall be rejected and removed from the site of the work. When a test shows the air content outside the ±1% tolerance from the Engineer's specified air content, immediate corrections to subsequent batches shall be made.
>
> All concrete shall be produced with an air content within the limits listed below:

Maximum Size Aggregate Contained In Concrete	Air Content		
	Min. %	Suggested %	Max. %
No. 1	6.0	8.0	10.0
No. 2	6.0	7.5	9.0
No. 3	4.0	5.5	7.0

The experimental results showed that specified tolerances of ±1 percent for air content were seldom achieved and that strict interpretation of the specifications would have resulted in rejection of a major portion of the concrete delivered to each project. As can be seen from Table 6, these tolerances were exceeded roughly from 20 to 76 percent of the time (sample sizes of 50 are not large enough for precise estimates of the fractions defective). The high degree of non-compliance resulted from the inconsistency between the variation implied in the specification tolerances, and the variations actually attained. The ±1 percent tolerances imply a standard deviation of less than 0.33 percent for air content. The standard deviations actually measured varied from 0.7 to 1.6 percent. Assuming that air content is normally distributed, compliance with these tolerances could only be achieved in 85 percent of the cases, even if the lowest standard deviation measured were always achieved. This indicates that the ±1 percent tolerances are inconsistent with the product's inherent variability.

The maximum and minimum limits for air content provide a range of 3 or 4 percent, depending on aggregate size. Assuming that the lowest standard deviation measured -- 0.7 percent -- can always be


A. AIR (PRESSURE)
B. AIR (CHACE)
NEW YORK
OTHERS
OCCURRENCES
0
2
4
0.4
0.6
0.8
1.0
1.2
1.4
1.6
STANDARD DEVIATION, % AIR CONTENT
P
C
T
U
P/C
E. CYLINDER STRENGTH
F. CORE STRENGTH
NEW YORK
OTHERS
OCCURRENCES
0
2
4
6
8
10
12
400
500
600
700
800
900
STANDARD DEVIATION, PSI

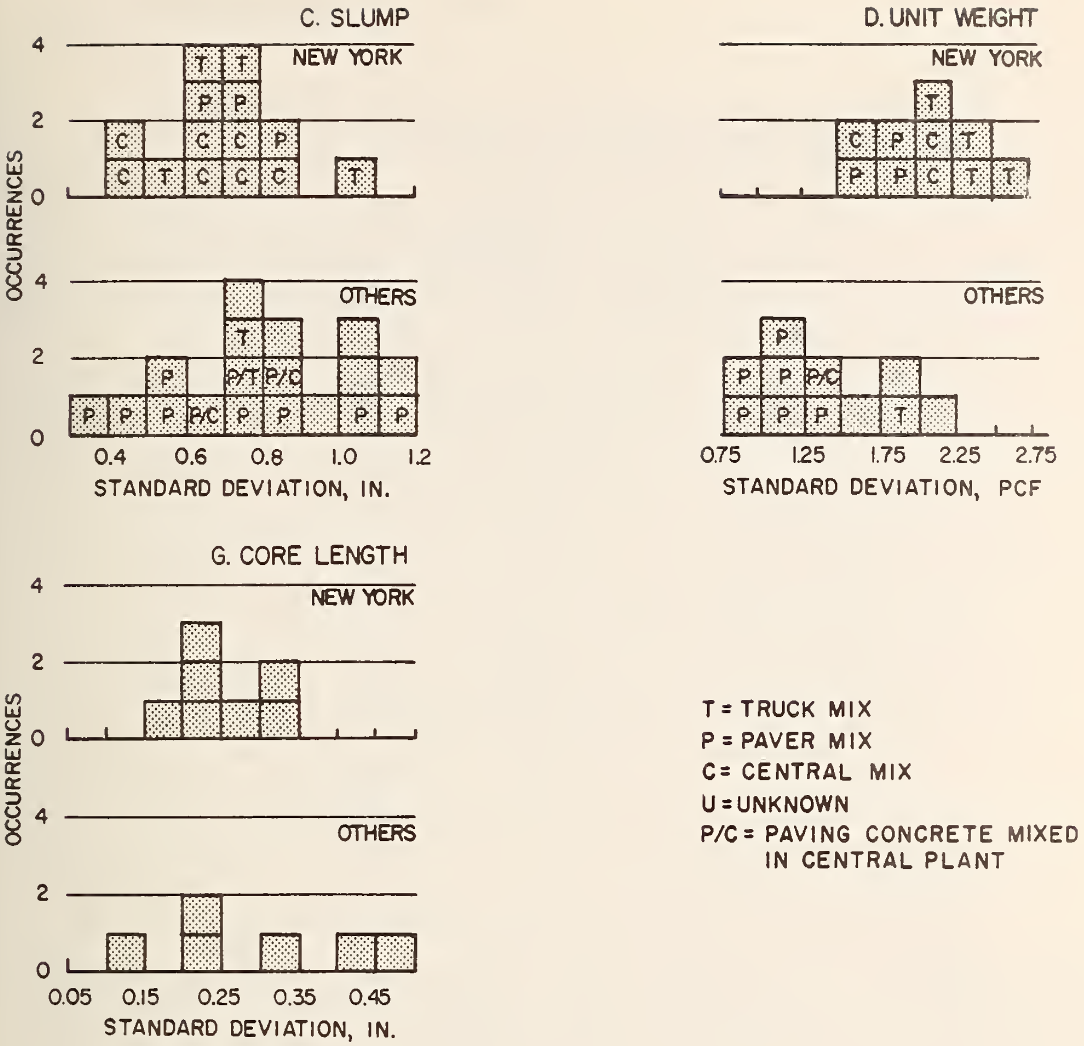


Figure 7. Overall standard deviations obtained in New York and elsewhere.

TABLE 6
OBSERVED PROPORTION OF BATCHES
REQUIRING ACTION UNDER 1967 SPECIFICATIONS

		Total Batches, percent		
Job	Variable	Exceeding Max. or Min. Limits	Exceeding Tolerance But Not Max. or Min. Limits	Total Batches Requiring Action
1	Slump	10	80	90
	Air Content	14	40	54
2	Slump	16	42	58
	Air Content	2	18	20
3	Slump	36	56	92
	Air Content	4	16	20
4	Slump	2	22	24
	Air Content	18	16	34
5	Air Content	46	24	70
6	Slump	12	82	94
	Air Content	30	46	76
7	Slump	32	66	98
	Air Content	8	24	32
8	Slump	28	58	86
	Air Content	10	42	52
9	Slump	84	14	98
	Air Content	26	14	40
10	Air Content	50	14	64

achieved, the range of 4 percent would make compliance possible in approximately 98 percent of the cases, and the range of 3 percent in about 96 percent. Because with some process control, a standard deviation of 0.7 percent or less for air content is believed achievable, these limits seem reasonable. The Table 6 results show that these limits were exceeded from 2 to 50 percent of the time. This, however, was not only due to large variations, but also to the mean not approaching the middle of the range.

The 1967 specifications for slump read as follows:

> Slump shall be determined in accordance with the Standard Method of Slump Test for Consistency of Portland Cement Concrete, ASTM Designation C 143.
>
> Concrete Paving and Culvert Headwalls. Concrete for pavements or culvert headwalls shall be placed at a slump as nearly consistent with the Desired Slump shown in the following table as is practicable:

Type of Placement	Desired Slump	Maximum Slump
Concrete Pavement	1"	2-1/2"
Culvert Headwalls	2"	3"

> Concrete that exceeds the maximum slump shown in the above table shall be rejected for use in the work.
>
> Reinforced Concrete Structures. The Contractor shall be responsible for the production of concrete having the slump ordered by the Engineer for each placement. The Engineer will, in general, order the minimum slump at which the concrete can be readily consolidated and properly finished. The column in the following table captioned "Suggested Slump" is presented as a guide to the Engineer. It is anticipated the slump ordered will, only in unusual situations, exceed the suggested value. A tolerance of one-half inch above or below the slump ordered by the Engineer will be permitted; however, in no case shall this tolerance or the ordered slump permit a slump to exceed the maximum slump listed below. Concrete failing to conform to either of these requirements will be rejected for use in the work.

(Type)	Suggested Slump	Maximum Slump
Footings	2"	3" max.
Pedestals	2"	3" max.
Walls over 1'6" thick	2"	3" max.
Walls less than 1'6" thick	2-1/2"	4" max.
Deck Slabs	2"	3" max.
Box culverts throughout	2-1/2"	4" max.
Bridge sidewalks	1-1/2"	3" max.
Rigid frames or arches	2-1/2"	4" max.
Piers	2"	3" max.
Tremie concrete 6" min.	7"	8" max.

> The specified slump shall be uniformly maintained throughout the placement.

Comparison of variations observed for slump with the variations implied in the specifications is difficult unless it is assumed that the suggested slump is the specified mean. If that assumption is not made, the mean can be chosen any place between zero and the suggested slump, to allow for more variation. If the suggested slump is assumed to be the specified mean, the upper half range is determined and the allowable variation is well defined.

If this assumption is made, the upper half range for slump is either 1 or 1-1/2 in., depending on concrete class and use. The 1-in. range implies an overall standard deviation of approximately

1/3-in., and the 1-1/2-in. range a standard deviation of about 1/2-in. Table 6 shows that the maximum limit for slump was exceeded from 2 to 84 percent of the time. This high degree of non-compliance resulted from a combination of larger variation than allowed, and the fact that the mean often exceeded the suggested slump.

F. Specification Revisions

As noted, the measured variations for slump and air content exceeded those implied in the specification limits, and resulted in a high degree of non-compliance, which regardless of causes is undesirable. Unenforced specification limits have no meaning and should be revised or enforced. Compliance can be attained by reducing the variability of the product through process control, by screening the material accepted, by changing the specification limits to accommodate all measured variability, or by a combination of these actions. The proper solution hinges on the appropriateness of existing specification limits and the feasibility of reducing variations. Widening the specification limits is appropriate only if limits are tighter than those required by the engineering demands imposed on the product, or "engineering limits."

Properly set specification limits imply that if these limits are exceeded, the product loses its ability to perform its intended function. Setting of specification limits, therefore, is strictly an engineering function. Specification limits should be set by design engineers who know the engineering demands imposed on the product and should be based on performance -- the ultimate engineering judgment. Once properly set, these limits should remain unchanged as long as materials or products are required to perform the same functions. Changing these limits is appropriate only if performance proves they were improperly set in the first place, the intended engineering function of the material changes, or the design is revised to require a different degree of performance. Changing the limits to accommodate product variability, when it leads to exceeding the limits required for the product to perform as intended, is usually unwise. In such cases, changes in specification limits to accommodate product variability are valid only when the variability proves irreducible and the economic consequences of accepting materials whose properties exceed the engineering limits are small compared to the cost of screening the materials accepted. Thus, in considering changes in specification limits, it is necessary to decide whether the existing limits are desirable limits or engineering limits.

Accordingly, data available in this Bureau and from the Materials Bureau were reviewed to arrive at the engineering limits given in Table 7. The reason for the small changes in slump specification limits was workability. The changes in limits for air content were based on data found in the literature on spacing factors and

TABLE 7
ORIGINAL AND REVISED SPECIFICATION LIMITS FOR SLUMP AND AIR CONTENT

Concrete Type	1967 Limits Minimum	1967 Limits Design	1967 Limits Maximum	Revised Limits Minimum	Revised Limits Design	Revised Limits Maximum
A. SLUMP						
Pavement	--	1"	2-1/2"	--	1-1/2"	3"
Sidewalks	--	1"	3"	--	2-1/2"	3-1/2"
Deck Slabs	--	2"	3"	--	2-1/2"	3-1/2"
Footings, Headwalls, Piers, Pedestals, Walls over 1'6" thick	--	2"	3"	--	2-1/2"	4"
Walls less than 1'6" thick, Box Culverts throughout, Rigid Frames and Arches	--	2-1/2"	4"	--	3"	4"
B. AIR CONTENT						
Class A	6.0%	7.5%	9.0%	4.0%	6.0%	8.0%
Class B	4.0%	5.5%	7.0%	3.0%	5.0%	7.0%
Class C	4.0%	5.5%	7.0%	4.0%	5.5%	7.0%
Items containing maximum size No. 1 aggregate	6.0%	8.0%	10.0%	5.0%	7.0%	9.0%

on limited performance data accumulated by this Bureau (5). From the available data, it was concluded that air contents between 4 and 8 percent provide adequate protection for concrete and that the required air content should reflect the mortar content. Accordingly, the minimum and maximum limits for air content of Class A concrete -- the richer mix -- were set at 4 and 8 percent, respectively. The limits for Class C concrete were set at 4 and 7 percent to account for the lower mortar content. The limits for Class B concrete were set at 3 and 7 percent. It was reasoned that Class B concrete is usually placed in unexposed portions of foundations and therefore some concrete with only 3-percent air could be tolerated.

The reasons for choosing these limits, although believed sound, may be disputed, *but the important point for the purpose of this report is that they were set as the engineering limits, providing boundary conditions for both process control and acceptance sampling.* The new limits eliminate the tolerance given inside the specification limits of ±1/2 in. and ±1 percent around the design slump and air content limits, respectively. But the revised limits themselves set at different levels, allow for approximately the same variations as those in the specifications in force in 1967. Thus, unless variation is reduced by process control or inspection increased, a large degree of non-complaince is likely to persist. With the revised means, the specification range for air content is either 3 or 4 percent. Assuming normality, for production to be acceptable the standard deviation of a property would be one-sixth the specification range. For air content this means obtaining a standard deviation of either 0.5 or 0.67 percent, depending on concrete class. Referring to Figure 7A, it can be seen that this does not happen very often under the present level of process control. Similarly, for slump the standard deviation would have to be 1/2-in. to achieve compliance. Figure 7C shows that standard of 1/2-in. or less was seldom encountered.

Since the specification limits cannot be changed, two alternatives are left to achieve compliance -- 1) reduction of variation through process control, and 2) screening. The latter is very expensive and is to be avoided if variation can be reduced by process control. This leads to the question of whether the variation of concrete properties can be economically reduced to achieve compliance. The analysis of variance results in Table 2 show that for slump and air content, the materials variance accounts for the bulk of the total variance. The percentages of materials variance for slump ranged from 54 to 85 percent. For air content, the materials variances accounted for from 70 to 94 percent of the total. The high materials variances suggest that with some process control, total variation can be reduced considerably. A good example of this is the air content of Job 10. Referring to Figure 5C, it can be seen that three batches of concrete had air content of less than 3 percent. As discussed, this indicated that no air-entraining agent was introduced into these batches -- an assignable cause. By eliminating these batches from the calculations, the total variance was reduced from 1.86 to 1.07 -- a reduction of 42 percent. Similar results were obtained for Job 9, (Fig. 5B), which included three concrete batches with air contents of less than 2 percent. By eliminating these three batches, the total variance drops from 2.58 to 1.26 -- a reduction of 51 percent. These examples strongly suggest that variation of air content can be drastically reduced. But the proof must ultimately involve trial runs.

Such runs or case studies were made in this investigation under relatively loose process control. The case study results showed that the standard deviations implied in the revised specifications, now in use, are achievable with some process control. Details of these case studies are given in Chapter II, after discussion of process control.

II. PROCESS CONTROL CHARTS

Acceptance sampling will reject most or all of a process output if a manufacturing process cannot be controlled to manufacture a product having the specified properties and property levels. Thus, a prospective seller must assure that his process output can consistently and economically satisfy the buyer's requirements. It would be foolish to attempt production without prior assurance that market demands can be met. To assure that the specified product can be economically manufactured requires capability studies. To assure that the product constantly meets the imposed requirements requires process control.

Process capability studies are necessary only when new products are needed, specification limits of current products are changed, or new processes are employed to manufacture established products. Because their aim is to determine the feasibility of manufacturing a given product under a given set of conditions, capability studies are occasional undertakings. By contrast, process control aims at obtaining a uniform output from a proved process and is a function that must last the duration of manufacturing. Process control encompasses constant sampling and testing, continuous analysis of test results, and physical manipulation of the production process itself. Specifically, it requires knowing what changes to make, how to produce those changes, and when to make changes to obtain the results desired. Capability studies provide the information necessary to decide what changes to make and how physically to accomplish them. The information needed to determine when process changes are necessary must come from continuous testing and instantaneous analysis of test results, usually accomplished graphically by plotting test results on control charts -- more specifically, on process control charts.

Both capability studies and process control involve the physical manipulation of machines and materials, and for this reason are the manufacturer's responsibility. The buyer can observe these functions, but seldom can effectively perform either function for two reasons. First, the buyer and his inspectors may be removed from manufacturing processes. Second, his inspectors rarely possess the many skills necessary for effective process control, which requires sharpening tools, calibrating measuring devices, blending materials, operating machinery, controlling temperature, and a knowledge of applied statistics. Inspectors cannot be expected to be that knowledgeable. Even if such men could be found, their success would depend greatly on the cooperation of the production process manager because most decisions involved in process control are outside the inspector's authority and greatly affect the economics of production.

This, however, does not mean that the buyer should ignore process control, or that he is at the mercy of the producer. The buyer should encourage process control and should understand it well in order to analyze process control data. Correct interpretation of process control data allows the buyer to predict which suppliers are likely to deliver products of poor quality, and to take appropriate protective steps. The knowledgeable buyer, depending on market conditions, can avoid those suppliers or increase the amount of acceptance sampling to assure the desired quality levels.

Process control data are usually available in the form of control charts. For this reason, the buyer or his representative must be familiar with the various types of control charts and the concepts on which these charts are based. The purpose of this chapter, then, is briefly to review the concepts and types of control charts in common use, so that engineers may make better use of control data.

A. Concepts and Nomenclature

In simplest terms, control charts are graphical tests of hypotheses. These charts are based on the idea that in a manufacturing process variations are inevitable, but that these variations can be minimized by eliminating causes of large variations. According to control chart theory, total variation in process output is composed of two elements: 1) variation due to assignable causes and 2) random variation. The former is that variation that can be assigned to specific factors and can be identified and eliminated by removing these causes. The latter is the variation that cannot be attributed to any single factor and cannot be economically eliminated. When all variation due to identifiable causes is removed from the process output and only random variation remains, the process is said to be in control. Control charts are used to test graphically the hypothesis that differences in properties of the process output are due only to random variation and no evidence of assignable variation exists, i.e., that the process is in control.

When the process is in statistical control, variation is a minimum and computed statistics of the output properties assume predictable patterns, which in most cases can be characterized closely by known frequency distributions. Control charts make use of these facts in a simple, systematic way. To set up a control chart, variation due to assignable causes is eliminated, magnitude of the random variation is computed, and frequency distribution of the properties of the process output is determined. Then, the sample size and statistic to be used in testing the hypothesis of control are chosen along with a confidence interval. Knowing the expected random variation, the sample size, the statistic to be used, and the confidence interval for the statistic chosen and its distributional form, critical values for the hypothesis that only chance (random) variation exists are computed and plotted. The result is a control chart in which the limiting lines correspond to critical values the controlled statistic cannot exceed, in order *not to reject* the hypothesis of control for the given sample size. Control charts can be set up for any property and for different statistics. Since different statistics follow different frequency distributions, control charts limits vary both according to the chosen confidence intervals and frequency distributions involved.

Control charts can be used for different purposes and based on a number of statistics. Control charts are commonly used for process control, for process acceptance, and for analysis of past data. These uses give rise to the nomenclature of process control charts, acceptance control charts, and control charts to analyze past data. Besides taking their names from their intended function, control charts are also named for the statistic used. The process control charts most commonly used, which take their name from the statistics used, are:

1. Control charts for the fraction defective or p-charts,
2. Control charts for number of defects or c-charts,
3. Difference control charts,
4. Cumulative sum control charts or cusum charts,
5. Standard deviation control charts or σ-charts,
6. Control chart for sample ranges or R-charts, and
7. Control charts for sample mean or $\bar{X}$-charts.

B. Charts Suitable for Concrete

The choice of the statistic to be used in control charts depends on the nature of the product and process to be controlled, the nature and ease of testing, the reproducibility of the test method, and the expertise of the control chart-users. To decide which statistic is the most appropriate in controlling the manufacturing of concrete, the advantages and shortcomings of each type of control chart must be viewed in light of the concrete or concrete materials properties eligible for control.

In the case of concrete, producers could choose to control the same properties that concrete buyers measure for acceptance sampling -- slump, air content and cylinder strength. But their choice is not and should not be limited to these properties. Among other variables eligible for control are the amounts and quality of ingredients used in making concrete. For example, slump and strength could be indirectly controlled by controlling the amount of water and cement used, and air content by the amount of air entraining agent introduced into each batch, the mixing time, and the mixing energy. The choice of variables to be controlled depends on the producer and on his knowledge of the relationships between the variables chosen for control and the desired properties in the final product. In the case of concrete, it seems possible to control slump and strength by controlling the water-cement ratio. But the data reported in Appendix C indicate that the relationship between the dosage of air-entraining agent and the resulting air content may not be easily determined to control air content indirectly. Thus, the properties most likely to be chosen for control are: 1) slump, 2) air content, 3) strength, 4) amount of cement, and 5) amount of water. All these are variables which can be measured on a continuous scale, and statistics such as 1) mean, 2) standard deviation, 3) range, and 4) fraction defective can be computed. This means that $\bar{X}$-charts, σ-charts, R-charts and p-charts, as well as others, are all theoretically applicable to concrete production. But although applicable, not all offer the same advantages.

1. p-Charts

Charts to control the fraction defective or p-charts are desirable from the management point of view. They provide a continuous record of quality for economic studies and management decisions. But they are not very helpful to the quality control engineer when a production unit can be out-of-specification for more than one property, because by controlling the total fraction defective, he does not know which property is causing defects or in what proportion. Thus, essential information needed to prevent defects is not readily available. Moreover, p-charts require large sample sizes and unless testing is relatively inexpensive and non-destructive, they are economically undesirable. Because concrete testing is time-consuming and expensive, and because concrete can be defective for more than one property, p-charts are not the most appropriate for control of its production.

2. c-Charts

Number-of-defects-per-unit control charts are used when one single production unit can have a large number of defects that are not necessarily detrimental to performance of the production unit but are nevertheless undesirable -- for example, blemishes per square yard of cloth, scratches on a refrigerator, or minor defects in a car. This type of control chart requires that the testing be by attributes. It is not the most appropriate to control such variables as encountered in concrete production.

3. Difference Control Charts

Difference control charts are used to test the hypothesis that a process output is no different than material in a standard lot kept under the same environmental conditions as the process output being judged. They are employed when test results are sensitive to such conditions. The process is said to be in control if the control statistic of an output sample does not differ from the corresponding statistic computed from a sample taken from the standard lot by more than the difference expected due to sampling variations. For concrete, no standard lot can be kept because it hardens, and thus this type of control chart is not appropriate.

4. Cusum Charts

Cumulative sum control charts can be used for control of both the process average and fraction defective. These charts are statistically more discriminating than the corresponding Shewhart chart or $\bar{X}$- and p-charts. But their limits depend on the average run length (*ARL*) and are difficult to compute without the aid of a computer or such monographs and tables as those given by Kemp (6,7). Although in principle these charts are applicable to concrete production, it is believed that they will not be well received by the concrete industry. What is gained in statistical efficiency with cusum charts does not compensate for the simplicity and clear graphical display of the process operation lost by not using the corresponding classical Shewhart charts. The concrete industry, which except in rare cases has no formal quality control, is not likely to seek the most sophisticated tools, but rather the simplest.

5. σ-Charts

Standard deviation control charts are adaptable for control of the variability of output properties that can be measured on a continuous scale, such as those of concrete, and thus they could be used. However, they require large sample sizes. If the sample is less than 10, the range is preferable as a measure of variability. Since, in concrete testing, it is very difficult to sample 10 or more consecutive production units, σ-charts are not the most appropriate. The range control chart is more effective because it allows judging variability at more frequent intervals.

6. *R*-Charts

Control charts for sample ranges are widely used to control the variability of process output. They are applicable to concrete properties and are considered the most appropriate for control of concrete variability.

7. $\bar{X}$-Charts

Perhaps the most widely used and misused of all control charts are those for sample means. These are simple to construct and provide self-explanatory displays of process conditions with time. Because of their simplicity and because the theory for these charts is widely published, $\bar{X}$-charts are considered the most appropriate to control the levels of concrete properties.

R and $\bar{X}$-charts then emerge as the most logical choices to control concrete production. It remains to determine whether they should be used separately or concurrently. The ultimate goal of process control is the elimination of defective production units. In the case of concrete properties, defectives can be caused both by shifts in means and by increases in variability. Thus, for effective control of the process, R- and $\bar{X}$-charts must be used concurrently. To illustrate this point, consider air content. If mean air content coincides with that desired, but its variability as measured by the standard deviation is larger than allowed, some concrete batches will have air contents outside the desired limits. This is illustrated in Figure 8A, where σ_0 is the desired standard deviation, and $\bar{X}_0$ the desired mean. If the process mean and standard deviation coincide with those desired, no results will exceed the tolerances. However, if $\bar{X}_0$ approaches the desired but σ_0 increases to σ_1 , some results will exceed the limits as represented by the shaded areas. Similarly, if the mean of the process shifts to $\bar{X}_0 \pm \delta$ while the process standard deviation remains approximately equal the desired, results will again fall outside the limits as shown in Figure 8B. If the mean increases to $\bar{X}_0 + \delta$, some results will exceed the upper limit as represented by the Area A_2. If the mean shifts to $\bar{X}_0 - \delta$, a fraction of the results represented by Area A_1 will exceed the lower limit. Thus, to assure that the process output meets specification tolerances, both the process mean and the variability must be controlled, and R- and $\bar{X}$-charts must be used concurrently.

C. Information Required in Designing Charts

Choosing the types of control chart to be used is only the first step. Next comes the more difficult task of gathering the necessary information. Construction of R- and $\bar{X}$-charts requires knowing 1) the frequency distribution of sample means and sample ranges, 2) the frequency distribution of the control properties, 3) the desired process mean, 4) the standard deviations of the control properties when the process is operating at the level of control desired, 5) the probability of falsely looking for trouble in the process when none exists, and 6) the size of the rational subgroup to be used.

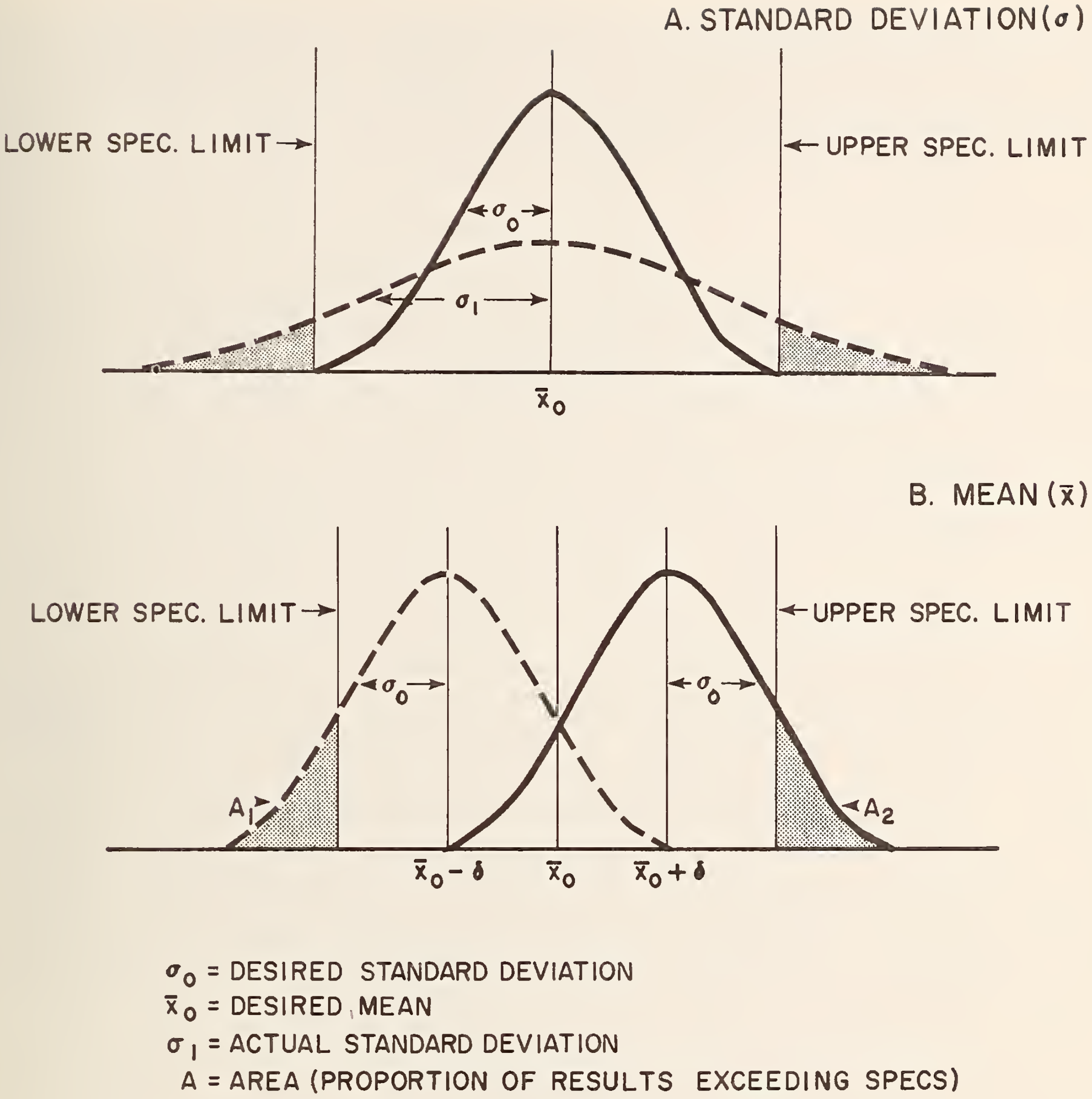


Figure 8. Effects on fraction defective of changes in standard deviation and shifts in mean.

The frequency distributions of sample means and sample ranges are well known. Their parameters are extensively tabulated in the quality control literature and present no problem. The literature and the data presented in the preceding chapter and tested for normality in Table 16 also indicate that the frequency distributions of concrete properties are approximately normal. The desired process mean is usually set in the specifications, and needs no attention at this stage, but the remaining parameters are not so readily obtainable. The selection of subgroup size and the probability of falsely looking for trouble in the process depend on costs, while standard deviations to be used must either be determined from given standards or obtained through process capability studies. For concrete, formalized quality control appears to be a rarity and it seems appropriate to discuss how this information may be obtained.

1. Producer's Risk

The choice of producer's risk or the probability of falsely looking for assignable causes when none may exist, depend on the economic consequences of not discovering assignable causes in those instances where they do exist. To stop the process and look for trouble adds to production costs, but so does rejection of production units. In choosing the probability of falsely looking for trouble, it is necessary to balance the cost of looking for assignable causes and discovering none against that of rejection resulting from assignable causes going undetected. If looking for assignable causes is inexpensive while the cost of rejections is high, the probability of falsely looking for trouble should be relatively large -- say 5 or 10 percent. However, if the cost of looking for trouble in the process is high while the cost of rejection is low, this probability should be chosen to be low. For concrete, the cost of rejections can be very high -- for example, rejection of a 6 cu yd load of concrete means a loss of at least $90. A few rejections can quickly dissipate a day's profit. But chasing non-existent assignable causes on 5 percent of the occasions when a sample is recovered from the process can be more expensive yet, especially if work must stop. This suggests that initially setting the probability of falsely looking for trouble at approximately 1 percent and using the customary 3σ limits is still appropriate. This risk could then be changed, based on actual cost data.

2. Selection of the Rational Subgroup

The choice of the rational subgroup must be consistent with the objective of control charting and must be based on both economic and process considerations. The objective of the range chart is continuous testing of the hypothesis that process variation does not differ from its variation when in control by more than expected due to sampling variation alone. For proper testing of this hypothesis, R-chart limits must be based on random variation alone, i.e., on the variation representing control. If other than random variation were included, the resulting limits would be wider, with loss of sensitivity of the R-chart to changes in process variation (3, pp. 13-49 through 13-50).

Similarly, the intent of an $\bar{X}$-chart is to detect shifts in process average greater in magnitude than those expected due to random variation. This is accomplished by continuously testing the hypothesis that the process mean at any time does not differ from that of the process when in control by more than expected due to random variation. The limiting values of the expected shifts, which are the $\bar{X}$-chart limits, depend on the random variation. Again, if these limits were computed based on a standard deviation including other than random variation, they would be wider and would result in loss in sensitivity of the $\bar{X}$-chart to shifts in process average.

Thus, for control chart purposes, it will be necessary to obtain a sample reflecting only random variation. Such a sample is also known as the rational subgroup, and most quality control books offer guidelines for its proper selection. These guidelines can be succinctly summarized -- include in the subgroup only consecutive production units manufactured with the same materials and under essentially the same conditions. It is reasoned that such a sample is most likely to reflect random variation alone because the process mean and variation are likely to remain stable over short periods.

Unfortunately, this golden rule cannot always be easily applied because recovery of samples from consecutive production units can be physically difficult or even impossible. For concrete, testing of consecutive production units is difficult because it takes about 20 minutes to sample and test one production unit. During this time, a plant in full production can mix at least ten batches, and concrete testing cannot be postponed. Thus, sampling consecutive units is a remote possibility unless more than one tester is provided or variables other than those inspected for acceptance sampling are used for control -- probably an unlikely case. However, if production of non-consecutive units occurred under essentially the same conditions and within a relatively short time, the variation between them might approach that of consecutive units. Under these circumstances, a sample approaching the rational subgroup would still be obtainable with one tester. Thus, when sampling for process control, care should be taken to assure that sampling is performed as quickly as possible, and during the sampling that aggregates, cement, admixtures, personnel, etc., remain unchanged. A subgroup consisting of consecutive or nearly consecutive production units is also desirable from a practical point of view, and preferable to a random sample. The practical importance is that such a sample facilitates the identification of assignable causes. For example, assume that a random sample of size n is recovered over 2 hours and has to be used for control chart purposes. Further assume that the mean computed from this sample shows a lack of control when plotted on the $\bar{X}$-chart, that the individual test results are not available, and the problem is to identify what caused the process mean to shift.

Under these circumstances, it is not known at what point during production the process first came under the influence of assignable causes. Thus, all factors present during the 2-hour period but not before must be suspected and investigated. In concrete production many things can change in 2 hours and the list of suspects may be large, making it difficult to isolate the culprit if it has not disappeared in the meantime. But if the sample consists of consecutive or nearly consecutive production units, the search would be limited to factors present during a very short time. The list of suspects would be shorter and the chance of the culprit disappearing would be minimized. Thus, random sampling that is *essential* for acceptance sampling is *undesirable* in sampling for *process control*.

3. Rational Subgroup Size or Sample Size

Another factor influencing the control limits is sample size. The larger the sample size, the tighter the limits and greater the sensitivity of the charts. Choice of subgroup size depends chiefly on economics. In choosing sample size, testing costs must be balanced against the consequences of producing defectives. Thus, the optimum sample size and sampling frequency should be determined for each individual process. However, findings of theoretical studies can serve as a guide until enough information is accumulated from case studies. A. J. Duncan made one such study of sample sizes for $\bar{X}$- and R-charts, (8, p. 398). He summarizes the findings as follows:

1. The customary sample sizes of 4 or 5 are close to optimum if the shifts to be detected are relatively large, e.g., if the assignable cause produces a shift of $2\sigma'$ or more in the process average. If it is the aim of the chart to detect shifts in the process average as small as one σ', sample sizes of 15 to 20 are more economical than sample sizes of 4 or 5.

2. If a shift in the process average causes a high rate of loss, i.e., high relative to the cost of inspection, it is better to take small samples quite frequently than large samples less frequently. For example, when the rate of loss is high; samples of 4 or 5 taken every half hour are better than samples of 8 or 10 taken every hour.

3. Under certain circumstances charts using 2σ or even 1.5σ limits are more economical than charts using the conventional 3σ limits. This is true if it is possible to decide very quickly and inexpensively

that nothing is wrong with the process when a point (just by chance) happens to fall outside the control limits, i.e., when the cost of looking for trouble when none exists is low. Contrariwise, it will be more economical to use charts with 3.5σ to 4σ limits if the cost of looking for trouble is very high.

4. If the unit cost of inspection is relatively high, the most economical design is one that takes small samples (say samples of 2) at relatively long intervals (say every 4 to 8 hours) with narrow control limits, say ± 1.5σ.

If the concrete properties conventionally tested are chosen for process control, which is likely to be the case, it is suspected that assignable causes would produce large shifts in process average. But this is only a conjecture, which cannot be substantiated with available data. The bulk of the data available to the author for the traditionally measured concrete properties were obtained under random sampling, which reflect random as well as assignable-cause variation, and preclude determining the magnitude of changes in process averages due to assignable causes. If producers chose other than conventionally tested properties for plant control, the magnitude of shifts in mean due to assignable causes would still be unknown, for data on other than conventionally tested properties are almost non-existent. This precludes choosing sample size on the basis of the expected magnitude of shifts in process mean, and the choice must be made on the basis of cost of rejection, which can be high. Thus, it appears desirable to test small subgroups at frequent intervals and samples of four taken at least every hour are a good starting point in accumulating data necessary for determining optimum sample size.

Regardless of the sample size ultimately chosen, its effect on sensitivity of $\bar{X}$- and R-charts can be shown with operating characteristic curves (8, pp. 391-395). The OC curves in Figure 9A show how sample size affects the $\bar{X}$-chart's ability to detect a shift in mean of a given magnitude. The OC curves in Figure 9B show how sample size affects the R-chart's ability to detect changes in process variation. It can be seen that the probability of not catching a shift in process average of the same magnitude increases as sample size decreases. For example, the probability of not detecting a shift in mean of magnitude $2.0\sigma'$, on taking the first sample after the shift has occurred, varies approximately from 0.04 for $n = 6$ to 0.55 for $n = 2$. Similarly, Figure 9B shows that the probability of not

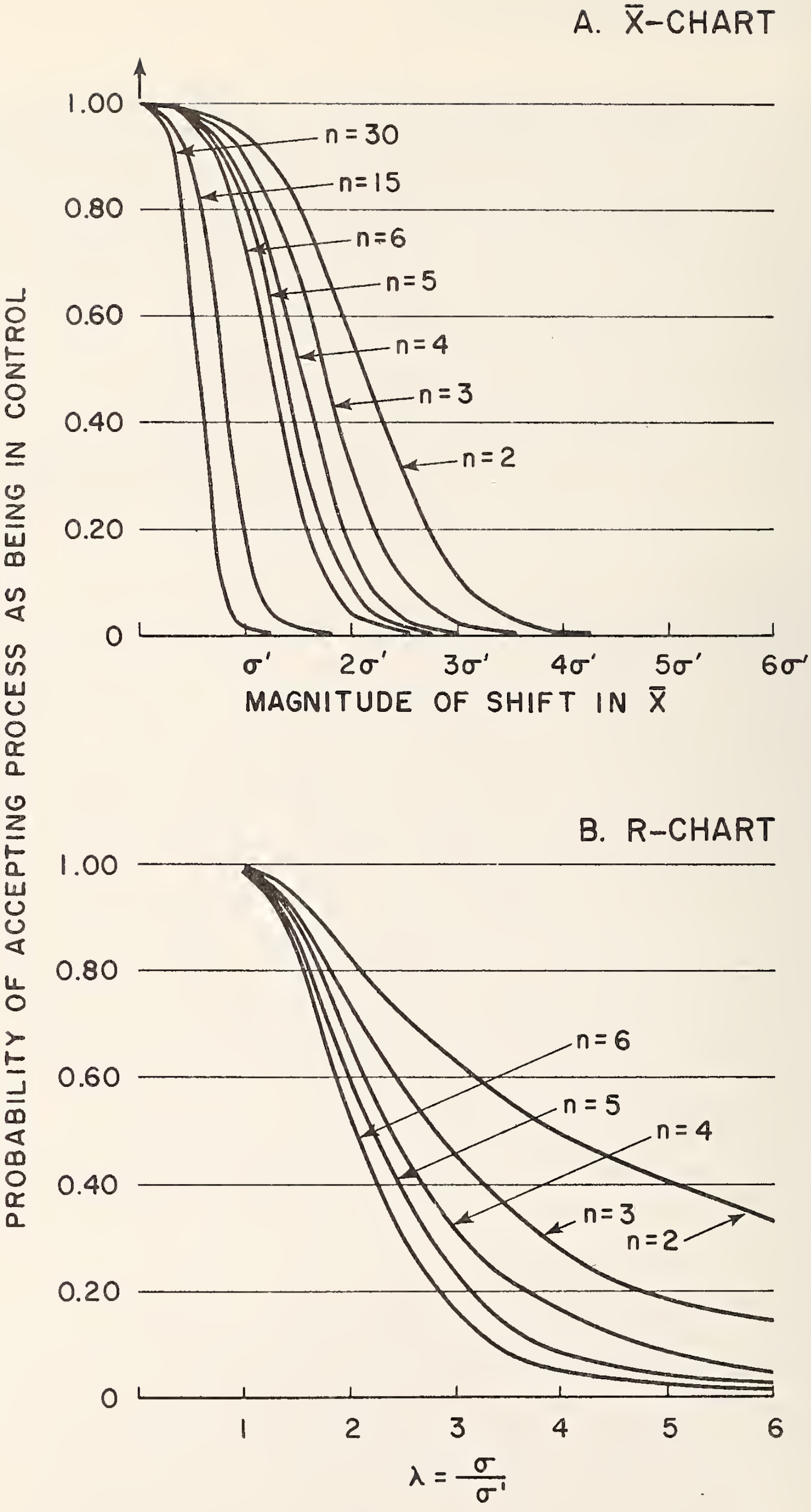


Figure 9. *OC* curves for process control charts for 3σ limits.

detecting a change in process standard deviation when it doubles ($\lambda = 2$) varies from 0.52 for $n = 6$ to 0.82 for $n = 2$. The OC curves also show that the probability of not catching large changes in process average and variation, upon taking the first sample after the changes have occurred, is relatively high with small subgroups. But the probability of not catching changes in process average and variation on the first and/or second sample after the changes have occurred is the product of the probability of not detecting the change for each individual sample, and generally the probability of not catching a change with any of g consecutive subgroups is P^g, where P is the probability of not catching the change with a single subgroup. Since P is a fraction, the probability of not detecting a change in either process average or process variation in any of g subgroups quickly becomes small even for moderate values of g. For this reason, sampling for process control should occur at frequent intervals.

4. Determination of Variation

Having chosen the producer's risk and rational subgroup size, the value of the standard deviation to be used must still be determined before R- and $\bar{X}$-charts can be constructed. This variation may be known from past experience, derived from given standards, determined through process capability studies, or approximated from recent process output. In setting up R- and $\bar{X}$-charts two situations can arise. In the first -- no standards given -- the minimum achievable process variation is unknown and the process must be brought into control with respect to itself. In the second -- standards given -- the process must be controlled to meet the standards but need not necessarily be brought into control with respect to itself. When no standards are given, the process must be manipulated until all assignable cause variations are eliminated and the properties controlled assume predictable patterns.

For correct determination of whether a process has reached a state of control, the decision must be based on test results from a large number of rational subgroups. But because testing is usually costly, it is customary to accept the hypothesis that a process has reached control solely on the basis of limited data from the recent past, and to estimate the standard deviation to be used in setting control limits from these same data. This procedure, which involves some risks, consists of the following steps (discussed further in 3, pp. 13-46 through 13-63):

1. Test a predetermined number of subgroups of size n,
2. Compute the mean and range for each subgroup,

3. Using appropriate formulae based on the data collected and the subgroup size, calculate upper and lower limits for both the *R*- and the $\bar{X}$-charts,

4. Plot the subgroup ranges and means respectively on *R*- and $\bar{X}$-charts, and

5. If all plotted points fall within the control limits, accept the hypothesis of control; otherwise reject it.

However, a process thus declared in control may in fact not have reached it. The probability of the process not having reached a state of statistical control depends on the number of rational subgroups and on the subgroup size. King (9,10) has computed these probabilities for a number of subgroups of size 5 as a function of shifts in process average. The resulting OC curves and OC curve upper bounds are shown in Figure 10. A process whose mean has shifted from the true but unknown process mean by $1.0\sigma'$ would be accepted as being in control with respect to itself only 3 percent of the time, if the $\bar{X}$-chart were based on 25 subgroups of 5. However, if the mean had shifted the same amount, the process would be accepted as in control approximately 90 percent of the time, if the mean control chart were based on only 2 subgroups of 5. Thus, of the OC curves shown, that based on 25 subgroups of 5 gives the lowest probability of declaring a process in control with respect to itself when in fact it is not. This is one reason why most books on quality control recommend basing the $\bar{X}$- and *R*-charts on at least 25 subgroups of 5. Another reason is that such control charts have been observed to work well in practice.

In some cases, particularly as producers accumulate data, the standard deviation is known and control limits can be easily determined. More frequently, standards are given and the standard deviation to be used for process control is derived from standards. This is usually the case when the sole objective of process control is to meet the buyer's (or inspecting agency's) requirements, and producers choose for control the same properties that will be tested for acceptance sampling. In those situations, the standard deviation to be used in constructing $\bar{X}$- and *R*-charts is taken as one-sixth the specification range for each control property. This approach assumes that the process is capable of producing outputs whose properties have standard deviations equal to or less than one-sixth the respective specification ranges.

For concrete properties, standards are given in the form of specifications, and bringing the process into control with respect to itself could appear to be superfluous work. But it is desirable to assure that the process is capable of meeting the specifications. When a process is brought into control

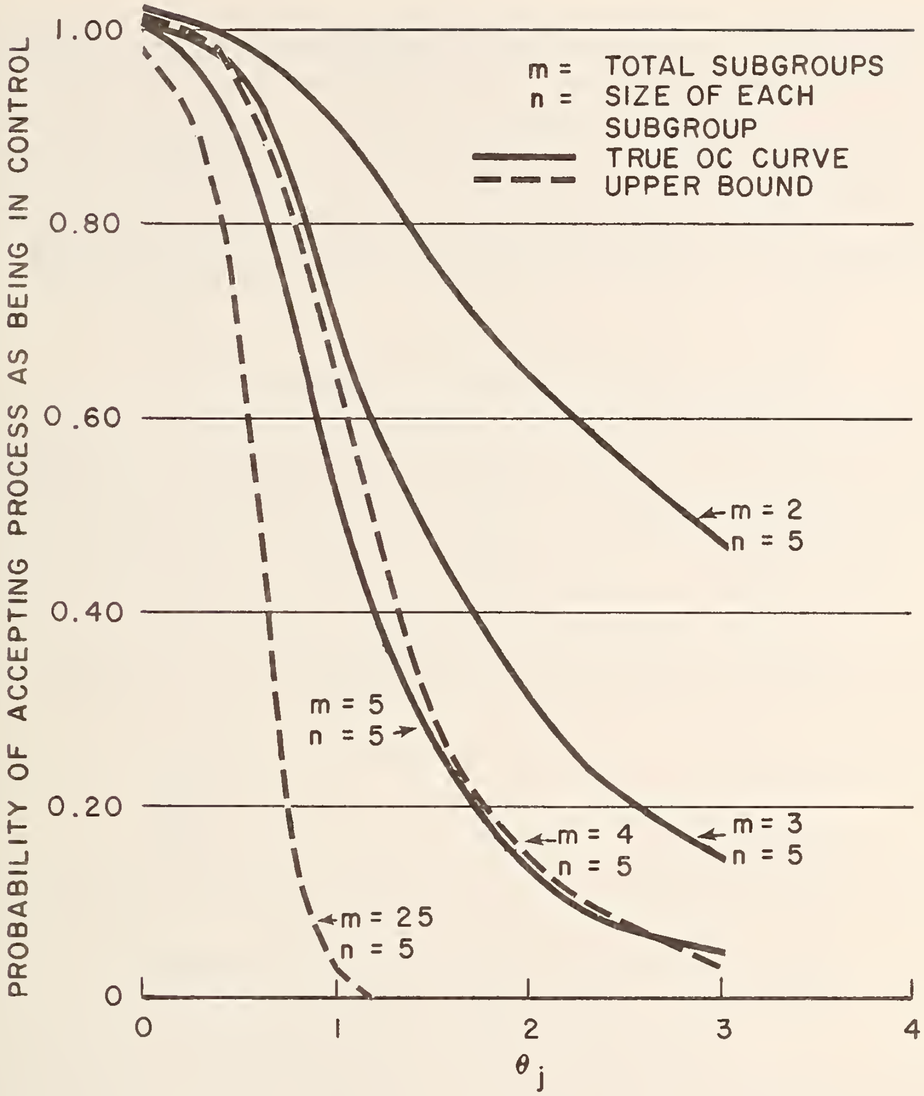


Figure 10. *OC* curves for control charts of sample mean (θ_j = size of shift in mean from the desired process mean in terms of the process standard deviation.

---

with respect to itself, the process mean may differ from that specified. But process variation is close to a minimum and represents the economically achievable. Thus, bringing the process into control with respect to itself will reveal whether the process can meet specifications, and if the process cannot meet the set requirements, defectives will result. Undertaking production under these circumstances is very risky if the buyer is using statistical acceptance sampling plans (see Chapter III) for inspection. Consequently, producers should bring the process into control with respect to itself whenever possible. However, because concrete testing is costly and difficult, it is suggested that initially producers use a minimum of 5 subgroups of 5 to establish control limits instead of the customary 25 subgroups of 5. This gives a probability of falsely declaring the process in control of approximately 50 percent (Fig. 10) when the mean shifts by $1.0\sigma'$. But producers could revise the limits and reduce this probability by using results from subgroups tested subsequently to monitor the process and falling within the control limits.

D. Chart Construction and Operation

Once necessary decisions have been made and essential parameters chosen, construction of R- and $\bar{X}$-charts is reduced to a simple step-by-step procedure. When no standard is given, R- and $\bar{X}$-charts must be based on observed data obtained from the process during the immediate or recent past. The necessary steps are as follows:

1. Take g rational subgroups of size n from the process, as close in time as possible, and compute the average range $\bar{R}$ as follows:

$$\bar{R} = \frac{\sum_{i=1}^{g} R_i}{g} \qquad (1)$$

where R_i = the range of each subgroup of size n, and

g = the number of subgroups of size n.

Then the upper control limit UCL and lower control limit LCL are computed as follows and the R-chart plotted as in Figure 11B:

$$UCL = \bar{R} + k \frac{d_3}{d_2} \bar{R} \qquad (2)$$

$$LCL = \bar{R} - k \frac{d_3}{d_2} \bar{R} \qquad (3)$$

where k = the number of the range standard deviations corresponding to one, minus the probability of falsely looking for trouble (usually taken as 2 or 3),

d_2 = the mean of the distribution of relative range R/σ' for sample of size n, and

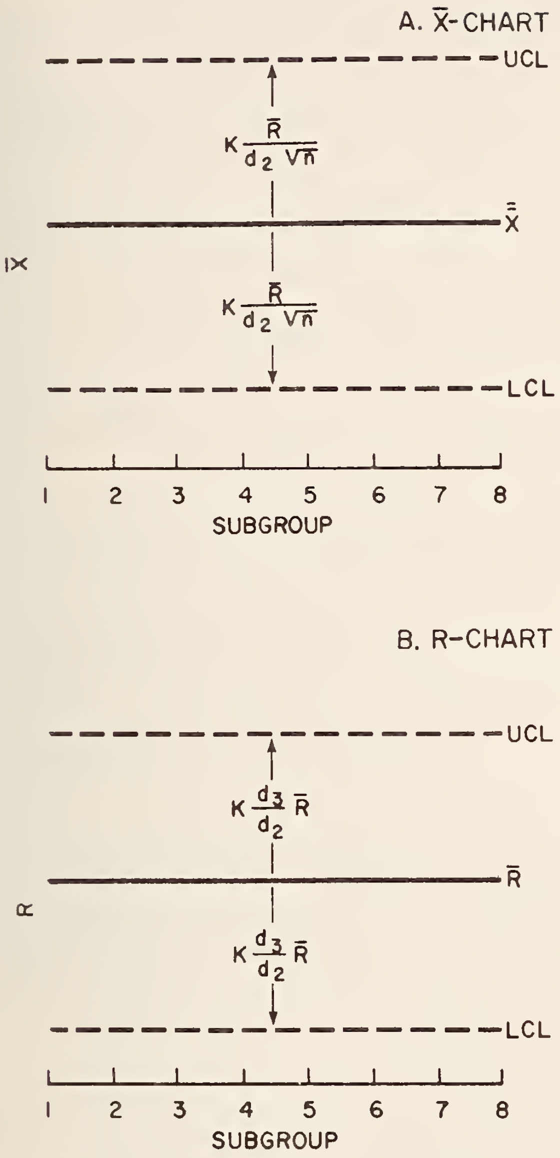


Figure 11. Example of process control charts when standards are not given.

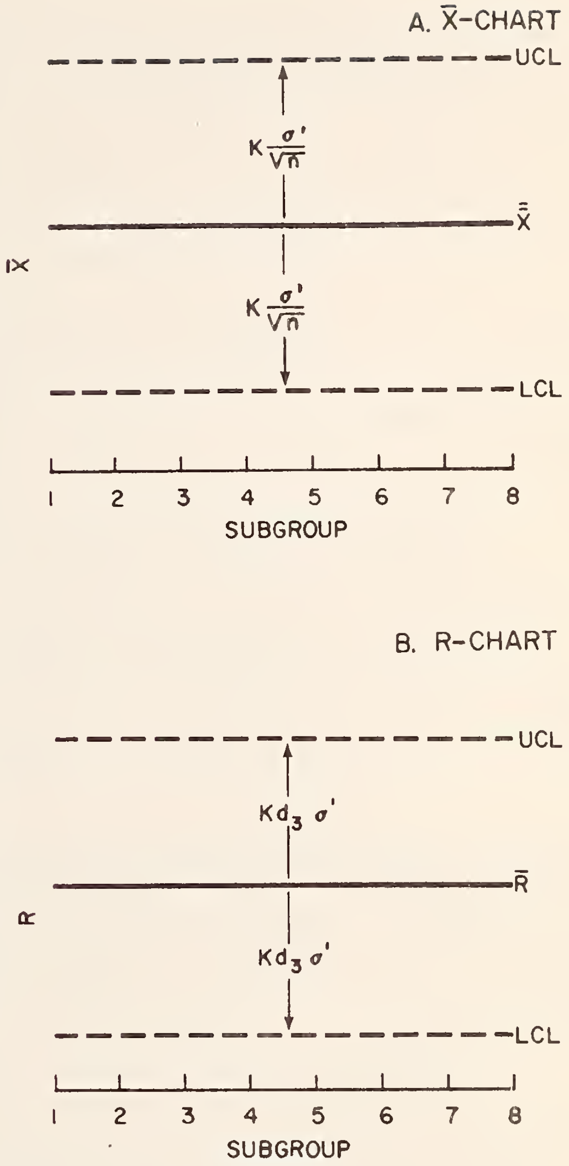


Figure 12. Example of process control charts when standards are given.

d_3 = the standard deviation for the distribution of the relative range for the sample size given (values of d_2 and d_3 are tabulated in most quality control books, but sometimes under different names; see 8, p. 908).

2. To set up the $\overline{X}$-chart from g subgroups, compute the grand mean $\overline{\overline{X}}$ as follows:

$$\overline{\overline{X}} = \frac{\sum_{i=1}^{ng} X_i}{ng} \tag{4}$$

where X_i = the individual test results,

g = the number of rational subgroups, and

n = the size of rational subgroup.

Then compute the upper control limit UCL and lower control limit LCL as follows and plot the $\overline{X}$-chart as in Figure 11A:

$$UCL = \overline{\overline{X}} + k \frac{\overline{R}}{d_2\sqrt{n}} \tag{5}$$

$$LCL = \overline{\overline{X}} - k \frac{\overline{R}}{d_2\sqrt{n}} \tag{6}$$

When the mean and standard deviation of control properties are known or derived from specification limits, the procedure for setting up R- and $\overline{X}$-charts for subgroups of size n also becomes very simple. Letting σ' equal the known or derived standard deviation, the procedure to construct R-charts reduces to the following steps:

1. Compute $\overline{R}$ by:

$$\overline{R} = d_2\sigma' \tag{7}$$

where $\overline{R}$ = the average range,

d_2 = the mean of the relative range, and

σ' = the known or assumed process standard deviation.

2. Compute the R-chart upper limit UCL and lower control limit LCL as follows:

$$UCL = d_2\sigma' + kd_3\sigma' \tag{8}$$

$$LCL = d_2\sigma' - kd_3\sigma' \tag{9}$$

3. Plot the R chart as shown in Figure 12B.

Similarly, letting the given or specification mean equal $\bar{\bar{X}}$, the procedure in setting up an $\bar{X}$-chart for standard given reduces to the following steps:

1. Compute the upper control limit UCL and lower control limit LCL as follows:

$$UCL = \bar{\bar{X}} + k\frac{\sigma'}{\sqrt{n}} \tag{10}$$

$$LCL = \bar{\bar{X}} - k\frac{\sigma'}{\sqrt{n}} \tag{11}$$

2. Plot the $\bar{X}$-chart as shown in Figure 12A.

Both the procedure for no standard given and that for standard given assume that the control properties are normally distributed. This assumption is usually of no great consequence, unless the properties controlled follow distributions deviating markedly from normality. As already mentioned, concrete properties can be taken to be approximately normally distributed and this assumption should not result in difficulties.

Consistent with the objective of control charts to test graphically the hypothesis that the control statistic does not fall outside the allowed interval, the operation of R and $\bar{X}$-charts reduces to three steps:

1. Sampling and testing rational subgroups,

2. Computing subgroup means and ranges, and

3. Plotting subgroup means and ranges on the appropriate control charts to see whether they fall within the chosen confidence intervals (which are represented graphically by the control limits).

If the values of subgroup ranges and means fall within the corresponding control charts limits, the process is considered in control and routine testing is continued. If either the mean or range of one or more subgroup plots outside the control limits, the process is taken to be out of control. When a point on either plot is out of control, assignable causes are sought, and if found they are identified and eliminated. During the search for assignable causes, testing frequencies are increased and continued until there is reason to believe that the process is back in control and likely to remain there. When there is evidence that control has been restored, routine testing is resumed until another point again plots out of control and the cycle starts again.

Experienced quality control personnel have refined the basic rules for operation of control charts in attempts to prevent the process from going out of control at all. They have established criteria for action before the assignable causes can cause trouble. For example, 2σ limits are often used as a warning limits for action when too many points approach or exceed those limits. Other criteria are also used. Duncan summarizes the common action criteria as follows (8, p. 347):

1. One or more points outside the control limits.

2. One or more points in the vicinity of a warning limit. This suggests the need for immediately taking more data to check on the possibility of the process being out of control.

3. A run (defined as successive items of the same class) of 7 or more points. This might be a run up or run down or simply a run above or below the central line on control chart.

4. Cycles or other nonrandom patterns in the data. Such patterns may be of great help to the experienced operator. Other criteria that are sometimes used are the following:

5. A run of 2 or 3 points outside 2σ limits.

6. A run of 4 or 5 points outside 1σ limits.

This multiplicity of criteria increases the chances of falsely looking for trouble, and the choice of action criteria should be based on economics.

E. Suggested Charts

The properties to be controlled, size of the rational subgroup, and probability of falsely looking for trouble are the producer's responsibility and choice. He should also choose whether to control each process with respect to given standards or respect to itself. To make these decisions, producers must rely on data systematically collected and properly analyzed. Unfortunately, few concrete producers in business in New York State practice formalized statistical quality control. This means that most producers probably do not have the necessary data to determine which is the best quality control setup for each of their particular plants. In fact, producers probably have no more information than state agencies, because under the present system the buyer usually does all the testing. For this reason, it seems appropriate that state agencies suggest process control charts for producers to use until they can accumulate enough information to set up quality control systems properly on an individual plant basis. Such suggestions, based on the points discussed in the two preceding sections, follow, and it is believed that they can provide a good starting point and yield good results:

1. Bring the process in control with respect to itself to assure that specification can be met,

2. When the process is in control with respect to itself and variation is consistent with the specification ranges, set up R- and $\bar{X}$-charts based on the standard deviation derived from the specifications, i.e., one-sixth the range for the property inspected,

3. Use a probability of falsely looking for assignable causes of approximately 1 percent by using 3σ limits control charts ($k = 3$), and

4. Use both R- and $\bar{X}$-charts to minimize rejections and a rational subgroup of size 4.

If these suggestions were taken, the resulting control chart limits to meet the present New York State concrete specifications would be as outlined here:

1. Slump

The current New York State specification for slump is summarized in Table 8A, where it can be seen that for sidewalk and deck slabs the upper half range for slump is 1.0 in. This means that σ' would have to be taken as 1/3-in. The same is true for walls less than 1 ft 6 in. thick, box culverts, and rigid frames and arches. But a standard deviation of 1/3-in. for slump is a rarity. The lowest standard deviation reported in the preceding chapter (Fig. 7C) is between 0.30 and 0.39 in. -- and this happened only once in 32 cases. Even granting that the standard deviations in Fig. 7C were computed from random samples and should be larger than within-group standard deviations, it is still a low figure which may be difficult to achieve. To avoid this problem the design slump should be changed to those in Table 8B for quality control purposes. This can be done because no lower limit is given for this property and would allow a σ' of 1/2 in. which, as will be seen, is easily achievable. This change would lead to the control chart limits in Table 9.

TABLE 8
DESIGN SLUMP AND RANGE AS NOW SPECIFIED
AND AS SUGGESTED FOR USE WITH PROCESS CONTROL CHARTS
TO ASSURE COMPLIANCE WITH SPECIFICATIONS

Concrete Type	Design Slump	Maximum Specified Slump	Upper Half Range	1/3 of Upper Half Range
A. SPECIFIED LIMITS FOR SLUMP				
Pavement	1-1/2"	3"	1-1/2"	1/2"
Sidewalks, Deck Slabs	2-1/2"	3-1/2"	1"	1/3"
Footings, Headwalls, Piers, Pedestals, Walls over 1'6" thick	2-1/2"	4"	1-1/2"	1/2"
Walls less than 1'6" thick, Box Culverts throughout, Rigid Frames and Arches	3"	4"	1"	1/3"
B. SUGGESTED DESIGN SLUMP AND RANGE				
Pavement	1-1/2"	3"	1-1/2"	1/2"
Sidewalks, Deck Slabs	2"	3-1/2"	1-1/2"	1/2"
Footings, Headwalls, Piers, Pedestals, Walls over 1'6" thick	2-1/2"	4"	1-1/2"	1/2"
Walls less than 1'6" thick, Box Culverts throughout, Rigid Frames and Arches	2-1/2"	4"	1-1/2"	1/2"

TABLE 9
SUGGESTED 3σ LIMITS FOR
R- and $\bar{X}$-CHARTS FOR SLUMP

	Limits for Range Process Control Chart ($n = 4$)			Limits for Mean Process Control Chart ($n = 4$)		
Concrete Type	*LCL*	$\bar{R}$	*UCL*	*LCL*	$\bar{\bar{X}}$	*UCL*
Pavement	0.00	1.03	2.35	0.75	1.50	2.25
Sidewalks, Deck Slabs	0.00	1.03	2.35	1.25	2.00	2.75
Footings, Headwalls, Piers, Pedestals, Walls over or less than 1'6" thick, Box Culverts throughout, Rigid Arches and Frames	0.00	1.03	2.35	1.75	2.50	3.25

TABLE 10
PRESENT SPECIFIED LIMITS FOR AIR CONTENT

	Air Content, percent				
Concrete Class	Min.	Design	Max.	Range	1/6 Ran
A	4.0	6.0	8.0	4.0	0.67
B	3.0	5.0	7.0	4.0	0.67
C	4.0	5.5	7.0	3.0	0.50
Items containing No. 1 max. size aggregate	5.0	7.0	9.0	4.0	0.67

TABLE 11
SUGGESTED 3σ LIMITS FOR
R- AND $\bar{X}$-CHARTS FOR AIR CONTENT

	Limits for Range Process Control Chart (n = 4)			Limits for Mean Process Control Chart (n = 4)		
Concrete Class	LCL	$\bar{R}$	UCL	LCL	$\bar{\bar{X}}$	UCL
A	0.00	1.38	3.15	5.00	6.0	7.00
B	0.00	1.38	3.15	4.00	5.0	6.00
C	0.00	1.03	2.35	4.75	5.5	6.25
Items containing No. 1 max. size aggregate	0.00	1.38	3.15	6.00	7.0	8.00

2. Air Content

The current specification limits for the various concrete classes are given in Table 10. Both upper and lower limits are given and the specification range is either 3- or 4-percent air content. The implied σ' for a range of 4 percent is 4/6 or 0.67-percent air. For the range of 3 percent, the implied σ' is 3/6 or 0.5 percent. These within-group standard deviations are believed achievable and the control charts limits necessary to meet the current specification are shown in Table 11.

3. Compressive Strength

The New York State general concrete specifications do not state a minimum strength. In most cases, the desired strength is attained by specifying the mix design and testing the concrete materials. In special cases, however, strength is also specified. Besides the state, many private concrete buyers give strength specifications. When strength specifications are given, producers must control this property in order to avoid rejection or penalties due to the discovery of defective concrete during acceptance sampling.

Concrete strength cannot be controlled directly with present testing procedures because to identify and eliminate assignable causes one must be aware of their existence. With present testing methods, the presence of assignable causes is not known until 28 days after concrete mixing. By this time, production conditions and materials may have changed many times, and in many cases production may even have ended. This makes it impossible to find or even look for assignable causes of relatively short duration. Because assignable causes in concrete production can change rapidly as materials are used and replaced from new sources, 28-day strength results cannot be effectively used for direct control of this property. This means that present testing methods must be abandoned in favor of quicker ones or strength must be controlled indirectly by controlling mix proportions and the quality of concrete materials.

If strength is to be controlled indirectly, valid correlations must be found between concrete components and strength. For quality control purposes, these correlations cannot be qualitative but must be quantitative. It is not enough to know how any factor or combination of factors affect strength. It is also necessary to know by how much. This presents a problem. Combinations of materials vary widely and a general correlation line, valid for all conditions, is unlikely. Correlation lines must be established for each materially different set of conditions and combination of materials. Consequently, no suggestion can be offered for general use.

Another possible way to control concrete strength is the use of so-called accelerated strength testing methods. Lately, researchers have been working on ways to predict concrete strength potential from the strength of concrete measured hours after mixing. Tiede (11) as well as others report success. These are now tentative ASTM standard testing procedures (12) and should be welcomed by quality control personnel. If ultimate concrete strength can be reliably predicted from early strength, the accelerated strength itself can be used as a standard and process control charts can be based on it. This reduces correlation between two similar variables -- early and potential strength -- and should simplify control of concrete strength. However, even in this case, correlation lines have been found to vary from job to job (12). For this reason, again, no general suggestion can be offered.

For long-duration projects and for general information, concrete producers may want to use control charts for the 28-day strength. In such cases, a σ' of 500 psi seems reasonable (Fig. 7E). Then, for $n = 4$, $k = 3$, the design value for the $\bar{X}$-chart would be 1500 psi above the minimum specified strength, and its lower limit would be 750 psi above the minimum strength. The central line on the R-chart would be 1030 psi and the upper limit 2350 psi.

F. Case Studies

Process control charts based on within-group variation, suggested in the preceding section, were used in three plants to control slump and air content. Sampling and testing were performed by personnel of the Materials Bureau with the consent and cooperation of the producers involved. The objective was to determine whether variations implied in the current specifications are actually achievable, and to demonstrate the feasibility of process control in concrete plants.

1. Case Study 1

The concrete being produced was modified Class C mix with No. 2 stone as the largest size of aggregate. The specifications to be met were the same as Class C for air content, but modified for slump to a maximum of 4 in. The slump specification was changed by the Materials Bureau in response to field complaints of slump changes between plant and job site. The appropriateness of the change may be disputed. However, the specification change does not invalidate the use of control charts. The upper limit for the slump $\bar{X}$-chart was changed to take into account this change in the maximum allowable slump, but was based on the same between-group standard deviation of 1/2-in.

Testing was performed at the plant over a four-day period. The results and testing times are presented in Table 12. The intervals for individual subgroups ranged from 20 minutes to over an hour. The first four subgroups consisted of almost consecutive units, because the plant was operating at a low rate waiting for trucks.

The results are plotted in Figure 13. As can be seen in Figure 13B, the slump mean was always in control but above the design mean in most cases. The slump range plotted out of control only once -- for a low slump of 1.75 in. Basically, slump range was in good control with most of the sample ranges around or below the mean range $\bar{R}$. This indicates that the within-group standard deviation of 1/2-in. implied in the current specifications is achievable.

TABLE 12
CASE STUDY 1 FIELD DATA

Test	Time	Slump, in.	Air Content, percent	Subgroup	Slump		Air Content	
					R	$\bar{X}_4$	R	$\bar{X}_4$
DAY 1								
1	9:26	2.00	6.2	1	1.00	2.63	1.40	6.80
2	9:36	2.75	7.2					
3	9:45	3.00	6.2					
4	9:54	2.75	7.6*					
5	10:02	2.00	6.2	2	2.00	2.94	1.20	6.90
6	10:14	2.50	6.8					
7	10:27	3.25	7.2					
8	10:36	4.00	7.4					
9	10:45	2.50	5.8	3	0.50	2.63	2.10	6.30
10	11:00	2.50	6.6					
11	11:09	2.50	5.4					
12	12:21	3.00	7.5					
13	1:00	2.25	6.0	4	1.00	2.44	1.70	6.30
14	1:08	2.50	6.4					
15	1:17	3.00	7.2					
16	2:03	2.00	5.5*					
17	2:17	3.00	6.6	5	1.00	2.50	1.80	6.60
18	2:25	2.50	7.5					
19	2:31	2.00	5.7					
20	2:38	2.50	6.6					
DAY 2								
1	7:44	2.00	5.1	6	1.00	2.50	1.50	5.60
2	7:51	3.00	6.6					
3	8:06	2.00	5.0					
4	8:15	3.00	5.7					
5	8:25	2.50	5.9	7	1.00	3.00	0.50	6.20
6	8:32	3.00	6.2					
7	8:58	3.00	6.4					
8	9:11	3.50	6.2					
9	9:25	3.50	6.0	8	2.50	3.19	0.60	5.70
10	9:32	1.75	5.4					
11	9:47	4.25	5.5					
12	9:52	3.25	5.7					
13	10:04	2.25	5.0	9	1.25	2.94	1.20	5.60
14	10:41	3.50	5.5					
15	10:50	2.50	5.7					
16	10:57	3.50	6.2					
17	11:08	3.25	6.2	10	0.75	2.94	0.70	5.80
18	11:24	2.50	5.7					
19	11:37	3.00	5.9					
20	11:45	3.00	5.5					
DAY 3								
1	7:40	2.00	5.2	11	1.50	2.50	0.60	5.30
2	7:54	2.00	5.0					
3	8:04	2.50	5.6					
4	8:13	3.50	5.4					
DAY 3 (cont.)								
5	8:24	3.50	5.2	12	1.50	2.63	0.80	5.20
6	9:09	2.00	5.0					
7	9:23	2.00	4.9					
8	9:32	3.00	5.7					
9	9:45	2.25	5.8	13	0.75	2.31	1.40	5.30
10	9:55	2.75	5.9					
11	10:18	2.25	4.5					
12	10:31	2.00	4.9					
13	10:37	3.50	6.2	14	1.00	2.75	1.40	5.70
14	10:46	2.50	5.7					
15	10:54	2.50	4.8					
16	11:32	2.50	6.0					
17	11:39	3.50	6.4	15	1.25	3.19	1.50	6.10
18	11:47	3.00	6.0					
19	11:53	3.75	6.8					
20	12:02	2.50	5.3					
DAY 4								
1	7:55	2.25	4.6	16	1.25	2.75	0.80	4.60
2	8:03	2.25	4.2					
3	8:13	3.50	5.0					
4	8:21	3.00	4.7					
5	8:29	3.25	5.5	17	0.50	3.06	4.80	4.60
6	8:52	3.25	1.2*					
7	9:07	3.00	6.0					
8	9:17	2.75	5.8					
9	9:24	3.50	6.0	18	0.75	3.13	0.70	5.70
10	9:30	3.00	5.7					
11	9:38	2.75	5.7					
12	9:45	3.25	5.3					
13	9:53	2.50	5.3	19	0.50	2.63	1.30	5.40
14	10:01	3.00	4.7					
15	10:11	2.50	6.0					
16	10:34	2.50	5.6					
17	10:41	2.50	5.6	20	1.75	2.75	1.00	5.90
18	10:50	2.75	5.6					
19	10:59	3.75	6.6					
20	11:08	2.00	5.8					
21	11:24	3.25	6.7	21	0.25	3.13	1.70	5.60
22	11:31	3.25	5.4					
23	11:39	3.00	5.0					
24	11:45	3.00	5.2					
25	11:57	2.25	4.9	22	1.00	2.56	1.80	4.70
26	12:08	2.50	5.1					
27	12:21	3.25	5.2					
28	12:29	2.25	3.4					

*Correction in amount of air-entraining agent made after this test.

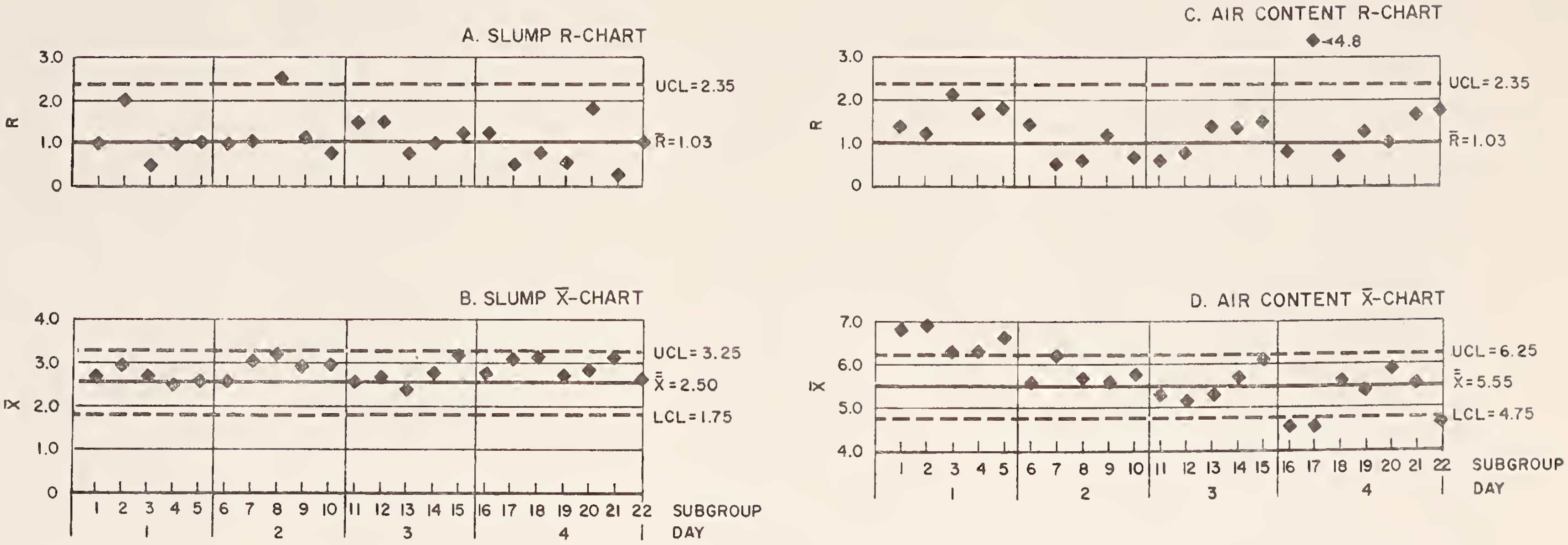


Figure 13. Control charts plotted from Case Study 1 data.

In contrast to mean slump, mean air content plotted out of control from the beginning. As can be seen from Figure 13D, the mean air content of the first subgroup plotted outside the upper control limit. The dosage of air entraining agent was adjusted but the mean of the second subgroup also plotted out of control. No adjustments were made after sampling the second subgroup but the mean of the third and fourth subgroups approached the upper limit. Then, the fifth subgroup mean went up again. At this point, the dosage of entraining agent was adjusted once more, and the mean air content came into control at the end of the first day. It remained in control for the second and third testing days. At the beginning of the fourth day, a low air content problem caused both the $\bar{X}$- and R-charts to show lack of control. After adjustment of the air-entraining agent, the mean air content remained in control until the end of testing.

The range chart for air content showed lack of control only once, when the mean chart also went out of control indicating that the high range was due to an assignable cause -- low dosage of air-entraining agent. The rest of the ranges plotted well within the upper limit in most cases. The range in fact remained in control while the mean air content was being lowered and variation between test results purposefully induced. This indicates that the within-group standard deviation of 1/2-percent air content implied in the current air content specification for Class C concrete is also achievable.

2. Case Study 2

Concrete was produced in a central plant and delivered to the paving train in open trucks. The plant was set up at the job site and the concrete produced was a modified Class C concrete, with No. 2 stone the maximum size of aggregate to be used in slipform paving. The specifications to be met for air content were those for Class A concrete, while the specifications for slump were the same as for Class C concrete. Testing lasted only one day and was performed at the plant by a three-man crew. The time spent sampling a subgroup (four production units) was about 30 minutes. The results are given in Table 13. The control charts are shown in Figure 14.

Referring to Figure 14B, it can be seen that for slump, the subgroup mean plotted out of control in three cases and was always above the design slump. No attempts were made to correct the situation, because it was claimed that slump at the grade was 3/4 to 1-1/4 in. less than at the plant. Thus the lack of control was intentional. Again this may be imputed as a violation of the specifications, but this brings home a point -- specifications are changed in the field.

TABLE 13
CASE STUDY 2 FIELD DATA

Test	Time	Slump, in.	Air Content, percent	Subgroup	Slump R	Slump $\bar{X}_4$	Air Content R	Air Content $\bar{X}_4$
1	7:40	2.50	7.0	1	1.00	2.75	1.40	6.80
2	7:55	2.50	6.2					
3	8:05	2.50	6.4					
4	8:15	3.50	7.6					
5	8:30	2.25	6.1	2	0.50	2.25	0.40	6.30
6	8:38	2.50	6.5					
7	8:45	2.00	6.5					
8	9:06	2.25	6.1					
9	9:15	2.75	6.7	3	0.75	2.94	1.10	7.30
10	9:25	2.50	7.8					
11	9:35	3.25	7.4					
12	9:45	3.25	7.1					
13	9:53	2.75	6.6	4	1.00	2.25	1.40	6.40
14	10:03	1.75	5.4					
15	10:10	2.50	6.8					
16	10:20	2.00	6.7					
17	10:35	1.75	6.4	5	0.75	2.00	0.60	6.40
18	10:45	2.00	6.0					
19	10:55	1.75	6.6					
20	11:05	2.50	6.6					
21	11:15	2.25	6.8	6	1.00	1.75	0.70	6.40
22	*	1.75	6.3					
23	*	1.25	6.2					
24	*	1.75	6.1					
25	*	2.25	6.6	7	0.50	2.12	0.50	6.40
26	*	2.25	6.5					
27	*	1.75	6.1					
28	*	2.25	6.5					
29	*	2.50	7.3	8	0.50	2.31	0.80	6.80
30	*	2.25	6.7					
31	*	2.00	6.5					
32	*	2.50	6.5					

*Time not recorded.

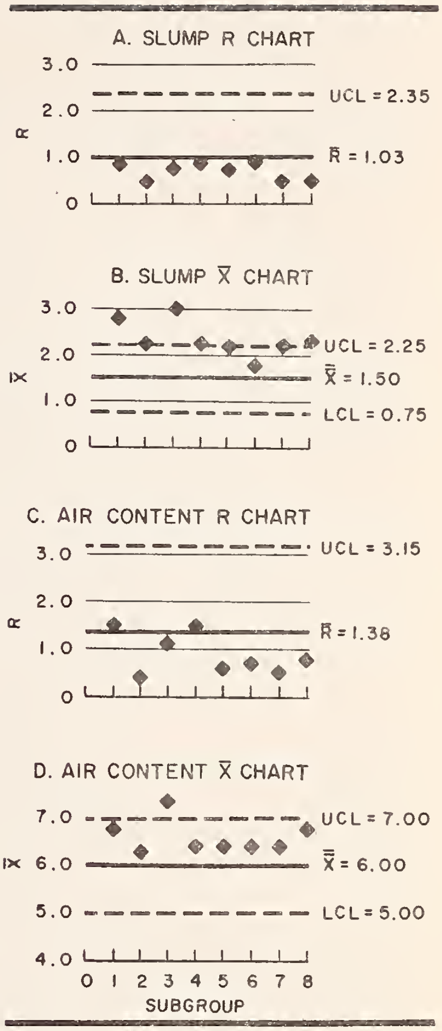


Figure 14. Control Charts plotted from Case Study 2 data.

The range chart for slump shows that variation within subgroups was very low, as indicated by the subgroup ranges never exceeding $\bar{R}$. As in the first case study, this indicates that the σ' of 1/2-in. implied in the current slump specification is easily achievable.

The $\bar{X}$-chart (Fig. 14D) shows that the mean air content of the third subgroup plotted out of control. This led to a search for assignable causes, and two were found. The solenoid controlling the flow of retarder was found to be defective and the dosage of entraining agent was high. The solenoid was replaced and the dosage of entraining agent reduced. With these changes, the sample mean remained in control for the rest of the day.

TABLE 14
CASE STUDY 3 FIELD DATA

Test	Time	Slump, in.	Air Content, percent	Subgroup	Slump		Air Content	
					R	$\bar{X}_4$	R	$\bar{X}_4$
DAY 1								
1	10:25	3.50	6.9	1	1.75	2.94	1.40	6.75
2	10:55	3.75	7.6					
3	11:50	2.00	6.2					
4	12:26	2.50	6.3					
5	2:00	2.25	5.7	2	0.75	2.56	0.60	6.40
6	2:26	3.00	6.7					
7	2:45	2.25	6.1					
8	3:20	2.75	6.1					
DAY 2								
1	7:31	2.75	5.4	3	1.25	2.75	0.80	5.85
2	8:00	2.50	6.2					
3	8:15	3.50	5.6					
4	9:10	2.25	6.2					
5	9:30	2.25	5.7	4	0.75	2.56	0.40	5.83
6	9:50	3.00	6.1					
7	10:02	2.50	5.8					
8	10:56	2.50	5.7					
9	11:10	2.75	6.4	5	0.50	2.81	1.40	5.63
10	11:25	3.00	5.0					
11	11:46	3.00	5.7					
12	12:05	2.50	5.4					
DAY 3								
1	7:39	2.00	5.5	6	1.25	2.69	2.30	6.40
2	8:00	2.75	6.2					
3	8:31	2.75	6.1					
4	8:50	3.25	7.8					
5	9:18	2.50	6.6	7	1.25	2.00	2.10	5.58
6	9:51	1.75	4.5					
7	10:13	1.25	5.0					
8	10:26	2.50	6.2					
9	10:53	1.75	5.6	8	0.75	2.19	1.10	5.75
10	11:28	2.50	6.4					
11	11:43	2.00	5.3					
12	1:05	2.50	5.7					
DAY 4								
1	6:20	2.25	4.9	9	1.25	2.75	1.90	5.80
2	7:05	3.50	6.8					
3	7:25	3.00	6.0					
4	7:45	2.25	5.5					
5	8:17	2.50	6.3	10	1.00	2.56	1.50	5.85
6	9:03	2.00	4.9					
7	9:37	2.75	5.8					
8	9:56	3.00	6.4					
9	10:20	2.00	5.3	11	0.50	2.13	1.50	5.65
10	11:44	2.50	6.7					
11	1:35	2.00	5.4					
12	2:23	2.00	5.2					
DAY 5								
1	8:04	2.50	6.6	12	1.50	3.25	1.20	5.88
2	8:39	2.75	6.1					
3	9:25	4.00	5.4					
4	10:27	3.75	5.4					
DAY 6								
1	8:40	2.75	6.8	13	0.75	2.63	0.70	6.95
2	9:42	2.50	6.6					
3	10:09	2.25	7.1					
4	10:46	3.00	7.3					
5	12:34	3.00	5.8	14	1.25	2.69	1.60	6.05
6	1:20	3.25	6.6					
7	2:10	2.50	6.7					
8	3:04	2.00	5.1					
DAY 7								
1	9:40	2.50	6.0	15	1.00	2.63	1.20	5.45
2	10:36	3.00	5.9					
3	11:57	3.00	5.1					
4	1:15	2.00	4.8					
DAY 8								
1	9:43	2.00	5.6	16	1.50	2.56	1.70	5.75
2	10:55	3.50	6.5					
3	11:34	2.75	6.1					
4	12:10	2.00	4.8					
DAY 9								
1	7:02	3.00	5.3	17	1.50	2.50	1.20	5.95
2	7:20	3.00	6.5					
3	8:09	2.50	6.1					
4	8:45	1.50	5.9					
5	9:10	3.25	6.5	18	1.00	2.63	0.20	6.38
6	10:00	2.25	6.3					
7	10:43	2.50	6.3					
8	11:20	2.50	6.4					
DAY 10								
1	7:23	2.50	8.1	19	1.00	2.75	3.00	6.40
2	8:01	2.25	5.8					
3	8:30	3.00	6.6					
4	9:02	3.25	5.1					
5	9:30	3.00	6.2	20	0.25	2.88	1.30	6.95
6	9:45	2.75	7.5					
7	10:25	3.00	6.8					
8	10:58	2.75	7.3					

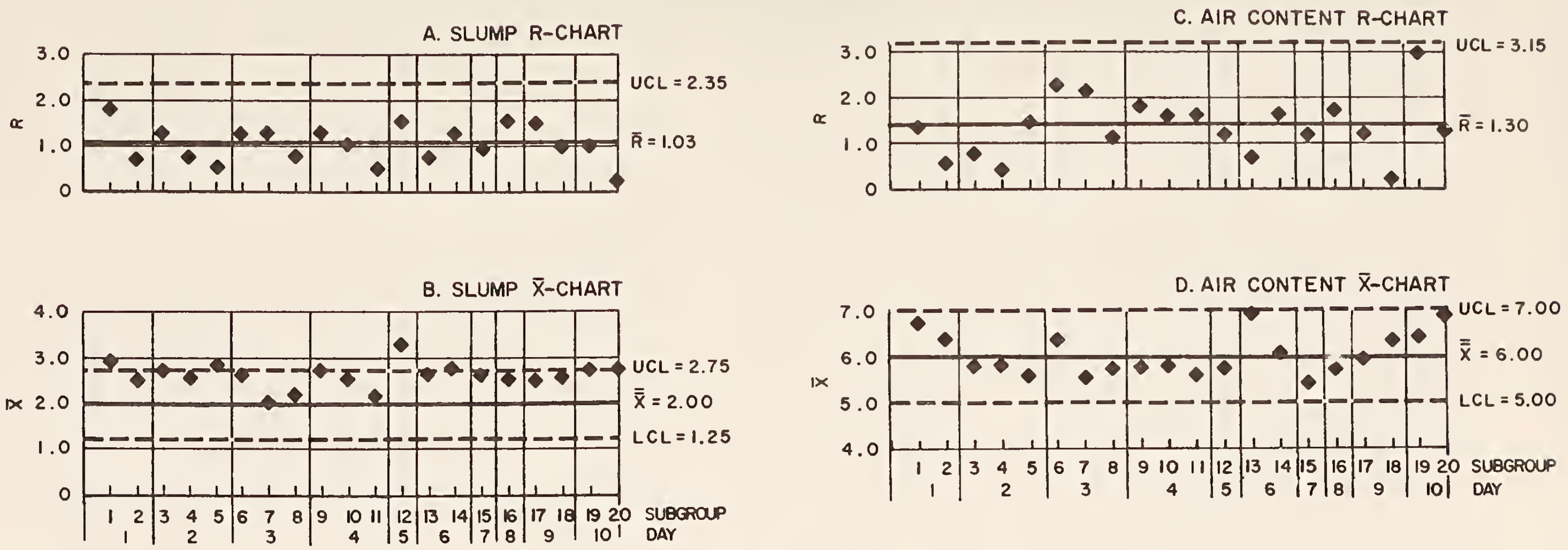


Figure 15. Control charts plotted from Case Study 3 data.

3. Case Study 3

The two preceding case studies indicate that within-group standard deviations allowed for slump and air content are easily achieved values, but the tested subgroups in both cases were recovered within approximately 30 minutes. This required three men -- two did the testing while the third kept the control charts and coordinated work with the contractor's personnel. The coordinator could be eliminated, but sampling four production units in 30 minutes under present sampling conditions would still require two men. With automatic sampling devices this could be reduced to one man, but such sampling devices are not now in use and in sampling men must climb trucks and perform other time-consuming, difficult maneuvers. Since it seems unlikely that concrete producers would use two men for quality control, it was decided to see what would happen if only one man did the testing. With one man doing the testing, the four units in the subgroup would be produced over a longer period, and it is questionable that the subgroups would be rational. Under these circumstances, the *R*-charts might show lack of control, because the range could reflect assignable cause variation. However, if the σ' chosen were larger than the inherent variability, control charts based on subgroups consisting of production units produced over long periods should still show control.

To see what would happen, data recovered more-or-less randomly on a bridge project were grouped in subgroups of four production units tested within 1 to 4 hours (Table 14). Using these subgroups and the specification for deck slabs, *R*- and $\bar{X}$-charts were derived and plotted as shown in Figure 15. The $\bar{X}$-chart for slump showed lack of control four times. But the *R*-chart for slump showed good variability control indicating that the 1/2" standard deviation is still achievable with quasi-random sampling. In Figure 15, both air content charts showed control throughout the 10-day testing period. As in the case of slump, the *R*-chart indicates that the σ' of 0.67 percent implied in the air content specifications for Class A concrete is achievable, even when the range reflects slight amounts of between-group or assignable cause variation.

Thus, the data presented in these three case studies strongly suggest that the within-group standard deviations implied in the current slump and air content specification are reasonable and achievable values. This means that with statistical quality control, these specifications could easily be met.

III. ACCEPTANCE SAMPLING

Data presented earlier in this report showed poor compliance with 1967 specifications for slump and air content. These specifications were revised and case studies undertaken to show that the revised limits could be met with some process control. But data collected since the specification revisions went into effect show that compliance, although improved, is still a problem.

Record sampling data collected during the period from November 1968 through October 1969 indicate average statewide percentage defectives of 6 percent for air content, 5 percent for slump, and 9 percent for compressive strength (if the informal strength specification is applied). Similarly, record sampling data for the period November 1969 through October 1970 indicates percentage defectives of 9 percent for slump, 6 percent for air content, and 4 percent for compressive strength if only Class A concrete is included. Furthermore, record sampling data indicate that slump and air content failures are independent of one another. For example, none of 375 concrete batches tested for the record sampling program during 1969-70 failed for both properties. This means that even if strength failures occurred exclusively in the same batches that fail for slump or air content, 10 to 15 percent of out-of-specification material is being accepted on a statewide average. On some jobs the percentage can be much higher.

This leads one to ask why non-compliance persists when it is known that with some effort toward process control the specification can be economically and easily met. This persisting high degree of non-compliance is believed ultimately to be due to present inspection procedures. Basically, present sampling rates are too low to discover defectives, and through rejection to force producers to pay attention to process control. Under the present inspection system, producers have no incentive to control their plants because the output is infrequently tested and rarely rejected.

Sound, statistically based acceptance sampling plans would correct this situation. But such plans are likely to increase the present testing rates and possibly inspection costs. Judging the worth of the increased inspection rests ultimately with top management. All that researchers can do is to provide facts and alternatives for managers to consider. With this in mind, an evaluation of the present inspection system will be attempted and alternatives suggested for consideration. Accordingly, acceptance sampling plans will be reviewed briefly to determine which are applicable to concrete inspection. Those deemed appropriate will be compared with present inspection schemes. Finally, the benefits and disadvantages of statistical sampling plans will be discussed.

A. The Acceptance Sampling Concept

It is widely recognized that buyers cannot perform quality control, which involves physical manipulation of manufacturing equipment, selection and handling of raw materials, and constant testing during manufacturing. For effective control of quality, buyers would have to replace the manufacturer's personnel with their own crews or assume managerial control of personnel on the supplier's payroll. Neither alternative is desirable and both are likely to provoke legal objections. No manufacturer is likely to let buyers manage his personnel or use his equipment. Faced with this situation, buyers have had to find some means to judge the quality of products manufactured outside their own facilities, or in case of a large company, in another department. The tools devised for this purpose are known as acceptance sampling plans, which are systematic inspection procedures to decide, with known risks, whether to accept or reject the product inspected.

As this definition implies, the objective of acceptance sampling is to determine a course of action and *not* to estimate or control quality. The courses of action include 1) acceptance of the inspected material if it meets sampling plan criteria, or 2) rejection of the inspected material if it does not meet the set criteria. The decision criteria for acceptance sampling plans are based on estimates of quality or probable fractions defective in material accepted. However, it must be emphasized that the purpose of acceptance sampling is not to estimate quality but to make a decision. Quality estimates are only incidental to the objective of acceptance sampling.

It should also be emphasized that acceptance sampling has no direct effect on quality. Controlling quality is the purpose of process control and control charts. Acceptance sampling and control charts have distinctly different objectives. The latter serve as guides to quality control engineers in producing a quality product. The objective of acceptance sampling is simply to decide whether to accept or reject the material inspected. The effects of acceptance on quality are important but indirect. As a result of acceptance sampling, suppliers are faced with two alternatives -- to control the process or to face rejection and loss of customers. Since smaller profits are usually better than no sales, acceptance sampling results in better quality.

B. Information Required in Designing Sampling Plans

To design any acceptance sampling plan it is necessary to know 1) the product's quality determinants, 2) the specification limits, 3) the desired quality levels, and 4) the allowable risks. Quality determinants dictate the choice of sampling plans to be used. The remaining information is necessary to establish the plan's decision criteria and to define its operating characteristic curve.

1. Quality Determinants

Quality determinants are the essential properties that a material or product must possess to satisfy its intended function. Quality determinants, depending on the product's use, can consist of physical properties or other characteristics such as shape, smoothness, and appearance. Identification of essential properties is usually the responsibility of the design engineer, who knows what qualities the material or product must have to satisfy its intended function. This is particularly true for engineering materials that have to stand predictable stresses.

2. Specification Limits

Specification limits are those within which the quality determinants can be allowed to vary without impairing the ability of the product to perform its intended function. Proper setting of specification limits implies that if they are exceeded the product loses its ability to perform its engineering task. For this reason, setting specification limits is strictly an engineering function. It should not be left to quality control engineers or statisticians, who would tend to set limits to include all variability in the product. The quality control engineer's function should be to devise process control schemes to meet the specification limits and not to change them to accommodate variation.

In some instances, specification limits are simple consequences of design. In others, these limits must be based on performance data -- the ultimate engineering judgment. An example of a specification limit as a direct consequence of design is the compressive strength of a steel column that must support a load. Loading condition, load size, and desired safety factor automatically determine the minimum compressive strength required. When specification limits do not automatically result from design, performance data must be used. An example is the air content of concrete. In that case, the only way to decide the optimum range of air content is to observe how concrete with different air contents, used in the same environment, performs over the years.

3. Quality Levels

Usually associated with lot sampling plans are two quality levels -- 1) the Acceptable Quality Level or *AQL*, and 2) the Lot Tolerance Fraction Defective or *LTFD*. In statistical terminology, the *AQL* is defined as the percentage defective for which the probability of acceptance is 1 - α. Similarly, *LTFD* is defined as the percent defective for which the probability of acceptance is β. These definitions are merely translations of corresponding engineering terms into statistical language, and it may be more meaningful to think of these quantities in engineering terms. In engineering language, *AQL* can be defined as the percent defective that can be easily tolerated without impairing performance, and *LTFD* as the maximum percent defective that can be tolerated and still meet the intended engineering requirements. The statistical definitions simply tell the statistician that the engineer wants to accept as many as possible of the lots having a fraction defective equal to or smaller than the *AQL*, and to accept only a small fraction of the lots with percents defective equal or greater than the *LTFD*. For continuous sampling, only one limiting quality level is specified. This is the Average Outgoing Quality Limit or *AOQL* which represents the maximum average percent defective that can be accepted under a continuous sampling plan.

The meanings of these quantities suggest that their choice is primarily an engineering function to be based on available performance data and good engineering judgment. The ideal combination of *AQL* and *LTFD* for lot plans and the *AOQL* for continuous sampling plans can be selected only after knowing the engineering and economic consequences of accepting any percentage of defective material.

4. Producer's and Consumer's Risks

The producer's risk α is the probability of rejecting materials of acceptable quality -- that is, the probability of rejecting products having a percent defective equal to the *AQL*. The consumer's risk β is the probability of accepting rejectable materials, or material containing a percent defective equal to the *LTFD*.

The choice of α and β depends on the consequences of a product's failure to perform its intended function, usually referred to as criticality. If a product's failure results in loss of life or in complete uselessness of the unit in which the product is incorporated, it is critical. In such cases β should approximate zero. When the product's failure causes minor consequences, β can be made large and α small. The proper combination of α and β must be based on engineering as well as economic considerations. For a non-critical product, α and β are usually chosen as 0.05 and 0.10, respectively. There is no particular reason for this choice except that these values have been found to work well in industry, as Dodge has noted (13).

C. Common Sampling Plans

To perform the function of acceptance sampling and testing, the buyer generally has three alternatives: 1) screening, 2) lot sampling, or 3) continuous sampling.

1. Screening

Screening consists of inspecting each identifiable portion of the product, or every item produced. It is very time-consuming and expensive. Because of its cost, screening is performed only when testing is relatively easy or when the quality of the product or items inspected is critical -- that is, when a defective would result in complete uselessness of the product or in loss of life.

2. Lot Sampling

Lot sampling consists of drawing a random sample from a lot and testing only the items in that sample to decide whether the entire lot is acceptable. Because manufactured goods -- such as nuts, bolts, light bulbs, etc. -- can be easily grouped into lots and because their quality does not change between the time of production and inspection, this method is widely used in industry. It has the advantage of requiring the testing of only a *sample* of items from the lot, and the greater advantage of allowing for rectification of lots when the product consists of identifiable, separate items or portions. When lot sampling is used, defective lots can be rectified by removing a portion of the defective items or by adding items of good quality, until the percentage defective in the lot is reduced to an acceptable level. Unfortunately, the advantage of rectification is lost when lot sampling is applied to a product that cannot be separated into identifiable units.

This can be illustrated by an example. Assume that asphalt concrete is being bought by a road builder and that he is testing for asphalt content. Further assume that the test for asphalt content can be made as the product is delivered to the job site, and that for the purpose of acceptance sampling the buyer designates a day's production -- 300 truck loads -- as the lot. After consulting the proper charts, the buyer determines that he must test 30 loads during the day, and that if out of that total 4 are defective, he must reject the lot. He gives his inspector the sampling plan and instructs him to enforce it. During the testing, two undesirable things can happen. The inspector may find the fourth defective truckload at his fifth test and reject the lot. But by that time, only one-sixth of the material designated as the lot may have been produced. Thus, theoretically, the inspection plan rejects five-sixths of the material in the lot even though it has yet to be produced. At the other extreme, the inspector may find the fourth defective to be the last truckload he tests at about the end of the production day. Again he rejects the lot. This time, unless reduced payment proportional to the decrease in lot quality can be arranged, the problem is worse. The whole day's production must be removed from the road surface because the good material cannot be separated from the bad to rectify the lot.

Four variations of lot sampling are generally used -- single-, double-, multiple-, and sequential-lot sampling.

a. Single-Lot Sampling

Single-lot sampling consists of taking one sample (*n* items of production) from a lot and making the acceptance or rejection decision on the basis of that one sample.

b. Double-Lot Sampling

Double-lot sampling consists of taking a smaller sample than needed to make a decision under the single-lot sampling plan. If the number of defectives in the sample is above or below pre-set limits, the lot is rejected or accepted. If the number of defectives lies within the preset limits, another random sample is drawn from the lot and combined with the first to form a single but larger sample, upon which the decision to accept or reject is then made. This plan has the advantage of reduced sampling in those cases where the decision about the lot can be made after testing the first sample.

c. Multiple-Lot Sampling

Multiple sampling plans work the same way as double sampling plans, except that more than two samples are drawn from the lot until the cumulative sample is large enough to make a decision. Here again the advantage is reduced sampling when the material tested contains more or less defectives than allowed and a decision can be made after the first or second sampling. If the number of defectives in the lot approximates the allowable limit, there may be no saving because of the large cumulative sample required.

d. Sequential-Lot Sampling

Sequential sampling differs from other lot plans in that production units are randomly inspected one-at-a-time until enough information is available to make the decision to accept or reject. With this plan, sampling could go on indefinitely if the number of defectives in the cumulative sample stays within the pre-set limits. Sequential sampling has the advantage of possible reduced inspection but has the same disadvantages of other lot sampling plans.

Lot sampling can be by variables or by attributes. Attribute sampling plans specify a sample size and the number of permissible defective units. The title "attribute" comes from the fact that the value of the production units inpsected are classified as having the attribute of being either defective or non-defective.

Sampling by variables, on the other hand, consists of 1) measuring one or more material properties on a continuous scale, 2) determining the distribution of each, 3) estimating the portion of material defective from the parameters computed for the distribution of each property, and 4) then judging the quality of the lot according to the estimated proportion defective. The principal advantage of this plan is the small sample size usually required for decision. Its major disadvantages include 1) having to devise a separate sampling plan for each property of the material tested, and 2) requiring knowledge of the distributions of the quality characteristics tested -- which may not be known and in some cases may not be easy to characterize. Sampling by variable, then, is most appropriate when the units in the lot have to be tested for only one property, and the distribution of this property is known.

3. Continuous Sampling

Continuous sampling consists of screening and random sampling used in combination and is strictly by attributes. There are many variations of continuous sampling plans (14). Some continuous sampling plans start with screening until a predetermined number of production units i are found free of defects and then revert to random sampling of a prechosen fraction f of production units until one defective is found. When one defective is found under random sampling screening is resumed and the cycle started again. Other continuous sampling plans begin with randomly testing a fraction f of production units and revert to screening when a defect is found under random sampling. The fraction of production units to be inspected randomly is kept constant for some plans, while varying for others. If the fraction inspected randomly is kept constant, the plans are said to be single-level continuous sampling plans.

If it is varied, the plans are said to be multilevel continuous sampling plans.

Of the continuous sampling plans that begin with screening, Dodge's CSP-1 is the most widely known (15). This is a single-level continuous sampling plan. It begins with screening (100-percent inspection) at the start of production until a specified number of consecutive production units is found to be free of defects. At this point, screening is discontinued and only a predetermined fraction of production units is inspected randomly. When a defective is found, screening is resumed and the cycle started again. The proportion of production units to be tested depends on the proportion of defective material that can be tolerated. To eliminate all defectives under this plan, 100-percent inspection is required. As the proportion of material inspected decreases, the maximum proportion of defective material that could be accepted increases.

Dodge's CSP-1 has led to many other variations of continuous sampling plans. Dodge and Torrey (16) have proposed continuous sampling plans under which screening is not resumed until two defective units are found within a space of k sample units. Lieberman and Solomon (17), following the general pattern of CPS-1, proposed the so-called multilevel continuous sampling plans which require less sampling than CSP-1 to guarantee the same degree of quality. These begin with screening and revert to random sampling of a fraction f as soon as i consecutive units free of defects are found, as in Dodge's CSP-1. However, when i consecutive units have been found free of defects, while sampling randomly one production unit for every $1/f$ number of production units, the sampling frequency is decreased to one unit for each $1/f^2$ production units (note that because $f < 1.0$, $1/f^2 > 1/f$). If i consecutive production units are found free of defects while sampling one unit in $1/f^2$ production units, the sampling frequency is reduced again to one unit for each $1/f^3$ production units, and so on. Theoretically, the sampling rate (inspection level) can be changed indefinitely, but in practice too many changes in inspection levels are likely to cause administrative problems. Wald and Wolfowitz introduced plans starting with random sampling, which were later modified by Girshick (18). His continuous sampling plans start with *lot* sequential sampling. Sequential sampling of lots of size N continues until the rejection line is crossed on the sequential sampling mask. When this occurs, a new lot is started and screening is initiated immediately. Screening is continued for all subsequent lots until one lot of N units is found free of defects. At this point, sequential lot sampling is resumed and the cycles started again. When the production process is under statistical control, this type of continuous sampling plan is likely to result in less inspection to assure the same quality as those which begin with screening. However, if the process is not under statistical control, these plans have no advantages over those that begin with screening.

4. Special Sampling Schemes

Two sampling schemes are often described in the literature -- lot-plot inspection plans (8, pp. 436-442) and chain sampling plans (3, pp. 13-96 through 13-97). Both apply to special situations. Neither applies to concrete and they need not be discussed here.

D. Characterization of Sampling Plans

All the sampling plans just discussed can be characterized by two curves -- the *OC* or operating characteristic curve, and the *AOQ* or average outgoing quality curve. In addition, complete description of multiple and sequential sampling plans requires knowledge of their *ASN* or average sample number curves. Similarly, complete characterization of continuous sampling plans requires knowing their *AFI* or average fraction inspected curves. Because of their importance, these curves are reviewed here, but *AOQ* curves are discussed only for continuous sampling. These latter curves have meaning only if rectification is possible for the product inspected; since concrete cannot be rectified when inspected under other than continuous sampling plans, *AOQ* curves would be meaningless. Thus, they are discussed only in terms of continuous sampling. For lot plans, most of the emphasis will be placed on *OC* curves, which are the most important.

OC curves of lot sampling plans and screening give the probability of accepting the lot as a function of lot fraction defective, and are the best indication available of the efficiency of lot sampling plans, and are essential in order to estimate the average quality of products accepted under any sampling plan. Since it is very unlikely that many concrete lots come from processes that are in complete control at all times and at the same level, the *OC* curves discussed are Type A -- that is, they give the approximate probability that a single lot of a given quality will be accepted or rejected under the sampling plan. (For discussion of the difference between Types A and B *OC* curves, see 8, p. 147-151.)

1. *OC* Curves for Screening

Assuming no testing errors, the *OC* curves for screening are vertical lines (Fig. 16). In screening, each item in the lot is inspected, and assuming no errors in testing, the true fraction defective in the lot is known. Knowing the true fraction defective, the lot is accepted or rejected, depending on whether the fraction defective exceeds that allowable. Thus, the probability of accepting the lot is either 0 or 1, as shown by the vertical lines in Figure 16.

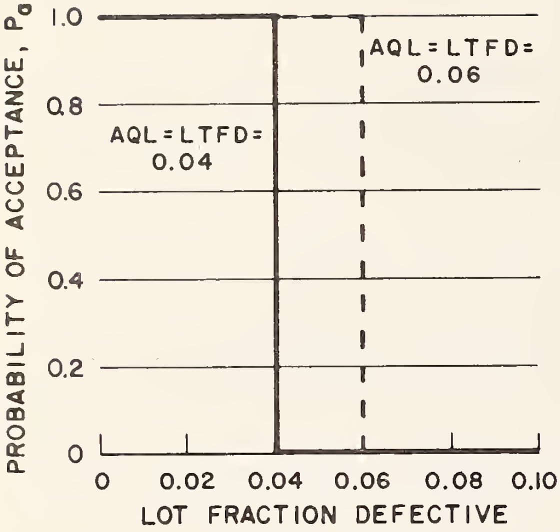


Figure 16. Screening operating characteristic curves for two allowable percents defective, P_1 and P_2.

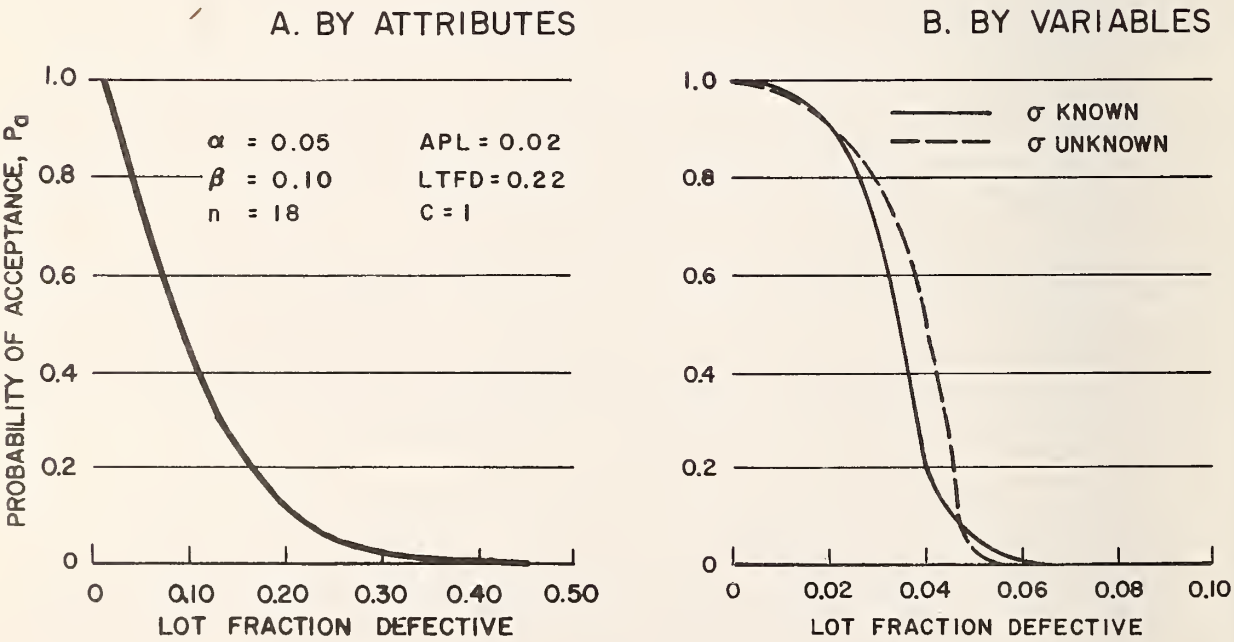


Figure 17. Operating characteristic curves for single-lot plans.

2. *OC* Curves for Single-Lot Plans

a. By Attributes

OC curves for single-lot plans by attributes give the probability of accepting the lot as a function of the lot fraction defective (Fig. 17A). To compute the ordinate P_a, the following quantities must be known: 1) producer's risk α, 2) consumer's risk β, 3) acceptable quality level *AQL*, 4) rejectable quality level, *LTFD*, and 5) lot size *N*.

Knowledge of *N*, β, *AQL*, and *LTFD* permits determination of sample size *n* and the acceptance number *c*, or the number of allowable defective units in the sample. Both *n* and *c* are usually obtained by use of charts or tables, rather than by actual computations. [The reader can find the theory in Dodge and Romig (19).] Knowing *n* and *c*, the probabilities of acceptance can be computed as the hypergeometric or binomial probabilities that a sample of size *n*, drawn from a lot having a given fraction defective, contains *c* or less defectives. Hypergeometric probabilities are used when the lot size cannot be considered infinite as compared to the sample size. If the lot can be considered infinite (in practice, when lot size is ten times greater than sample size), the probabilities of acceptance can be computed using binomial probabilities.

b. By Variables

These plans are also characterized by *OC* curves, which give the same information for single-lot sampling by variables as those for single-lot plans by attributes. However, for these curves, the probability of acceptance is obtained by computing the probability that the mean of a sample taken from a lot falls within a pre-established interval, assuming that the variable tested is normally distributed. To compute the probability of acceptance for these *OC* curves, α, β, *AQL*, *LTFD*, and *n* must be known, just as in the case of single-lot plans by attributes. The probabilities of acceptance can be computed, assuming a known standard deviation or assuming that the standard deviation is unknown. For the same conditions and sample size, the *OC* curve computed assuming known standard deviation is steeper or more discriminating, as shown in Figure 17B.

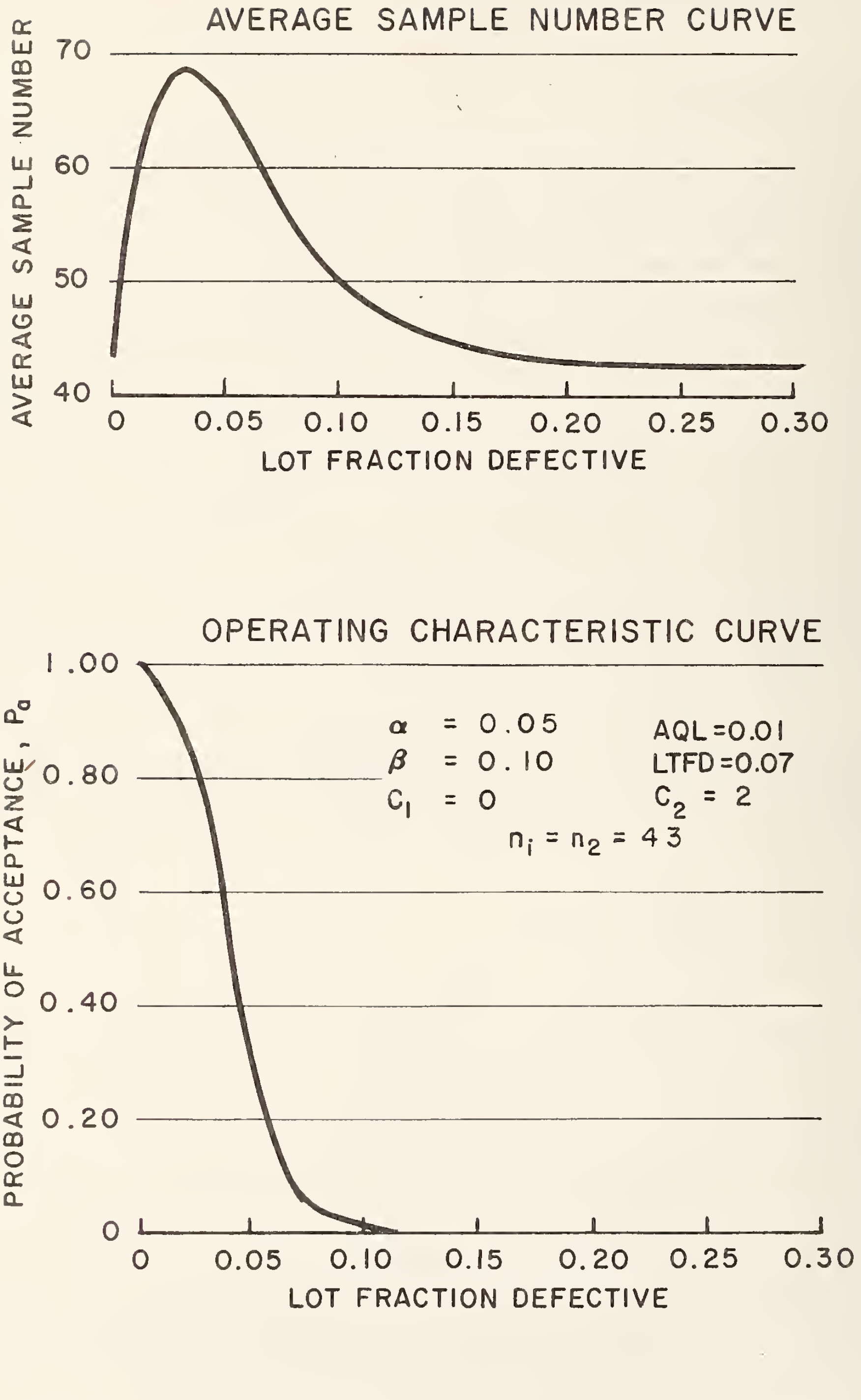


Figure 18. Characteristics of double sampling plan; the *ASN* curve is for uncurtailed inspection.

3. *OC* and *ASN* Curves for Multiple Sampling Plans

a. By Attributes

In contrast to single sampling plans by attributes, multiple sampling plans by attributes are characterized by two curves. One is the *OC* curve, which gives the same information as in the case of single sampling plans. The other is the *ASN* curve which gives the average number of production units to be inspected in order to reach a decision for each lot fraction defective. *OC* curves for multiple sampling plans by attributes have the same shape as those for single sampling plans, and their derivation requires the same information plus the number and sizes of samples to be recovered. However, the probability of acceptance for each lot fraction defective cannot be as easily computed. In computing the probabilities of acceptance for a multiple sampling plan, the probabilities of accepting, rejecting, or continuing to sample the lot after inspection of each sample must be taken into account and the computations become more involved. Computer Programs (20) or computation work sheets are usually used for this purpose.

The average sample numbers corresponding to each lot fraction defective are computed by making use of the probabilities of a decision, or the sum of the probabilities of acceptance and rejection after the inspection of each sample (8, pp. 182-186). Plotting the probabilities of acceptance and average sample numbers versus the lot fractions defective, the two curves that characterize a multiple sampling plan having given α, β, *AQL*, *LTFD*, *N*, and a specified number of samples k of a given size n, can be obtained as shown in Figure 18.

b. By Variables

Multiple sampling plans by variables are also characterized by *OC* and *ASN* curves. These have the same meaning as the corresponding curves for multiple lot plans by attributes. *OC* curves for these plans, however, are computed using probabilities tabulated for the normal distribution instead of binomial probabilities.

4. *OC* and *ASN* Curves for Sequential Sampling Plans

These plans are a special case of multiple sampling, for which the sample size is 1 and the number of samples is indeterminate. As multiple sampling plans by attributes, they are characterized by an *OC* curve and an *ASN* curve. These have the same gen-

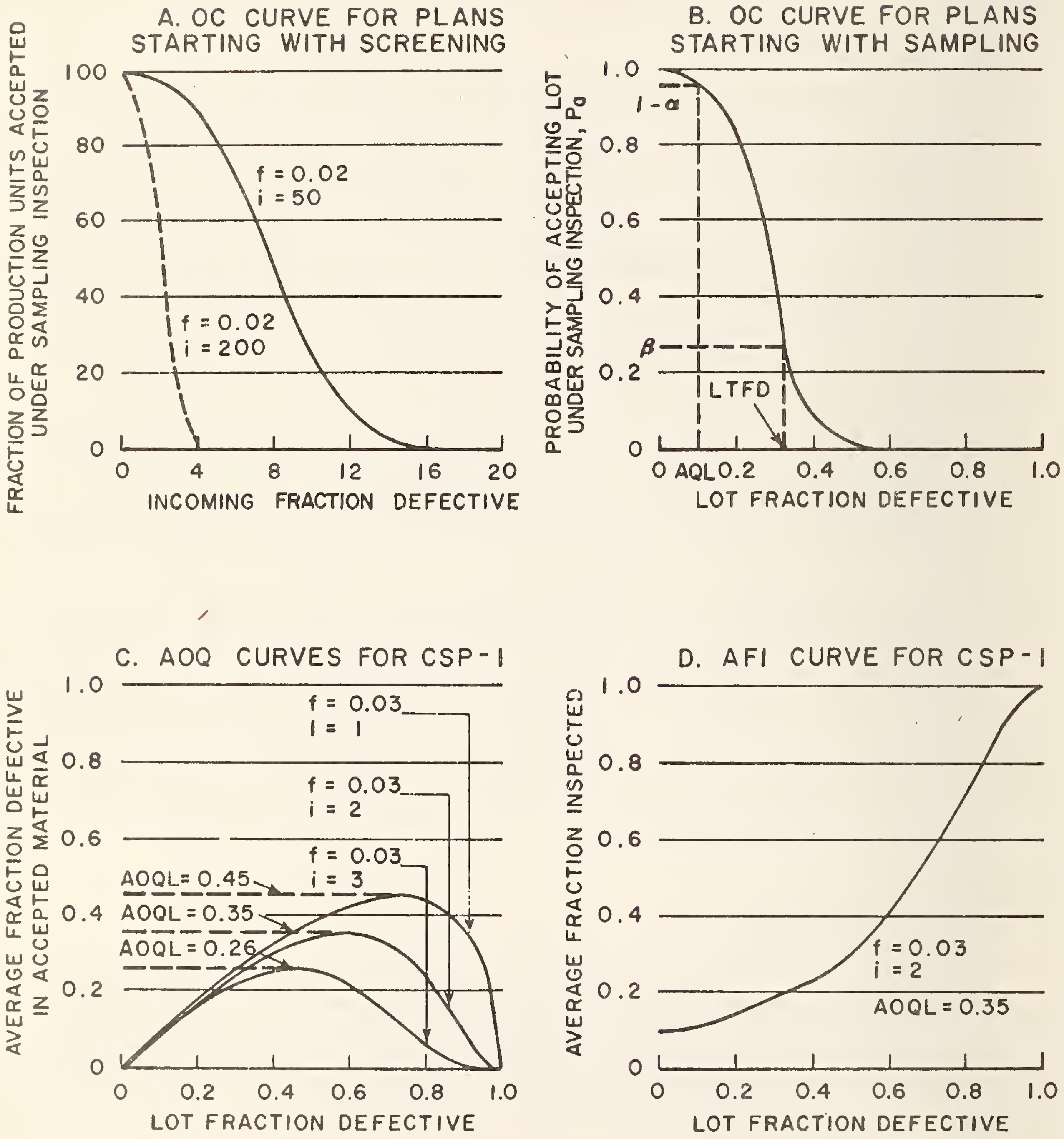


Figure 19. Curves characterizing continuous sampling plans.

eral shape as those for corresponding multiple lot plans and give the same information.

5. Curves Characterizing Continuous Sampling Plans

Three curves are usually employed to characterize continuous sampling plans -- the *OC* curve, the *AFI* or average fraction inspected curve, and the *AOQ* or average outgoing quality curve.

a. *OC* Curves

In the case of single and multi-level sampling plans that begin with screening, such as the Dodge type (15), the *OC* curve gives the proportion of production units accepted under sampling inspection versus the process fraction defective, as shown in Figure 19A. The shape of these *OC* curves depends on the fraction f of production units inspected during the sampling phase of continuous sampling, and on the number of production units i that must be found free of defects to switch from screening to sampling. These curves become more discriminating as f and i increase.

For continuous sampling plans that begin with sampling rather than screening, and for which the production output to be inspected is subdivided into lots, the *OC* curve gives the probability of accepting a lot under sampling inspection as a function of the lot fraction defective. The Girshisk plan (18) is such a plan. These *OC* curves depend on α, β, *AQL*, and *LTFD*, just as those for single, multiple, and sequential sampling by attributes. An example of such an *OC* curve is shown in Figure 19B.

b. *AOQ* Curves

AOQ curves for continuous sampling plans give the average fraction defective in the material accepted under both screening and sampling as a function of the incoming or lot fraction defective, as shown in Figure 19C. These curves are dependent on both f and i. As i and f increase, the average outgoing quality or average fraction defective accepted decreases, as shown in Figure 19C. Each of these curves has a maximum, usually known as the *AOQL* or average outgoing quality limit. It represents the maximum percent defective that could be accepted under a given sampling plan, assuming no error in testing.

The reader is reminded that these curves can be derived for all lot sampling plans. *AOQ* curves for lot sampling plans are used when overall average quality is more important than the quality of single production lots and when rectification is possible. For concrete, the quality of a single lot is important, and if inspected under lot sampling, concrete does not lend itself to rectification. However, under continuous sampling, rectification occurs during production while under the screening phase, and the *AOQ* curve is much more meaningful. Under continuous sampling, the fractions defective accepted can actually be limited to those indicated by the *AOQ* curves. For lot sampling, the average outgoing quality or average fraction defective in the material accepted can be computed, but cannot be attained because rectification is impossible for concrete after all the lot is inspected. For that reason, *AOQ* curves for lot sampling are not covered here.

c. *AFI* Curves

Another important curve associated with continuous sampling plans is the average fraction inspected curve. This gives the fraction of the total production that must be tested on the average to assure that the *AOQL* is not exceeded. This curve is very important when the cost of testing must be balanced against desired quality levels. As can be seen in Figure 19D, the average fraction inspected increases as the fraction of defective production units submitted for inspection increases.

IV. CONCRETE ACCEPTANCE SAMPLING

A. Present Sampling Schemes

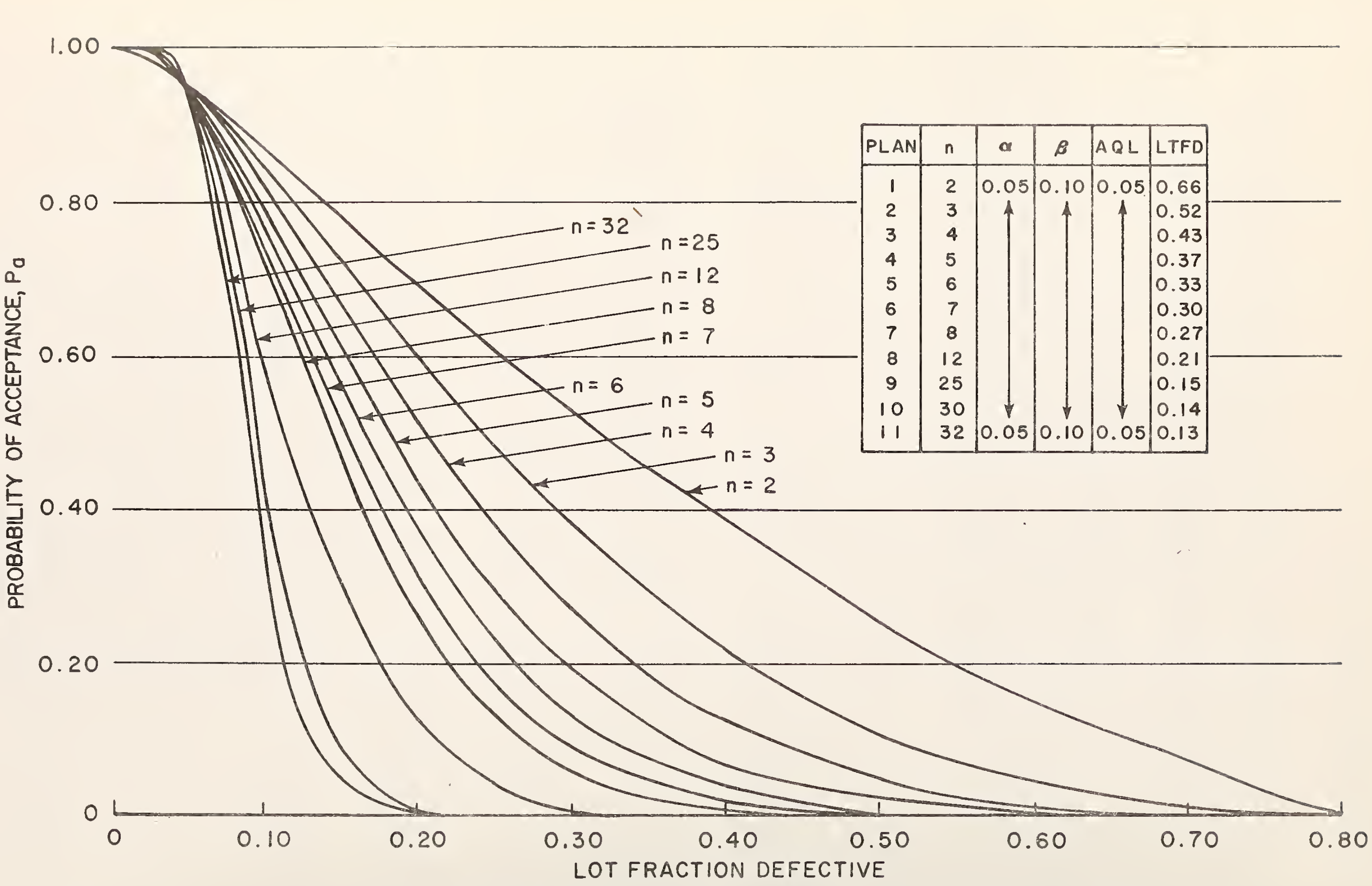
Having reviewed the common types of statistical sampling schemes, it is now appropriate to examine currently used sampling procedures to determine whether these schemes can be considered sampling plans, and what degree of protection against out-of-specification materials they afford concrete buyers. To do so, the Statistical Quality Control Task Force of the Federal Highway Administration was asked for information on concrete sampling rates used by the various states. These sampling frequencies were expressed as numbers of tests per specified number of cubic yards, square yards, or linear feet of pavement, as well as by number of tests per hour. To reduce these to a common basis, a production rate of 2500 cu yd per day was assumed and the sampling rates reduced to numbers of production units tested daily.

Review of these data revealed 1) that in most cases very few concrete samples were tested daily, 2) that testing rates changed for each property, and 3) that higher sampling rates were specified for air content. These variable sampling rates made it difficult to estimate the total degree of protection that each of these sampling schemes afforded, but it was generally low. To inspect for air content -- the concrete property most often tested -- 17 of 21 states for which information was available tested approximately 8 production units or less daily, as shown in Table 15.

To show what protection the Table 15 sampling rates would afford, *OC* curves must be constructed. Such curves require knowing the consumer's and producer's risks, and also the *AQL*, *LTFD*, or both. This presents some problems. A reasonable way to choose the *AQL* is to assume it equal to the statewide fraction defective, which represents what is accepted in practice, and approaches what in reality is considered an acceptable quality level. This average for air content is approximately 5 percent. Using an *AQL* of 5 percent and choosing consumer's and producer's risks of 0.05 and 0.10, respectively, the sample sizes in Table 15 would result in the *OC* curves shown in Figure 20, if lot sampling by variables with standard deviation known were used. The lot would be a day's production, because these are daily sampling rates.

TABLE 15
AIR CONTENT SAMPLING RATES

Approximate Tests per Day	Number Of States
2	5
3	2
4	5
5	1
6	1
7	1
8	2
12	1
25	1
30	1
32	1



PLAN	n	α	β	AQL	LTFD
1	2	0.05	0.10	0.05	0.66
2	3				0.52
3	4				0.43
4	5				0.37
5	6				0.33
6	7				0.30
7	8				0.27
8	12				0.21
9	25				0.15
10	30				0.14
11	32	0.05	0.10	0.05	0.13

Figure 20. *OC* curves for sampling plans by variables corresponding to the sampling frequencies given in Table 15.

These curves show that unless the sample size exceeds 25, the plans are not very discriminating (see App. A glossary). For example, for a sample size of 6, a lot having a fraction defective of 16 percent would have a 50-percent probability of acceptance. Even a sample size of 32 -- the highest in the table -- would result in a rather weak plan. For that plan, a lot containing 9-percent defective would have a probability of acceptance of 50 percent.

If lot sampling by attributes were assumed, the *OC* curves would be still less discriminating. The same would be true if one assumed lot sampling by variables with standard deviation unknown. Hence, under lot sampling, the concrete quality guaranteed by the sample sizes listed in the table, in most cases, is low.

Lot sampling was assumed to illustrate the degree of quality that could be guaranteed with lot sampling, but may be inappropriate in most cases. Most sampling procedures tacitly recognize the impossibility of rectification after the concrete is in place, and their wording often suggests continuous sampling of the CSP-1 type discussed earlier. New York, for example, has a sampling procedure that suggests continuous sampling of this type. The state's Materials Method 9.2-3, referring to acceptance sampling of structural concrete for slump and air content, states:

> One (1) set of tests shall be made from each placement regardless of size and thereafter at the rate of one per 50 cubic yards for the duration of the placement. This rate shall be increased by the engineer whenever indications of unacceptable concrete are noticed.

Assuming that concrete is delivered to structures in 7 cu yd batches, 14 percent of the production units would have to be routinely sampled for slump and air content to comply with this materials method (1 truck/7 trucks = 14 percent of production). Inferring that "this rate shall be increased by the engineer whenever indications of unacceptable concrete are noticed" means that after a failing test, all subsequent production units are tested until one acceptable result is obtained, this scheme would amount to a continuous sampling plan with $f = 0.14$ and $i = 1$. The average outgoing quality curve of this sampling plan is shown in Figure 21. As can be seen, the *AOQ* limit for this plan is 46.5 percent. This means that if the incoming fraction defective were 0.78, 46.5 percent of the concrete accepted under this plan would be defective (or out of specifications). If the incoming fraction defective were the expected 0.15, the fraction of defective material incorporated into the structure would be 12.5 percent of the material accepted.

Acceptance sampling for air content and slump of pavement concrete is even weaker. The same Materials Method 9.2-3, when referring to the testing frequency of pavement concrete for those properties, states:

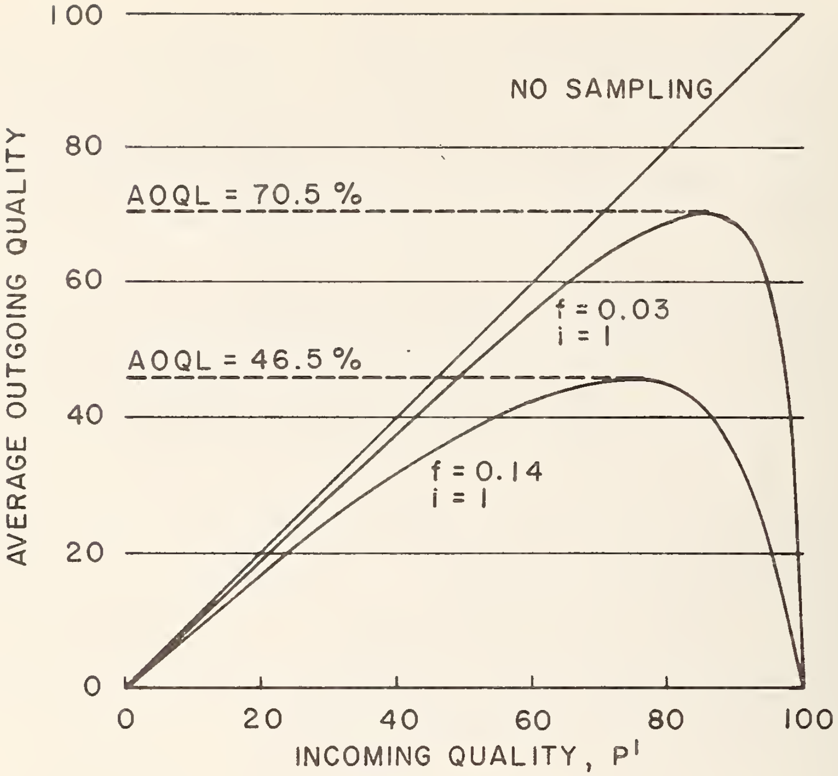


Figure 21. *AOQ* curves for the continuous sampling plans implied in New York State Materials Method 9.2-3.

> One (1) set of tests shall be made from the initial daily placement and thereafter at a rate of one set per 150 to 200 cubic yards of concrete placed. This rate shall be increased by the engineer whenever indications of non-acceptable concrete are noticed.

Assuming that pavement concrete is being produced by central mixers and delivered in 6 cu yd batches (probably an underestimate, but if batch size were assumed larger, the resulting sampling would not be much better), and that one test is being performed each 200 cu yd, the required sampling rate would normally be 3 percent. Again assuming that after a failing test, screening is performed until one acceptable production unit is found, the implication would be a continuous sampling plan having $f = 0.03$ and $i = 1$, with the *AOQ* curve shown in Figure 21. It is evident that this plan affords little protection when the incoming percentage defective is 20 percent or less. The *AOQ* limit for this plan is 70.5 percent, as compared with 46.5 percent for structural concrete.

The frequency of strength testing is specified in Materials Method 9.2-3, but with no provision for rejection of concrete having inadequate strength. Thus, strictly speaking, no acceptance sampling plan exists for this property.

Statistically, present sampling plans are weak. The *OC* curves in Figure 20 and *AOQ* curves in Figure 21 show that regardless of the type of sampling assumed, current testing frequencies are generally too low to guarantee a high degree of quality. However, the percentage of defective concrete being accepted in New York State is less than 15. This is much lower than the percentage that present sampling plans would accept if the incoming percentage defectives were higher. This relatively low percentage of concrete out of specifications reaching the final location is made possible by other safeguards rather than the discriminating power of the present sampling plans. Some major factors believed responsible for this low percentage are 1) the feedback provided by the inspector after discovery of concrete batches out of specifications, and 2) the inherent capability of the production process. In fact, because most production processes are capable of producing concrete having less than 20-percent defective, present sampling procedures lead concrete buyers to accept almost all concrete produced. It is only when the percent defective goes above 20 percent that these plans begin to become effective. This is especially evident for the continuous sampling plans in Figure 21, when the *AOQ* curves are compared with the straight line that represents no sampling at all. Moreover, present sampling procedures may be even less efficient than shown in Figures 20 and 21, because the sampling may not be random.

The preceding examination of current sampling trends leads one to conclude that if concrete containing small fractions defective is desired, testing frequencies must be increased and sampling must be systematized under statistical sampling plans.

B. Plans for Inspection of Plastic Concrete

1. Applicable Plans

The proper choice of acceptance sampling plan depends on the nature of the product, criticality of the characteristic measured, time-dependence of the inspected properties, and cost. For concrete, these are all important considerations and greatly limit the choice among plans that can be used.

a. Screening

Screening, or testing every production unit, is generally very expensive and appropriate only when the product is critical. Concrete's criticality changes with its uses, and structural concrete can be classified as critical. A structure's criticality may be judged by two standards: 1) the possibility of loss of life, and 2) the cost and convenience of maintenance. Piers, slabs, and other structural components qualify on both counts. Safety factors in general use are likely to eliminate the possibility of loss of life, but the cost and inconvenience of repairing piers and scaled structural slabs still warrant inspection of each concrete load placed for slump and air content. This is feasible because structural pours usually involve small numbers of production units, and screening for these properties can be accomplished by one inspector. However, screening for strength, using present inspection procedures, has no real advantage unless it is decided in advance that construction will proceed in stages, and that units constructed in each stage are to be subject to possible removal for at least 28 days. Screening for strength and stage construction become practical only if accelerated testing methods are employed for strength. With some of the proposed accelerated strength tests, the decision to accept or reject each construction stage can be made after 1 to 3 days. Since forms in most cases are kept in place for a couple of days, keeping each structural part available for removal would not interfere with production.

b. Lot Sampling

When production rates are very high, as in the case of concrete pavements, screening is very expensive. With present testing methods, about 20 minutes are required for proper sampling and testing of a production unit of plastic concrete. This limits the number that one inspector can sample and test daily to about 15. At that testing rate, 20 inspectors would be needed to screen 300 production units. Considering that 300 units a day may be a low placement rate for a good paving train, and that concrete testing cannot be delayed, the possibility of screening becomes remote. Where pavers are still used, screening becomes virtually impossible because a paver produces many small production units. In such cases, the only way to assure reasonable quality levels is to employ either lot or continuous sampling.

Unfortunately, lot sampling requires that the inspected concrete be lotted and each lot judged as a whole. This means that a lot of concrete can be judged only after part has already hardened in place. Since hardened concrete cannot be separated into identifiable acceptable and defective portions, rectification of the lot is impossible.

The impossibility of rectification has serious implications. In most cases when lots are rejected, the manufacturer can remove defective units from rejected lots and replace them with an equal number of acceptable units. Thus, economic loss is limited to the combined cost of the defective material discarded and the cost of testing. For concrete, identification and separation of acceptable and defective concrete -- or rectification -- are impossible. Consequently, enforcement of the rejection decision means removal and disposal of all concrete in rejected lots. This would be economically disastrous, and unless an alternative to concrete removal can be found to enforce the contractual agreement, lot sampling is undesirable.

A reasonable way to enforce the contractual agreement is reduced payment for reduced quality. This in fact is the only practical alternative to enforcing the contractual terms regarding quality. This is because strength cannot be measured before the concrete hardens, while production units are still identifiable and separable. Consequently, whenever strength is specified as a quality determinant, enforcement of specifications means either removal of all concrete or reduced payment. Removal is such a drastic measure that it would seldom be enforced, leaving reduced payment as the only practical means of enforcing the contractual terms regarding quality.

Having adopted the idea of reduced payment for reduced quality, some variations of lot sampling may be considered. Concrete properties are measured on a continuous scale, which means that lot sampling both by variables and by attributes may be used. However, not all variations of lot sampling are possible for concrete. In fact, those requiring the smallest sample sizes cannot be used due to the nature of concrete. The variation requiring least sampling is the sequential. This requires that production units be randomly selected and tested in the order drawn, which makes sequential sampling impossible. Having to test production units in the random order in which they are drawn leads to two undesirable situations -- either the production unit chosen for testing is yet to be produced, or it has already hardened in place, precluding testing with present methods.

Next to sequential sampling, multiple sampling requires the smallest sample sizes, but it also is impossible with present testing methods. As has been discussed, multiple sampling plans require that additional samples be taken when the decision to accept or reject the lot cannot be made on the basis of the first sample. Because concrete hardens, no additional samples can be tested after the last unit in the first sample has been tested. Thus, multiple-lot sampling plans cannot be used with standard

TABLE 16
RESULTS OF CHI-SQUARE (χ^2) AND KOLMOGOROV-SMIRNOV (K-S)
GOODNESS-OF-FIT TESTS FOR NORMALITY
Hypotheses Tested at 0.05 Level of Significance

	χ^2 Test					K-S Test			
Job	Deg. of Freedom*	Calc. Value	Critical Value	Accept H_0:**	Reject H_0:	Calc. Value	Critical Value	Accept H_0:**	Reject H_0:
A. SLUMP									
1	4	8.63	9.49	✓		0.07	0.19	✓	
2	3	5.52	7.82	✓		0.08	0.19	✓	
3	3	5.45	7.82	✓		0.08	0.19	✓	
4	3	4.21	7.82	✓		0.10	0.19	✓	
5	4	12.28	9.49		✓	0.12	0.19	✓	
6	4	11.18	9.49		✓	0.09	0.19	✓	
7	3	4.71	7.82	✓		0.10	0.19	✓	
8	3	1.84	7.82	✓		0.05	0.19	✓	
9	4	3.55	9.49	✓		0.06	0.19	✓	
10	3	14.44	7.82		✓	0.12	0.19	✓	
B. AIR PRESSURE									
1	4	9.99	9.49		✓	0.13	0.19	✓	
2	4	5.79	9.49	✓		0.06	0.19	✓	
3	4	6.41	9.49	✓		0.07	0.20	✓	
4	3	6.36	7.82	✓		0.07	0.19	✓	
5	4	3.90	9.49	✓		0.06	0.19	✓	
6	4	6.59	9.49	✓		0.08	0.19	✓	
7	4	3.72	9.49	✓		0.05	0.19	✓	
8	3	12.06	7.82		✓	0.12	0.19	✓	
9	4	10.57	9.49		✓	0.11	0.19	✓	
10	3	5.12	7.82	✓		0.11	0.19	✓	
C. AIR (CHACE)									
1	4	2.07	9.49	✓		0.05	0.19	✓	
2	4	3.82	9.49	✓		0.06	0.19	✓	
3	3	0.44	7.82	✓		0.04	0.20	✓	
4	3	4.77	7.82	✓		0.09	0.19	✓	
5	4	5.13	9.49	✓		0.07	0.19	✓	
6	3	3.73	7.82	✓		0.07	0.19	✓	
7	4	8.96	9.49	✓		0.09	0.19	✓	
8	3	6.60	7.82	✓		0.10	0.19	✓	
9	4	5.62	9.49	✓		0.11	0.19	✓	
10	3	3.30	7.82	✓		0.07	0.19	✓	

*Degrees of freedom for χ^2 test is the number of subgroups minus 3.
**H_0: sample distribution approximates a normal distribution.

testing methods. This narrows the list of lot sampling plans applicable to fresh concrete to single-lot sampling by variables or attributes.

(1) Single-Lot Sampling by Variables

This variation of sampling, in principle, is applicable to concrete inspection. It requires a sampling plan for each quality determinant, knowledge of the frequency distribution of each property tested, and determination of whether variations can be assumed known.

For plastic concrete, the properties to be inspected can be assumed to be normally distributed. The χ^2 and Kolmogorov-Smirnov goodness-of-fit tests, performed on the experimental data presented in Chapter I, are given in Table 16. The results show that for slump and air content, the assumption of normality could not be rejected in most cases when tested at the 5-percent level of significance. This is certainly not conclusive proof of normality because for effective testing of normality with the χ^2 test -- the more conclusive of the two -- one should use 200 points or more (21). The data sets in Table 16 contain about 50 values each, but the facts that data were obtained from processes not in control, and that these tests did not reject the hypothesis of normality, indicate that the distributions do not markedly deviate from normality. Cylinder strengths were not checked for normality due to lack of data, but the literature (22,23,24) indicates that the assumption of normality for this property is justified.

The data analyzed in Chapter I provide a measure of the variability of concrete properties. The variances in Table 2 were tested for homogeneity and found to be non-homogeneous. This indicates that standard deviations of concrete properties change from project to project. Consequently, the more appropriate version of single-lot sampling by variables for concrete is with standard deviation unknown.

(2) Single-Lot Sampling by Attributes

Assuming adoption of reduced payments, single-lot sampling by attributes is also adaptable to concrete. It requires a larger sample size to assure the same degree of quality as a corresponding plan by variables, but has the administrative advantage of requiring one sampling plan for all properties of the same degree of criticality.

c. Continuous Sampling

(1) Plans Beginning with Sampling

As discussed, continuous sampling plans can start with screening or with sampling. The latter have an advantage of reduced sampling when the production process whose output is being inspected is

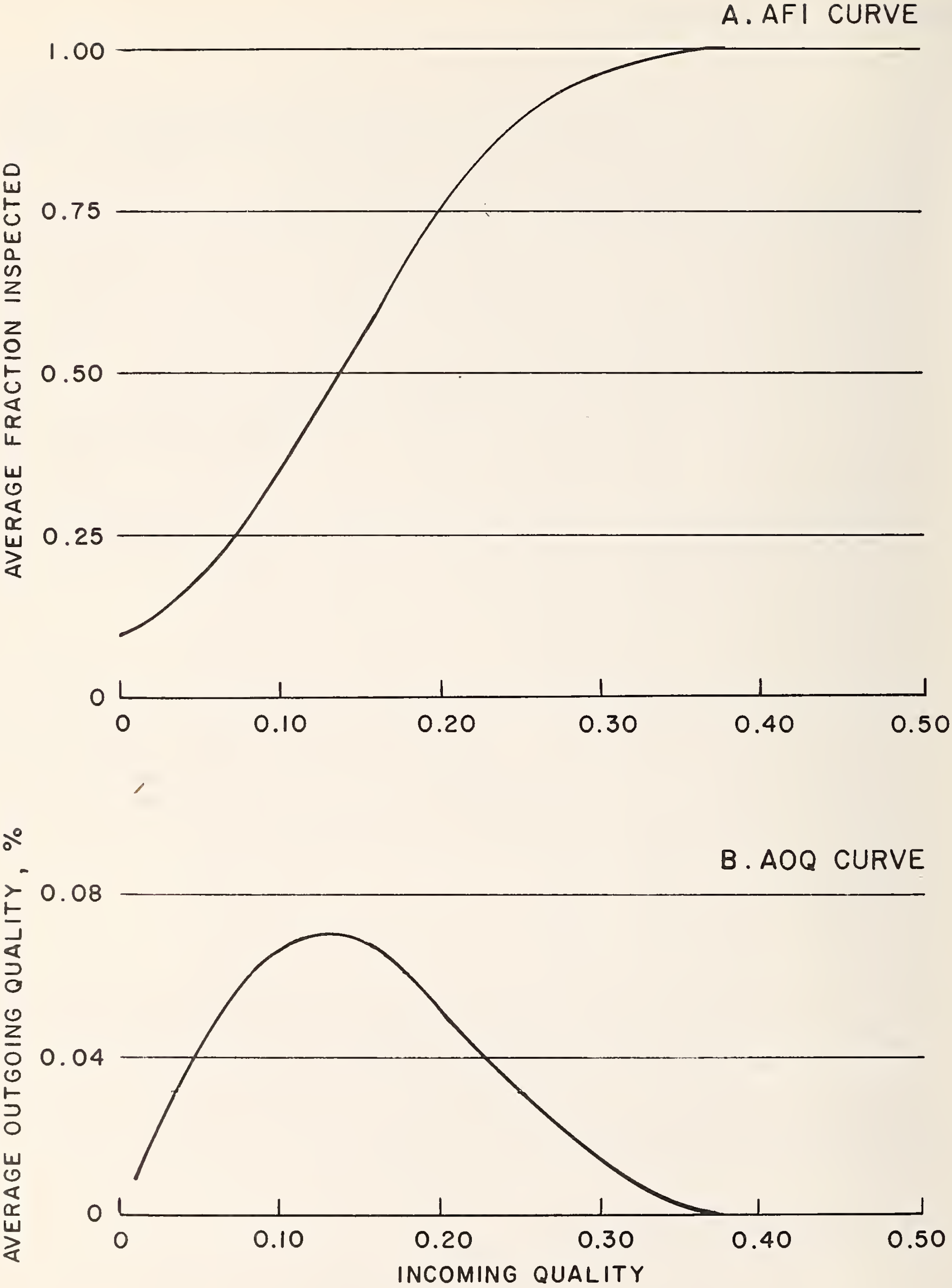


Figure 22. *AOQ* and *AFI* curves for a continuous sampling plan f = 0.10, i = 10, *AOQL* = 0.07.

known to be in statistical control. When it is, this saving in sampling results from omitting the screening of the first *i* production units (see Section C, Chapter III). When the process is not in statistical control, random sampling will soon discover a defective unit, triggering the screening. This results in little saving in the amount of sampling. Statistical quality control in concrete plants is not common, and continuous sampling that begins with random sampling does not seem desirable. It is also undesirable because (as also discussed in Section C of Chapter III) it requires artificial lotting.

(2) Plans Beginning with Screening

Because concrete-producing processes may not be in statistical control at the start of production, continuous sampling that starts with screening is ideal for concrete, except for cylinder strength. In its final location, concrete cannot be separated into distinct and identifiable portions, and therefore defective concrete cannot be conveniently separated from non-defective. That is, concrete cannot be rectified, and if lot sampling were used without provision for reduced payment, rejection of a lot would mean the waste of large portions of acceptable material. This would follow from the fact that with lot sampling, the decision to accept or reject the lot cannot be made until all the lot has been produced and most is hardened in place. With continuous sampling, concrete defective because of slump and air content can be prevented from reaching its place of use, leaving only concrete defective because of strength to reach that place. But since strength accounts for only a small fraction of the total percent defective, this would minimize the administrative headache of what to do with defective material in place.

This means that for inspection of concrete, such as pavement concrete for which strength is not a specified property, a continuous sampling plan of the CSP-1 type is the most desirable. Unfortunately, this type requires a large amount of sampling. For example, to assure that no more than 7 percent of the concrete incorporated in the work is defective for slump and air content, a continuous sampling plan having an *AOQL* of 0.07 would be required. A CPS-1 plan approximating this *AOQL* is Plan $f = 0.10$, $i = 10$. Its *AOQ* and *AFI* curves are shown in Figure 22. The latter curve shows that the minimum average fraction inspected is approximately 0.10, and occurs when the incoming fraction defective approaches zero. The same curve shows that if the incoming percent defective were the 10-to-15 percent being produced, the fraction inspected would jump to 36 and 56 percent, respectively. Assuming a production rate of 300 units a day, this would mean sampling and testing 108 or 168 concrete batches daily. If a smaller *AOQL* were desired, which is likely to be the case, the number of production units to be tested per day would be still greater. Considering that with present testing procedures, one inspector can test about 15 production units daily, such rates are too large for implementation.

Besides the increase in amount of sampling, continuous sampling has other disadvantages. Some sampling occurs during the plan's screening phase. This requires that enough inspection personnel be available at a moment's notice to perform screening, and makes the inspection load variable and unpredictable, so that allocation of inspection manpower is difficult. Moreover, because it takes about 20 minutes to sample and test a production unit, during screening production rates would have to be drastically reduced to permit testing of every batch produced.

2. Suggested Sampling Plans

For the reasons discussed earlier in this chapter, screening and single-lot sampling are the only practical alternatives for concrete inspection. Screening is appropriate for critical concrete and small pours. Single-lot sampling is desirable for inspection of large volumes produced and placed at high rates. Critical concrete is usually concrete in structures. High placement rates are usually encountered in pavement construction. This makes it necessary and desirable that different inspection procedures be used for structural and paving concrete.

a. Inspection of Structural Concrete

(1) Air Content and Slump

Maintenance of structural concrete is expensive and inconvenient. Some of this maintenance results from improper air entrainment and high water-cement ratios. While it is recognized that bridge deterioration is caused by many factors, researchers (25,26,27,28,29) agree that improper air entrainment is the most probable cause of scaling. For that reason alone, it is

TABLE 17
SAMPLE SIZES NECESSARY
TO APPROXIMATE AN *AQL* OF 0.006
WHEN LOT SIZES ARE SMALL
Based on Hypergeometric Distribution Tables

Lot Size	Sample Size	Acceptance Number, *c*	α	β	*AQL*, %	*LTFD*, %
15	13	0	0	0.133	0.6	6.6
19	18	1	0	0.105	↑	↑
25	24	1	0	0.080	|	|
30	28	1	0	0.130	|	|
35	28	1	0	0.096	↓	↓
50	34	1	0	0.091	0.6	6.6

desirable to screen structural concrete for slump and air content. Screening is also desirable for other reasons. First, lot sampling would result in screening for most structural pours, and second, concrete in those pours could not be considered to form rational lots.

For example, using the same plan by variables that will be recommended for inspection for slump and air content of paving concrete (*AQL* = 0.003, *LTFD* = 0.036, $\alpha = 0.05$, $\beta = 0.10$) would result in a sample size of 35. This sample size is independent of lot size, and would in most cases exceed the number of production units in structural pours, because these usually consist of less than 150 cu yd delivered in large truck mixers or agitators. Assuming truck capacities of 6 or 8 cu yd, a pour of 150 cu yd would involve either 19 or 25 truckloads -- much less than 35.

Using lot sampling by attributes, sample sizes would be reduced in inspecting small lots, but not enough to warrant sampling instead of screening. For example, simultaneously to satisfy the *AQL* and *LTFD* for both slump and air content in the plan by variables just suggested for each of these properties, a plan by attributes would be required having approximate *AQL* and *LTFD* of 0.006 and 0.066, respectively (as discussed in the preceding chapter, failures for slump and air content seem to be independent). Aiming for $\alpha = 0.05$ and $\beta = 0.10$, the plans most closely satisfying these conditions for various lot sizes are those in Table 17. These sample sizes were obtained by using probabilities from the hypergeometric distribution, to take into account the small lot sizes. But even with this refinement, sampling has no advantage over screening for lot sizes up to 30.

Screening is also desirable because structural concrete is not likely to be from a rational lot -- a necessary assumption of lot sampling. Plant production rates far exceed the number of units used in a structural pour. This means that plants might be producing concrete for other uses at the same time they are supplying a structural pour, and that plant settings might be constantly changing. Under those circumstances, it is hard to conceive of concrete in structural pours as rational lots.

For these reasons, screening of all structural concrete for slump and air content is recommended, and if adopted this should not in most instances increase the present cost of inspection. The usual bridge span or pier and cap requires less than 15 truckloads or 90 cu yd of concrete. The inspector normally assigned to structural pours should be able to handle this testing load.

But screening for slump and air content would require deviation from the standard ASTM sampling procedure for truck-mixed concrete and concrete delivered in truck mixers. If that procedure were followed, test results would not be available until at least a portion of the concrete was already in the forms. This would defeat the very purpose of screening, which is to eliminate defective production units before use. This situation can be corrected without serious consequences by testing only the first portion of the truck's discharge. The ASTM procedure would give a more precise estimate of property levels in the tested batch, but screening is basically a sampling plan by attributes, and the loss of precision would not greatly affect the results, because in sampling by attributes it is more important to know whether any single production unit possesses the necessary quality determinants within the specified limits, than to estimate the precise level of the quality determinants in that unit. This point is discussed further in Appendix E.

(2) Strength

In New York, contractors are allowed to begin pier construction seven days after foundation concrete has been poured. Similarly, steel placement can begin seven days after the pier or cap has been poured. Under present testing procedures, the presence of concrete defective for strength is discovered only 28 days later. At that stage, foundations may be supporting piers which in turn may be supporting steel. Enforcement of screening would mean removal of steel or both steel and piers, causing such havoc that screening is likely never to be enforced. In this respect, screening is impractical, but it is desirable for some of the same reasons that make it desirable for air content and slump. First, structural concrete is not likely to come from rational lots. Second, any sampling plan that assures reasonable quality levels would require testing more than the number of production units usually included in structural pours making screening unavoidable in most cases. Thus, screening should also be used for strength, but its implementation should be accompanied by reduced payment rather than removal, as for slump and air content, unless the strength is so low as to impair the structure. The cutoff point for removal could be adopted to coincide with ACI Standard 214-65, which allows 10 percent of concrete in structures to be under-strength.

Another possible way to use screening is to adopt an accelerated testing procedure. With some of these procedures, the decision to accept or reject a structural pour could be made two or three days after pouring concrete. Since forms are left in place for at least two days, keeping structural units available for removal would not delay construction, and screening could be implemented by removal of all concrete.

Regardless of the testing procedure used, screening for concrete could be accomplished by casting one cylinder from each production unit. This is also a departure from ASTM standard procedures, but is desirable. Two cylinders give a more precise estimate of the strength of one production unit than one cylinder. But in acceptance sampling, it is more important to assure that as many production units as possible comply with specifications than precisely to estimate the level of a given property in a single production unit.

Metcalf (30) has illustrated that for acceptance sampling, it is better to increase the number of production units tested than to improve the precision of the estimate of the property level of a single unit. In his analysis of a British Standard specification, he compares the *OC* curve obtained when three cylinders are cast from each of ten production units with that obtained when one is cast from each of 30 production units. The resulting curves show that the latter *OC* curve is much more discriminating.

b. Inspection of Pavement Concrete

Concrete delivery rates to paving projects are so high that screening and continuous sampling become too expensive to implement, leaving lot sampling as the only practical alternative to assure reasonable quality levels. As discussed, the variations of lot sampling applicable to concrete are single-lot sampling by attributes and by variables, the latter being preferable because it requires smaller sample sizes. Lot sampling by variables requires a plan for each of the inspected properties, and definition of the lot and inspection unit, as well as choosing risks and quality levels.

For plastic concrete, the logical choice of inspection unit is the production unit. The lot chosen should be large enough to reduce sampling and small enough to allow frequent judgment of the product, and should consist of production units made under essentially the same conditions. A day's production of paving concrete approximates these conditions. It allows judging concrete quality daily, and consists of 300 to 500 production units -- a large enough number to reduce sampling. Paving absorbs all concrete from a plant, and thus plant production is likely to be continuous and plant settings, materials, and personnel are likely to remain unchanged during the day, resulting in homogeneous lots.

Failure of paving concrete is not likely to cause loss of life or to have other drastic consequences, and the producer's and consumer's risks can be chosen as the customary 0.05 and 0.10, respectively. This leaves the task of choosing appropriate quality levels. Valid choices of quality levels can be based only on performance data. Unfortunately, objective data relating quality levels and concrete performance are scarce. This makes their choice difficult and somewhat arbitrary.

Lacking performance data, two approaches appear reasonable. The first is to assume that concrete produced in the past has rendered satisfactory performance, and to choose the *AQL* to coincide with the average fraction defective of past production. The second is to decide what quality levels are economically achievable, and to design sampling plans accordingly. For concrete, the first approach appears inappropriate. Lack of process control has resulted in concrete having a high average fraction defective. Using this approach would lead to *AQL*'s of approximately 5 percent for each individual property. This implies that concrete having a combined *AQL* for all properties of about 15 percent is acceptable material that can be used without adverse consequences. Considering that to keep the sample size at manageable levels, the *LTFD* must be at least three times the

AQL, this would mean that some concrete lots would be allowed to contain 40-percent out-of-specification material. This does nothing to improve present concrete quality, and would sanction production of concrete without process control, as in the past. But more important, this approach would retain the high fraction defective that many engineers blame for pavement scaling and other difficulties, and thus is inadvisable.

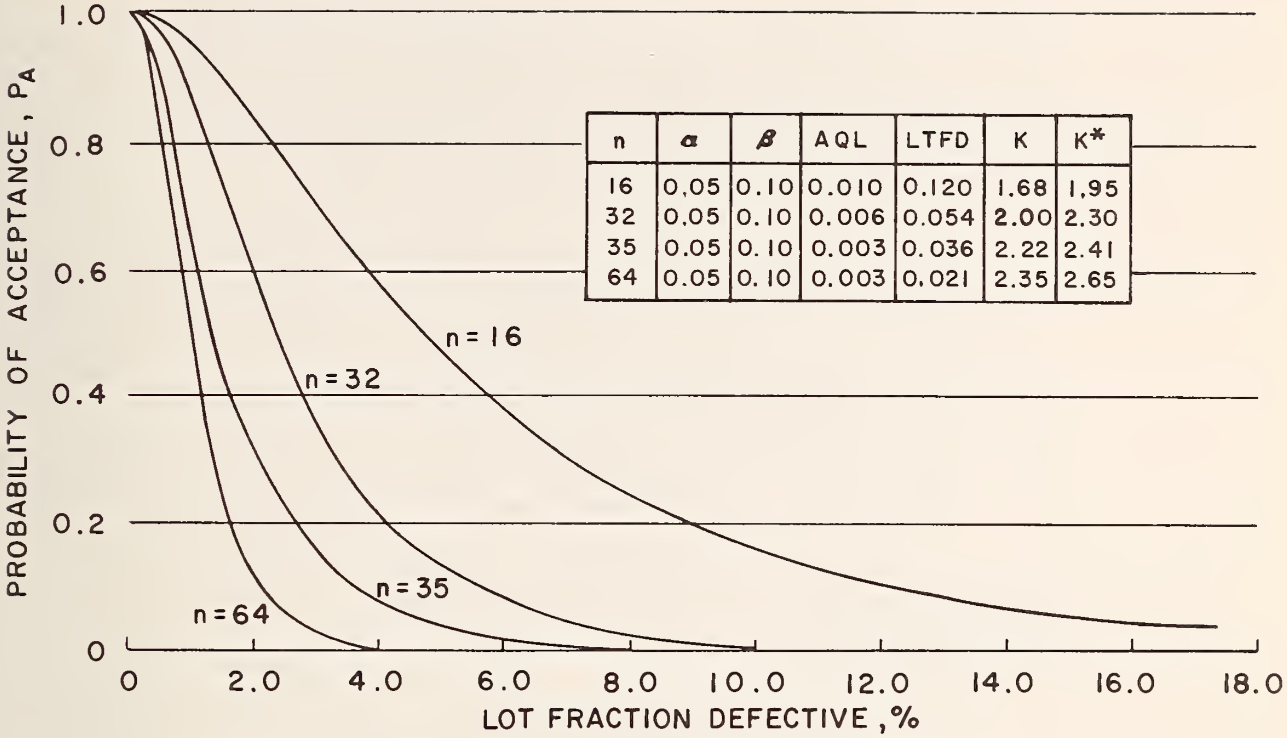
The second approach seems more appropriate. The Chapter II case studies show that at least for slump and air content, current New York State specifications can be met without great effort in controlling the process. It thus seems reasonable to assume that the fraction defective for these properties can approach zero. Using this approach, quality levels can be chosen that will result in improved concrete quality, and at the same time keep sample sizes at manageable levels. If the *AQL* is chosen as zero, the resulting plan will be screening. To avoid that difficulty, a small *AQL* and *LTFD* can be chosen to keep the sample size reasonable. To see how sample sizes vary, a number of sampling plans were obtained using a number of *AQL* and *LTFD* combinations. The results are presented in Table 18. Each line represents two sampling plans having the same *OC* curve, but requiring different sample sizes depending on whether the standard deviation is known. It can be seen that for the same *AQL*, sample sizes decrease as the ratio *LTFD/AQL* increases. It can also be seen that for the same *OC* curve, the sample sizes required when the standard deviation is known are much smaller than those necessary when it is unknown. The great differences in sample sizes stress that plans with standard deviation known should be used whenever possible. Unfortunately, assuming known standard deviations for concrete is inappropriate. First, most plants do not practice statistical quality control, and the standard deviation of the process changes as assignable causes come into play. Second, even if process control were practiced, the standard deviation would probably change from plant to plant. Moreover, there is no way to assume a safe standard deviation. If a large one is assumed, producers of uniform product would be penalized and those of variable product rewarded. If a small standard deviation were assumed, the percentage defective accepted under the plan would depend primarily on the mean, and would remove any incentive for producers to reduce variability, as Figure 8 illustrates. This would lead to acceptance of more defective material than the plan would theoretically indicate. Both consequences are undesirable, and it is believed that plans with known standard deviations cannot be used for concrete inspection. This leaves adjustment of the *AQL* and *LTFD* as the only means of choosing a reasonable sample size.

TABLE 18
VALUES FOR SEVERAL SAMPLING PLANS
R = LTFD/AQL

Plan	α	β	*AQL*, %	*LTFD*, %	*n* for Standard Deviation Known	*k*	*n* for Standard Deviation Unknown	Maximum Standard Deviation For Spec. Ranges of		*k**
								3 Units	4 Units	
R = 7										
1	0.05	0.10	0.3	2.1	17	2.350	64	0.57	0.76	2.646
2			0.4	2.8	16	2.233	56	0.60	0.80	2.499
3			0.5	3.5	15	2.142	50	0.62	0.82	2.420
4			0.6	4.2	15	2.071	46	0.63	0.84	2.375
5			0.7	4.9	14	2.005	42	0.66	0.88	2.277
6			0.8	5.6	13	1.948	37	0.68	0.91	2.190
7			0.9	6.3	13	1.897	35	0.69	0.92	2.184
8	0.05	0.10	1.0	7.0	12	1.851	33	0.71	0.94	2.117
R = 9										
9	0.05	0.10	0.3	2.7	13	2.288	48	0.61	0.82	2.442
10			0.4	3.6	12	2.171	41	0.63	0.84	2.374
11			0.5	4.5	12	2.080	36	0.64	0.86	2.333
12			0.6	5.4	11	2.003	32	0.65	0.87	2.302
13			0.7	6.3	11	1.932	29	0.67	0.89	2.236
14			0.8	7.2	10	1.875	27	0.68	0.91	2.202
15			0.9	8.1	10	1.824	25	0.71	0.95	2.112
16	0.05	0.10	1.0	9.0	9	1.772	23	0.73	0.97	2.064
R = 10										
17	0.05	0.10	0.3	3.0	12	2.260	43	0.62	0.82	2.434
18			0.4	4.0	11	2.143	37	0.62	0.82	2.426
19			0.5	5.0	11	2.052	32	0.64	0.85	2.356
20			0.6	6.0	10	1.975	28	0.67	0.90	2.225
21			0.7	7.0	9	1.904	26	0.68	0.91	2.200
22			0.8	8.0	9	1.847	24	0.71	0.95	2.113
23			0.9	9.0	9	1.790	22	0.74	0.99	2.015
24	0.05	0.10	1.0	10.0	8	1.739	20	0.77	1.02	1.954
R = 12										
25	0.05	0.10	0.3	3.6	10	2.215	35	0.62	0.83	2.410
26			0.4	4.8	9	2.098	29	0.64	0.85	2.340
27			0.5	6.0	9	2.001	26	0.69	0.93	2.157
28			0.6	7.2	8	1.919	23	0.70	0.93	2.156
29			0.7	8.4	8	1.847	21	0.72	0.95	2.095
30			0.8	9.6	8	1.791	19	0.73	0.97	2.065
31			0.9	10.8	7	1.734	17	0.76	1.01	1.979
32	0.05	0.10	1.0	12.0	7	1.682	16	0.77	1.03	1.950

(1) Slump and Air Content

These properties can be tested using concrete from the same sample. To reduce the sampling and simplify administration, it is desirable to use the same sampling plan for both properties. Doing so will result in a slightly looser plan for slump, because the *AQL* consists of material exceeding only the upper limit. But if the *AQL* is small, the difference is negligible and presents no problem. Since the fraction defective for these properties can reach zero, a small *AQL* can be chosen and the same plan used for both properties. It then remains to choose a plan.



n	α	β	AQL	LTFD	K	K*
16	0.05	0.10	0.010	0.120	1.68	1.95
32	0.05	0.10	0.006	0.054	2.00	2.30
35	0.05	0.10	0.003	0.036	2.22	2.41
64	0.05	0.10	0.003	0.021	2.35	2.65

Figure 23. *OC* curves for sampling plans with standard deviation unknown.

The first sampling plan with standard deviation unknown in Table 18 requires 64 samples. It is fairly discriminating, as its *OC* curve in Figure 23 shows. But to test 64 production units of concrete for slump and air content would require four inspectors or more. The last plan with standard deviation unknown requires 16 samples, but referring to Figure 23, it can be seen that its *OC* is not discriminating. Under this plan, a lot having a fraction defective due to one property alone of 5 percent would have a probability of acceptance of 0.45, and a lot having 12 percent defective would have a probability of acceptance of 0.10. Thus, this plan is not adequate to guarantee low fractions defective. In Figure 23, it can be seen that the plan marked n = 35 gives an *OC* approximating that of plan n = 64, and that the *OC* curve of plan n = 32 is much less discriminating than that of plan n = 35. For three samples, it appears desirable to take plan n = 35, which has an *AQL* of 0.003 and *LTFD* of 0.036, with α = 0.05 and β = 0.10. This means that if a producer delivered a continuous stream of lots having 3.6-percent defective for either property, 90 percent of the lots would be rejected. It also means that if he delivered lots with 0.3-percent defective for one property, 5 percent of the lots would be rejected. Thus, to avoid economic losses, producers would have to control their processes to produce fractions defective less than 0.3 percent for each property. This plan is discriminating enough, requires only two or three inspectors, and is recommended for use.

It is recognized that many engineers accustomed to thinking of inspection strictly as a visual task, and exposed to literature claiming that statistical sampling should reduce the amount of testing, will think of these testing rates as perhaps unrealistic. But they are necessary if current specification limits are to have the meaning traditionally attributed to them and compliance is desired. They are unrealistic only if compliance with specification limits is meaningless.

Implementation of the suggested plan reduces to a step-by-step procedure, as follows:

(a) Slump

1. Sample and test 35 production units randomly selected throughout the day.

2. Compute the mean slump $\bar{X}$ and standard deviation S from the sample.

3. Accept the lot if the following condition is satisfied:

$$\frac{\bar{X} - L}{S} \geq K$$

where $\bar{X}$ = mean slump,

S = sample standard deviation,

L = specification limits, and

K = standard deviate that must be equalled or exceeded in order for the lot to be acceptable, from Table 18.

4. If either or both of these conditions are not satisfied, reject the lot. In this case, rejection would mean computing an appropriate reduced payment.

(b) Air Content

1. Sample and test 35 randomly selected production units recovered throughout the day. These can be the same 35 sampled for slump.

2. Compute the sample mean $\bar{X}$ and standard deviation S.

3. Accept the lot if the following conditions are met:

$$S \leq MSD$$

$$\frac{U - \bar{X}}{S} \text{ or } \frac{\bar{X} - L}{S} \geq K$$

where MSD = maximun allowable standard deviations for appropriate specification range from Table 18,

L = upper or lower specification, and

K = standard deviate that must be equalled or exceeded in order for the lot to be acceptable, from Table 18.

It is to be emphasized that when the population standard deviation is unknown, both criteria must be satisfied. If only the mean were controlled, it could be in the acceptable range, but the standard deviation could be so large as to lead to a higher fraction defective than the plan allows.

It also should be noted that the sampling plans just recommended are not the only ones available for inspection of concrete for slump and air content. For example, if enough inspection manpower were available, sampling by attributes could be used. Keeping the lot a day's production, one plan by attributes that would assure about the same quality levels as the two plans by variables just suggested is $\alpha = 0.05$, $\beta = 0.10$, $AQL = 0.006$, and $LTFD = 0.066$. The acceptance number for this plan is 1 and the required sample size is 59. Its OC curve is that shown in Figure 24. Its implementation requires no calculations and is as follows: sample and test 59 randomly selected production units from the lot, and if two or more of the tested units

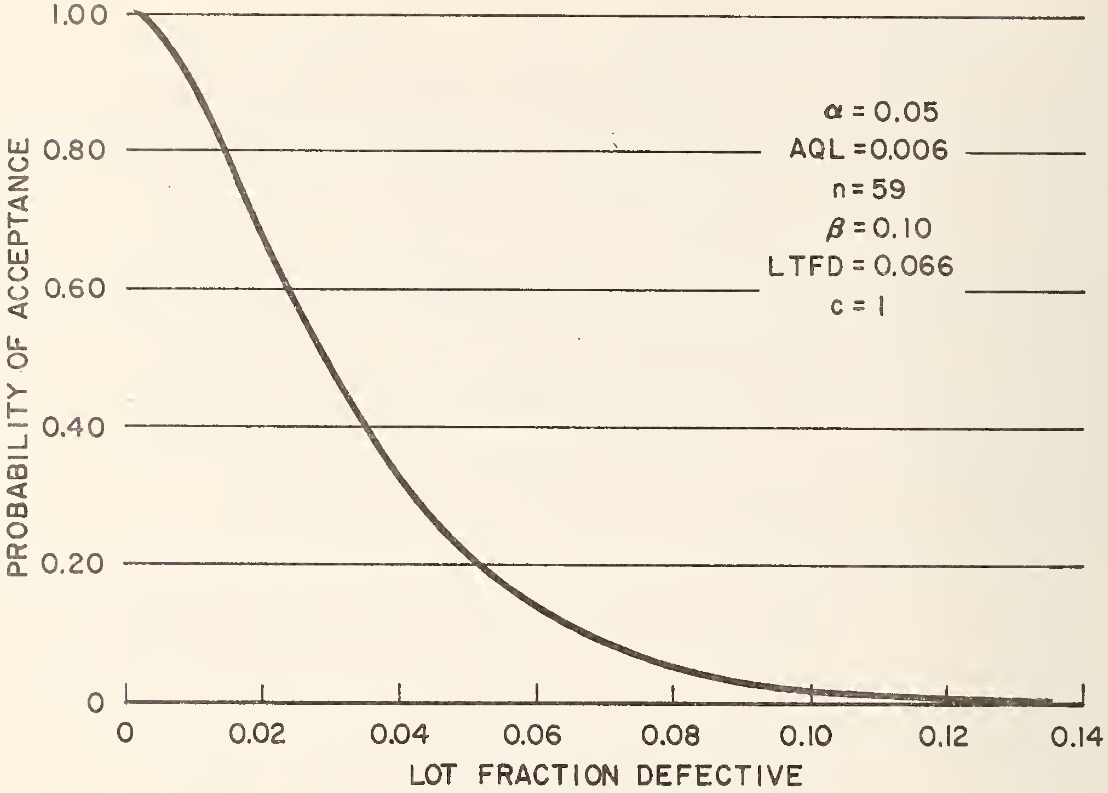


Figure 24. *OC* curve for attributes single sampling plan.

fail to meet specifications for either slump or air content or both, reject the lot -- otherwise accept it. But sampling and testing an additional 24 production units seems a very high price to pay to avoid calculations.

(2) Strength

Currently, New York State does not formally specify compressive strength for concrete pavements, which means that no acceptance sampling plan is required for this property. However, if the now-informal specifications were to be formalized, a sampling plan is necessary. The type appropriate is again lot sampling by variables with standard deviation unknown. But to design such a sampling plan, one must know the desired quality levels and must define the lot and inspection unit. The lot, inspection unit, and risks can be chosen as for slump and air content, but the choice of quality levels is difficult because objective performance data again are lacking.

Until performance data are systematically accumulated on a large scale, a reasonable sampling plan for paving concrete compressive strength can be designed by assuming that the pertinent ACI specification is correct. "Specifications for Concrete Pavements and Concrete Bases" (ACI 617-58) states:

> The average compressive strength for use with the allowable stresses for design of dowels and tie bars as recommended in "Recommended Practice for Design of Concrete Pavements" (ACI 325), shall not be less than 4000 psi at 28 days, when specimens are molded and tested in accordance with Section 304.

If 4000 psi is the minimum average strength and 3000 psi is taken as the minimum strength, the lower-half specification range would be 1000 psi. This means that to eliminate nearly all defectives, the standard deviation would have to be 1000/3 or 333 psi. But such a small standard deviation is a rarity, as can be seen from Figure 7E. ACI considers coefficients of variation from 0.10 to 0.15 to represent good production control. This means that for a 4000-psi average, the attainable standard deviation should range from 400 to 600 psi, with a midrange value of 500 psi. Thus, ACI implicitly allows a fraction defective or an *AQL*. However, the magnitude of the implied *AQL* is a matter of speculation, because it depends on the magnitude of the standard deviation, and on the interpretation given to "shall not be less than 4000 psi" in the quoted specification. If this means that concrete is acceptable as long as its average strength does not fall below 4000 psi, even if every average is exactly 4000 psi, the *AQL* can be inferred by assuming an achievable standard deviation. If, however, it means that it is desirable to have a higher average most of the time, such an inference is impossible.

Assuming the first interpretation, the *AQL* would depend on the magnitude of the standard deviation, and becomes a matter of choosing an achievable value. As discussed, a reasonable value achievable with some process control is 500 psi. Assuming normality, this standard deviation implies an *AQL* of approximately 2.3 percent. To design a plan, an *LTFD* must still be determined. The results in Table 19 show that to keep the sample size at a manageable number of cylinders, the *LTFD* must be at least 11.5 percent. This would result in Plan 3, requiring 30 samples per lot (a day's

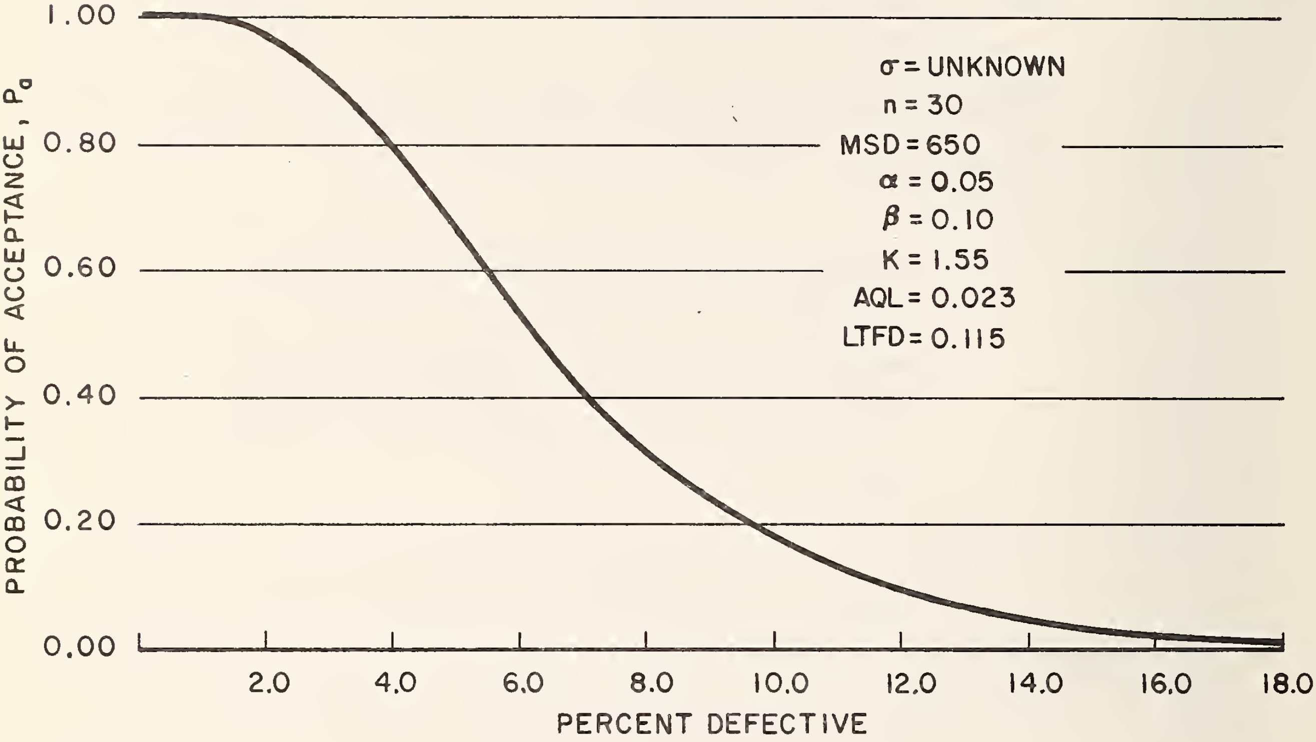


Figure 25. Suggested sampling plan by variables for strength, with standard deviation unknown.

TABLE 19
SAMPLING PLANS FOR
CONCRETE STRENGTH WITH STANDARD DEVIATION UNKNOWN
R = LTFD/AQL

Plan	α	β	*R*	*AQL*, %	*LTFD*, %	*n*	*k*
1	0.05	0.10	3	2.3	6.9	79	1.71
2	0.05	0.10	4	2.3	9.2	44	1.62
3	0.05	0.10	5	2.3	11.5	30	1.55

production). This is a reasonable plan and is recommended. This plan's *OC* curve is shown in Figure 25, and its acceptance criteria are given in Table 19.

The plan's implementation is basically the same as for those suggested for slump and air content. The steps are as follows:

1. Sample 30 production units of concrete randomly selected throughout the day's production and cast one cylinder from each unit. These

can be the same production units sampled for air content and slump, or different ones.

2. Compute the sample mean $\bar{X}$ and sample standard deviation S after testing for 28-day compressive strength.

3. If the following condition is met, accept the day's production as regards strength:

$$\frac{\bar{X} - 3000}{S} > K \text{ or } 1.55$$

4. If this condition is not satisfied, compute an appropriate reduced payment for strength.

This sampling plan assumes that the mix design can yield concrete having an average 28-day strength of 4000 psi. This means that if acceptance sampling for strength is introduced, producers should be allowed to design their own mixes in order to avoid conflicting requirements. If buyers specify a mix design, the mix might not attain the required strength. Then rejection under the sampling plan would be prevalent and would lead to administrative headaches. Producers would make every effort to blame the buyer's mix design to avoid the contractual consequences of low strength.

Again, it realized that 30 cylinders a day is a sizeable inspection load. But the plan suggested here is relatively loose. Requiring less sampling would lead to a very large *LTFD*. Thus, if strength is important, this amount of testing is necessary. The only possible ways to reduce sampling and still assure good quality would be to assume a known standard deviation or to increase lot size. Both seem inappropriate. Assuming a known standard deviation is unrealistic, as can be seen from Figure 7E. Increasing the lot size could result in irrational lotting and would require judging a large amount of material on relatively few samples (only 30). Choosing the lot size as several days' production or as the whole job is appropriate only if it can be shown that variation between production units produced on different days equals or is only slightly larger than variation between production units produced within the same day. If that were the case, increased lot size would be desirable. But this should be done only after making sure that variation remains constant with time -- i.e., that the process is being controlled successfully.

c. Inspection of Concrete for Miscellaneous Uses

Concrete is often used in small quantities, for which the only applicable inspection procedure is screening. Sampling plans would not be appropriate for two reasons. First, any reasonable sampling plan would require sampling all production units supplied to a small pour. Second, such pours would absorb only a minute part of plant's production, and it is unlikely that the production units included in small pours constitute a rational lot. For these reasons it is recommended that concrete used in sidewalks, gutters, curbs, drop inlets, headwalls, and other small pours be screened at least for slump and air content.

3. Reduced Payment

Ideally, reduced payment should equal the economic consequences of reduced quality. But devising ideal payment schedules is difficult for concrete. First, performance data are not available relating the percentage of defective material to the maintenance-free life of concrete structures and pavements. Second, the small sample sizes usually used in acceptance sampling, including those just suggested, are too small for a precise estimate of the fraction defective in the concrete represented by the sample. But the main function of reduced payment is to provide an alternative to enforce the contractual agreement. Thus ideal payment schedules, although desirable, are not essential.

When a sampling plan rejects a lot, the probability that the lot fraction defective exceeds the *AQL* is very high. This means that the buyer is not receiving the quality specified in the contract. Under these circumstances the buyer's first responsibility should be to assure that suppliers do not gain financially by delivering poor quality and have an incentive to adhere to the contractual quality requirement. To provide this incentive, reduced payment schedules do not necessarily have to be ideal as long as they are reasonable. One such schedule is shown in Table 20. Many others can achieve the same objective and are equally worthy of consideration. The ultimate validity of any proposed payment schedule must be judged on its ability to satisfy its intended purpose. This judgment can be made only after implementation. The approach taken in preparing Table 20 was to give stiff penalties after the percent defective exceeded that believed easily achievable, to encourage producers to pay attention to process control.

TABLE 20
SAMPLE REDUCED PAYMENT SCHEDULE
FOR PAVING CONCRETE

Estimated Lot Fraction Defective, %	Pay Factors (% of bid price) for Properties Shown		
	Slump	Air Content	Strength
0.0-0.3	100	100	100
0.3-1.0	100	100	100
1.0-2.0	97	97	100
2.0-3.0	94	94	100
3.0-4.0	90	90	95
4.0-5.0	85	85	90
5.0-6.0	80	80	85
7.0-10.0	70	70	80
10.0-12.0	Remove Pavement		70
Over 12.0	Remove Pavement		

> The lot fraction defective in all three cases would have to be estimated from the sample mean and standard deviation, assuming normality because the sampling is by variables. Once the fraction defective for each of the three properties is known, the reduced payment is simply calculated by multiplying the bid price by all three pay factors consecutively, as long as removal is not necessary for any or all properties. For example, if a lot had estimated lot fraction defectives of 1, 2, and 3 percent, respectively for slump, air content, and compressive strength, the unit price to be paid would be 1.00 x 0.97 x 1.00 x bid price, or 97 percent of the bid price. Of course, reduced payment schedules must be part of the contractual agreement or they become meaningless.

C. Plans for Inspection of Hardened Concrete

Acceptance of the final product causes the least interference with production, and thus is administratively ideal. With final product inspection, the buyer need never see the manufacturer's plant or operation. His task reduces to testing a sample of production units and deciding whether to accept the material the sample represents. This procedure is simple and minimizes friction between buyers' and suppliers' representatives. But it is appropriate only if product quality determinants can be measured after production, and testing does not result in destruction. For concrete, two of the quality determinants -- air content and strength -- can be measured after production but slump cannot. However, slump is measured as a clue to potential concrete strength, and if strength is measured, slump can be omitted without great loss of information. In this respect, final product acceptance is possible. But in most cases, destructive sampling and testing precludes its adoption for concrete. The only exception is pavements. Determining the quality of structural concrete would mean loading structural units to failure, or recovering many cores from each structural component. Both are undesirable, since the first would result in destruction, and the latter in structural damage and unsightly patches. This makes testing hardened concrete in structures inappropriate and undesirable. By contrast, cores are routinely recovered from concrete pavements to measure thickness without apparent ill effects. These cores can also be used to measure both strength and air content. Thus, testing hardened pavement concrete is possible, and in some respects desirable, because 1) it eliminates interference with production, 2) it permits optimum allocation of inspection manpower, 3) it allows judging pavements only once, and 4) it can result in possible reduction in total inspection costs.

Inspecting the highway after the paving operations end prevents inspectors from attempting to assume the foreman's duties, and sampling and testing does not delay operations. It allows scheduling inspection independently of paving, thus allowing both systematic scheduling of inspection tasks and better use of available manpower. Another advantage is in providing results to judge the pavement once, instead of four times as under the present system. Concrete pavements are now inspected for concrete quality determinants, for roughness, for skid resistance, and for thickness, and judgments are made after each inspection. Testing hardened concrete would require waiting 28 days for a decision on concrete quality determinants. Since waiting is necessary to judge the concrete quality, the other properties could be measured in the meantime and the pavement judged only once on the basis of both concrete and pavement properties.

Under certain circumstances, sampling and testing could be reduced by increasing the lot size. If concrete producers successfully employed process control, the uniformity of concrete produced on different days would not change significantly. In such cases, the lot could be chosen as all concrete in a pavement. This would result in fewer but larger lots, and a reduced overall amount of testing. This would be so because the *OC* curves for lot plans by variables are independent of lot size, and *OC* curves for lot plans by attributes become independent of lot size as it increases to about 10 times the sample size. Choosing a large lot would not change the *OC* curve or quality assured, but because the sample size is the same for small as for large lots, this would reduce the overall testing. Considering that with more than one high-pressure meter (Appendix C), one man could make as many air content determinations each day using cores as he could testing plastic concrete, and that testing for strength is relatively fast, this offers great potential under the right circumstances. The only drawback of increased lot size is having to pass judgment on large amounts of materials which represent large investments. If the assumption of process control proves false, producers could be faced with large losses, both in reduced payment and removal. But even if increasing the lot size to include more than a day's production does not prove feasible, the other advantages still make inspection of hardened pavement concrete desirable and worthy of consideration.

1. Assuming Unsuccessful Process Control

Inspecting concrete after hardening does not change its criticality or its performance. Thus, risks and quality levels can be chosen as for the sampling plan recommended for the inspection of plastic pavement concrete. Because when process control is not exercised, large amounts of plant output are likely not to result in rational lots, the lot should again be chosen as a day's production. But the inspection unit can no longer be chosen as a production unit, because after concrete is mixed in a pavement it is difficult to distinguish between batches. The inspection unit in this case must be a core.

In deriving sampling plans for the inspection of hardened concrete, then, the only change is the inspection unit. But *OC* curves for sampling plans by variables (which would still apply) are independent of the inspection unit and of the magnitude of the standard deviation. Hence, the resulting plans will be those given for air content and strength for the inspection of plastic concrete, having the *OC* curves in Figures 23 and 25. But the acceptance criteria for air content will have to be adjusted slightly for concrete pavements placed in conventional forms.

Appendix C indicates that air contents of hardened concrete measured by the high-pressure method and those of corresponding plastic concrete have statistically equal variances but statistically different means, with mean air content of cylinders slightly lower than that of the corresponding fresh concrete. Other laboratory data for cores substantiate these results. Thus, the standard deviation of air content measured by the high-pressure meter can be assumed to be the same as for plastic concrete (0.5 percent). Moreover, the slight difference in means can be taken into account by increasing the level of air content in fresh concrete or by changing the lower limit of the air content specification when using high-pressure meter results, while keeping the difference between the lower and upper limits constant. This means that in those cases where one core represents one production unit of plastic concrete, the resulting sampling plan and acceptance criteria are identical to the sampling plans given for plastic concrete in Figures 23 and 25. This is the case for concrete pavement placed with slipform paving. In such paving, the concrete is placed full-depth and the reinforcing mesh vibrated into place. A core taken from a pavement so placed is very likely to represent only one batch of fresh concrete. However, by contrast, in conventional paving, the slab is constructed in two layers, with the concrete in the bottom layer usually coming from a different batch than that of the top layer. The limited data in Table 21 illustrate this point. As would be expected, air content may differ greatly between the top and bottom halves of the same core. This necessitates a slight adjustment in allowable specification range, and thus in acceptance criteria, because the values are averages from samples of Size 2.

TABLE 21
AIR CONTENT IN TOP AND BOTTOM PORTIONS OF CORES TAKEN FROM VARIOUS CONVENTIONALLY PAVED JOBS

	Linear Traverse Air Content, %		
Core	Top	Bottom	Difference
1	6.35	7.93	1.58
2	4.20	4.87	0.67
3	0.78	5.81	5.03
4	5.45	3.77	1.58
5	5.20	4.20	1.00
6	4.70	5.20	0.50
7	7.72	11.41	3.69
8	10.11	1.54	8.57
9	14.61	15.41	0.80
10	8.19	8.71	0.52
11	10.51	6.79	3.72
12	9.20	11.46	2.26
13	8.58	9.35	0.77

Assuming normality, the standard deviation of sample means of Size 2 is approximately the standard deviation of the single values, divided by $\sqrt{2}$. For practically all production to be acceptable, the range of the average from samples of Size 2 must be $6(\sigma/\sqrt{2})$. The value of σ for a single value is the same as for the air content of plastic concrete, or 0.5 percent. Thus, the range becomes $(6 \times 0.5)/\sqrt{2}$, or 2.16. The mean value, however, remains unaffected and is still 5.5 percent.

Using this new range the maximum standard deviation *MSD* can be computed as follows:

$$MSD = \frac{U' - L'}{2K^*} = \frac{2.16}{2 \times 2.41} = 0.45$$

where U' and L' are the upper and lower specification limits for sample means of Size 2. Then the lot is acceptable if the following conditions are met:

$$\frac{\bar{X} - 4.42}{S_{\bar{X}}} \geq 2.22 \quad \text{and} \quad \frac{6.58 - \bar{X}}{S_{\bar{X}}} \geq 2.22$$

where $\bar{X}$ = the mean air content computed from the core results, and

$S_{\bar{X}}$ = the standard deviation of core air content.

For strength, the acceptance criteria are given with the *OC* in Figure 25 still applying. It may be argued that in this case the strength of cores represents the strength of two concrete batches. But until data are available to substantiate this, it is sufficient to assume that each core strength represents a single rather than an average value. If false, this assumption increases the consumer's risk, because the specification limit for single values is lower than that for average values. However, since conventional paving is being replaced with slipform paving, this situation is not likely to occur very often.

2. Assuming Successful Process Control

Again risks and quality levels remain unchanged and the inspection unit can still be chosen as a core. But if process control is successfully used by producers, the lot can be chosen as the whole pavement. As discussed, the *OC* does not depend on lot size and the resulting plans would still be those in Figures 23 and 25 with the necessary modifications explained for pavement constructed with conventional form paving.

This approach reduces sampling and is recommended, but with caution. If the assumption of successful process control proves false, making the whole pavement the lot can have serious consequences, because rejection of a pavement can cause considerable financial losses regardless of whether the penalty is reduced payment or removal. For example, review of the

New York State Department of Transportation "Construction Contract Status" summary issued in October 1970 showed that 62 of the 77 concrete pavement projects listed exceeded 8 lane miles in length. Data for other years indicate similar trends. Judgment of such amounts of material should be based on substantiated assumptions. One way to substantiate the assumption of process control is to inspect the producer's quality control system and records. If the records show lack of control or no control at all, it may be desirable to inspect the concrete while plastic or to avoid his product.

It may be aruged that in judging these large amounts of concrete, the sample size should be increased. In doing so, however, the *OC* curves would become more discriminating, because they are independent of lot size and would in fact mean changing the required quality levels (31). It would also result in foregoing the advantage of reduced sampling and testing. In fact, rather than increase the sample size, it may be more advantageous to reduce lot size and judge concrete more frequently, to provide the producers with feedback regarding their concrete quality. Rejection, however infrequent, would provide a real incentive for process control while control can still result in quality.

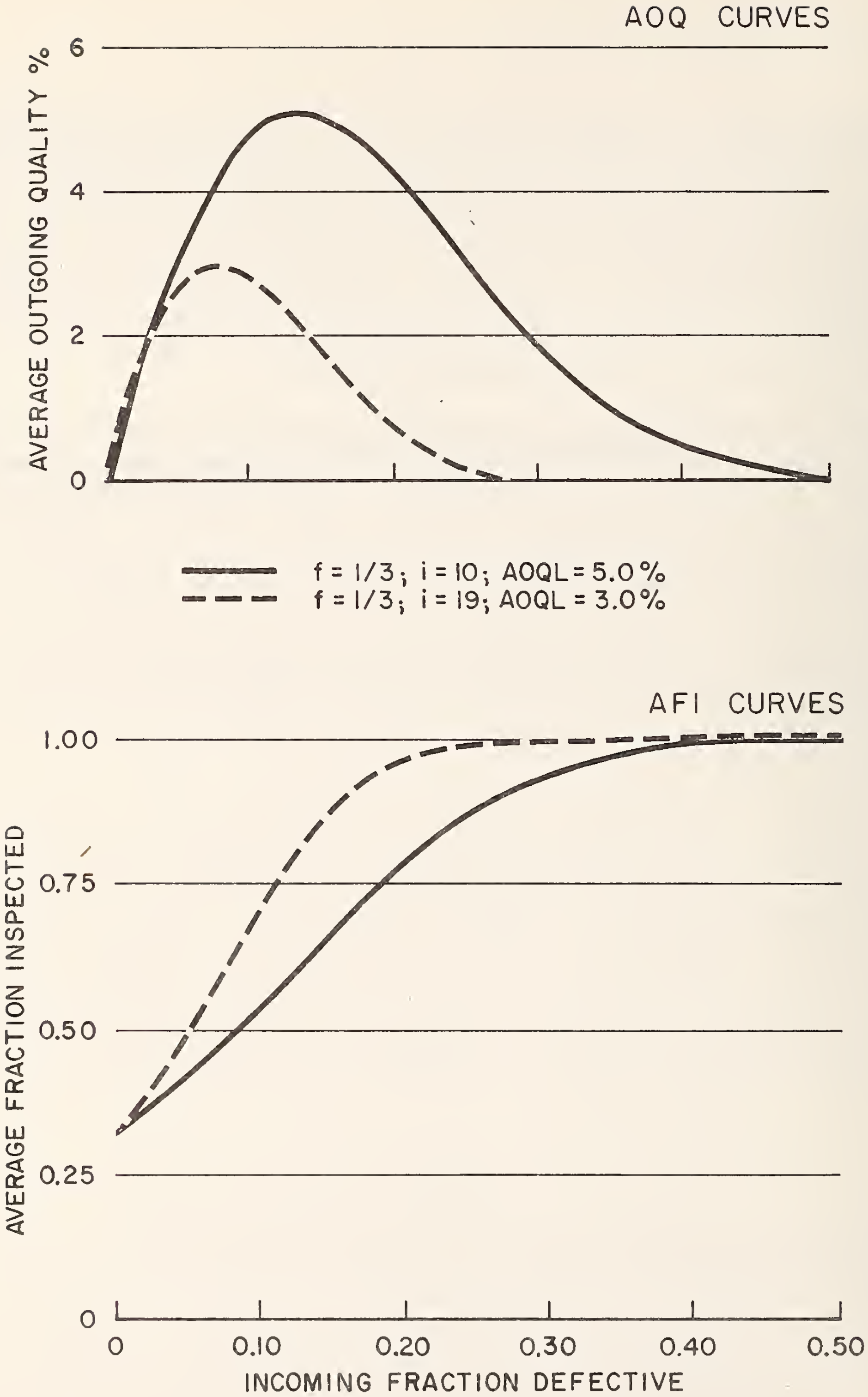


Figure 26. *AOQ* and *AFI* curves for two continuous sampling plans approximating lot plans by attributes.

V. DESIRABILITY OF STATISTICAL SAMPLING PLANS

A. Advantages

All the sampling plans recommended require more sampling and testing than is now performed in New York. If continuous sampling were used for slump and air content, the sampling would increase still more. For example, guaranteeing *AQL*'s of 5 and 3 percent for slump and air content alone would require testing at least one-third of all production units (Fig. 26). This leads one to ask -- why statistical acceptance sampling plans? Many valid reasons, separately or in combination, warrant the adoption of statistical acceptance sampling plans. Some are administrative, and others involve quality considerations. Seven of the more important will now be discussed.

1. Delineation of Responsibility for Process Control and Acceptance Sampling

Acceptance sampling plans allow the buyer to specify a product's quality determinants and the limits within which those properties can be allowed to vary, and then to assure with known risks that the product meets the specifications without interferring with the production process. This is a desirable situation that can be achieved only in this way. In the absence of acceptance sampling, the buyer has two other options -- 1) to accept all products regardless of quality, or 2) to rely on process control to assure quality. It is the second option that most concrete buyers choose and almost all concrete producers resent. Table 22 shows that for concrete, most state highway departments specify both mix design and production method and then reserve the right to reject concrete that does not meet the specification when tested. This draws complaints from producers who cannot change the mix design or deviate from the specified production methods, but must accept the economic loss of rejected concrete. In some instances, this point is probably overstressed by contractors who follow neither the specified production method nor mix design, but blame both for poor results. In many other cases, however, the complaints may be justified. Regardless of their validity, this setup causes friction between inspectors and foremen, which can easily be eliminated by adopting acceptance sampling. With acceptance sampling, the concrete buyer can assure the desired quality levels regardless of what the process is doing. This allows the buyer to leave process control to the

TABLE 22
STATUS OF STATE SPECIFICATIONS*

State	Date of Spec.	Specification Type Method Only	Final Product Inspection Only	Method and Final Product Inspection	State	Date of Spec.	Specification Type Method Only	Final Product Inspection Only	Method and Final Product Inspection
Ariz.	1969		✓		Mont.	1959			✓
Ark.	1959			✓	Neb.	1955			✓
Calif.	1971		✓		Nev.	1968		✓	
Colo.	1967			✓	N. H.	1960			✓
Conn.	1969		✓		N. J.	1961		✓	
Del.	1965			✓	N. M.	1970			✓
D. C.	1957			✓	N. Y.	1962			✓
Fla.	1966			✓	N. C.	1959			✓
Ga.	1956	✓			N. D.	1961			✓
Hawaii	1969		✓		Ohio	1971			✓
Idaho	1967			✓	Oreg.	1970			✓
Ill.	1968			✓	Pa.	1970			✓
Ind.	1971			✓	R. I.	--			✓
Iowa	1960			✓	S. C.	1955			✓
Kans.	1960			✓	S. D.	1969			✓
Ky.	1956			✓	Tenn.	1968			✓
La.	1966		✓		Tex.	1962			✓
Maine	1968			✓	Utah	1970			✓
Md.	1968			✓	Vt.	1964			✓
Mass.	1967		✓		Va.	1970			✓
Mich.	1970		✓		Wash.	1957			✓
Minn.	1968			✓	W. Va.	1969			✓
Miss.	1956			✓	Wis.	1968			✓
Mo.	1961			✓	Wyo.	1960			✓

*Data not available for Ala., Alaska, Okla.

producer, where most people would agree it belongs, thus eliminating the source of complaints.

Process Control as a manufacturer's function has been long recognized by both industry in general and the U. S. Department of Defense. This recognition is based on the experience that no one can control materials and machines better than the operators. The industry experience has been such that even within the same company the functions of acceptance sampling and process control are separated. Traditionally, within the same company, the inspection department has been given the responsibility for product acceptance and the manufacturing department the responsibility for process control (3, pp. 6-2 through 6-21). Lately there have been some suggestions that both functions should be placed in the hands of the inspection department in companies with large inspection forces (32). But these suggestions are based on economic considerations rather than on the adequacy of the traditional approach. However, in industry, there is no confusion when a vendor-vendee relationship such as that between state agencies and concrete producers arises. When one company buys from another, the vendor controls the process; the buyer performs acceptance sampling. Companies are well aware of the motto "quality cannot be inspected into a product, but must be built into it," and that only the builders can control quality. They are equally aware that statistical acceptance sampling quickly discovers both good and bad quality.

These seem to be the very two points missed by state highway agencies. State agencies seem to believe that inspectors can control concrete making processes without actually operating the plants, and have so little faith in acceptance sampling as to fear that if inspectors do not control the process, no one will. They miss the point that even relatively weak acceptance sampling plans will discover out-of-specification material in most cases and that the contractor would have to exercise process control or face disastrous economic losses. To illustrate this point it is enough to refer to the *OC* curve in Figure 25, which is a relatively weak one, and see what would happen. The curve shows that if the percent defective in the incoming lots were 6 percent, the plan would reject approximately 50 percent of the lots. Similarly, if the percent defective in the incoming lots were 10 percent, 84 percent of the lots would be rejected. If the percent defective in lots were the allowable 2.3 percent, the plan would eventually reject 5 percent of the lots. This means that to assure no rejection, the material submitted for inspection must have a percentage defective less than that allowable. Under these circumstances, the contractor has no choice but to control the process.

Another hurdle for acceptance sampling seems to be the fear of increased prices resulting from improved sampling. This fear is often overemphasized and perhaps not warranted. If quality levels are set at levels achievable with present equipment, there is no reason to believe that the contractor will incur rejection very often. This would happen only if the quality levels required were drastically different from those possible with the present equipment.

2. Delineation of Acceptable Quality Level

To design any acceptance sampling plan, quality levels must be known. This means that substantial compliance is uniquely defined when the sampling plan is formulated. Substantial compliance is attained if the plan accepts the material. It is not attained if the plan rejects the material. With acceptance sampling, deciding compliance becomes automatic and uniform from project to project. This provides the bidders with a common basis for bidding. Thus, "bidding the engineer" and possible abuses of "engineering judgment" are eliminated. Moreover, since what constitutes substantial compliance is known by both the engineer and his counterpart even before the project begins, the opportunity for misunderstanding and litigation is minimized, if not eliminated.

3. Information for Reduced Payments

Concrete does not lend itself to rectification. This means that to enforce the contractual agreement regarding quality, construction units containing a percentage of out-of-specification concrete exceeding that allowable would have to be removed and destroyed. This measure is so drastic that it is almost never enforced. Removal of defective structural units is often impossible without affecting other parts of the structure, causes the waste of large amounts of concrete meeting the specifications, retards construction, and leads to legal disputes. Thus state highway agencies and concrete buyers in general are reluctant to enforce the quality agreement through concrete removal. This assures contractors of full payment for almost all concrete placed regardless of quality and provides no incentive for the contractor to improve concrete quality. A less drastic but more effective way to enforce the contractual terms regarding quality is reduced payment for reduced concrete quality. Reduced payment requires estimates of concrete quality. Acceptance sampling plans provide a uniform and valid basis for such estimates. With acceptance sampling the sample size remains constant (for concrete), allowing only one payment schedule based on the sample size of the adopted plan.

4. Uniformity of Decision Criteria

Acceptance sampling plans, regardless of type, specify in detail the criteria to be used in judging the material. This avoids confusion and disputes, because contractors and buyers are aware of the criteria before the project even begins -- at the contract signing. The engineer is no longer responsible for establishing quality rules and changing them as he sees fit, but is bound to enforce a contractual agreement. He becomes the *chief inspector*. This sounds like an infringement on engineering judgment, but in reality it is not. This is only a way of applying the best available engineering judgment uniformly on all projects. In acceptance sampling, engineering judgment still plays a vital role in choosing the boundary conditions that sampling plans must satisfy -- the *AQL*, *LTFD*, α, and β. But once used, it cannot be arbitrarily changed without valid reasons.

5. Information for Vendor Rating

Poor quality is undesirable in nearly all respects. It increases inspection costs, reduces profits, causes waste, delays manufacturing, and requires constant unpleasant decisions. To avoid all this, the Department of Defense stipulates that defense contractors must be capable of controlling their manufacturing processes and have adequate quality control facilities and personnel, as outlined in Military Standard Q-9858A. Industry takes an even dimmer view of poor quality. Companies avoid doing business with habitual producers of poor quality. To decide which vendors should be avoided, buyers in industry employ "vendor ratings" (3,33). Information for the ratings comes from the results of acceptance sampling. Whether vendor rating is appropriate for the concrete industry is a managerial decision, but if rating concrete producers is desirable, acceptance sampling provides the only systematic way to accumulate the necessary information.

6. Data for Proper Choice of Quality Levels

The previous five reasons for adopting acceptance sampling are primarily administrative, but they are not the only justifications for adopting sampling plans. Acceptance sampling is an essential element of quality assurance and its adoption is primarily warranted for quality reasons. Ideally, the main purpose of acceptance sampling is to assure the necessary degree of quality -- all others are secondary. However, in the case of concrete, there is another, equally important reason for adoption of acceptance sampling -- systematic collection of data for objective relating of quality and performance. Quality levels for concrete are not based on valid relationships between quality and performance, but on what is believed achievable or acceptable. This is a far-from-ideal way of setting quality levels, because it precludes an economic analysis to choose the most economical quality levels. To choose them for concrete, one must know how quality affects maintenance, service life, and overall performance of structures. Unfortunately, under the present inspection system, sample sizes are generally too small to judge concrete quality. As a result, the performance of innumerable structures can be observed but cannot be correlated to concrete quality. Acceptance sampling could effectively be used to correct this situation. Acceptance sampling plans require either a random sample large enough to judge quality or to limit the percent of defective material accepted to a known maximum. Thus, the results could be recorded and eventually used to establish how quality affects performance. To establish this relationship will require some time. But it is very desirable, and would provide a means to building the most economical structures, ending subjective and arbitrary setting of quality levels, and enhancing the chances for uniformity of specifications among state agencies.

7. Uniform Quality

Another advantage of acceptance sampling is uniformity of concrete from project to project. Under the present inspection

> system, sample sizes are generally too small for valid judgment of concrete quality and as a result all concrete is usually accepted. Thus, the percent of out-of-specification concrete placed depends on process output and varies from project to project, and even from structure to structure within the same project. But with acceptance sampling plans, if the percent defective in process output does not lie between the *AQL* and *LTFD* and does not approach the *AQL*, a large amount of concrete will be rejected. Regardless of whether the consequence of rejection is reduced payment or concrete removal, large economic losses result. To avoid them, producers are likely to control their processes to achieve quality levels close to the *AQL*. Thus, quality levels in time will approach the *AQL* on all projects.

B. Monetary Value

The major advantages of statistical acceptance sampling have just been enumerated and discussed. The next logical point for consideration is whether these advantages are worth the required increase in testing and sampling. Unfortunately, a valid answer to this question is impossible without first adopting some sort of acceptance sampling.

The added cost of increased inspection is worth incurring only if the quality levels assured result in savings larger than the cost of inspection. In the case of concrete, savings may possibly be realized from the reduced maintenance cost, reduced structure size, or longer service life that could accompany higher quality levels. But valid data relating quality level and performance are now limited and a comparison between costs and possible benefits of obtaining a given quality is impossible without first establishing these necessary relationships. Objectively to relate performance and quality, it is necessary that both performance and quality levels be measured. Quality can be objectively estimated only by screening or some other type of statistical sampling plan that allows valid inferences about the material from which the sample is drawn. Thus, ironically, an objective judgment of the worth of inspection must await the implementation of some sort of valid statistical sampling plans.

However, if higher quality levels can be assured without additional inspection personnel, it appears desirable regardless of the exact monetary savings that may result. Higher sampling rates and assurance of higher quality levels without more inspection personnel are not beyond the realm of possibility. Increased testing rates with present inspection forces could be achieved by reallocation of present manpower, or by deviating from present standard sampling and testing procedures. In the first case, plant inspectors could be dispatched to projects with high concrete placement rates. In the second, the air content of fresh concrete could be measured with the Chace indicator, and one rather than two cylinders could be cast from each sampled production unit. These changes would not result in poorer quality and should be considered.

The results obtained in this investigation show that Chace indicator measurements estimate the air content of concrete approximately 1.5-percent higher than corresponding high-pressure meter measurements. Other researchers have found similar results, although the discrepancy has changed from project to project. This means that if the discrepancies for particular projects are initially evaluated so that they can be taken into account, the Chace indicator could be used in conjunction with the pressure meter to increase testing rates.

Another way to improve utilization of manpower for concrete inspection is to inspect the final product in which concrete is used whenever possible. This approach, as discussed, is impractical for the inspection of structures where one structural concrete unit is interconnected with others, but seems very appropriate for the inspection of concrete pavements and feasible with present technology. Cores could be used to measure thickness, strength, and air content. The same tests now performed for roughness and texture could be retained. Strength measurements would replace the slump test, and as discussed in Appendix C the air content could be measured by the high-pressure method. On an experimental basis, the high-pressure method takes about 1/2-hour per test, but if more than one meter were available and the method were adopted for production, the time might be cut in half. This is a new approach and a marked departure from present inspection philosophy, but should nevertheless be given due consideration.

VI. ACCEPTANCE CONTROL CHARTS

In some instances, the specification range is much greater than six process standard deviations and the process output can meet specification limits even when the process mean is shifted out of control. When this occurs, there is little chance of producing defectives and it may be desirable to use the types of $\bar{X}$-charts known as acceptance control charts.

Construction of acceptance control charts requires that the process standard deviation be known. Its control limits *do not coincide* with those of $\bar{X}$-charts for process control based on the same sample sizes and standard deviations. Unlike process control charts, acceptance control charts are not designed to detect lack of process control. Their only goal is to assure with known risks that the percentage of defective output is limited to pre-established levels. In this respect, acceptance control charts resemble acceptance sampling plans with standard deviation known. In fact, double-limit acceptance control charts and double-limit sampling plans by variables with standard deviation known both require approximately the same sample size to assure the same quality levels with the same risks.

Although statistically almost identical to variable sampling plans with known standard deviation, acceptance control charts differ in concept. Acceptance control charts accept or reject a process, while acceptance sampling plans accept or reject lots. The course of action required to implement decisions of acceptance control charts is *to do something about the process*; the action required to implement decisions made with acceptance sampling plans is *to reject or accept individual lots* -- a limited amount of production. Another basic difference between acceptance control charts and acceptance sampling plans lies in the sampling. Acceptance sampling *requires* random sampling, while for acceptance control charts it is *desirable* to take the necessary sample all at once. This is because if the sample is recovered over a particular period, a change in the process that has taken place during that time may be covered up by the averaging of sample results (8, p. 435).

The limits for acceptance control charts depend on: 1) acceptable process level *APL*, 2) rejectable process level *RPL*, 3) producer's risk α, 4) consumer's risk β, 5) subgroup size n, and 6) process standard deviation σ'. These quantities have meanings analogous to the parameters necessary to design sampling plans by variables. Specifically, the *APL* is the process fraction defective that can be accepted with no adverse consequences; the *RPL* is the process fraction defective that can be barely tolerated; producer's risk is the probability of rejecting a process that is producing a fraction defective equal the *APL*; consumer's risk is the probability of accepting a process producing a fraction defective equal the *RPL*; subgroup size n is the number of consecutive units that should be tested to assure that the conditions set by specifying the *APL*, *RPL*, α, and β are met; and the process standard deviation is that used to determine the limits, approximating the standard deviation of the process measured when no assignable causes are present.

Sample size n is set once α, β, *APL*, and *RPL* are chosen, and must be calculated from these quantities before the limits for acceptance control charts can be derived. The value of n is independent of the magnitude of the standard deviation and is computed as follows:

$$n = \left[\frac{Z_\alpha + Z_\beta}{Z_{APL} - Z_{RPL}}\right]^2$$

where Z_α = the normal deviate corresponding to α,

Z_β = the normal deviate corresponding to β,

Z_{APL} = the normal deviate corresponding to *APL*, and

Z_{RPL} = the normal deviate corresponding to *RPL*.

Knowing n and σ, control charts limits are computed as follows:

$$UCL = U - Z_\alpha \sigma' - Z_\beta \frac{\sigma'}{\sqrt{n}}$$

$$LCL = L + Z_\alpha \sigma' + Z_\beta \frac{\sigma'}{\sqrt{n}}$$

where UCL = the upper control limit,

LCL = the lower control limit,

U = the upper specification limit,

L = the lower specification limit,

Z_α = the normal deviate corresponding to α,

Z_β = the normal deviate corresponding to β,

σ' = the known standard deviation, and

n = the sample size.

These limits are derived by using the specification limits as reference points, while the reference point for process control charts limits is the design or specified mean (34). This is not by accident. It is consistent with the assumption that the specification range should be greater than $6\sigma'$ to use acceptance control charts, and that the process mean can shift about the design mean without producing defectives as long as the standard deviation remains unchanged.

A. Applicability to Concrete

The objective of acceptance control charts is to reject processes whose output equals or exceeds the rejectable process level *RPL*. This objective limits their applicability to only those properties that can be measured immediately after manufacturing. For those properties that cannot be measured then, use of acceptance control charts leads to two difficulties. First, if the process shifts to the rejectable level, defectives will be produced during the time lag between production and testing. Depending on the time elapsed, this can result in accepting substantial amounts of inferior product. Second, a process operating at a rejectable process level at the time the sample is produced can shift back to an acceptable level while waiting for test results. When this happens, rejecting the process on the basis of the last available data leads to rejecting an acceptable process and causes unnecessary manufacturing delays.

Concrete properties that can be measured immediately after mixing are slump and air content, and in principle acceptance control charts can be used for these properties provided that sampling and testing are performed at the plant site. But, although applicable in theory, the use of acceptance control charts for slump and air content is neither practical nor desirable. They are impractical because no saving in testing is realized, and undesirable because conditions for the use of such charts do not exist. Acceptance control charts are desirable when: 1) the specification range is wide enough to accommodate shifts in process averages of considerable magnitude without resulting in defectives, 2) the process standard deviation is known and stable, 3) the production units included in the subgroup represent consecutive production, and 4) the decision of rejection can be enforced.

Neither slump nor air content meet these conditions. For slump and air content, the specification range approximate six standard deviations and the advantage of acceptance control charts is lost.

1. Specification Range

Acceptance control charts are used to give producers of uniform product an advantage when the specification range is considerably greater than six times the process standard deviation. If the specification range is very large, compared to the six standard deviations needed to meet the specifications, the process average can be allowed to shift considerably without producing defectives. Under these circumstances, both acceptance sampling and process control can be relaxed, as is illustrated in Figure 27. The specification range is twelve process standard deviations, but for most properties the specification limits need provide only a range of six standard deviations to eliminate nearly all defectives. Thus, the process average in Figure 27 can shift to $\pm 3\sigma'$ from the nominal design value without producing any defectives. Only when the process average moves outside the shaded area in Figure 27 does production of defectives begin, and the chances for their production are almost non-existent. Because their production is unlikely, the producer need not be very meticulous about process control. He only needs to prevent very large shifts in the process average, which usually take very little effort to avoid. Similarly, the buyer is not likely to receive defectives, and can afford to accept the material as long as the process is monitored to prevent large shifts in process level. To assure this, he can reply on acceptance control charts, using his own data or the producer's data. For slump and air content the specification ranges are six standard deviations and using acceptance control charts is not desirable.

2. Standard Deviation

In the preceding discussion, it was tacitly assumed that the process standard deviation was known. In fact, for slump and air content it changes from plant to plant, and there is no impartial way to assume a safe value. If small standard deviations are assumed, producers of unacceptable quality are rewarded and buyers penalized. If large standard deviations are assumed, producers of uniform quality will suffer unnecessary and unfair rejection. This means that to be fair in setting up acceptance control charts, the standard deviation should be determined for each concrete plant and that control charts limits would have to change from plant to plant. This would be an administrative nightmare, but to make it worse, the histograms in Figures 7A and 7C show that the needed standard deviations of 0.5 for slump and 0.5 or 0.67 for air content are seldom achieved without process control. Contractors might exercise process control during determination of the standard deviation and relax it during usual production. If that were the case, concrete buyers would accept large fractions of out-of-specification material. This would happen because accept-

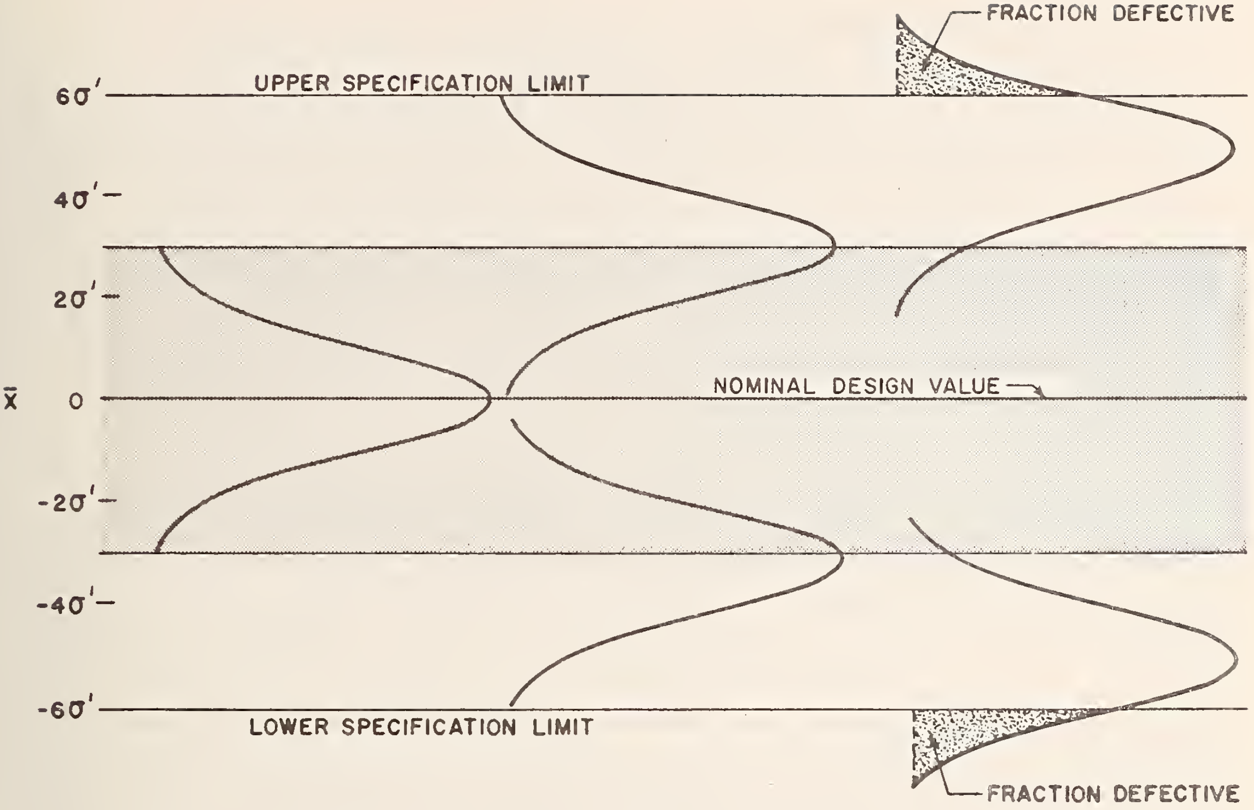


Figure 7. Magnitude of shifts in mean necessary to produce defectives.

---

ance control charts control *only* the process average while defectives can be caused by both a shift in process level and an increase in standard deviation (Fig. 8). For these reasons, acceptance control charts are again undesirable.

3. Subgroups

For acceptance control charts, the sample should consist of consecutive production units. Recovering such a sample is a difficult task in the case of concrete, even if the sample size is small. The sample for acceptance control charts depends on α, β, *APL*, and *RPL* and can be relatively large. For example, for an *APL* of 0.003 and *RPL* of 0.036, the necessary sample size is 10 if $\alpha = 0.05$ and $\beta = 0.10$. If higher quality levels were required, the sample size would be larger. These relatively large sample sizes make the recovery of samples consisting of consecutive or almost consecutive production units a difficult task. This is another drawback for acceptance control charts.

4. Enforcement of the Rejection Decision

As already noted, acceptance control charts accept or reject a process and not a finite amount of material. If the concrete buyer uses acceptance control charts, he can run into difficulties in enforcing rejection. When a process is rejected, a producer could refuse to look for assignable causes. In such a case, the buyer could not enforce his rejection decision. He could stop buying the product, but having a less demanding market the producer may not care. Since the decision does not involve material, but rather doing something totally under the producer's control, the buyer must depend on the producer's cooperation.

Within a single company, acceptance control charts can work because the producer and those responsible for process acceptance report to the same manager. In such cases disputes can be quickly resolved with no necessity for litigation. But in a vendor-vendee relationship, this arrangement can lead to problems.

5. Amount of Sampling

Finally, from the standpoint of testing and sampling, there is no advantage in using acceptance control charts. To assure the same quality levels with the same risks, acceptance control charts and sampling plans by variables with standard deviation known require the *same* sample size. In fact, the sample size is computed with the same formula. But for an acceptance sampling plan, the sample size *must* consist of a random sample. This is an advantage because sampling of consecutive concrete production units is difficult. Hence, if a point on an acceptance control chart represents the same amount of material as a lot, acceptance sampling plans by variable with standard deviation known are preferable to acceptance control charts.

B. Chart Limits for Slump and Air Content

Acceptance control charts cannot replace screening. Consequently, they could be used only for pavement concrete and then only for slump and air content. This limitation, combined with the other reasons discussed, make acceptance control charts very undesirable and they are *not* recommended for concrete inspection. However, if they are used, derivation of limits and implementation criteria should be noted.

As discussed earlier in this chapter, limits for acceptance control charts are computed as follows:

TABLE 23
ACCEPTANCE CONTROL CHART LIMITS

Property	α	β	APL, %	RPL, %	σ'	n	U	L	UCL	LCL
Air Content	0.05	0.10	0.3	3.6	0.5%	10	7%	4%	5.82%	5.14%
Slump	0.05	0.10	0.3	3.6	0.5"	10	3"	--	1.82"	--

$$UCL = U - Z_{\alpha} - Z_{\beta}\frac{\sigma'}{\sqrt{n}}$$

$$LCL = L + Z_{\alpha} + Z_{\beta}\frac{\sigma'}{\sqrt{n}}$$

The first task is to compute n and assume a value for the process standard deviation σ'. The latter varies from plant to plant and choosing a specific value is misleading. But, in the interest of a numerical illustration, σ' will be chosen as 0.5 for both slump and air content. Sample size n depends on the *APL*, *RPL*, α and β. The different mode of inspection does not change the criticality nor the required quality levels. For concrete pavement, then, risks and quality levels can be chosen as for the acceptance sampling plan recommended for inspection of slump and air content of concrete pavement. This leads to the values in Table 23, where the *APL* represents the *AQL* and the *RPL* the *LTFD* of the suggested sampling plan for slump and air content.

Using these values of α, β, *APL*, and *RPL*, n can be computed, and knowing n, the upper and lower control limits can be found as just outlined. For slump, there is only the upper limit:

$$UCL = 3.0 - 1.96 \times 0.5 - 1.28 \times \frac{0.5}{\sqrt{10}} = 1.82$$

For air content, both limits are necessary:

$$UCL = 7.0 - 1.96 \times 0.5 - 1.28 \times \frac{0.5}{\sqrt{10}} = 5.82$$

$$LCL = 4.0 + 1.96 \times 0.5 + 1.28 \times \frac{0.5}{\sqrt{10}} = 5.14$$

The control limits are shown in Figure 28B and the *OC* curve for these control charts is given in Figure 28A. This curve applies to both properties because the *APL* is very small.

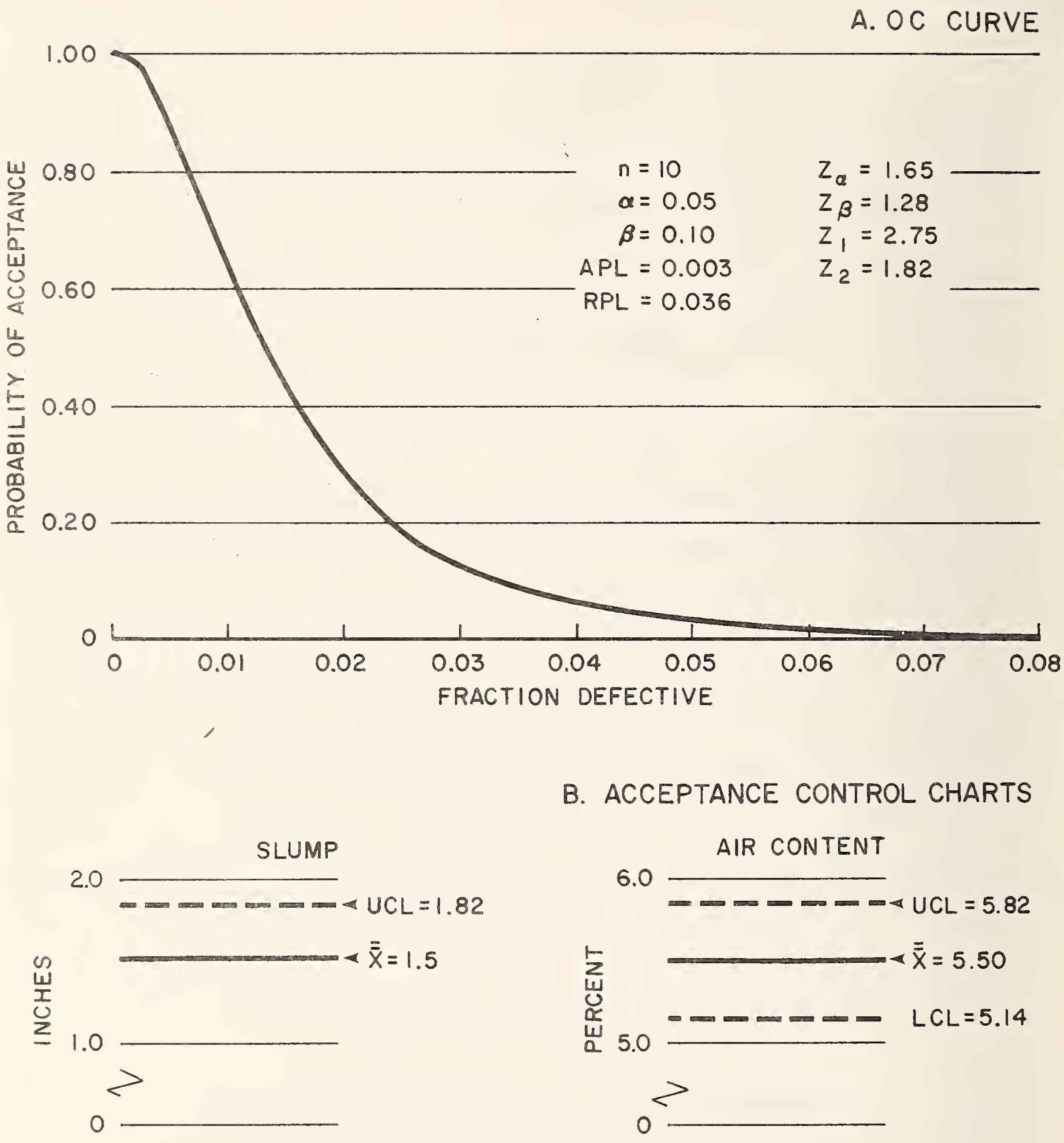


Figure 28. Limits and *OC* curves for acceptance control charts equivalent to the sampling plan n = 35 (Fig. 23), if standard deviation is known.

Implementation should consist of the following steps:

1. Sample and test 10 consecutive or quasi-consecutive production units of concrete.
2. Compute the mean slump and air content.
3. Plot the results on the appropriate control chart in Figure 28B, and
4. If a point on either or both charts plots outside the control limits reject the process and otherwise continue production.

It is stressed that the sample size cannot be chosen arbitrarily, but determined by the desired quality levels and risks. An arbitrary value of n does not -- except by coincidence -- assure the desired quality levels. This point is often missed by those who suggest acceptance control charts for construction materials, but assume a sample size that often leads to very loose *OC* curves and compute the limits in the same way as for process control charts.

CONCLUSIONS AND IMPLEMENTATION CONSIDERATIONS

A. Conclusions

Many conclusions have been drawn in the preceding chapters. Some are specific and of limited value to anyone but researchers. Others are far-reaching, deserve the attention of management, and are restated here in an order corresponding to the report's chapters.

1. Concerning Analysis of Variance

a. Both testing and sampling variations are significant parts of the total variation measured for concrete properties. But although significant, combined sampling and testing variations account only for 40 percent or less of the total. The major portion of variation in concrete properties is materials variation and is therefore reducible through process control.

b. The overall variability of concrete properties changes sufficiently from project to project that a constant standard deviation cannot be assumed for all concrete produced.

c. Concrete produced in a single central mixer or paver is likely to be more uniform than concrete mixed in several truck mixers.

2. Concerning Process Control

a. Statistical process control is the function of producers.

b. No statistical process control was in use in any concrete plant visited during this study.

c. Three case studies showed that the current New York State specifications for slump and air content can be met. This is evidence that with statistical process control most plants should be able to comply with these specifications.

3. Concerning Acceptance Sampling

a. The acceptable quality level or allowable fraction defective implied in the New York State specification for concrete is zero. According to sampling theory, to assure that no defective concrete is accepted, every production unit should be tested. Since Materials Method 9.2-3 calls for testing only a small proportion of production units, this constitutes a contradiction.

b. Of the most commonly used sampling plans, lot sampling and screening appear the most appropriate for concrete inspection. Lot sampling is most appropriate for inspection of pavement concrete, while screening is most appropriate for structural concrete and concrete placed in small quantities.

c. Continuous sampling is theoretically ideal for the inspection of slump and air content, but very difficult to implement because it requires more sampling than lot plans, and at times requires screening, making the inspection load variable and unpredictable which leads to difficulties in allocating inspection personnel.

d. Theoretically the sampling rates specified in New York State Materials Method 9.2-3 are almost equivalent to no sampling at all when the fraction defective due to one or more properties in the concrete inspected is less than 15 or 20 percent (Fig. 21).

e. Statistically based sampling plans for pavements, and screening for structures and small pours, must be used if low percentages defective (say, less than 5 percent) are to be assured.

f. The sampling procedure described in Appendix A of Materials Method 9.2, which calls for sampling only the first portion of truck mixers, is superior to the ASTM standard sampling procedure for truck mixers because it allows judging concrete before its placement into inaccessible forms. As explained in Appendix E, the loss of precision associated with sampling only one portion of truck mixers can lead to misjudging an additional minute fraction of the total production inspected. But this fraction is negligible and the practice of sampling only one portion of truck mixers should be retained.

g. Lack of maintenance cost data, combined with lack of materials quality data, make determination of the effect of materials quality on performance virtually impossible. This in turn makes impossible any determination of the monetary value of concrete inspection (see p. 110).

4. Concerning Acceptance Control Charts

Acceptance control charts are not appropriate for concrete inspection.

B. Implementation Considerations

This review of quality assurance of portland cement concrete reveals deficiencies in the philosophy and/or implementation of concrete inspection. Chief among these deficiencies are: 1) the emphasis placed on process control instead of acceptance sampling, 2) relegating inspection to engineering personnel with inadequate training in statistical quality assurance, and 3) lack of knowledge of what constitutes acceptable quality levels. To correct these deficiencies the Department must do the following:

1. Consider implementing screening for all structural concrete and small pours,

2. Consider implementing sampling plans of the type recommended in this report for the inspection of pavement concrete,

3. Institute a rational system of data analysis in order to correlate inspection results with performance, so that quality levels can be chosen on the basis of economic analysis,

4. Consider adopting a strength specification and inspecting hardened concrete pavements to take advantage of flexibility of scheduling afforded by inspecting the final product,

5. Review the present method of allocating manpower for concrete inspection to eliminate process control functions assumed by the state,

6. Provide formal training in quality assurance for project engineers and inspectors, and

7. Continue research to improve present testing tools, and develop new testing instruments and techniques to reduce testing time and cost.

REFERENCES

1. Miller, I., and Freund, J. E. Probability and Statistics for Engineers. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1965, p. 335.

2. The Statistical Approach to Quality Control in Highway Construction. Bureau of Public Roads, U. S. Department of Transportation, April 1965.

3. Juran, J. M., Seder, L. A., and Gryna, F. M., Jr. Quality Control Handbook. Second Edition. New York: McGraw-Hill Book Co., 1962.

4. Davies, O. L. The Design and Analysis of Industrial Experiments. New York: Hafner Publishing Co., 1956, pp. 106-115, 286-289.

5. Amsler, D. E., Eucker, A. J., and Chamberlin, W. P. Techniques for Measuring Air Void Characteristics of Concrete. Research Report 11, Engineering Research and Development Bureau, New York State Department of Transportation, January 1973.

6. Kemp, K. W. "The Average Run Length of Cumulative Sum Charts When a V-Mask is Used." Journal of the Royal Statistical Society, Series B, Vol. 23, No. 1 (1961), pp. 149-153.

7. Kemp, K. W. "The Use of Cumulative Sums for Sampling Inspection Schemes." Applied Statistics, Vol. 11, No. 1 (March 1962), pp. 16-31.

8. Duncan, A. J. Quality Control and Industrial Statistics. Third Edition. Homewood, Illinois: Richard D. Irwin, Inc., 1965.

9. King, E. P. "The Operating Characteristics of the Average Chart." Industrial Quality Control. Vol. 9, No. 3 (November 1952), pp. 30-32.

10. King, E. P. "The Operating Characteristic of the Control Chart for Sample Means." Annals of Mathematical Statistics, Vol. 23, No. 3 (September 1952), pp. 384-395.

11. Smith, P., and Tiede, H. "Early Determination of Concrete Strength Potential." Highway Research Record 210, 1967, pp. 29-66.

12. Miller, R. W., and Chamberlin, W. P. *Prediction Equations for Accelerated Strength Testing of Concrete*. Engineering Research and Development Bureau, New York State Department of Transportation, November 1970. Submission to ASTM Committee C-9, Subcommittee II-i (Accelerated Strength Testing).

13. Dodge, H. F. "Notes on the Evolution of Acceptance Sampling Plans, Part I." *Journal of Quality Technology*, Vol. 1, No. 2 April 1969), pp. 77-88.

14. Banzhaf, R. A., and Brugger, R. M. "MIL-STD-1235 (ORD), Single- and Multi-Level Continuous Sampling Procedures and Table for Inspection by Attributes." *Journal of Quality Technology*, Vol. 2, No. 2 (January 1970), pp. 41-53.

15. Dodge, H. F. "A Sampling Inspection Plan for Continuous Production." *Annals of Mathematical Statistics*, Vol. 14, No. 3 (September 1943), pp. 264-279.

16. Dodge, H. F., and Torrey, M. N. "Additional Continuous Sampling Inspection Plans." *Industrial Quality Control*, Vol. 7, No. 5 (March 1951), pp. 7-12.

17. Lieberman, G. J., and Solomon, H. "Multi-Level Continuous Sampling Plans." *Annals of Mathematical Statistics*, Vol. 26, No. 4 (December 1955), pp. 686-704.

18. Albrecht, L., Gulde, H., MacLean, A., and Thompson, P. "Continuous Sampling at Minneapolis-Honeywell." *Industrial Quality Control*, Vol. 12, No. 3 (September 1955), pp. 4-9.

19. Dodge, H. F., and Romig, H. G. *Sampling Inspection Tables: Single and Double Sampling*. New York: John Wiley & Sons, 1944.

20. Stralkowski, C. M. *Computer Solutions for the OC and ASN Curves of Six Sampling Plans*. Unpublished master's thesis, University of Wisconsin (Madison), 1964.

21. Cochran, W. G. "The χ^2 Test of Goodness of Fit." *Annals of Mathematical Statistics*. Vol. 23, No. 3 (September 1952), pp. 315-345.

22. "Evaluation of Compression Test Results of Field Concrete Reported by ACI Committee 214." Title 52-17 (Proc. Vol. 52), *Journal of the American Concrete Institute*, Vol. 27, No. 3 (November 1955), pp. 241-257.

23. Abdun-Nur, E. A. "How Good is Good Enough?" Title 59-2 (Proc. Vol. 59), Journal of the American Concrete Institute, Vol. 59, No. 1 (January 1962), pp. 31-46.

24. Abdun-Nur, E. A. and Waddell, J. J. "Control of Concrete Mixes." Title 55-61 (Proc. Vol 55), Journal of the American Concrete Institute, Vol. 30, No. 9 (March 1959), pp. 947-962.

25. Larson, T. D., Cady, P. D., and Price, J. T. "Review of a Three-Year Bridge Deck Study in Pennsylvania." Highway Research Record 226, Highway Research Board, pp. 11-25.

26. Finney, E. A. "Preventive Measures for Obtaining Scale-Free Concrete Bridge Structures." HRB Bull. 323, 1962, pp. 26-42.

27. Grieb, W. E., Werner, G., and Woolf, D. O. "Resistance of Concrete Surfaces to Scaling by De-Icing Agents." HRB Bull. 323, 1962, pp. 43-62.

28. Concrete Bridge Deck Durability. NCHRP Synthesis of Highway Practice 4, Highway Research Board, 1970.

29. Durability of Concrete Bridge Decks: A Cooperative Study, Report 3. Publication T107C, Portland Cement Association, 1967.

30. Metcalf, J. B. The Specification of Concrete Strength: Part I: The Statistical Implications of Some Current Specifications and Codes of Practice. Report LR 299, Road Research Laboratory, Ministry of Transport, 1970.

31. Juran, J. M., and Gryna, F. M., Jr. Quality Planning and Analysis From Product Development Through Usage. New York: McGraw-Hill Book Co., 1970, pp. 344-346.

32. Tucker, R. J. W. "Quality Improvement By Elimination of the Traditional Inspection Department." Transactions, 24th Annual Technical Conference, American Society for Quality Control, 1970, pp. 237-242.

33. Papariella, F. A. "Quality Suppliers are V.I.P." Transactions, 24th Annual Technical Conference, American Society for Quality Control, 1970, pp. 225-228.

34. Freund, R. A. "Acceptance Control Charts." Industrial Quality Control, Vol. 14, No. 4(October 1957), pp. 13-23.

APPENDICES

A. Glossary
B. Uniformity Study Field Project Summary
C. Air-Content Correlations
D. High-Pressure Air Meter Operating Procedure
E. Sampling Truck-Mixed Concrete

A. GLOSSARY

To avoid digressions, some abbreviations, symbols, and terms were introduced in the text of this report without dwelling on their meanings. This appendix provides definitions to help engineers relate engineering terms to the statistical terminology. For this reason, some definitions do not strictly conform to usage among statisticians. For more precise definitions of these terms, the interested reader is referred to American Society for Quality Control Standards A1, A2, and Proposed Standard A3, as well as to Juran (3) and Duncan (8).

1. Symbols

α = producer's risk

β = consumer's risk

c = acceptance number

c_1 = acceptance number for the first sample of a multiple-sampling plan

c_2 = acceptance number for the combined first and second sample of a multiple sampling plan

d_2 = adjustment factor used to estimate the universe standard deviation σ in the formula $\sigma = \frac{R}{d_2}$

d_3 = reciprocal of the standard deviation of the relative range; a tabulated value used in control charts

H_0: = null hypothesis

H_1: = alternate hypothesis

K = critical value in a single-limit acceptance sampling plan by variable

K^* = critical value in a double-limit acceptance sampling plan by variables

m = number of subgroups

n = sample size or subgroup size

N = lot size

P_a = probability of acceptance

r = coefficient of correlation

R = range

$\overline{R}$ = average range

S = sample standard deviation

S^2 = sample variance

σ = universe standard deviation

$\bar{X}$ = mean

$\bar{\bar{X}}$ = grand mean, or weighted mean of several means.

2. Terms and Abbreviations

ACCEPTABLE PROCESS LEVEL (*APL*) -- The process level farthest from standard still yielding product quality, in terms of either percent defective or deviations from standard, that is desired to accept $1-\alpha$ of the times it occurs.

ACCEPTABLE QUALITY LEVEL (*AQL*) -- The average fraction defective considered acceptable by the consumer. The fraction defective for which the probability of acceptance is $1 - \alpha$.

ACCEPTANCE CONTROL CHART -- A chart used for acceptance of a process.

ACCEPTANCE NUMBER -- The number of defective items, which cannot be exceeded in the sample, in lot sampling by attributes.

ACCEPTANCE SAMPLING -- Sampling by the consumer to determine whether the product meets his specifications.

ACCEPTANCE SAMPLING PLAN -- An acceptance sampling procedure that specifies a) sample size, b) sampling procedure, c) acceptance procedure, and can be characterized by its operating characteristic curve.

ACCURACY -- As used here, the deviation of an estimate from the true value (see also PRECISION).

ANALYSIS OF VARIANCE (*ANOVA*) -- A technique to determine whether the means of a specified classification differ significantly.

ASSIGNABLE CAUSES -- Factors contributing to variation in quality that can be detected by statistical methods. Commonly assignable causes must be identified and removed to attain statistical control.

AVERAGE FRACTION INSPECTED (*AFI*) -- The average number of production units that would have to be tested in continuous sampling for a given incoming material's quality.

AVERAGE OUTGOING QUALITY LIMIT (*AOQL*) -- The maximum percentage defective in a product accepted under a continuous sampling plan.

AVERAGE OUTGOING QUALITY LEVEL CURVE -- A curve giving the average fraction defective in the accepted product as a function of the percent defective in material submitted to inspection.

AVERAGE SAMPLE NUMBER (*ASN*) -- The average number of production units that must be tested, under multiple sampling, to reach a decision for a given lot fraction defective; each lot fraction defective has an average sample number.

AVERAGE SAMPLE NUMBER CURVE -- A curve giving the average number of production units needed to reach a decision in multiple sampling as a function of the lot fraction defective.

BINOMIAL PROBABILITIES -- The probabilities associated with a binomial distribution; tabulated in most mathematical tables.

CAPABILITY STUDY -- see PROCESS CAPABILITY STUDY

CHI-SQUARE (χ^2) GOODNESS-OF-FIT TEST -- A statistical procedure used to test hypotheses about the distribution forms of groups of data.

CONFIDENCE INTERVAL -- A interval designating the chance of including the universe value.

CONSUMER'S RISK (β) -- The probability that a lot of rejectable quality will be accepted.

CONTROL CHART -- A graphical chart with the control limits and plotted values of a given statistical measure, such as the mean, standard deviation, range, etc.

CRITICALITY -- Classification of various specified product properties according to their effects on performance, safety, and product durability.

CURTAILED INSPECTION -- Ending sampling immediately after the number of permissible defectives in the sample is exceeded.

DEFECTIVE -- Not conforming to specifications.

DISCRIMINATING -- A sampling plan's ability to discriminate between *AQL* quality and *LTFD* quality; the steepness of the *OC* curve.

DISTRIBUTION FORM -- The form of density distribution that a variable assumes.

EFFICIENCY (of a sampling plan) -- The ability of the plan readily to discover lots containing a lot fraction defective greater than the acceptable quality level (*AQL*).

ENGINEERING LIMITS -- As used in this report, engineering limits are those within which a property can be allowed to vary for the product still to perform the intended engineering function.

F-TEST -- A statistical test used to determine whether two variances or two components of variance are significantly different.

FIXED PROPORTION SAMPLING -- Sampling only a fixed proportion of production units or total production; a type of sampling that for small lots results in a very inefficient *OC* curve.

FRACTION DEFECTIVE -- The percentage defective (out-of-specifications) in a lot or universe.

HOMOGENEITY OF VARIANCE -- Statistical equality of variance.

HYPOTHESIS -- An assumption about the form of a population or its parameters.

INSPECTION -- Acceptance sampling.

INHERENT VARIABILITY -- Variability that cannot be attributed to any one assignable cause and cannot be eliminated.

LOT -- A number of production units or an amount of material designated to be represented by a sample for inspection purposes.

LOT FRACTION DEFECTIVE (*LFD*) -- The percentage defective in a lot.

LOT TOLERANCE FRACTION DEFECTIVE (*LTFD*) -- The maximum fraction defective considered by the consumer to be allowable in a lot; that lot fraction defective for which the probability of acceptance is β.

MEAN -- The arithmetic average of a variable.

NORMAL DISTRIBUTION -- A continuous frequency distribution, defined by its mean $\bar{X}$ and its standard deviation σ, according to the equation

$$f(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left(\frac{(X-\bar{X})^2}{2\sigma^2}\right)}$$

NORMALITY -- The condition for which a variable assumes the form of a normal distribution for its density distribution.

OPERATING CHARACTERISTIC (*OC*) CURVE (of a sampling plan) -- A curve relating the probability of acceptance for each lot to the lot fraction defective.

PARAMETER -- A constant or coefficient of a universe that describes some characteristic of its distribution.

POOLING OF VARIANCE -- The combining of sums of squares to get a combined mean square; the statistical averaging of variances.

POWER OF A SAMPLING PLAN -- The ability of a sampling plan to discriminate between good and bad quality; see also EFFICIENCY.

PRECISION -- The standard deviation of a variable -- the smaller the standard deviation, the higher the precision.

PRECISION OF MEASUREMENTS -- The standard deviation of repeated measurements; the smaller the standard deviation, the higher the precision (note that this is not the same as ACCURACY).

PREDICTION LIMITS -- Limits for the actual value of a dependent variable Y_i associated with a specific value of the independent variable.

PROCESS CAPABILITY STUDY -- As used in this report, this is the physical manipulation of a process as well as the analysis of data necessary to remove all assignable causes of variation from the process output. For a more restrictive definition, see Juran (3, pp. 11-15).

PROCESS CONTROL -- Both the physical manipulation of machines and materials and the data analysis necessary to keep the process in a state of statistical control.

PRODUCER'S RISK (α) -- The probability that a lot of acceptable quality may be rejected.

PROPERTY LEVEL -- The mean of the measured property.

QUALITY ASSURANCE -- The activities necessary to provide assurance that the overall quality control job is in fact being effectively carried out; it involves continued evaluation of the effectiveness of the quality control program and includes process control, acceptance sampling, and the feedback necessary for proper setting of specification limits.

QUALITY DETERMINANTS -- Those properties determining the quality of the product, or those properties that the product must possess to function as intended.

QUALITY LEVEL -- The fraction defective in the material or product; the lower the fraction defective, the higher the quality level.

RANGE -- The difference between the smallest and largest measurement in a set of data.

RANDOM SAMPLE -- A sample drawn from a universe in such a manner that each unit in the universe has an equal chance of being chosen.

RATIONAL SUBGROUP -- A number of consecutive production units among which the variation is as small as possible; a number of production units whose variation is not due to assignable causes.

RECTIFICATION OF A LOT -- Screening defective units or material from a lot, and replacing them with an equal number of acceptable units or an equal amount of material.

REJECTABLE PROCESS LEVEL (*RPL*) -- The process level closest to standard that yields product quality desired to reject $1-\beta$ of the times it occurs.

SAMPLE -- see RANDOM SAMPLE.

SAMPLING PLAN -- see ACCEPTANCE SAMPLING PLAN.

SAMPLING SCHEME -- A sampling procedure that cannot be classified as a statistical acceptance sampling plan.

SEQUENTIAL SAMPLING MASK -- The boundary lines associated with the acceptance or rejection decision for a sequential sampling plan.

SCREENING -- Inspecting each production unit or identifiable portion of material in the lot (universe).

SELECTIVE RESAMPLING -- Resampling only when a defective is found.

SIGNIFICANCE LEVEL or LEVEL OF SIGNIFICANCE (α) -- The probability of committing Type I error.

SKEWED -- Asymmetrical.

STANDARD DEVIATION -- A measure of variability.

STANDARD ERROR OF ESTIMATE -- The standard deviation of the deviations from a curve plane or surface of regression; not to be confused with standard deviation.

STATISTIC -- A quantity computed from a sample to estimate a parameter; see also PARAMETER.

SUBGROUP -- see RATIONAL SUBGROUP.

SUBGROUP SIZE -- The number of production units included in a rational subgroup.

SUBGROUPING -- Dividing production segments into rational subgroups.

t-TEST -- A statistical test to determine if the differences between means are significant.

V-MASK -- A V-shaped mask used in cusum charts.

WARNING LIMITS -- Limits on control charts to warn when the process is drifting out of control, usually 2σ limits.

VARIANCE -- The square of the standard deviation.

TABLE 24
GENERAL JOB INFORMATION

Job	Paving Dates (1967)	Concrete Use	Mixing Method	Transport		Delay Time in Discharging	Placement Method	Air-Entraining Agent	Other Admixture	Aggregate Types		Temperature Ranges, F	
				Method	Distance					Fine	Coarse	Concrete	Air
1	6/5 to 6/19	pavement	central-mixed	truck mixers	5 mi	10 min	screw spreader (two layers)	Sika	none	natural sand	crushed gravel (2-1/2" max)	74 to 92	60 to 98
2	6/21 to 6/29	pavement (base course)	central-mixed	side dump truck	1 to 2 mi	none	slide spreader	Sika	none	natural sand	crushed dolomitic limestone (2-1/2" max)	74 to 88	68 to 70
3	7/6 to 7/13	pavement	E-34 paver	dry-batch, dump trucks to mixer	1 to 2 mi	5 min	screw spreader	Airecon	none	red & grey sandstone	crushed sandstone (2-1/2" max)	70 to 86	65 to 86
4	7/20 to 8/7	structural	central batch, truck-mixed	truck mixers	1 mi	5 min	crane bucket	Sika	Plastiment	granitic weathered sandstone, shale, Graywacke sandstone	crushed stone (1-1/2" max)	74 to 89	70 to 93
5	8/11 to 8/17	structural	central batch, truck-mixed	truck mixers	6 to 7 mi	5 min	crane bucket	Daravair	none	sandstone, dolomite, limestone	crushed gravel (1-1/2" max)	76 to 92	72 to 96
6	8/15 to 8/23	pavement	central-mixed	dump trucks	1 to 2 mi	10 min	slipform paver	Daravair	none	sandstone, dolomite, limestone	crushed gravel (1-1/2" max)	70 to 83	75 to 92
7	8/29 to 9/1	pavement	E-34 paver	dry batch, dump trucks to mixer	9 to 10 mi	5 min	screw spreader	Sika	none	quartz sand	crushed traprock (2-1/2" max)	64 to 92	57 to 86
8	9/9 to 9/20	pavement	E-34 paver	dry batch, trucks to mixer	20 to 25 mi	5 min	screw spreader	Daravair	carbon black	quartz sand	crushed limestone (2-1/2" max)	72 to 84	68 to 90
9	9/25 to 10/2	pavement	central batch truck-mixed	truck mixers	9 to 10 mi	5 min	slide spreader	unknown	unknown	weathered shaley-slatey material	crushed limestone (2-1/2" max)	70 to 83	62 to 78
10	10/17 to 10/24	structural	central batch, truck-mixed	truck mixers	4 to 5 mi	10 min	crane bucket	Daravair	none	granitics & schists	crushed gravel (1-1/2" max)	56 to 77	38 to 75

B. UNIFORMITY STUDY FIELD PROJECT SUMMARY

1. Testing Procedure

Except as noted, sampling and testing of concrete was performed according to applicable ASTM standard sampling and testing procedures or the manufacturer's recommendations, as follows:

Property	Test Procedure
Slump	ASTM C 143-66
Air content (pressure method)	ASTM C 231-62
Compressive strength	ASTM C 39-66
Unit weight	ASTM C 138-63
Air content (Chace indicator)	Manufacturer's recommendations

The determination of unit weight complied with ASTM C 138-63 in all respects, except that a 1/4 cu ft container was used instead of the 1/2 cu ft recommended by ASTM.

2. General Job Information

Table 24 briefly summarizes production methods and other job conditions for each job visited.

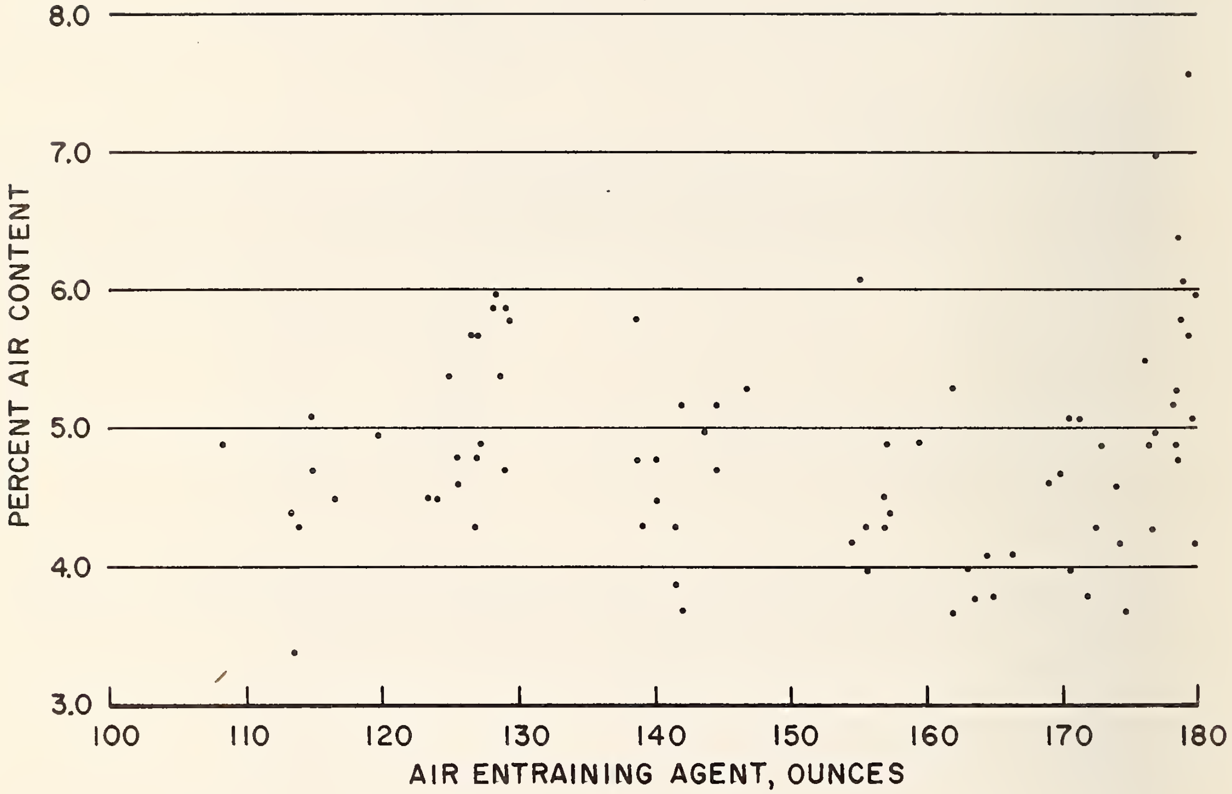


Figure 29. Plastic concrete air content related to amount of air-entraining agent introduced into each batch (all batches were 8-1/2 cu yd).

C. AIR-CONTENT CORRELATIONS

1. Air Content Versus Dosage of Air Entraining Agent

The air-entraining agent introduced into a concrete batch is usually measured faster and more easily than the air content of the corresponding fresh concrete. If the air content could be predicted from the dosage of air-entraining agent, a larger number of production units could be tested with the same manpower. The increased testing would increase the efficiency of inspection. To investigate the possibility of predicting air content from air-entraining agent dosage, data were collected on a paving project during 1968 construction season. The project was supplied with central-mixed concrete delivered by side-dumps. The amount of air-entraining agent was measured for 82 batches of 8-1/2 cu yd each at the plant. Air content of the same batches was measured in the field after discharge from the hauling units, following the appropriate ASTM sampling and testing procedures. In addition, slump, temperature, and sand and cement content were recorded for each batch. Plots were prepared relating air content to dosage of air entraining agent, slump, temperature, sand content, and cement content. The plots suggested no correlation between air content and any one of these variables. The plot showing air content versus the dosage of air entraining agent is shown in Figure 29. This plot is typical of all the others and is included as a sample. Consistent with the plots, a multiple correlation analysis revealed no correlation between the variables mentioned and air content. The partial coefficients of correlation were poor and the multiple correlation coefficient was 0.019. Figure 29 suggests that the response to a dosage of air entraining agent ranging from 13.0 to 21.2 oz per cu yd is approximately the same. This means that to detect any difference in air content, the dosage must be increased drastically, but it does *not* mean that the dosage of air entraining agent has no effect on the air content. Correlation between these two variables was not found only because the dosage of air entraining agent was not varied over a large enough range.

Knowledge of the relationship between dosage of air entraining agent and corresponding air content, developed under varied conditions, is the key to production of concrete with proper air entrainment. It deserves more than the passing attention given to it here. But determination of this relationship is likely to require well-designed capability studies that could not be included in the scope of this study. New York State does not own concrete plants that can be operated on a non-production basis for experimental purposes, and it is doubtful that any producer would participate in studies of this nature knowing that production would be impaired. It is also questionable that state agencies should undertake studies that mainly benefit the producer. Ideally, all that state agencies have to do is to specify the air content desired and leave its production to the manufacturer. But, in reality, inspection agencies would reap large benefits if inspection manpower could be reduced to assure a given quality level.

2. Air Content of Hardened Concrete and Corresponding Plastic Concrete

It is often argued that inspection (a term here synonymous with acceptance sampling) of the final product is preferable to the inspection of component materials or of the product in an intermediate state in the manufacturing process. For concrete, final product inspection implies determining air content of the hardened concrete. The air content of set concrete can be determined by microscopic techniques, such as the linear traverse or the high-pressure method. Microscopic techniques are too time-consuming to be used effectively for inspection. The high-pressure method is much faster. With this method, 15 specimens can be processed in an 8-hour day by two men and it shows some promise as an inspection tool. For this reason, it is of interest to determine the degree of correlation between results obtained by this method and those obtained by the conventional pressure meter used to measure air content of fresh concrete.

To investigate the correlations between results obtained by the two methods, cylinders from the original ten jobs (Chapter I) that were not too badly damaged in strength determinations were processed with the high-pressure meter, and the air contents thus measured were correlated with those of corresponding fresh concrete. The procedure and the equipment for determining cylinders air content are described in Appendix D, where a sample sheet is given showing the necessary calculations. (No ASTM standard exists for this method of air content determination.)

The results of the correlations ranged from mediocre to excellent. Figure 30 shows the results obtained for Jobs 6 and 9 -- the worst and best -- and those for other jobs are given in Table 25. The variances obtained for air content by the two methods can be considered equal for all but Job 8. Mean air contents as measured by the two methods, however, are not statistically equal for most jobs and the average hardened concrete air content is lower in eight out of ten cases (Table 25).

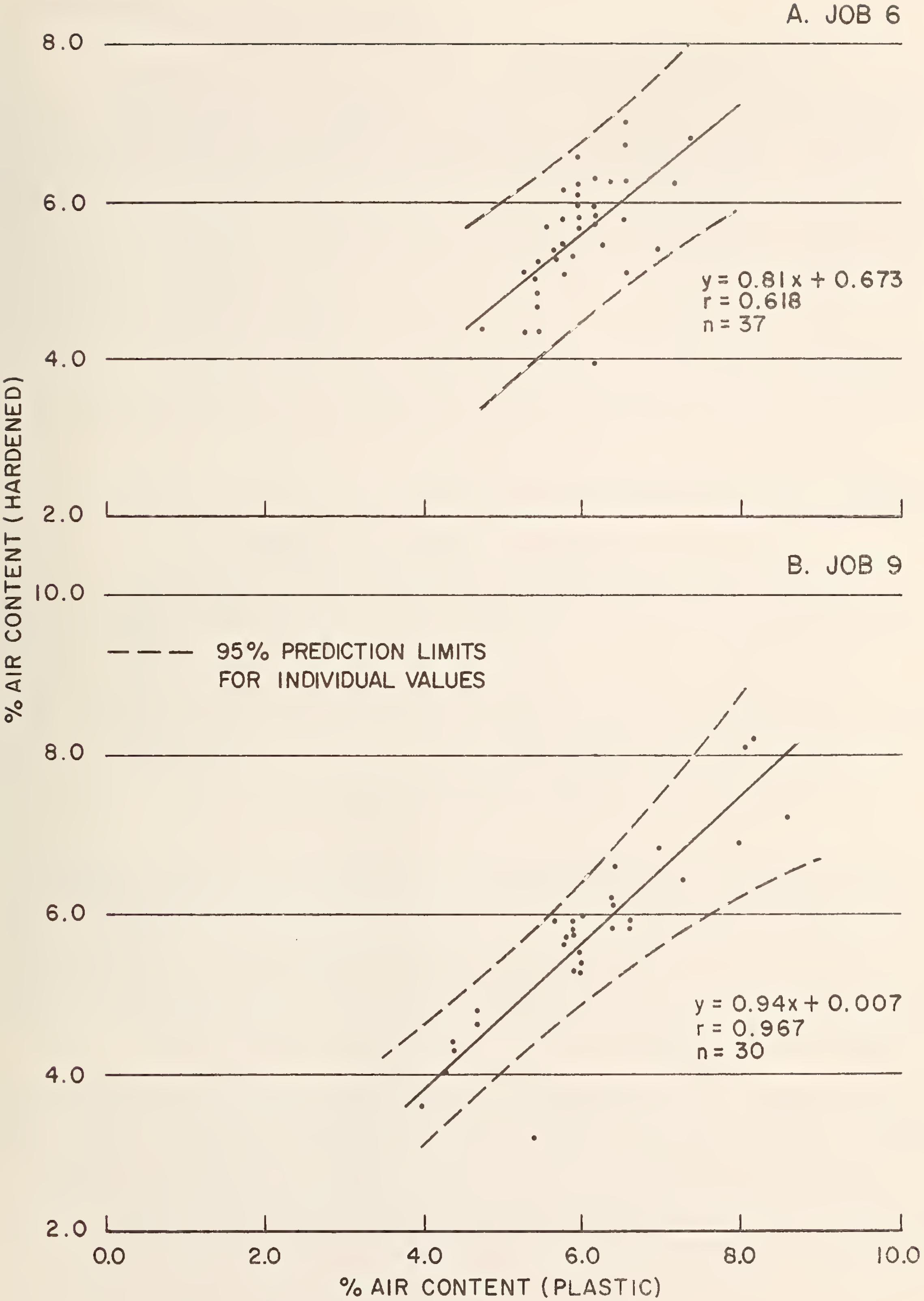


Figure 30. Air content of plastic concrete measured by the pressure method and hardened concrete measured by the high-pressure air meter.

TABLE 25
CORRELATIONS OF AIR CONTENTS OF CYLINDERS AND PLASTIC CONCRETE
Hypothesis Tested at 0.05 Level of Significance

Job	Data Points	Plastic Concrete Air Content (Pressure Method)		Cylinder Air Content (High-Pressure Air Meter)		F-Test*		t-Test**		Correlations	
		Mean	Variance	Mean	Variance	Cannot Reject H_0:	Rejects H_0:	Cannot Reject H_0:	Rejects H_0:	Correlation Coefficient r	Fraction of Variance Removed r^2
1	38	4.45	0.5712	4.42	0.5010	✓		✓		0.8061	0.6499
2	38	6.26	0.6998	7.16	0.7094	✓			✓	0.7766	0.6031
3	39	5.04	0.8882	5.62	1.0703	✓			✓	0.6722	0.4518
4	36	8.35	1.5791	7.36	1.7650	✓			✓	0.7897	0.6236
5	40	6.03	1.4237	5.71	1.3756	✓		✓		0.7701	0.5930
6	37	6.02	0.3040	5.54	0.5204	✓			✓	0.6183	0.3823
7	29	5.65	0.8066	4.89	0.6332	✓			✓	0.6919	0.4787
8	26	4.95	0.5106	4.36	1.0687		✓		✓	0.7623	0.5811
9	30	5.94	2.2363	5.65	2.1594	✓		✓		0.9665	0.9342
10	35	6.22	0.7952	5.61	1.1446	✓			✓	0.8571	0.7346

*H_0: variance of cylinder air content = variance of fresh concrete air content.

**H_0: mean of cylinder air content = mean of fresh concrete air content.

TABLE 26
COMPARISONS OF AIR CONTENTS OF CORES AND PLASTIC CONCRETE
Hypothesis Tested at 0.05 Level of Significance

Job	Core Air Content (High-Pressure Air Meter)			Plastic Concrete Air Content (Pressure Method)			F-Test*		t-Test**		
	n	Mean	Variance	n	Mean	Variance	Cannot Reject H_0:	Rejects H_0:	Cannot Reject H_0:	Rejects H_0:	$X_1 - X_2$
1	12	5.85	0.6942	54	4.55	0.9110	✓			✓	1.30
2	21	8.77	1.8986	50	5.91	0.6661		✓		✓	2.86
3	20	5.95	2.5680	50	5.14	0.5958		✓		✓	0.81
6	14	6.19	2.8920	50	6.17	0.5033		✓	✓		0.03
7	15	6.29	0.6790	50	4.94	0.5460	✓			✓	1.35
8	17	5.38	0.7585	50	4.82	0.7876	✓			✓	0.56
9	10	7.37	2.3685	50	5.80	2.5752	✓			✓	1.57

*H_0: mean of core air content = mean of plastic concrete air content.

**H_0: variance of core air content = vairance of plastic concrete air content.

No substantiated explanations can be offered for the varied degrees of correlation obtained for the different jobs, and only two conclusions appear valid. First, it seems that variability of the measurements is the same for both methods, implying that the two tests are of the same order of precision. It also appears that the mean air content of fresh concrete can be approximated within 0.5 percent by measuring the air content of a random sample of cylinders.

Table 26 shows the differences in mean air content obtained by measuring fresh concrete and a random sample of cores from the same site by the high-pressure meter. The differences are larger than for the cylinders, but this would be expected since the cores came from different batches than the concrete tested while plastic and there is no one-to-one correspondence.

D. OPERATING PROCEDURE FOR THE HIGH-PRESSURE AIR METER

1. Sample Preparation

a. Dry the sample in an oven at 290 F ± 10 deg for at least 72 hours.

b. Cool the sample at room temperature for 4 hours.

c. Weigh the sample and record the weight on the first line of the data sheet (Fig. 31).

d. Immerse the sample in water for at least 48 hours.

e. Weigh the sample in water and record the weight on the third line of the data sheet.

f. Weigh the sample in air at a saturated surface dry (SSD) condition and record the weight on the second line of the data sheet.

f. Test the sample in the high-pressure air meter.

2. Operation of the High-Pressure Air Meter

a. Close the following:

(1) Pump on-off control.

(2) Chamber drain valve.

(3) Chamber overflow valve.

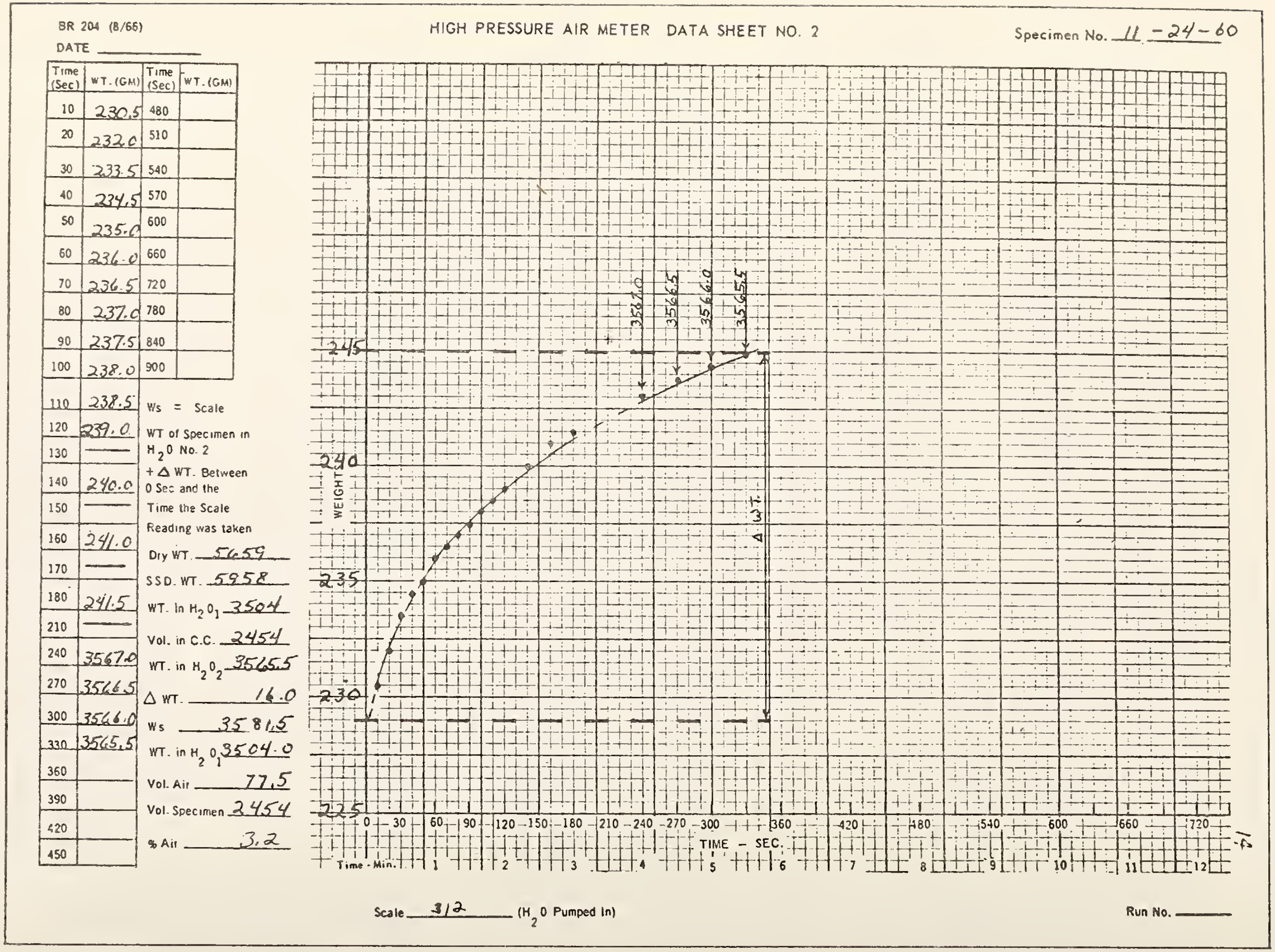
(4) Water inlet valve.

All other valves should be *open*.

b. Connect all water, air, and overflow lines.

c. Fill the chamber half-full of water.

d. Place the sample and the metal sample holder in the chamber.



BR 204 (8/66)

DATE

HIGH PRESSURE AIR METER DATA SHEET NO. 2

Specimen No. 11 -24-60

Time (Sec)	WT. (GM)	Time (Sec)	WT. (GM)
10	230.5	480	
20	232.0	510	
30	233.5	540	
40	234.5	570	
50	235.0	600	
60	236.0	660	
70	236.5	720	
80	237.0	780	
90	237.5	840	
100	238.0	900	
110	238.5		
120	239.0		
130	——		
140	240.0		
150	——		
160	241.0		
170	——		
180	241.5		
210	——		
240	3567.0		
270	3566.5		
300	3566.0		
330	3565.5		
360			
390			
420			
450			

Ws = Scale WT of Specimen in H_2O No. 2 + Δ WT. Between 0 Sec and the Time the Scale Reading was taken

Dry WT. 5659

SSD. WT. 5958

WT. In H_2O_1 3504

Vol. in C.C. 2454

WT. in H_2O_2 3565.5

Δ WT. 16.0

Ws 3581.5

WT. in H_2O_1 3504.0

Vol. Air 77.5

Vol. Specimen 2454

% Air 3.2

Scale 3/2 (H_2O Pumped In)

Run No.

Figure 31. Typical data sheet and calculations.

e. Secure the lid and gasket and tighten with a torque of 135 ft-lb.

f. Open the overflow valve.

g. Open the pump on-off control valve for a few seconds to make sure the pump lines are filled with water, then close the valve.

h. Open the water inlet valve.

i. When water is flowing freely from the overflow hose, and no air bubbles are noted, close the overflow valve and then the water inlet valve.

j. Place the water source for the pump on the left platform of the scale and tare to 0.

k. Open the pump on-off control valve and allow the pressure in the system to stabilize at 4000 psi.

l. Close the pump on-off control valve when the interval between pump strokes is greater than 5 min.

m. Record the amount of water used from the pump water source on the twelfth line of the right column on the data sheet.

n. Remove the pump water source from the scale.

o. Place a 1-1/2 qt bowl on the right platform of the scale and tare to 0.

p. Arrange the overflow line to flow into the bowl without its touching either the scale or the bowl.

q. Open the overflow valve and reduce the pressure in the system to 0 psi in 10 seconds. Do not open the valve too rapidly or it will splash out of the bowl.

r. From the time that 0 psi is reached, record the weight of water in the bowl in the left column at the prescribed time intervals for 6 minutes.

s. At 6 minutes from 0 psi, remove the sample from the chamber and obtain its weight in water as soon as possible. Record the time and weight on the appropriate line in the right column.

t. Take at least three more weights in water at intervals of 30 seconds and record them on the appropriate lines in the right column.

u. Calculate the percent of air in the sample.

3. Calculations (Fig. 31)

a. Percent of absorption $= \dfrac{\text{SSD weight} - \text{dry weight}}{\text{dry weight}} \times 100$

or

$$\frac{\text{Line 2} - \text{Line 1}}{\text{Line 1}} \times 100$$

b. Volume of sample $= \text{SSD weight} - \text{initial weight in } H_2O \times 1 \text{ cc/g}$

or

$$\text{Line 2} - \text{Line 3} \times 1 \text{ cc/g}$$

c. Volume of air in sample

(1) Choose an appropriate scale for the Y-axis on the data sheet.

(2) Plot the time-weight curve for the data recorded in the data sheet's left column.

(3) Extrapolate the curve on both ends.

(4) Plot the first value in the right column on the extrapolated portion of the curve above the appropriate time on the X-axis.

(5) Plot the subsequent three values in the right column relative to the first value in that column.

(6) Calculate Δ weight by subtracting the weight on the Y axis at 0 time, from the weight on the Y axis of the last recorded time on the X axis.

(7) To find W_s, add Δ weight to weight in H_2O_2 which is the last recorded weight in the right column (Line 7 = Line 5 + Line 6).

(8) Find the volume of air or water in the sample by subtracting the initial weight in water from W_s (Line 9 = Line 7 - Line 8).

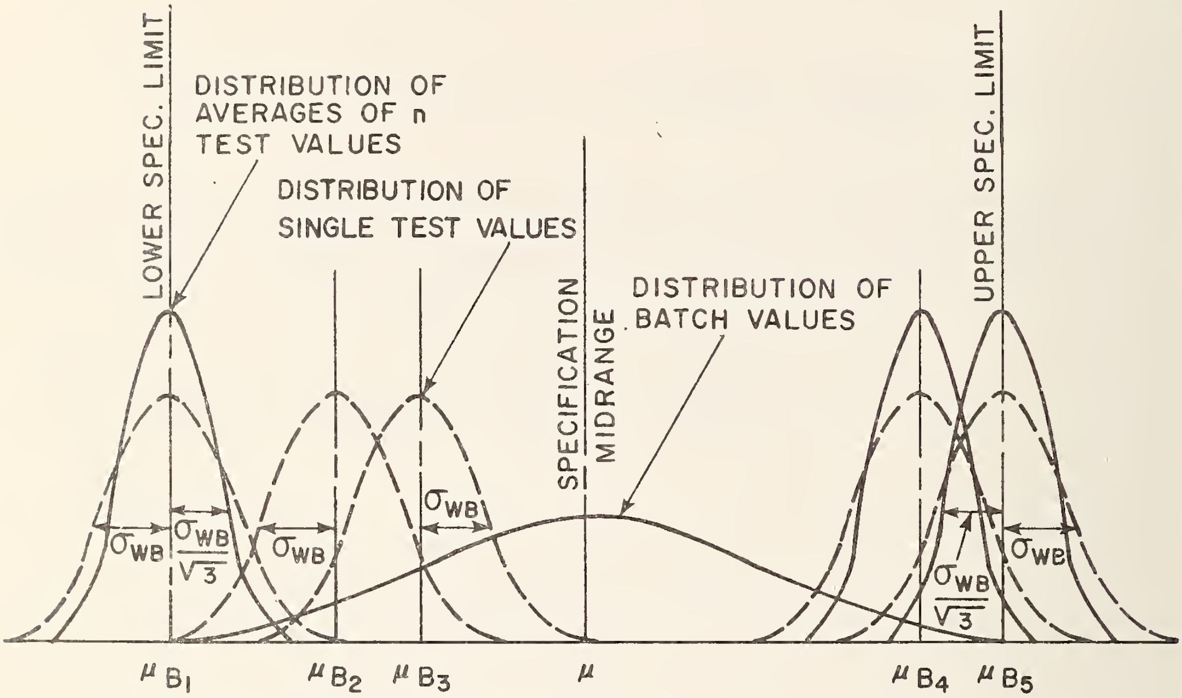
d. Percent of air in sample $= \dfrac{\text{volume of air}}{\text{volume of sample}} \times 100$

or

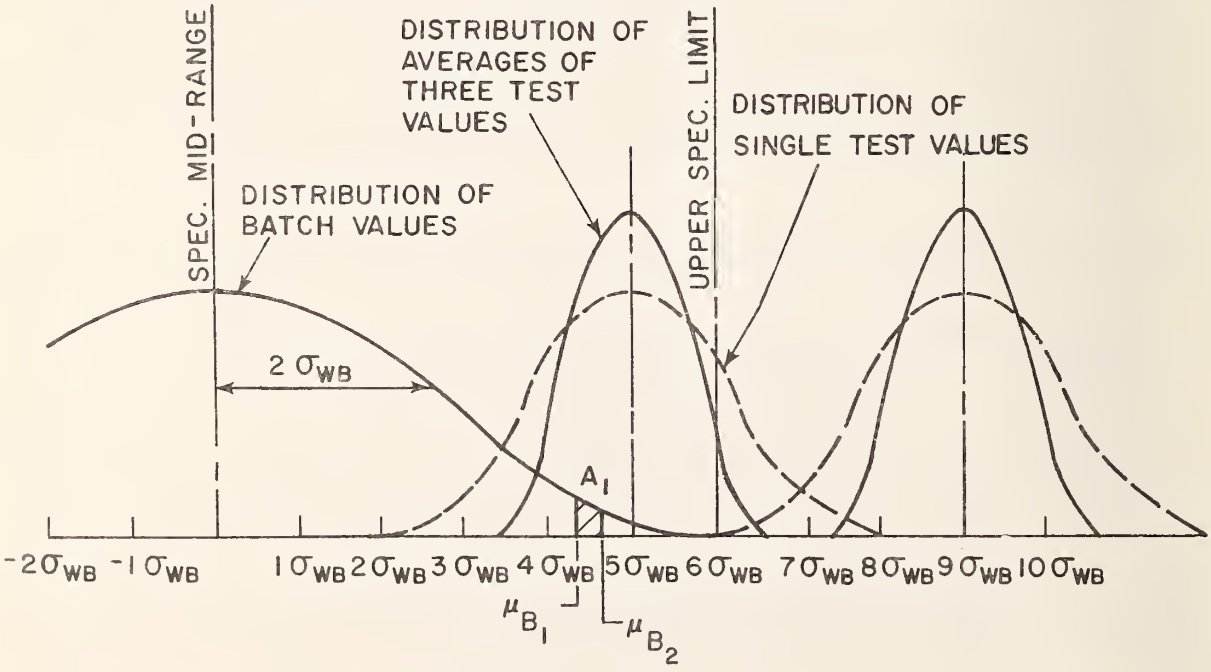
$$\frac{\text{Line 9}}{\text{Line 10}} \times 100$$

e. The percent air may also be calculated by using the amount of water pumped into the system (Line 12) and a calibration curve. This method is less accurate than that just noted.

A. POSSIBLE DISTRIBUTIONS OF SINGLE AND AVERAGE TEST VALUES RELATED TO SPECIFICATION LIMITS AND BATCH VALUES



B. EXAMPLES OF PROPORTIONS OF AREAS EXCEEDING UPPER SPECIFICATION LIMIT FOR DISTRIBUTIONS OF SINGLE AND AVERAGE TEST VALUES



jure 32. Relationship of distributions of batch and test val-
; to specification limits.

E. SAMPLING TRUCK-MIXED CONCRETE

Until 1968, the ASTM procedure for sampling truck mixers recommended that the sample be taken "at three or more regular intervals throughout the discharge of the entire batch, except that samples shall not be taken at the beginning or end of discharge." In 1968, the ASTM sampling procedure for sampling revolving drum truck mixers and agitators was changed to read: "Sample the concrete at two or more regularly spaced intervals during the discharge of the middle portion of the batch." The change may have been introduced to alleviate the problem of delayed test results, but in that respect the two procedures are not much different. Taking the samples from at least two intervals from the middle of the discharge still results in most concrete being in the forms by the time test results are known. Thus, in Chapter IV it was recommended that the ASTM sampling procedure be abandoned when screening structural concrete, without extended discussion. In this appendix, illustrations are given to show that sampling only the first portion of a truck-mixer does not greatly affect inspection results. This point is illustrated by Figure 32 and Tables 27 and 28, assuming that three portions should be recovered from the middle of batch discharge when using the ASTM procedure. If the sample were recovered from only two portions from the middle of the truck discharge, the same argument would hold but the number of batches misjudged would increase.

In Figure 32A, the distributions shown by dashed lines represent some of those that would be obtained by testing all possible test portions in a concrete batch. Similarly, the distribution shown by solid lines and having for standard deviation $\sigma_{WB}/\sqrt{3}$ represents distributions of batch estimates that would be obtained by averaging three test results from samples taken from three different portions of the truck-mixer discharge. Comparing these two distributions, it can be seen that the estimates obtained by averaging three test results cluster closer to the true batch means μ_B than the single test values. This means that testing concrete taken from three locations in the batch and averaging the results gives a more precise estimate of true batch values than testing only concrete from one portion of the truck discharge. This, in fact, is what happens when sampling from three portions of the truck according to the ASTM procedure, because the mixing of the sample consists of physically averaging the properties of the concrete in the sample.

TABLE 27
APPROXIMATE NUMBERS OF BATCHES IN A LOT OF 10,000
MISJUDGED BY SAMPLING ONLY ONE PORTION OF A TRUCK MIXER AND
USING THE ASTM SAMPLING PROCEDURE HAVING A SINGLE SPECIFICATION LIMIT

μ_B	Z_1	P_1	Z_2	P_2	A_1	$10^4 A_1 P_1$	$10^4 A_1 P_2$
$0\sigma_{WB}$	6.00	0.000	10.4	0.000	--	0.00	0.00
$1\sigma_{WB}$	5.00	0.000	8.7	0.000	--	0.00	0.00
$2\sigma_{WB}$	4.00	0.000	6.9	0.000	--	0.00	0.00
$3\sigma_{WB}$	3.00	0.000	5.2	0.000	--	0.00	0.00
$3.5\sigma_{WB}$	2.50	0.006	4.3	0.000	0.0090	0.54	0.00
$3.75\sigma_{WB}$	2.25	0.013	3.9	0.000	0.0080	1.04	0.00
$4.0\sigma_{WB}$	2.00	0.023	3.5	0.000	0.0030	0.69	0.00
$4.15\sigma_{WB}$	1.85	0.032	3.2	0.000	0.0020	0.64	0.00
$4.25\sigma_{WB}$	1.75	0.040	3.0	0.000	0.0050	2.00	0.00
$4.50\sigma_{WB}$	1.50	0.067	2.59	0.005	0.0020	1.34	0.10
$4.65\sigma_{WB}$	1.35	0.088	2.38	0.009	0.0010	0.88	0.01
$4.75\sigma_{WB}$	1.25	0.106	2.16	0.015	0.0010	1.06	0.15
$4.85\sigma_{WB}$	1.15	0.125	1.99	0.023	0.0010	1.25	0.23
$4.95\sigma_{WB}$	1.05	0.147	1.83	0.034	0.0005	0.07	0.02
$5.00\sigma_{WB}$	1.00	0.159	1.70	0.045	0.0010	1.59	0.45
$5.10\sigma_{WB}$	0.90	0.184	1.56	0.060	0.0003	0.05	0.02
$5.15\sigma_{WB}$	0.85	0.200	1.47	0.071	0.0007	0.14	0.05
$5.25\sigma_{WB}$	0.75	0.230	1.30	0.097	0.0006	0.14	0.58
$5.35\sigma_{WB}$	0.65	0.260	1.13	0.131	0.0005	0.13	0.07
$5.45\sigma_{WB}$	0.55	0.291	0.96	0.170	0.0003	0.09	0.05
$5.50\sigma_{WB}$	0.50	0.309	0.90	0.185	0.0002	0.06	0.04
$5.55\sigma_{WB}$	0.45	0.326	0.78	0.218	0.0002	0.07	0.04
$5.60\sigma_{WB}$	0.40	0.345	0.69	0.245	0.0002	0.07	0.05
$5.65\sigma_{WB}$	0.35	0.363	0.61	0.270	0.0002	0.07	0.05
$5.70\sigma_{WB}$	0.30	0.382	0.52	0.300	0.0001	0.04	0.03
$5.75\sigma_{WB}$	0.25	0.401	0.43	0.334	0.0003	0.12	0.10
$5.85\sigma_{WB}$	0.15	0.441	0.26	0.398	0.0002	0.09	0.08
$5.90\sigma_{WB}$	0.10	0.460	0.17	0.433	0.0001	0.05	0.04
$5.95\sigma_{WB}$	0.05	0.490	0.09	0.464	0.0001	0.05	0.04
$6.00\sigma_{WB}$	0.00	0.500	0.00	0.500	0.0000	0.00	0.00
$6.05\sigma_{WB}$	-0.05	0.490	-0.09	0.464	0.0000	0.00	0.00
$6.10\sigma_{WB}$	-0.10	0.460	-0.17	0.433	0.0000	0.00	0.00
$6.15\sigma_{WB}$	-0.15	0.441	-0.26	0.398	0.0000	0.00	0.00
				Total Batches Misjudged		$\sum = 12.13$	$\sum = 2.20$

μ_B = batch mean in terms of within-batch standard deviation.

P_1 = Probability of misjudging batch by sampling only one portion of truck discharge.

P_2 = Probability of misjudging batch, sampling according to ASTM procedure.

A_1 = Percentages of batches having means between the tabulated values of μ_B

N = 10,000.

TABLE 28
BATCHES MISJUDGED BY TWO SAMPLING METHODS
FOR VARIOUS SHIFTS IN PRODUCTION MEANS

Deviation of Actual Mean from Specification Mean		Batches Misjudged in 10,000		Additional Batches Misjudged by Sampling Only First 1/3 of Truck	
		Sampling by ASTM Procedure	Sampling Only First 1/3 of Truck	Number	Percent
Lower Spec. Limit	-3.0σ	5,000	5,000	0	
	-2.5σ	857	1,361	504	5.04
	-2.0σ	589	969	380	3.80
	-1.5σ	322	555	233	2.33
	-1.0σ	132	251	119	1.19
	-0.5σ	46	96	50	0.50
Spec. Midrange	0.0σ	3	13	10	0.10
	0.5σ	46	96	50	0.50
	1.0σ	132	251	119	1.19
	1.5σ	322	555	233	2.33
	2.0σ	589	969	380	3.80
	2.5σ	857	1,361	504	5.04
Upper Spec. Limit	3.0σ	5,000	5,000	0	

Thus, if testing errors are assumed negligible, the results approximate the values that would be obtained by testing three concrete portions taken from three different parts of the truck and averaging the results. Consequently, the estimates of batch means thus obtained give more precise estimates of the true batch means than those that would be obtained by sampling only a portion of the truck-mixer or batch. But this added precision and the smaller probability of misjudging a batch associated with it are not worth the price of postponing judging the truckload until most or all has been placed in the forms.

Referring to Figure 32A, it can be seen that for batches having as mean μ_{B_2} and μ_{B_3}, the wrong decision in judging the acceptability of the batch cannot be made regardless of whether one or three portions of the batch are sampled. In these instances, both the distribution of single values and that of average values fall within the specification limits and the batch cannot be misjudged. This is also true for all other batches not shown, whose distributions of single test values are entirely included in the specification range. In these cases, sampling more than one portion of the truck discharge would give a more precise estimate of the true level of the property tested but would not influence the judgment of the batch. In acceptance sampling by attributes or screening, the correct judging of the production unit is what is important. Thus, in these cases the added precision serves little or no purpose. When the batch means coincide with one of the limits, there is still no advantage in sampling more than one portion of the truck. In such cases, the probability of misjudging the batch is 0.50, both with average values and with single values. This can be readily seen by observing that half of the area of both the dashed-line and solid-line distributions straddling the upper limit in Figure 32A lies on both sides of the upper limit. Since areas represent probabilities, the probability of misjudging the batch, or saying that its value is greater than the limit when in fact it equals the limit, is 0.50 regardless of the precision of the batch estimate.

The advantage of the added precision given by the ASTM procedure comes into play only when batch means approach the specification limits. Referring to Figure 32B, it can be seen that when the batch mean approaches the upper limit the *proportion* of area under the curve beyond the upper specification limit is greater for the distribution of single test values than for the distribution of the average of three tests. Because areas represent probabilities, the batch has a greater probability of being misjudged when only one portion is tested.

Thus, at first glance, the ASTM procedure appears desirable. But further evaluation reveals that its advantage is not very great. The ASTM procedure would be preferable if the means of all batches produced approached the limits. But in most cases this is not the situation. In a majority of instances the producer is aiming at the specification midrange as an overall mean, and only a very small percentage of the total batches produced will approach the specification limits. Since the batch cannot be misjudged unless its mean approaches these limits, the practical advantage of the ASTM procedure is drastically reduced. This is illustrated by the results in Table 27, which were obtained by assuming 1) that the distributions involved are normal, 2) that the batch-to-batch standard deviation σ_B is twice the within-batch standard deviation σ_{WB} as shown in Figure 32B, and 3) that the process is geared to meet the specifications. The first two assumptions are believed approximately correct. Data presented in Table 16 show that in most cases the assumption of normality for batch values is warranted. Normality for single test values can be assumed because the variation among the results is due mainly to sampling and testing errors which are likely to occur randomly. The distribution of the average of three test results can be taken to be normal because it is a distribution of averages. Other data available to the writer and recorded while evaluating the mixing efficiency of truck mixers show that within-batch variation is about half batch-to-batch variation, making the second assumption warranted. The third assumption -- that the process is geared to meet the specifications -- is not always valid. However, the invalidity of this assumption would only change the numerical results in Table 27. If the process were out of control or not geared to meet the specifications, the number of batches musjudged would increase under both types of sampling, and the number of additional batches misjudged by not using the ASTM procedure would also increase as the difference between the true mean and the specified midrange increases until the true mean coincides with one of the limits as shown in Table 28 which was prepared by repeating the computations in Table 27 many times. But if producers were not trying to achieve the specified mean, a large proportion of production would be rejected under either sampling procedure, and continuing production would become economically unfavorable.

In Table 27, μ_B is the distance of the batch mean from the specification midrange in terms of within-batch standard deviation; P_1 is the greatest probability of misjudging the batches whose means lie between two successive values of μ_B by sampling according to the ASTM procedure; A_1 is the percentage of batches expected to fall between two successive values of μ_B assuming that the producer is trying to achieve the specification midrange as an overall mean; $10^4 A_1 P_1$ is the number of batches out of 10,000 that would be expected to be misjudged sampling only a portion of the batch and the batch-means fall between two successive values of μ_B as indicated in the table; $10^4 A_1 P_2$ is the expected number of misjudged batches out of 10,000 under the ASTM sampling procedure if the batch means fell between the same two successive values of μ_B; the summations in the last line represent the approximate total number of batches out of 10,000 that would be misjudged under each sampling procedure. The greatest probability of misjudging the batches represented by the area between two successive values of μ_B is the theoretical probability of misjudging the batch whose mean lies closest to the limit. Referring to Figure 32B, this would be the probability of misjudging the batch having as mean μ_{B_2} limiting A_1. Since all other batches represented by the area A_1 have a smaller probability of being misjudged, the results in Table 27 represent the worst that can be obtained. It can be seen in this table that even assuming the worst probability of misjudging batches falling in a given interval and for a single-limit specification, 2 batches out of 10,000 would be misjudged under the ASTM procedure while 12 out of 10,000 would be misjudged by sampling only a portion of the truck. If the specification had double limits, these numbers would double. (The limits are assumed symmetrically located about the specification midrange.) Thus, for a double-limit specification -- the worst case -- only 19 more batches in 10,000 would be misjudged by abandoning the ASTM procedure. This is equivalent to approximately 2 batches in 100. Having to postpone judgment of all batches until all or most of the concrete in each batch is already in the forms is too high a price to avoid misjudging two batches in 100. It does not seem justified. This is especially true for most structural pours, which take less than 30 truckloads of concrete and for which the probability of misjudging more than one batch is very low.

☆U.S. GOVERNMENT PRINTING OFFICE: 1973 730-401/90 1-3

FHWA
R&D