Pl. return to
Walter A. Shewhart,
463 West. St, N.Y.C.

Introduction

The potential contributions of statistics to the science of engineering are an important national asset; an asset of interest to all of us because its use makes possible the most efficient and effective use of natural resources to satisfy human wants; an asset, however, that has for years remained frozen and is only now beginning to be liquidized; an asset that can be completely liquidized only when engineers and others learn to use it as they have learned to use the product of the scientist of the past.

Much has appeared in the literature to indicate some of the contributions of statistics to date in the field of engineering and manufacturing. My object today is not so much to review what has been done as to survey the potential contributions of statistics to the science of engineering. In doing this, I shall follow the old advice that the easiest way to reach the top is to go to the bottom of things and I shall go to the bottom of the difference between engineering with, and without, statistics.

Science of Engineering

Made possible by Scientific Method

Statistics contributes to each of 3 steps.

Scientific Method

Hypothesis	Statistical Hypothesis
Experiment	Statistically designed experiment
Test of Hypothesis	Test of Statistical hypothesis

Statistically scientific method is in fact a fundamental discipline that includes all of customary scientific method based upon the concept of exact laws as a limiting case that would be valid if repetitive operations of any given kind always gave identically the same results.

Basic Engineering Problem

Engineer's job is to develop OPERATIONS that if performed will produce a thing having quality characteristics within previously specified tolerance ranges.

Repetitions of an operation give (1)

$$X_1, X_2, X_3, \ldots, X_i, \ldots, X_n, X_{n+1}, \ldots X_{n+j} \ldots$$

supposed to fall within previously specified tolerance range L_1 to L_2 (Fig. 1).

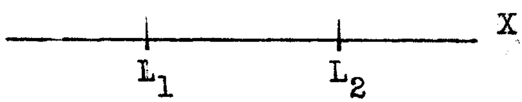

Fig. 1

Failure means
- Loss because of rejection
- Loss of property as when fuse fails
- Loss of life as when shell explodes at wrong time

Probability enters fundamental problem because one cannot be certain that a repetion will give value of X within tolerance range.

1. Maximum assurance
2. Minimum tolerance range.

55	[illegible]	45	27	33	59	[illegible]	[illegible]	[illegible]	62	39	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	36	45	26	[illegible]	[illegible]	42	38	[illegible]
[illegible]	[illegible]	23	18	34	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	58	[illegible]	[illegible]	[illegible]	17	71	28	53	38	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	75	38	28	[illegible]	23	42	35	[illegible]
[illegible]	[illegible]	[illegible]	26	59	56	[illegible]	37	[illegible]	[illegible]	35	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	43	39	[illegible]	43	[illegible]	43	22	50
[illegible]	21	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	58	52	[illegible]	30	[illegible]
[illegible]	[illegible]	23	[illegible]	[illegible]	[illegible]	[illegible]	33	43	39	24	90
[illegible]	34	68	[illegible]	59	43	[illegible]	62	31	48	[illegible]	96
30	36	53	48	48	33	27	27	36	58	54	80
[illegible]	41	26	34	[illegible]	[illegible]	27	[illegible]	[illegible]	[illegible]	73	[illegible]

TABLE 1 - Thickness in Arbitrary Units of In[illegible]
on 144 Relay Springs

BASIC CONTRIBUTION OF

CLASSICAL STATISTICAL THEORY

Basic Hypothesis

Some repetitive operations (called random) exist in nature that obey laws of chance. The probability that such a random operation will give a previously specified event, as, for example, the occurrence of a value of X within a previously specified tolerance range, is a definite number associated with that event.

Basic Experiment or Operation (Drawing at Random)

If we can learn to know a random physical operation as such (not by what it gives) then we can know when theory can be used. This is to be contrasted with quantitative criteria for randomness considered later in control theory.

Table 1

How draw random sample of 10? May we take every fourteenth one?

We can draw from a bowl
- a. Thoroughly mix
- b. Draw with eyes shut.

Basic Test of Hypothesis

Classical theory usually assumes law of distribution f(x)dx. Then distribution theory enables one to compute limits within which samples may be expected to fall.

Hence can test hypothesis that statistic of random sample of size n came from an assumed population.

Test of hypothesis as a rule of procedure.

Errors of 1st kind.

Fundamental Types of Application

If hypothesis that an operation obeys law f(x) of probability is established, then:

1. Estimation

   Can choose efficient, unbiased estimates that approach "true" values as statistical limits.

2. Distribution

   Can compute $f(\theta)$ for samples of size n. For example, $f(\overline{X})$ rapidly approaches normality irrespective of f(x) for most f's found in practice.

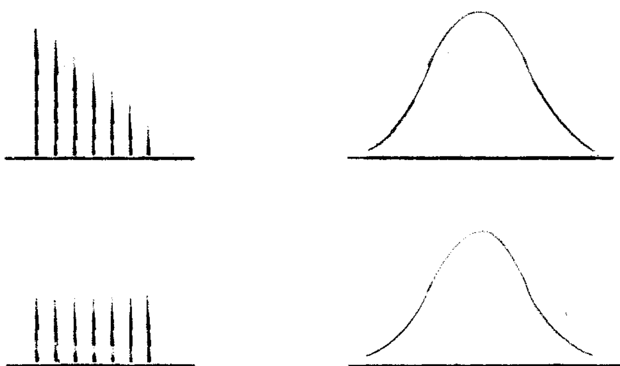

Fig.

Importance in setting overall tolerances.
For $Y = X_1 + X_2 + \ldots + X_n$,

$$\sigma_y = \sqrt{n}\sigma_x \quad \text{and not } n\sigma_x.$$

Also important in control chart.

Knowledge of an operation of randomization and means of testing statistical hypotheses make it possible to

3. <u>Introduce random operation into design of experiment</u>

n objects < Treatment A / Treatment B

↑

Randomize

4. <u>Curative Sampling</u> (cf. Simon's paper)

Screen to meet < Consumer risk / Producer risk

But does not improve product.

BASIC CONTRIBUTION OF STATISTICAL CONTROL THEORY

Ounce of prevention is worth a pound of cure.

Classical theory — Assumes laws of probability to be discovered

But it was discovered in 1924 in mass production that such laws do not exist.

Basic Hypothesis

Hypothesis IIa: The maximum attainable degree of validity of prediction that an operation will give a value X lying within any previously specified tolerance limits is that based upon the prior knowledge that the probability of this event is q' or more generally upon the prior knowledge of the law of chance underlying the operation.

Hypothesis IIb: The maximum degree of attainable control of the cause system underlying any repetitive operation in the physical world is that wherein the system of causes produce effects in accord with a law of probability.

Hypothesis IIc: It is assumed that some criterion or criteria may be found and methods developed for their application to the numbers obtained in a sequence of repetitions of any operation such that whenever a failure to meet the criterion or criteria is observed, an assignable cause of variability in the results given by the operation may be discovered and removed from the operation. It is further assumed that, as these causes are removed, a state of statistical control is approached where the results of repetitions of the operation behave in accord with a law of chance.

Basic Experimental Data

X_1,	X_2,	...,	X_i,	...,	X_n,	X_{n+1},	...,	X_{n+j}, ...
C_1	C_2		C_i		C_n	C_{n+1}		C_{n+j}

Past | Present | Future

Screening looks at X's of the past

Control looks toward causes controlling X's of the future.

Emphasis on f(x)

Emphasis on order

Three Conditions

(1) $C_i =\!\circ\!= C_j$

Drawing from a bowl. Repetitions under the same essential conditions

Keep order of occurrence

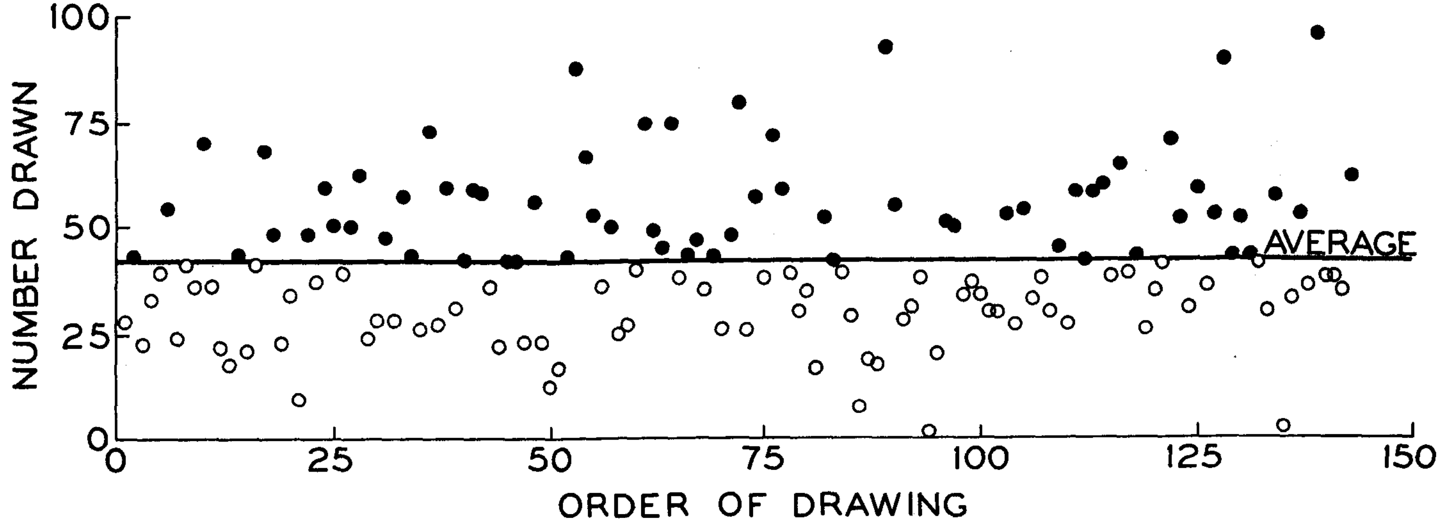

FIG. 1a SEQUENCE OF NUMBERS DRAWN AT RANDOM

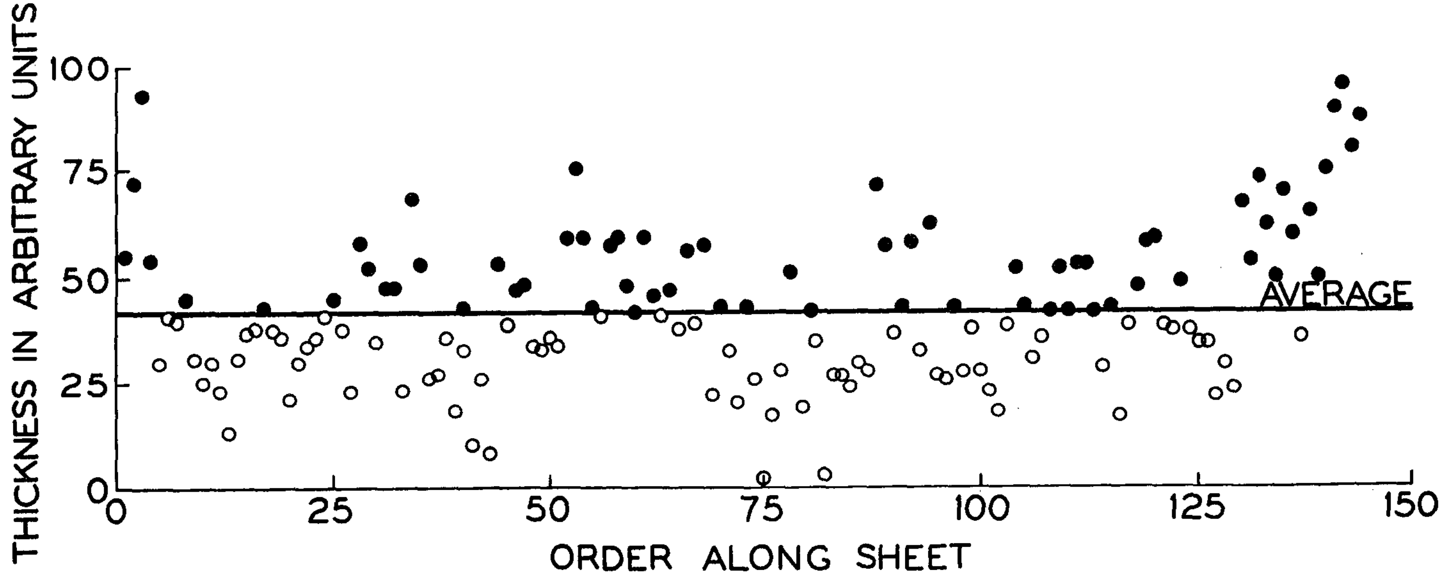

FIG. 1b SEQUENCE OF VALUES OF THICKNESS OF INLAY ON RELAY SPRINGS

2. $c_i \neq c_j$

Experimentalist may surmise erratic effects or trends.

Table 1 again

graph Original data

Describe way data were obtained.

Experimentalist must suggest orders.

3. A priori reason for expecting assignable groups

Examples:

- Source of material
- Different rollings
- Different heat treatments

Basic Test of Statistical Control Hypothesis

Operation of Statistical Control

1. Specify in a general way how an observed sequence of n data is to be examined for clues to the existence of assignable causes of variability. For example, it is essential that the order in an observed sequence always be tested for randomness whenever $C_i \Rightarrow C_j$ or $C_i \not\Rightarrow C_j$.

2. Specify how the original data are to be taken and how they are to be broken up into subsamples upon the basis of human judgments about whether the conditions under which the data were taken were essentially the same or not.

3. Specify the criterion of control that is to be used and indicate what statistics are to be computed for each subsample and how these are to be used in computing action or control limits for each statistic for which the control criterion is to be constructed. Three of the conditions that such criteria should satisfy are as follows: the limits in the criteria should be as nearly independent as possible of the functional form of the law of chance when the state of statistical control is attained; the criteria should in so far as possible minimize the error of accepting the hypothesis when false and should keep the error of rejecting the hypothesis when true less than some prescribed value fixed by economic considerations; and should indicate as closely as possible the condition under which the assignable causes enter the operation and as much as possible about the nature of these causes.

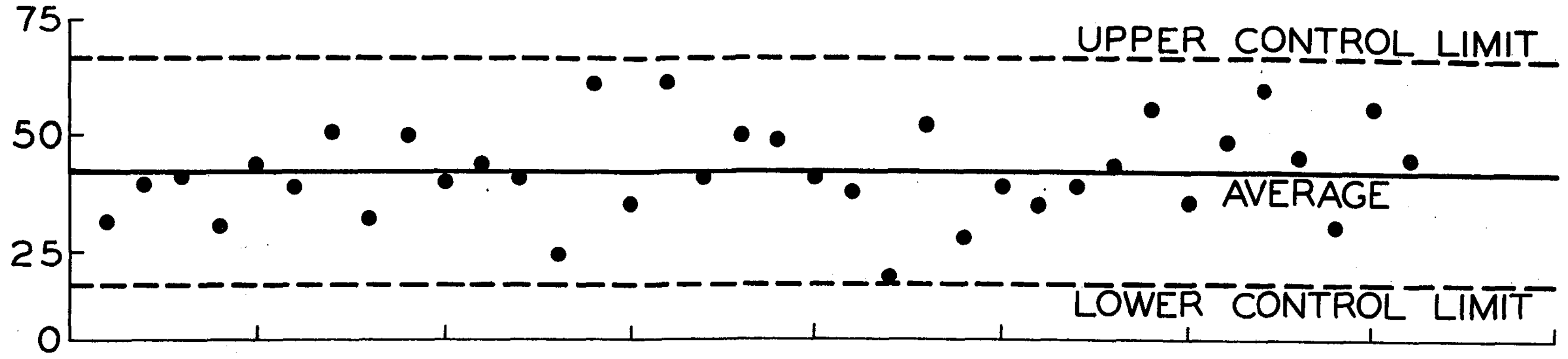

FIG. 2a CONTROL CHART FOR AVERAGES OF FOUR-RANDOM SEQUENCE OF FIG. 1a

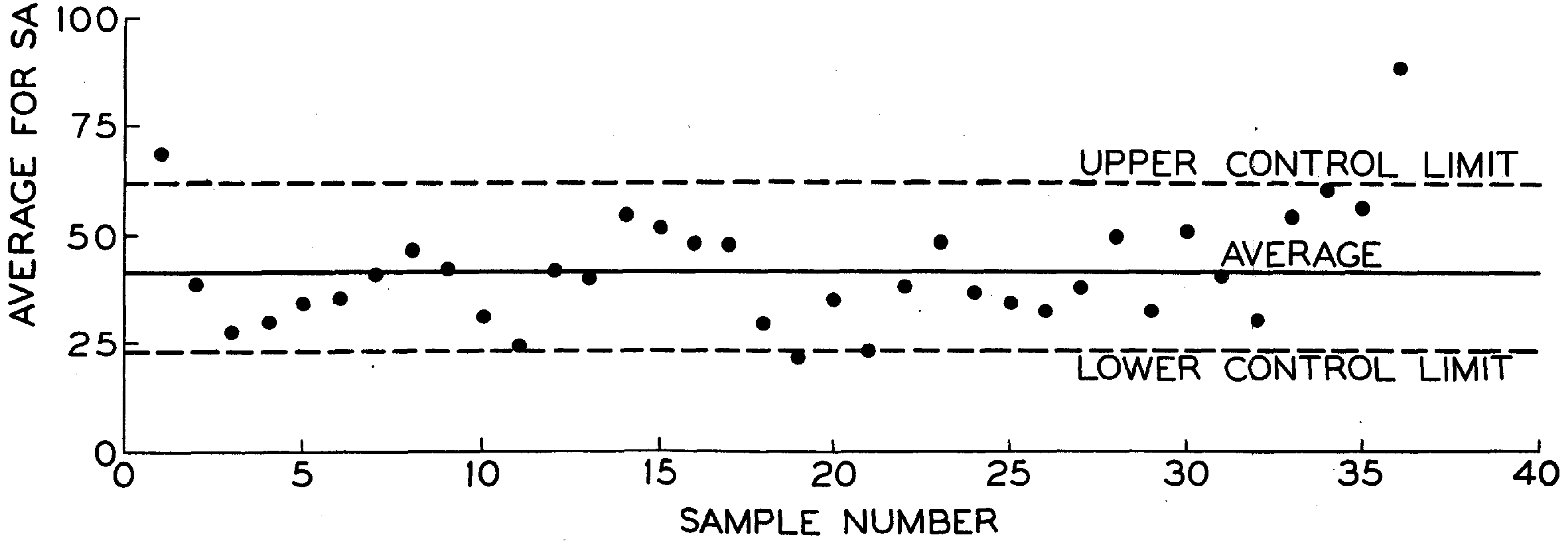

FIG. 2b CONTROL CHART FOR AVERAGES OF FOUR SEQUENCE OF THICKNESS OF FIG. 1b

Table 2

Length of Runs	Drawings from Bowl: Runs above and below Average: Observed Number	Drawings from Bowl: Runs above and below Average: Expected Number	Drawings from Bowl: Runs up and down: Observed Frequency	Drawings from Bowl: Runs up and down: Expected Frequency	Measurements of Inlay Thickness: Runs above and below Average: Observed Number	Measurements of Inlay Thickness: Runs above and below Average: Expected Number	Measurements of Inlay Thickness: Runs up and down: Observed Frequency	Measurements of Inlay Thickness: Runs up and down: Expected Frequency
1	42	37			[illegible]	[illegible]		
2	18	18	65	61	13	18	50	56
3	10	9	22	27	3	9	28	24
4	4	4	10	8	5	4	8	7
5	4	2	1	2	0	2	2	2
6	0	1	0	0	4	1	1	0
7	0	0	0	0	4	0	0	0
8	0	0	0	0	1	0	0	0

4. Specify the action that is to be taken when an observed statistic falls outside its control limits. The general action required is to look for assignable causes whenever the criteria are not satisfied.

5. Specify the quantity of data that must be available and found to satisfy the criterion of control before the engineer is to act as though he had attained a state of statistical control.

Example

Control chart for averages of 4.

More about causes - Combination of

Runs up and down	Kermack and McKendrick
Runs above and below average	Cochran and others.

Table 2

Figure with ups and downs

Clue to nature of causes:

Where they enter
not in rolling mill
slippage at cleavage planes

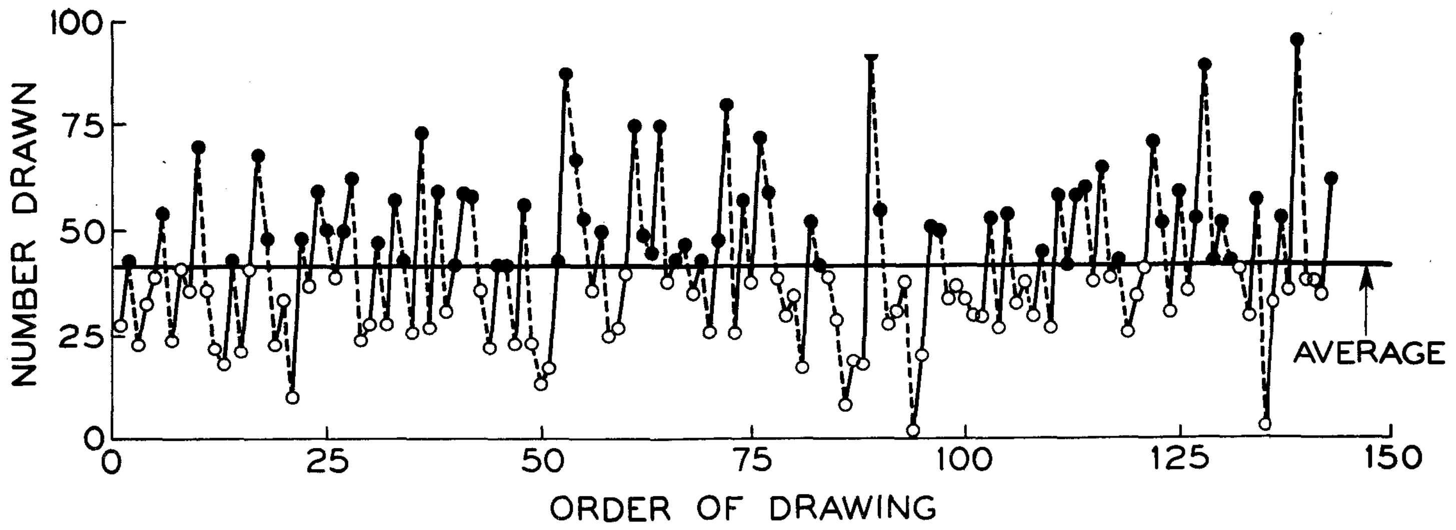

FIG. 3a SEQUENCE OF NUMBERS DRAWN AT RANDOM

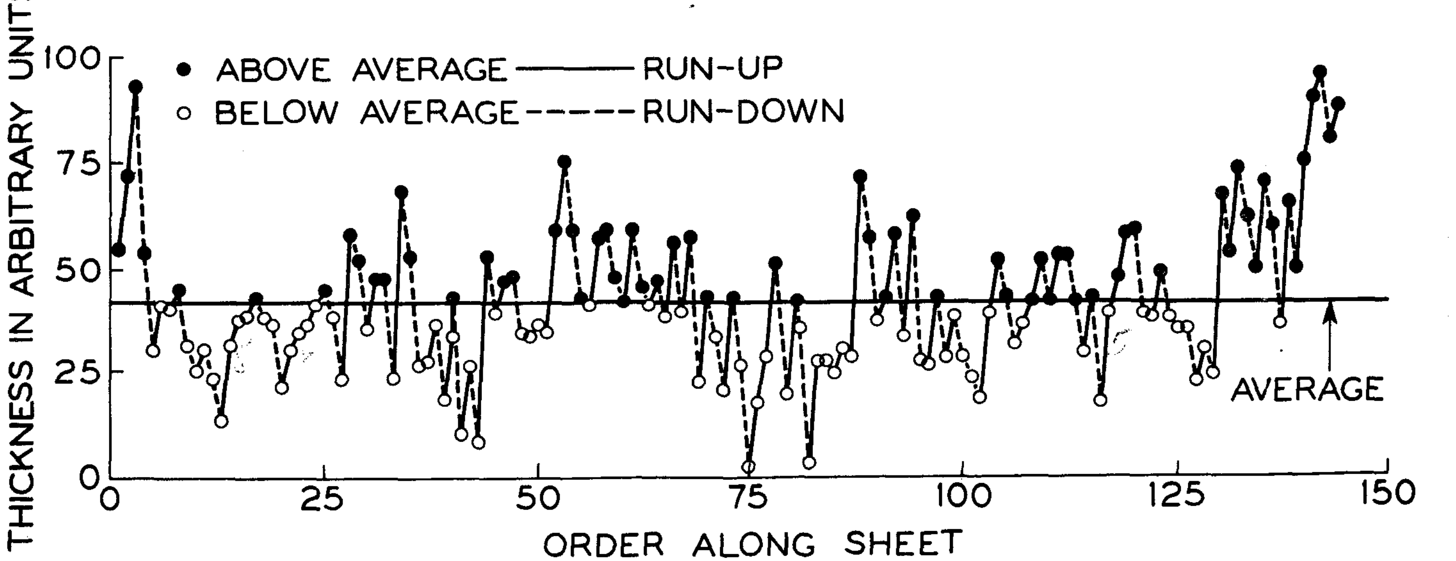

FIG. 3b SEQUENCE OF VALUES OF THICKNESS OF INLAY ON RELAY SPRINGS

SUMMARY OF POTENTIAL CONTRIBUTIONS OF STATISTI

TO THE SCIENCE OF ENGINEERING

The basic contribution of statistics to the science of engineering is an improved scientific method to fit the world of probability in which we live. Classical theory contributes the hypothesis of a repetitive operation obeying a law of chance, the knowledge of which combined with the knowledge of formal mathematical distribution theory enables an engineer to make valid predictions of the outcome of future repetitions of the operation. Statistical control theory contributes the hypothesis that it is humanly possible to remove assignable causes of variability in the repetitive operations of the engineer until such operations approach a state of maximum control and obey laws of chance, a knowledge of which provides maximum assurance that the results of repeating an operation will fall within previously specified tolerance limits. To test these two hypotheses, statistical theory provides the necessary experimental techniques outstanding among which are (a) the operation of randomization and (b) the operation of statistical control.

Broadly speaking, statistical theory treats of repetitive operations and provides the engineer with a method of regulating such operations to his best interest. Fundamentally, the engineer's job is to devise operatior that, if carried out, will give results within previously specified tolerance limits. Sometimes the operation, like that of building a bridge, is to be carried out only once or at most a few times and sometimes the operation, like that of mass production, is to be carried out an indefinitely large number of times. Inherently all such operations are potentially repetitive, and in this sense, differ only in the number of repetitions carried out. It has long been recognized that one of the most revolutionary principles ever introduced into manufacturing was that of interchangeability dating back at least to Eli Whitney in 1798. The introduction of that principle prompted the engineer to consider the advantages of introducing repetitive operations into production processes and the contribution of statistics to engineering may be thought of as a means of maximizing the advantages to be attained by interchangeability.

The basic contributions of statistics to scientific method make possible the attainment of the following objectives that are not otherwise attainable and that are of interest to all of us:

(1) Even before a repetitive operation has reached a state of statistical control, it makes possible the establishment of sampling plans that will screen at minimum cost the output of such an operation so as to meet previously specified tolerance requirements and previously specified producer and consumer risks.

(2) It provides efficient experimental techniques based upon the operation of randomizing the results obtained from one operation or production process that is not in a state of statistical control before submitting them to two or more subsidiary operations or treatments for the purpose of comparing the effects of these subsidiary operations, as in the case of field tests

and the like. Such a procedure minimizes the chance of concluding that observed differences in the effects of the subsidiary operations are significant when in fact they came about because of assignable differences in the results of the first operation.

(3) The operation of statistical control provides an experimental technique for minimizing tolerance ranges and maximizing the assurance that the product turned out by a given process will meet its tolerance requirements. Such an operation makes possible the most efficient use of limited quantities of raw material and provides the maximum degree of refinement attainable by any production process. Preliminary studies indicate that the operation of statistical control also provides a useful technique for eliminating assignable causes of variability in certain kinds of human effort, as, for example, typing and other forms of transcription. Both strategically and commercially, industrial groups and even nations often need every increment of efficiency in the use of limited quantities of raw materials and human effort that can be provided through the application of the operation of statistical control. Likewise they often need maximum refinement in quality through elimination of assignable causes not only in pursuit of the arts of peace but also in time of war. As one example, the attainment of maximum homogeneity and hence minimum tolerance ranges in the properties of raw and fabricated materials may extend the potential carrying capacities of ships in the air and on the sea. Needless to say both the engineer and the consumer of the engineer's or manufacturer's products stands to gain through the increased assurance that the products will be found to meet their tolerance requirements.

(4) The operation of statistical control provides a technique for modifying and coordinating the three fundamental steps in the process of mass production, namely, specification, manufacturing, and inspection, so that the maximum number of pieces of product having a quality within specified tolerance limits can be turned out at given cost. It does this by showing how to minimize the cost of inspection and the cost of rejection.

In conclusion it may be said that statistical theory plus mass production provides a means of maximizing our physical comforts in time of peace and our strategic factors in time of war.