STATISTICAL VIEWPOINT
IN
ENGINEERING EXPERIMENTATION

by

Walter A. Shewhart
Bell Telephone Laboratories, Inc.

Address to be given before Mathematics Colloquium, Carnegie Institute of Technology, Pittsburgh, Pa., 3/9/50.

CLASSICAL STATISTICAL VIEWPOINT

In many quarters, even today, the statistician is likely to be thought of as one engaged in the collection, tabulation, and reduction of data, particularly of a financial or economic character. He is thought of as the fellow who divides up the pie-chart to show where each dollar of income goes: the fellow who draws the cute little figures of one kind or another to get across the facts that he collects. Statisticians have even been called the show-girls of the market place because they live by their figures.

Statisticians themselves are largely responsible for the fact that the potential contributions of statistics are only now coming into recognition. In the first place, they have often indicated that statistical method was to be used only under very restrictive conditions. Thus, when the Royal Statistical Society of London was founded more than a century ago its stated objective was "to collect, arrange, digest, and publish facts illustrating the condition and prosperity of society in its material, social, and moral relations, ..." Here we have the limitation that the "facts" are to be from the field of social science. Likewise one of the best modern introductions to the theory of statistics states that "by statistical methods we mean methods specially adapted to the elucidation of quantitative data affected by a multiplicity of causes". Just what this limitation implies is not very clear but it has usually been taken to mean social and economic phenomena.

Now, a fundamental distinction between science and other forms of organized knowledge lies in the concern of science with the possibility of accurate prediction. Also in many quarters, particularly among natural scientists and engineers, social and economic phenomena are considered to be essentially unpredictable and beyond the reach of experiment, thus falling outside the scope of scientific engineering.

Hence in so far as statisticians have explicitly or implicitly implied that the use of statistical method was limited to a particular class of data that might even lie outside the pale of scientific method, they have lost caste with natural scientists and engineers.

A still further way in which statisticians, particularly business statisticians, down through the past century and up to more recent times have fenced themselves in is by laying stress on the collection, reduction, and presentation of data for the use of management in making inferences as though these steps could be taken without first giving thought to whether said collection, reduction, and summarization is valid. Historically the business statistician got himself recognized as the company specialist in dressing up great masses of data for the management to use in making decisions without assuming any responsibility for providing management with a scientific basis for making valid decisions involving always an element of prediction.

In this sense, the statistician often reminds one of Pat who had been in this country only a few months when he met his old friend Mike. To Mike's question: How are you doing, Pat?, Pat replied: Just fine, I have an easy job, I carry the bricks and mortar up seven flights and the man up there does the work". Pat assumed no responsibility for the quality of what he carried up.

MODERN STATISTICAL VIEWPOINT

Since the turn of the century the picture of the statistician's role has been revolutionized. In every field of science it is now recognized that valid scientific predictions can only be made on a statistical basis. Statistical method should no longer be viewed as something apart from scientific method, involving hypothesis, experiment, and test of hypothesis but as a scientific method in which each step is adjusted to take account of the fact that valid predictions can only be made in the probability sense within limits. No longer is statistical method relegated to the subject matter of the social sciences - instead it should play a fundamental role wherever valid scientific prediction is possible. This means that the statistician must be a co-partner of every scientist interested in making valid predictions in terms of probability limits.

The really important contribution of statistics to the science of management can be glimpsed therefore not so much from any set of detailed applications and methods for solving special problems as from certain fundamental ideas, hypotheses, principles, and concepts underlying statistical theory and the understanding and appreciation of statistics not as something apart from scientific method but as an improved scientific method adjusted to fit the world in which we live.

ENGINEERING EXPERIMENTATION

The engineer's job is to devise and develop operations that, if carried out, will produce things that people want. The engineer's job is fundamentally an act of control - one of four fundamental acts of Man.

1. Act of rational abstraction.

   Logic and mathematics.

2. Act of Measurement.

   Fundamental role of science is to discover the laws of nature in terms of which we may predict the future in terms of the past.

   Repetitive measurements assumed to be possible.

3. Act of Human Evaluation.

   Basic "meter" is human being.

4. Act of control.

4.1 Act of control is a culminating act resting upon the other three and in addition involves the setting of goal.

4.2 The fundamental distinction between science and other forms of organized knowledge lies in the concern of science with the possibility of accurate and valid prediction.

Pure science differs from engineering science in at least the four following ways:

4.2.1 The engineer is largely concerned with making predictions about repetitive operations.

4.2.2 The pure scientist is primarily interested in valid prediction within confidence limits, the engineer is interested in valid prediction within tolerance limits in k space.

4.2.3 The predictions of the engineer are almost certain to be tested.

4.2.4 If the predictions are invalid, the engineer may løøse his shirt.

TWO KINDS OF FACTORS TO BE CONTROLLED

S 30569

1. Statistics really grew up in the outer ring. Here we are concerned primarily with a study of prediction of time series - perhaps the most difficult of all predictive problems. This is the area of evolving processes.

   The modern approach in this area is through the study of the stochastic process classified in general as stationary time series.

2. Some question whether the repetitive type of experiment can be made in the outer ring. In all science it is recognized that it is not possible to conduct an experiment in which the operation of measurement does not in some way influence the phenomenon to be observed. This is particularly true in the measurement of preference.

S 36019

3. The repetitive act is possible in the inner ring.

THE REPETITIVE OPERATION

In all fields of scientific enquiry there are instances where the operations to be controlled are repetitive in character. That is to say, the operation is one that is to be repeated again and again under the same conditions. At each repetition our attention is focussed on the result of the repetition which can be described in terms of one or more characteristic features.

Measurement

Perhaps the simplest and most fundamental or basic type of repetitive operation in all scientific fields is that of measurement. For example, in measuring a "constant" of nature like the velocity of light by a given method, we are free to repeat the operation of measurement again and again at will under presumably the same essential conditions. From such a series of n repetitions, we get a sequence of measurements

$$x_1, x_2, \dots x_i, \dots x_n$$

not all of which are the same. Scientists accept the fact that even when they are measuring a so-called "constant" of nature the only kind of observable constancy is a kind of statistical variability following some law of probability. What the scientist usually does when he gets to the place where he thinks that the measurements are made under the same essential conditions is to summarize his data and treat it as though it did follow a law of probability, totally unconscious of the fact that if he were to put his observed sequence through some modern statistical machinery, he would likely conclude that the data do not follow the assumed law of probability.

In other words, he is likely to find that without the use of the theory of statistical inference, he cannot summarize his data in a way that valid predictions within limits can be made.

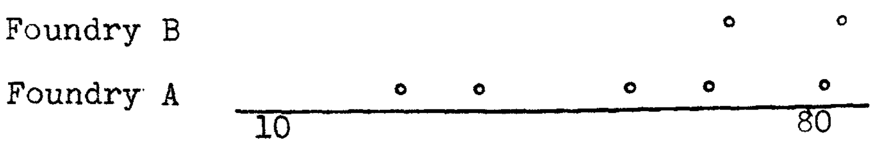
Act of Abstraction

Random variable - A sequence of variates $x_1, x_2, x_3, \ldots x_n$ is said to be a random series, or to satisfy the condition of randomness, if $x_1, x_2, x_3, \ldots x_n$ are independently distributed with the same distribution: i.e., if the joint cumulative distribution function (c.d.f.) of $x_1, x_2, x_3, \ldots x_n$ is given by the product $F(x_1)F(x_2)F(x_3) \ldots F(x_n)$ where F(x) may be any cdf.

The mathematical statistician has provided methods of testing the hypothesis that some specified F(x) exists and also that some F(x) exists.

Much stress has recently been laid on small sample tests.

Brownlee, Industrial Experimentation, 1949, p. 35.



Brownlee concludes that there is no evidence that the outputs are different upon basis of t-test.

Fig. 2

Brownlee, 1949, p. 38

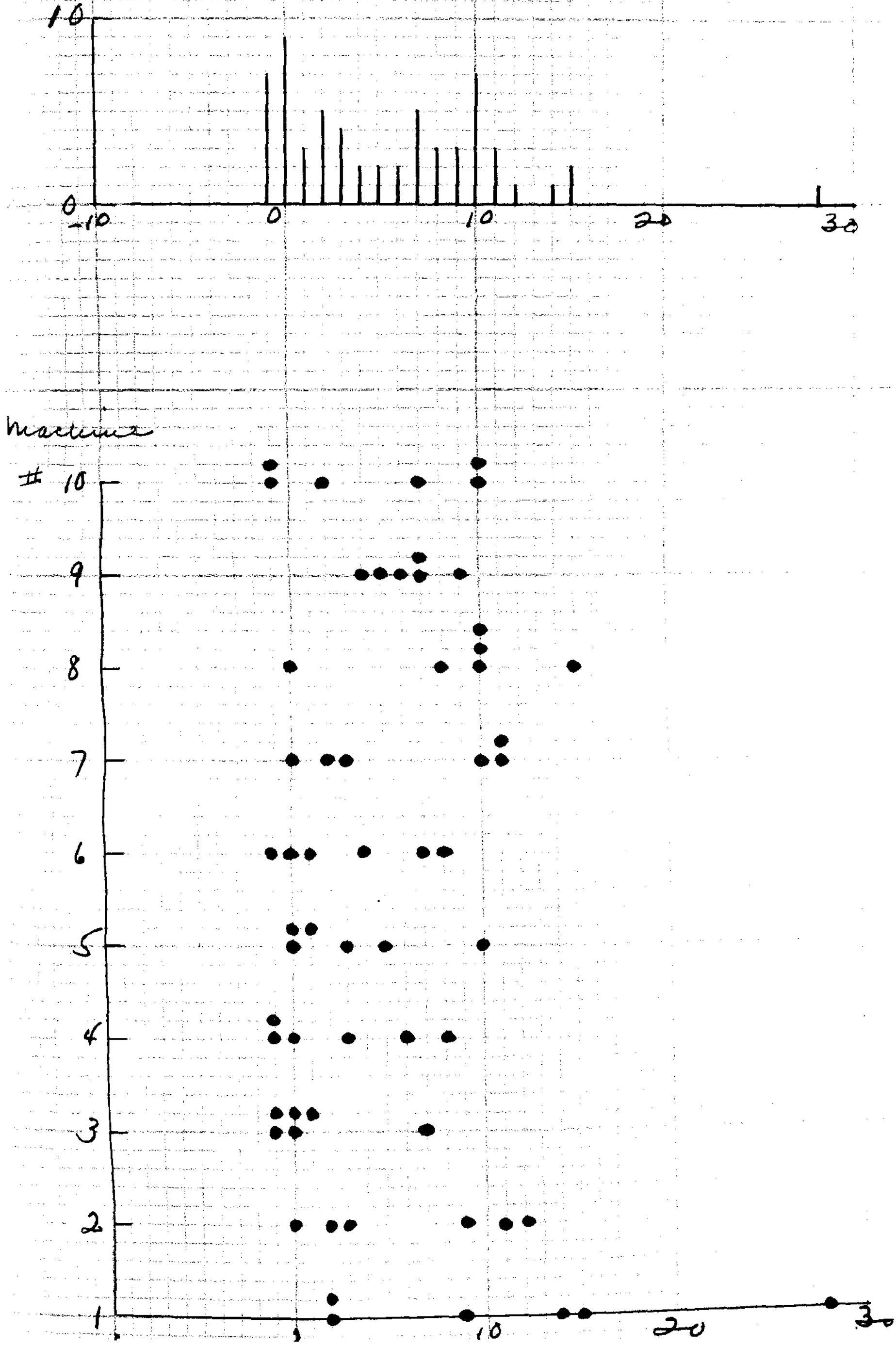
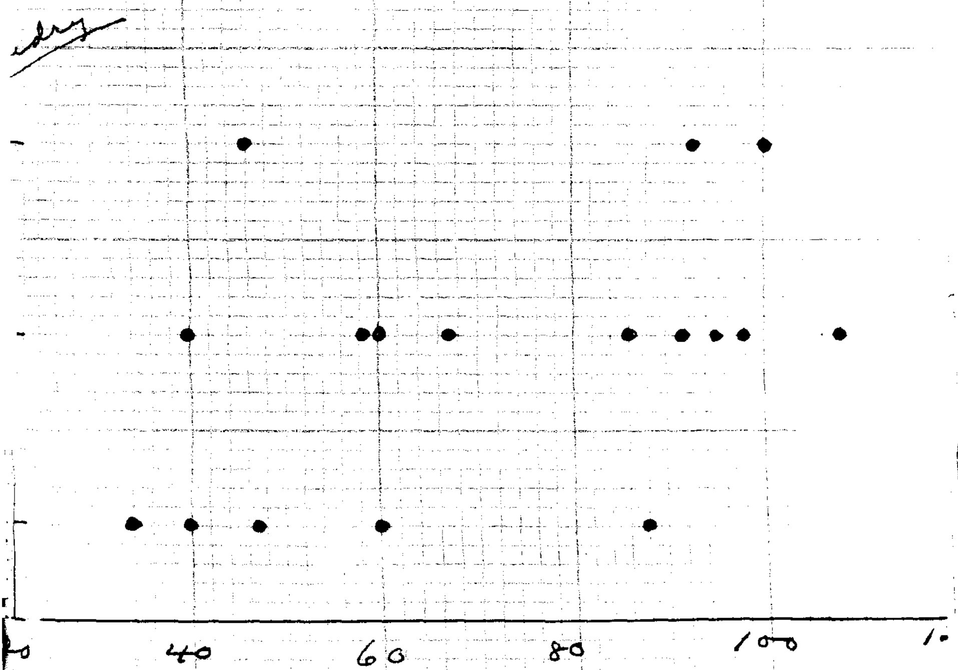


Fig. 3

Brownlee, p. 39



Brownlee concludes that differences in variance are entirely nonsignificant.

All of these tests valid <u>only</u> if <u>separate small samples are random</u>. No attention is given to such tests by Brownlee.

CRITERIA FOR RANDOMNESS AND CONTROL

1. The problem

S 25438

Past Present Future

2. Seven sequences of 144 numbers.

S 25405

S 25407

3. Randomness a necessary but not sufficient condition for control.

4. Criteria

   4.1 Accuracy-precision chart.

   4.2 The gap test.

   4.3 Runs above and below percentile.

   4.4 The eta chart.

S 36394

   4.5 Serial correlation.

EXAMPLES

1. 18 values of g.

S 28184

S 30527

S 30530

2. 144 observed values of thickness of inlay.

Data	S 21612
Distribution	S 30568
Runs	S 30528
Eta	S 28088
Eta (random)	S 28085

3. Birge data

Data	S 31019
Runs	S 30669
Eta	S 38086
Eta (random)	S 30535

4. Chemical measurements

Clarke	S 30981	30983	30982
Iron alloy	S 28594		
" "	S 30584		
(eta)	S 28592		

HOW LARGE SAMPLE REQUIRED FOR TESTING FOR RANDOMNESS

The answer to this depends on experience and cannot be given by theory.

Following slides show that it is certainly greater than 10.

S 33171

S 25408

S 25406

Practical answer is at least 100.

In factorial analysis it is usually assumed that you know a priori the k factors to be investigated

$$F_1, F_2, \ldots, F_i, \ldots, F_k.$$

My experience shows in all cases that I have studied that important unknown factors are always present.

If we know a priori the k factors, $F_1 \ldots F_n$, then we can follow customary theory of Design of Experiment.

SETTING TOLERANCE LIMITS AFTER STATISTICAL STATE OF CONTROL HAS BEEN OBTAINED

If we set tolerance limit to cut off 1% of product with a probability that only once in 100 times will more than 1% be rejected, and if the variable is normally distributed we require

140 observations

The more general problem of setting tolerance limits on the important quality characteristics is one in discriminatory analysis.

FUTURE OF STATISTICS IN ENGINEERING EXPERIMENTATION