# ANNUAL SURVEY OF STATISTICAL TECHNIQUE.

## DEVELOPMENTS IN SAMPLING THEORY

By W. A. Shewhart

### INTRODUCTION

It was with considerable hesitation that I accepted the invitation to write this introductory survey on statistical technique, realizing the extensive annual additions on the formal or mathematical side and the serious problems of induction involved in appraising the usefulness of these formal developments. My experience as a member of a recent committee of the American Statistical Association on Research in Statistics had done much to make me sensitive not only to these difficulties, but also to those arising from the divergence of opinion among statisticians as to what constitutes statistical technique. In fact, two members of this committee attached to the committee's report[1] statements emphasizing the fact that statistical technique in its broader sense should include not only that of mathematical analysis, but also techniques of questionnaire drafting, sampling under field conditions, tabulation of data, computation, and planning.

One of the first things that I undertook was to survey the rather extensive literature that has grown up on the application of statistics in economics and some of the associated sciences, for the purpose, among other things, of getting a better picture of what economists include under the subject of statistical technique. Beginning at least as far back as 1925, the annual addresses[2] of the past presidents of the American Statistical Association, including those of Chaddock, Ayres, Day, Snyder, Wilson, Rorty, and Ogburn, were concerned to a large extent with the discussion of some phases of methodology. In 1930 the program of the annual meeting of this Society contained numerous papers on statistical methodology contributed by men in various fields.

After reading this series of addresses and papers, together with some similar ones appearing in other statistical journals, I have become more convinced than ever that *the* answer to the question, What is statistical technique? cannot be given. For example, Ogburn in his presidential address, 1931, states that statistical method must vary according to subject matter, and that statistics must be taught in various departments where it is applied. He also says that, for these reasons, statistics tends to break up into separate disciplines somewhat as philosophy did. In others of these articles, we find evidence of somewhat similar beliefs.

[1] Cf. *Proc. Amer. Stat. Assoc.*, March, 1932, pp. 252–259.

[2] Printed in the March issues of the *Jour. of the Amer. Stat. Assoc.*, 1926 to 1932, respectively.

Furthermore, such terms as common sense, logic, judgment, and scientific method, are used by some of these authors as though they were apart from, and of greater significance than, statistical method or technique. In fact, Rorty, 1930, states that a statistician should be instinctively and primarily a logician and a scientist in the broader sense, and only secondarily a user of specialized statistical techniques.

It seems desirable, therefore, that this first review of statistical technique should attempt to throw some light on the developments taking place in the conceptual relationship between logic and scientific method on the one hand, and statistical technique on the other, and should indicate in a general way the lines of development in attacking those problems which may perhaps be considered as belonging exclusively to the field of statistical methodology. This need for unification of the theoretical-quantitative and the empirical-quantitative approach to economic problems is emphasized as one of the objectives of the Econometric Society.

## THE FUNDAMENTAL PROBLEM OF INDUCTION

Both natural and social scientists are interested in the *discovery* of so-called facts, relationships, and causes. Since the existence of such entities must be induced from observed data, scientists are interested in gathering, presenting, and interpreting data, both quantitative and qualitative. This problem is common to the fields of scientific method, logic (including the theory of knowledge), and statistical method. To this extent the three fields are closely related.

It is, of course, common to distinguish between two kinds of induction: intuitive and *rational*. By mutual agreement, the discussion of technique in any one of these three fields is usually limited to that having to do with rational induction. It is also true that modern students of induction such as Keynes, Nicod, Ramsey, Broad, Johnson, Lewis, Whitehead, Russell, and Eddington, as well as scientists in general, are in agreement that rational induction can never lead to certainty. In other words, they are in agreement that no matter how many data we gather, analyze, and interpret, we can never do more than say that such and such is probably a fact, or such and such relationship probably exists, or such and such is probably the cause of some event.

The end of research in this sense is the establishment of a judgment that such and such is a fact, relationship, or cause, in which we can believe with a certain degree of rational belief or probability. We shall assume that if the inference or judgment $P$ is connected to the evidence $Q$ through some probability relation, there is an objective rational degree $p_b'$ of belief in $P$ upon the evidence $Q$. To the extent to which this assumption is justified, the ultimate object of research in any field is the establishment of the rational degrees of belief to be associated with such judgments.

Furthermore, it is generally agreed that the objective degree of belief $p_b'$, in an inference or judgment $P$, is not an intrinsic property like truth, but inheres in the inference or judgment through some relation to the evidence $Q$. In other words, we are not justified in saying that $p_b'$ is the degree of rational belief or probability in the judgment $P$. We must add that this is the degree of rational belief upon the evidence $Q$.

Limiting ourselves to the quantitatively measurable characteristics of a phenomenon, it is customary to distinguish conceptually between the phenomenon which exhibits under presumably identical conditions the same magnitudes of physical characteristics, and the phenomenon which under presumably identical conditions exhibits stability of the measurable characteristics only in the statistical sense. For a considerable period of time phenomena in the field of physics were assumed to behave in accord with the first concept. Contrasted with the "exactness" of physical science there grew up the concept of the "inexactness" of most phenomena in the field of social science, and in such other fields, for example, as biology and psychology. For example, in the first edition of Yule's classic, *An Introduction to the Theory of Statistics*, 1910, he characterized inexact phenomena as those arising from a multiplicity of causes, and stated that statistical methods are those adapted to the elucidation of quantitative data affected to a marked degree by such a multiplicity of causes.

Perhaps it is traceable to such distinctions, made not only by Yule but by many other writers before and since, that many have come to consider that statistical methods have to do with only a limited class of data, and that they are not necessarily and inherently associated with scientific method as such. Today, of course, we appreciate as never before that even in the field of physics, it is not possible to duplicate phenomena exactly. Accordingly, in all fields of science we have been forced at one stage or another to introduce the statistical concept of constancy, involving the following assumption: If a sequence of events happens under the same essential conditions, where an event is characterized in terms of $c$ quality characteristics $X_1, X_2, \cdots, X_i, \cdots, X_c$, then the ratio $p$ of the number of events in the sequence of $n$ such events having characteristics falling within the respective ranges $X_1 \pm \frac{1}{2}dX_1, X_2 \pm \frac{1}{2}dX_2, \cdots, X_i \pm \frac{1}{2}dX_i, \cdots, X_c \pm \frac{1}{2}dX_c$, to the total number $n$ of such events, approaches a definite limit[3] $p'$, as the

[3] Of course, this limit is inherently different from a mathematical one in that we never reach a value $n_0$ of $n$ such that for $n \geqq n_0$, the difference $|p - p'|$ becomes and remains less than some previously assigned positive quantity $\epsilon$.

number $n$ increases indefinitely, where this limit is termed a statistical probability.

That things are not reproducible in the exact sense of the older natural sciences, complicates the problem of induction, and at the same time opens up a new field of investigation which seemingly should be considered as the field of statistical methodology, or technique. From this viewpoint we readily distinguish three characteristic statistical problems: specification of objective distributions, deduction of distributions of statistics of samples of size $n$ drawn at random from a given specified universe, and statistical induction.

In other words, instead of assuming the objective constancy of any variable $X$ observed under essentially the same conditions, we assume an objective probability distribution which may be either continuous or discontinuous. For the purposes of illustration, let us assume that the objective probability distribution is a continuous function $f'$ of $X$ and $m'$ parameters, or, in other words, that the statistical probability $dp$ that an observed value of $X$ will lie within the range $X \pm \frac{1}{2}dX$ is

$$dp = f'\ (X, \lambda_1', \lambda_2', \cdots, \lambda_{m'}')dX. \tag{1}$$

The variable $X$ may, of course, be an observed regression coefficient or estimate of a parameter in a physical law as well as a direct measurement. A fundamental problem is the discovery or inference of (1) from an observed set of data.

### SPECIFICATION

The problem of specification is essentially that of fixing upon the functional form $f'$ involving the $m'$ parameters of unknown[4] magnitude to be used as a hypothetical basis for the solution of the other two problems—distribution and induction. In other words, it is the second step in the application of scientific method involving observation, hypothesis, deduction, and experimental verification. Obviously, the scientific investigator must in any field construct a mental model of phenomena he observes, and then test its consistency with itself, and its concordance with the results of further experiment. An hypothesis so conceived is not something fixed, but something subject to change if at any stage in an investigation the experimentalist becomes convinced that the observed data are not consistent therewith.

This concept of the rôle of hypothesis has thoroughly permeated the field of natural science—particularly the field of physics. However, only within the last few decades has it begun to influence the work of the statistician, as may be illustrated by reference to the treatment of

[4] Of course, as a basis for some statistical inferences we assume both the form $f'$ and the parameters to be known.

frequency curves in texts on applied statistics. Such treatments often start by discussing methods of collecting and tabulating observed frequency distributions, and of summarizing these in terms of measures of central tendency, dispersion, and skewness, as though these processes were entirely separate from interpretation. Obviously, however, the interpretation of an observed distribution will depend upon whether or not there is justification for believing that it was produced at random from a constant system of chance causes. Furthermore, the statistics to be used in summarizing the observed data in respect to central tendency, dispersion, and similar characteristics, will depend upon the assumption adopted in respect to the constancy of the chance cause system and the functional form of the objective distribution function. The older treatment of frequency curves was more from the viewpoint of graduation of observed distributions than from that of establishing the necessary hypothetical basis for testing the assumptions of constancy of the underlying chance cause system and of giving efficient ways of summarizing the essential information which recent developments in statistical technique emphasize.

One of the earliest attempts at specification is that having to do with the establishment of the so-called law of error. Later the development of generalized frequency curves was attacked from several directions[5] by Gram (1879), Thiele (1889), Charlier (1905), Pearson (1895), Edgeworth (1896), Fechner (1897), and Bruns (1897). Out of these efforts belonging to the nineteenth century have crystallized two general systems—Pearson's system of closed curves and the Gram-Charlier open series. During the last two decades, however, there has been no material development from this angle except by way of appraising the comparative advantages[6] of these two major systems of specification.

By taking enough terms in the open Gram-Charlier series, we may theoretically secure a perfect fit to an observed distribution, but even a superficial knowledge of the fluctuations to be expected in a series of observed frequency distributions of samples of a given size taken under essentially the same conditions leads one to question the significance of closeness of fit thus obtained. It was not until 1900 that statisticians were given the necessary tool in the form of Pearson's $\chi^2$ test to enable them to test the hypothesis that an observed distribution could have arisen from some specified universe with known parameters; not until

[5] Cf. Rietz, H. L., *Mathematical Statistics* (Chicago: The Open Court Publishing Company, 1927), Chap. III.

[6] One such interesting appraisal is that of J. F. Steffensen in *Some Recent Researches in The Theory of Statistics and Actuarial Science* (Camb. Univ. Press, 1931), pp. 35–48. He gives reasons for believing that Pearson's curves are preferable.

later that the $\chi^2$ test was extended to allow for estimates of the parameters from a sample.

DISTRIBUTION

In order to test the hypothesis that one or more values of a variable taken under presumably the same essential conditions constitute a random sample from a postulated universe, it is necessary to have deduced the nature of the variation that may be expected from observation to observation, or from some function of a group of observations to the same function of another group, so that the observed may be compared with the theoretical variability. Now, it is customary to attach significance to both confirmation and infirmation as a basis for inference. However, there is little object in studying confirmatory evidence unless we can feel assured that we have succeeded in maintaining the same essential conditions and that the observed values are free from constant errors.

As an illustration, we may consider the measurement of an objective constant such as the charge on an electron. The literature certainly reveals the tendency of the physicist to emphasize first of all the need for gaining assurance that the observations are free from constant errors, and that the differences between observations are not significant. In fact, the physicist seldom tabulates more than five or ten observations after he has satisfied himself that he has maintained the same essential conditions and eliminated constant errors. Of course, Millikan in his original report on the measurement of the charge on an electron presented fifty-eight values, but his final estimate was based upon a comparatively small group because he felt that the others gave evidence of bias.

Since it is essential first to determine whether or not observed differences are significant, it is apparent that the technique of testing the hypothesis of constancy must be such that it can be used in testing significant differences between small groups of observations, if it is to be of much importance to the research man who realizes, as does the physicist, the lack of significance of multiplication of instances until assignable differences and constant errors have been eliminated, in so far as it is humanly possible to do so.

Keeping this situation in mind, we are in a position better to appreciate the practical significance of the developments in distribution theory taking place since 1900. Although prior to that time, statisticians were using various measures of central tendency, dispersion, skewness, kurtosis, and correlation, nevertheless comparatively little was known about the distributions of such measures for samples of any given size $n$ drawn from even such a simple specified universe as the normal distribution. Of course, the distribution of the averages of samples had been known for a long time prior to 1900; so also had the distribution of the standard deviations been discovered by Helmert in 1876, although this work was overlooked by most statisticians until some time after it was rediscovered empirically in 1908 by "Student," and rigorously in 1915 by R. A. Fisher. It remained for theoretical statisticians of the present century to deduce the distribution functions of various statistics as a function of sample size $n$, first under the assumption that the samples had been drawn at random from the normal law, and then to investigate the effects of functional forms of the universe upon the distribution of these statistics.

It should be emphasized that the object of such studies is not to remove the necessity of taking so-called large samples, but rather to put into the hands of the research man a useful tool for detecting significant differences applicable to small samples of observations which the careful scientist recognizes must be done before significance can be attached to large samples.[7]

Since 1900, exact distribution functions for any sample size for many statistics of samples drawn from a normal universe have appeared in rapid succession, largely at the hands of R. A. Fisher and his associates at the Rothamsted Experimental Station. Preliminary studies have been made, in particular by Rider and by Egon Pearson and some of his associates, to determine the effect of the functional form of the universe upon the distribution functions of some of the more important statistics. An excellent survey of much of this work up to 1930 has recently been given.[8] Similar reviews by Irwin for the years 1930 and 1931 have appeared in the *Journal of the Royal Statistical Society* for 1931 and 1932, respectively. Prospective students of this subject will find the extensive bibliographies of Rider and Irwin of great help. It may be of interest that Irwin lists 89 articles appearing in 1931.

Not only has marked progress been made in deducing exact distribution functions of many important statistics, but also in showing how many of these distribution functions are related to four compara-

[7] On this important point, Egon Pearson makes the following pertinent comment: "Critics of small sample theory are inclined to argue, 'What can you infer from two samples of five? Give us two samples of 100 and we may tell you something!' They do not realize that the distinction is often not of this kind at all. The comparison lies between a technique which can deal with *both* (a) 40 samples of 5 (or even 100 of 2), and (b) 2 samples of 100, and a technique which can only deal accurately with (b)." "A Survey of the Uses of Statistical Method in the Control and Standardization of the Quality of Manufactured Products," read before the Royal Statistical Society, December 20, 1932.

[8] Rider, Paul R., "A Survey of the Theory of Small Samples," *Annals of Mathematics*, Second Series, XXXI (October, 1930), 577–628.

tively simple distributions, namely, the normal law, Karl Pearson's $\chi^2$, "Student's" ratio of the error of the average of a random sample drawn from a normal universe to the observed standard deviation, and R. A. Fisher's variable $z$, defined by the relation $e^{2z} = (s_1^2/s_2^2)$, where $s_1^2$ and $s_2^2$ are estimates of the variance of a normal population derived from two samples from that population. Largely to the work of R. A. Fisher do we owe this contribution, and as early as 1924 he gave the following table[9] to indicate the applications of the four chief cases of the $z$ distribution:

| I<br>Normal curve | II<br>$\chi^2$ | III<br>Student's | IV<br>$z$ |
|---|---|---|---|
| Many statistics from large samples | Goodness of fit of frequencies<br>Index of dispersion for Poisson and Binomial samples<br>Variance of Normal samples | Mean<br>Regression coefficient<br>Comparison of means and regressions | Intraclass correlations<br>Multiple correlation<br>Comparison of variances<br>Correlation ratio<br>Goodness of fit of regressions |

Since 1924, Fisher, Hotelling, Egon Pearson, McKay, and others have succeeded in relating certain other exact distributions to one or the other of these four.

The practical importance of being able to relate the exact distribution functions of many statistics to a comparatively small number of distributions is that it reduces the number of tables required for testing hypotheses in respect to various statistics. For example, it was pointed out by Fisher in 1924 that the distribution of the ratio of the observed variance $s^2$ for a sample of $n$ drawn from a normal universe with standard deviation $\sigma'$ to the square of the standard deviation $\sigma'$ is distributed as $\chi^2$, where $\chi^2$ is put equal to $n_1 s^2/\sigma'^2$, and where $n_1$ is the number of degrees of freedom, being one less than the size $n$ of the sample.

Thus, for example, we might ask the question: In a sample of sixteen drawn from a normal universe, what is the probability of occurrence of an observed standard deviation 20 per cent greater than the unknown true standard deviation of the parent population? Karl Pearson[10] solves this problem through the use of his tables giving indirectly the integral of the distribution of the standard deviation and obtains the probability 0.08329. We may, however, make use of the relationship of

[9] "On a Distribution Yielding the Error Functions of Several Well-known Statistics," presented at the International Mathematical Congress, Toronto, Canada, 1924. This paper outlines the framework around which Fisher's *Statistical Methods for Research Workers* has been built.

[10] *Tables for Statisticians and Biometricians*, Part II, page civ.

$\chi^2$ to the ratio $n_1s^2/\sigma'^2$ and through the use of the well-known $\chi^2$, $P$ tables obtain directly the probability 0.0834.

## STATISTICAL INDUCTION

*A. Estimation.*—Assuming that a sample is a random one from a population of a given functional form but with unknown parameters, there arises the important practical problem of estimating the magnitudes of the parameters. Thus for the case of one variable specified by (1) where $f'$ is assumed known but the $\lambda''$s are assumed to be unknown, we have the problem of estimating the $\lambda''$s from the available information $Q$. Now, of course, we may limit $Q$ to the evidence given by the sample, or we may include in addition as a part of $Q$ evidence available before the sample was taken.

In 1922, Fisher[11] proposed three criteria of estimation: (a) The criterion of consistency, which states that when applied to the whole population, the derived statistic should be equal to the parameter; (b) The criterion of efficiency, which states that in large samples, when the distributions of statistics tend to normality, that statistic should be chosen which has the least probable error; (c) The criterion of sufficiency, which states that the statistic chosen should summarize the whole of the relevant information supplied by the sample. This criterion of sufficiency is interpreted by Fisher as meaning that if $\lambda_i'$ be the parameter to be estimated, $\Theta_{i1}$ a statistic which contains the whole of the information as to the value of $\lambda_i'$ which the samples supplies, and $\Theta_{i2}$ any other statistic, then the surface of distribution of values $\Theta_{i1}$ and $\Theta_{i2}$, for a given value of $\lambda_i'$, is such that for a given value of $\Theta_{i1}$, the distribution of $\Theta_{i2}$ does not involve $\lambda_i'$. The researches of Fisher have led him to the conclusion that an efficient statistic can, in all cases, be found by what he terms the method of maximum likelihood, assuming no other information than that the sample has been drawn at random from a population of given form but with unknown parameters.

When a sample is so large that it would be generally agreed by students of the subject that the information given by the sample overshadows all pertinent information available prior to the taking of the sample, these criteria, I believe, are generally accepted as constituting a rational basis for deciding upon estimates of parameters. On the other hand, as previously noted, it is generally agreed by students of induction that any induction or judgment should be consistent with all available pertinent information. Certainly, in the case of small samples, this generally recognized criterion must be considered in conjunction

[11] *Trans. Roy. Soc. of London*, Series A, Vol. CCXXII, 309–368.

with those for the consistency, efficiency, and sufficiency, of an estimate based upon a sample. Recent discussions bearing upon this general subject have been given by Molina and Wilkinson,[12] Karl Pearson,[13] and others. In fact, Pearson introduces in contrast with the "most likely" estimate based solely upon a sample, what he calls the "most reasonable" estimate which takes into account prior information. In this connection, Pearson argues that in almost all cases there is some, even though vague, *a priori* experience existing that should be used, together with the information given by the sample, in finding what he calls a most reasonable estimate.

This argument certainly has weight, because in a practical case we seldom, if ever, know *a priori* that a sample has been drawn at random from a universe of given functional form, and it usually works out in such instances that before a practical man is willing to believe that the given method of measurement will lead to the sought-for objective value as a statistical limit simply through increasing the size of the sample, he has already formed some judgment that the objective value lies within some agreed-upon range and experienced what he considers a rational degree of belief in this judgment that must certainly be considered when interpreting a small sample.

In any case, however, a knowledge of the distribution of a given statistic as a function of sample size enables the experimentalist to determine the relative significance of increasing the number of observations under statistically controlled conditions, and shows that, at least for efficient statistics, the degree of precision attained is inversely proportional to the number of observations. Furthermore, a knowledge of the distribution in any given case enables one to appreciate the significance of the variability introduced in the process of sampling. As an illustration, suppose we are given a sample of four known to be drawn from a normal universe of unknown average $\overline{X}'$ and standard deviation $\sigma'$. Based upon this sample, what shall we take as estimates of $\overline{X}'$ and $\sigma'$? The knowledge that for all practical purposes the observed standard deviation for a sample of four may vary all the way from zero to more than twice the true standard deviation $\sigma'$ must of necessity make one cautious in placing too much reliance on any estimate based solely upon a sample of four.

*B. Testing hypotheses.*—Given a random sample $\Sigma$, an important problem is that of determining whether or not it is likely to have come from a certain population $\pi$ which may be either completely or par-

[12] "The Frequency Distribution of the Unknown Mean of a Sampled Universe," *Bell System Tech. Jour.*, VIII (Oct. 1929), 632–645.

[13] *Loc. cit.*, pp. clxxi–clxxx.

tially specified. Two distinct methods of approach are open: (a) To start from the population $\pi$ and seek the probability that a sample $\Sigma$ should have been drawn therefrom; (b) To start from $\Sigma$ and seek the probability that $\pi$ is the population sampled. Two recent papers[14] by Neyman and Egon Pearson break new ground in providing a critical basis for developing and using statistical criteria. In treating the first of these two methods they note two important requirements: (1) We must be able to reduce the chance of rejecting a true hypothesis to as low a value as desired, (2) The test must be so devised that it will reject the hypothesis tested when it is likely to be false. Pearson and Neyman were the first to emphasize the latter point and to provide a criterion or test which takes both into consideration.

In two later papers[15] they extend their studies to the problems of two and $k$ samples. In the case of two samples they fix upon the set of admissible hypotheses concerning the populations $\pi_1$ and $\pi_2$ from which two samples $\Sigma_1$ and $\Sigma_2$ have been drawn, restricting this by the further assumption that $\pi_1$ and $\pi_2$ are both normal. They develop a test for the hypothesis $H$ that $\pi_1$ and $\pi_2$ are identical. They then show how R. A. Fisher's $z$ test may be used in testing the hypothesis $H_1$ that the two samples $\Sigma_1$ and $\Sigma_2$ have come from unknown normal populations with the same variance but with means having any values whatsoever. Similarly, they show that if it be assumed that the populations $\pi_1$ and $\pi_2$, besides being normal, have the same variance, the "Student" $t$ test is the appropriate one for testing the hypothesis $H_2$ that the means of these populations are the same. They then argue that their criterion for testing hypothesis $H$ is more crucial than tests of $H_1$ or $H_2$ taken separately although they point out that in many problems the hypothesis to be tested will present itself in one of the two forms, either $H_1$ or $H_2$.

These results are of fundamental importance in testing for assignable causes of variation or lack of homogeneity, which tests should always be made before one places too much significance on the contributions of a single sample no matter how large. The importance of this work, however, is not alone in the new criteria provided, but also in the fact that these papers present an excellent critical introduction to some of the philosophical and logical, as well as the mathematical, problems involved in developing rational techniques for testing hypotheses.

[14] "On the Use and Interpretation of Certain Test Criteria for Purposes of Statistical Inference. Parts I and II," *Biometrika* XXA (July and December 1928), 175–240 and 263–294.

[15] "On the Problem of Two Samples" and "On the Problem of $k$ Samples," *Bull. de l'Académie Polonaise des Sciences et des Lettres* 1930, pp. 73–96, and 1931, pp. 459–481.

In trying to appraise contributions in statistical technique, we must keep in mind that the ultimate judgment or inference cannot be stated with certainty, and that even though we were in a place to give an accurate estimate of the degree of rational belief $p_b'$ associated with a given judgment, this degree of rational belief is based upon certain evidence $Q$. Anyone passing upon the significance to be attached to a judgment based upon a given set of data must, of course, attempt to *weigh*[16] the significance of each of the contributory quantitative and qualitative elements going to make up the pertinent information $Q$.

### Some Important Books on Technique, 1931–32

1. *Tables for Statisticians and Biometricians, Part II*, edited by Karl Pearson. London: Cambridge University Press, 1931. 512 pages. England 33*s*, export $7.30.

Contains not only the tables issued in *Biometrika* during the past seventeen years, but also 250 pages of valuable and critical discussion of the use of these tables.

2. *The Methods of Statistics*, by L. H. C. Tippett. London: Williams and Norgate, Ltd., 1931. 222 pages. 15*s*.

An excellent introduction to many of the recent developments in statistical technique made possible through developments in distribution theory during the past three decades.

3. *Contributions to the History of Statistics*, by Harold Westergaard. London: P. S. King & Son, 1932. vii+280 pages. 12*s*. 6*d*.

Valuable as a survey of the evolution of statistical technique up to 1900.

4. *Statistical Methods for Research Workers*, by R. A. Fisher. 4th edition. London: Oliver & Boyd, and New York: G. E. Stechert & Co., 1932. xi+307 pp. 15*s*.

Revised and enlarged edition, bringing up to date the survey of R. A. Fisher's major contributions to statistical technique from 1912 to 1932.

5. *An Introduction to the Theory of Statistics*, by G. U. Yule. 10th edition. London: Charles Griffin and Co., Ltd., and Philadelphia: J. B. Lippincott Company, 1932. xv+434 pages. 12*s*. 6*d*.

Principal item of revision is that of bringing the excellent bibliography in this classic up to date.

[16] Cf. Keynes, J. M., *A Treatise on Probability* (New York: Macmillan Co., 1921), p. 77 *et seq.*

6. *Business Statistics*, by Riggleman and Frisbee. New York and London: McGraw Hill Book Company, Inc., 1932. xix+707 pages. $4.00.

Part I gives an interesting introduction to statistical methods in respect to techniques of gathering, tabulating, and presenting data, and calculating averages, index numbers, dispersions, and other simple statistics, in so far as this can be done without giving due consideration to the developments of the present century as here reviewed in respect to the problems of specification, distribution, and induction. Part II, comprising approximately one-half of the book, constitutes an excellent introduction to some of the complex business and economic problems faced by the industrialist.

The student of statistical technique will doubtless find it interesting to start with Westergaard's history, review Yule's interesting summary and the practical discussions of applied technique in the field of business as presented by Riggleman and Frisbee. He will note the emphasis on techniques of collecting and presenting data, and those of calculating certain characteristics of observed data. If now he studies the above-mentioned texts by Pearson, Fisher, and Tippett, he will experience a definite change in emphasis and will begin to appreciate that the rational interpretation of data depends not only upon a knowledge of the statistical techniques involved in testing statistical hypotheses, but also that the techniques of collecting and presenting data inherently depend upon the techniques and the hypotheses that are going to be used in the interpretation thereof.

*Bell Telephone Laboratories*