Journal of the
Operations Research
Society of Japan

VOLUME 3 *March 1961* NUMBER 4

SILVER JUBILEE OF MY DISPUTE WITH FISHER[1]

JERZY NEYMAN

Statistical Laboratory, University of California, Berkeley.

(Received Feb. 1, 1961)

1. INTRODUCTION

After publishing Sir Ronald Fisher's article [1], the Editors of the Operations Research Society of Japan called my attention to Sir Ronald's uncomplimentary references to my work and offered to publish a reply.

The first expressions of disapproval of my work were published by Fisher [2] in 1935. During the intervening quarter of a century Sir Ronald honored my ideas with his incessant attention and a steady flow of printed matter published in many countries on several continents. All these writings, equally uncomplimentary to me and to those with whom I was working, refer to only five early papers, [3] to [7], all published between 1933 and 1938.

In general, scientific disputes are useful even if, at times, they are somewhat bitter. For example, the exchange of opinions and the studies surrounding the definition of probability given by Richard von Mises, clarified the thinking considerably. On the one hand, this dispute brought out the superiority of Kolmogoroff's axiomatization [8] of the theory. On the other hand, the same dispute established firmly von Mises' philosophical outlook on "frequentist" probability as a useful tool in indeterministic studies of phenomena. There are many similar examples in the history of science. Thus, at least in its early stages, my

1) Prepared with the partial support of the National Science Foundation Grant G-14648.

dispute with Sir Ronald was probably useful. Unfortunately, from the very start it has been marred by Sir Ronald's unique style involving torrents of derogaotry remarks. Here is a sample collected from the ten page silver jubilee article [1] of Fisher: "In expressing this opinion ...Neyman was...some hundred years out of date", "...a curious misapprehension...", "it is astonishing to find this elementary error, incorporated in the teaching...", "Neyman's doctrine challenged...", "many young men...have been partly incapacitated by the crooked reasoning...", "I believe sanity and realism can be restored...", etc. As to the arguments supporting these and similar assertions, the concluding paragraph of the article is indicative:

> "As I have already spoken at some length, perhaps you will forgive me if I do not on the present occasion enter into the detailed mathematics of any such example. I shal hope to expand these matters further in Conferences with my mathematical colleagues."

This is an admission that the paper does not contain any scientific material.

Because of my admiration for the early work of Fisher [5], his first expressions of disapproval of my ideas were a somewhat shocking novelty and I did my best to reply and to explain. Later on, the novelty wore off and I found it necessary to reply only when Fisher's disapprovals of me included insults to deceased individuals for whom I felt respect. My last paper [9] in reply to Fisher was written mainly because of Sir Ronald's particularly virulent attack [10] on Abraham Wald, published soon after the latter's untimely death. Subsequent polemical writings of Fisher, including a book [11], I left without reply. The purpose of the present paper is to mark the silver jubilee of the dispute, to reiterate my appreciation of the early scientific work of Fisher, to present regrets that currently it has been largely replaced by futile polemics, and to express my hope that in the years to come we shall see some more of Fisher's true research comparable to that he did in the prewar years.

2. MY POINT OF VIEW OF FISHER

Fisher's earlier excellent work, earned him general recognition [12], was concerned with the theory of experimentation of which he is the undisputable founder and leader. It is in connection with this domain and with his work at the Rothamsted Experimental Station that Fisher noticed the necessity for solving certain very difficult distribution problems and actually solved a surprising number of them. Here again Sir Ronald has no rivals. His first large paper, concerned with the distribution of the correlation coefficient [13], marked an epoch and it is regrettable that this paper is not included in the big volume of Fisher's collected works [14]. However, even so, this volume contains a great number of Fisher's writings illustrating his remarkable activity. The paper linking the distributions of χ^2, of Fisher's z and of Student's t is one of them. Then there are papers giving the distribution of the multiple correlation coefficient, the limiting distribution of the largest member in a sample, the paper on the distribution of gene ratios for rare mutations, etc., etc.

In addition to works on the theory of experimentation and on innumerable problems of distribution, Sir Ronald wrote a number of papers concerned with the foundations of statistical theory. Here again Fisher's paper were very useful and left an indelible imprint on statistical literature. However, in this third domain of his work, Fisher was somewhat less successful than in the two others: his mastery of manipulative mathematics was far ahead not only of that of his contemporaries, but also of his own mastery in operating with concept. At one time he seems to have realized this himself for both his large papers on foundations of the theory contain passages ([15], p. 323, [16] p. 700—701) indicating regrets about his inability to prove certain statements (subsequently found inaccurate), beliefs that certain other statements could be proved rigorously, and admissions of having no intention to do so. In my opinion, Sir Ronald's difficulties were not limited to proofs of stated propositions. They were, and still are, more pronounced in his efforts to deal with abstract concepts. This is reflected in Sir Ronald's tendency to "postulate away" the difficulties he encounters by formulating dogmas. One such dogma is Professor Fisher's "measure of confidence or diffidence" in the form of "likelihood". Possibly, it is this tendency to dogmatism that is at the root of the unparalleled violence of Sir Ronald's reactions to other schools of thought.

3. THE ESSENCE OF THE DISPUTE

The subject of the dispute may be symbolized by the opposing terms "inductive reasoning" and "inductive behavior". Professor Fisher is a known proponent of inductive reasoning. After a conscientious effort to find the exact meaning of this term, I came to the conclusion that, at least in the sense of Fisher, the term is empty, except perhaps for a dogmatic use of certain measures of "rational belief" such as the likelihood function and the fiducial probability.

The term "inductive behavior" means simply the habit of humans and of other animals (Pavlov's dog, etc.) to adjust their actions to noticed frequencies of evnts, so as to avoid undesirable consequences. From the point of view of inductive behavior [7], mathematical statistics is the discipline concerned with the frequency properties of various rules of behavior based on observable random variables (for example: rejection or acceptance of hypotheses, methods of estimation, etc.) and, in particular, with the search for rules (or "strategies") that satisfy some pre-assigned conditions (most powerful tests, unbiased minimum variance estimators, maximin ordering policies, etc.). The most charitable of the opinions of Professor Fisher about this point of view of statistics is that, while possibly useable in industry and by mentalities "geared to technological performance", such as Russian and American [10, p. 70], it is inconsistent with the process of drawing correct conclusions in Natural Science.

This is the general background of the dispute. Its specific phases concern the theory of testing statistical hypotheses which Egon S. Pearson and I developed in the 1930's [3], [4], and with the theory of estimation, by confidence intervals, on which I did some work at about the same time [5], [6], [7]. Sir Ronald heartily disapproves of both.* On my own part, I consider that Sir Ronald's "fiducial argument", as distinct from the theory of confidence intervals with which it is occa-

* An exception in this respect seems to be the concept and the problem of similar regions, intervening in both theories. The first formulation of the problem and its first partial solution were given in 1933[3]. Sir Ronald later gave the problem the picturesque label "the problem of the Nile" [7] and since 1936 refers to it from time to time [11, p. 163]. Without mentioning its original source or its original solution.

sionally confused, is no more than a misconception born out of an early mistake [18].

This mistake is conceptual and consists in confusing two kinds of probabilities. One of them is the probability of arriving at a correct conclusion as to the value of a parameter θ, if the process of estimation is that by interval $[\theta_1(X), \theta_2(X)]$, the confidence interval, appropriately defined before making the observations on X. The other probability is that the value of the parameter θ will lie within the limits $[\theta_1(x), \theta_2(x)]$ where x is the already observed value of the random variable X. In the simple cases for which the problem of confidence intervals is solved, the value of the former probability may be set in advance and the limits $\theta_1(X)$ and $\theta_2(X)$ determined accordingly, without any reference to probabilities *a priori*. Also, methods were developed to determine the confidence limits $\theta_1(X)$ and $\theta_2(X)$ so as to satisfy certain conditions of optimality. As to the second probability, it must depend upon probabilities *a priori*.

Originally, there was the presumption [5] that Sir Ronald's favorite phrase "fiducial distribution of a parameter in the light of a sample observed" is no more than a slip of the tongue and is not meant to refer to the second kind of the probability just discussed. In fact, it is fair to say that Sir Ronald himself warned his readers not to make this mistake. Accordingly, it was thought the confidence intervals and the fiducial argument are two slightly different aspects of the same theory. However, in his treatment of the so-called Behrens-Fisher problem [19], Sir Ronald operated on "fiducial distributions" as if they were ordinary probability densities, some *a priori* and some *a posteriori*, and this established his conceptual difficulty [20], [21]. When this difficulty was pointed out, Fisher was emphatic in declining to recognize his mistake. Instead, he insisted that his result was obtained correctly and that those who doubt fail to understand the essence of fiducial argument. From then on, in the minds of many statisticians, including myself, the theory of confidence intervals and the fiducial argument of Fisher became two distinct entities. In addition, quite a few statisticians agree with me that fiducial argument is simply a misconception.

4. CURRENT POLEMICS

Interesting as these things are, a quarter of a century is a long period for their discussion. As a consequence, in his uninterrupted stream of published material, Sir Ronald is forced to hash and to rehash the same items. The problem of the test for the two-by-two contingency table is one example. The problem is completely trivial and, if treated from the point of view of the theory of testing statistical hypotheses, its solution is found [22] as easily as, for example, the solution of the problem of the minimum variance unbiased estimator of the mean of a normal distribution. Yet, the two-by-two table has been discussed recently by Sir Ronald at least three times [10, p. 72], [11, p. 85], [1, p. 7]. The content of these discussions is more or less the same, but there is a variety of invectives.

As another example of replication in Fisher's polemical writings, the case of the so-called single event probability may be mentioned. Briefly, it is concerned with the probability $P\{X<Y\}$ where X and Y stand for two perfectly determined numbers, at least one of which is unknown. For instance, Y may be equal to 5 and, to use my own example [6], X may mean the 1000th decimal in the expansion of $\pi=3.14159\cdots\cdots$, the value of which I do not know. The question under discussion is about the value that the probability $P\{X<5\}$ may possess. Naturally [6], the answer depends upon the theory of probability with which one works. If the theory is such as that of Sir Harold Jeffreys, dealing with intensities of belief, then $P\{X<5\}$ may have any value between zero and unity. On the other hand, within the theory of probability with which I like to work, intended as an idealization of relative frequency, the assumption that both X and Y have perfectly determined values implies that the reference set is composed of just one element and, therefore, the probability $P\{X<Y\}$ can be only zero or unity. Under different guises this problem was discussed recently by Fisher at least twice, in [10, p. 75] and [1, p. 4]. In both cases the suggestion that the single event probability can be either zero or unity is treated most unfavorably. Sir Ronald's opinions are: "A complementary doctrine of Neyman violating equally the principles of deductive logic……" and "It is truly astonishing to find this elementary error incorporated in teaching of many

mathematical departments in the United States……". One might think that Sir Ronald adheres to the intensity-of-belief theory of probability. However, on the very next page 5 of his silver jubilee article [1] he endorses the system of probability of Kolmogoroff. And it so happens that the example of a reference set E with just one element ξ is considered on p. 2 of Kolmogoroff's famous book [8]. There are only two possible subsets. One of them is the set E itself and the other the complementary empty subset. The probability of the first is unity. The probability of the second is zero. Did Sir Ronald read this but failed to understand its relevance, or did he endorse Kolmogoroff's theory without reading it?

While the number of scientific questions in Fisher's polemical writings is small, Sir Ronald has an inexhaustible supply of invectives, some of them quite colorful. The following two are my favorites, [10, p. 73] and [11, p. 141].

> "The phrase 'Errors of the second kind', although apparently a harmless piece of technical jargon, is useful as indicating the type of mental confusion in which it was coined."
>
> "A distinction without a difference has been introduced by certain writers who distinguish 'Point estimation', meaning some process of arriving at an estimate without regard to precision, from 'Interval estimation' in which the precision of the estimate is to some extent taken into account. 'Point estimation' in this sense has never been practiced either by myself, or by my predecessor Karl Pearson,……, or by his predecessor Gauss of nearly one hundred years earlier, who laid the foundations of the subject. The distinction seems only to be made in order to support the claim, which is not indeed historical, to the effect that the authors have made in this matter an original contribution. It shows great confidence in the ignorance of students to put such a claim forward."

In addition to the variety of invectives, Fisher's polemical writings abound in unconventional treatments of facts. Some such instances have been recorded before [9]. Ordinarily, misrepresentations are of offensive character and impute to other scholars either incomptence or malice or both. Occasionally they are fantastic. For example, soon after Karl Pearson's death, Sir Ronald expended some ingenuity in order to convince his readers that Pearson was incompetent to the point of doubting the accuracy of the familiar formula npq. Currently we may record a "defensive" misrepresentation.

On p. 56 in [11] Sir Ronald informs the reader that the list of individuals recognizing "the rational cogency of fiducial argument" includes Kolmogoroff. In particular, Sir Ronald refers to Kolmogroff's paper [23] and specifically to the section "Fisher's fiducial limits and fiducial probability". Unfortunately, the information given by Sir Ronald is incomplete.

On page 20 of his article Kolmogoroff describes a "wide spread error". According to him, this error consists in confusing the probability that a ***future application*** of a rule in estimating the parameter θ by indicating the limits $\theta_1(X) \leqq \theta \leqq \theta_2(X)$ will lead to a correct result, with another probability, namely the probability ***a posteriori*** that the parameter θ will lie within the limits $\theta_1(x) \leqq \theta \leqq \theta_2(x)$ where x is the ***already observed*** value of the random variable X. Footnote 14 attached to this description asserts:

"This error has been committed, in some of his works, by Fisher."

This is hardly an unreserved endorsement of "cogency of fiducial argument". It will be noticed that the error pointed out by Kolmogoroff coincides exactly with the error I impute to Sir Ronald since 1934 and which is one of the main subjects of our dispute. It is remarkable that an incident of this kind should happen to a man of Sir Ronald's repute and achievements.

5. CONCLUSIONS

Sir Ronald Fisher's postwar polemical writings represent an unusual phenomenon, probably unparalleled in the history of science. However, because the scientific matters considered are somewhat out of date and have been repeatedly discussed, the interest of this phenomenon lies outside of the general sphere of statistical research. During the intervening period, mathematical statistics develoved several new branches and many new ideas, with quite a few problems still unsolved. It may be hoped that some of these problems will attract Sir Ronald's interest and attention and that, in due course, he will use his outstanding research talent to provide the solutions. The Journal of the Operations Research

Society of Japan may be useful as a source of inspiration.

REFERENCES

1. Sir Ronald Aylmer Fisher, "Scientific Thought and the Refinement of Human Reasoning." *J. Oper. Res. Soc., Japan*, Vol. 3, No. 1 & 2, 1960, pp. 1—10.
2. R. A. Fisher, "Discussion of Dr. Neyman's paper" *J. R. S. S.* Suppl., Vol. 2, 1935, pp. 154—157.
3. J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses." *Philos. Trans. Roy. Soc., London, Ser. A*, Vol. 231, 1933, pp. 289—337.
4. J. Neyman and E. S. Pearson, "The testing of statistical hypotheses in relation to probabilities *a priori*." *Proc. Camb. Philos. Soc.*, Vol. 29, 1933, pp, 492—510.
5. Jerzy Neyman, "On the two different aspects of the representative method ......" *J. R. S. S.*, Vol. 97, 1934, pp. 558—625. (Spanish version of this paper appeared in *Estadística, J. Inter-Amer. Stat. Inst.*, Vol. 17, 1959, pp. 587—651.)
6. J. Neyman, "Outline of a theory of statistical estimation based on the classical theory of probability", *Philos. Trans. Roy. Soc. London*, Ser. A, Vol. 236, 1937, pp. 333—380.
7. J. Neyman, "L'éstimation statistique traitée comme un problème classique de probabilité." *Actualités Sci. Industr.*, No. 739, 1938, pp. 25—27. (Russian version of this paper appeared in *Uspekhi Mat. Nauk*, Vol. 10, 1944, pp. 207—229.
8. A. Kolmogoroff, "Grundbegriffe der Wahrscheinlichkeitsrechnung." Springer, Berlin, 1933.
9. Jerzy Neyman, "Note on article by Sir Ronald Fisher." *J. R. S. S.* Ser. B, Vol. 18, 1956, pp. 288—294.
10. Sir Ronald Fisher, "Statistical methods and scientific induction." *J. R. S. S.* Ser. B, Vol. 17, 1955, pp. 69—78.
11. Sir Ronald A. Fisher, "Statistical methods and scientific inference." Oliver and Boyd, London, 1956.
12. Jerzy Neyman, "Indeterminism in science and new demands on statisticians," *J. A. S. A.*, Vol. 55, 1960, pp. 625—639.
13. R. A. Fisher, "The frequency distribution of the values of the correlation coefficient......" *Biometrika*, Vol. 10, 1915, pp. 507—521.
14. R. A. Fisher, "Contributions to mathematical statistics." Wiley, New York, 1950.
15. R. A. Fisher, "On the mathematical foundations of theoretical statistics." *Philos. Trans. Roy. Soc., London*, Ser. A, Vol. 223, 1922, pp. 309—368.

16. R. A. Fisher, "Theory of statistical estimation." *Proc. Camb. Philos. Soc.*, Vol. 22, 1925, pp. 700—725.
17. Ronald Aylmer Fisher, "Uncertain inference." *Proc. Am. Nat. Acad. Arts. Sci.*, Vol. 71, 1936, pp. 245—258.
18. R. A. Fisher, "Inverse probability." *Proc. Camb. Philos. Soc.*, Vol. 24, 1930, pp. 528—535.
19. R. A. R. Fisher, "The fiducial argument in statistical inference." *Annals of Eugenics*, Vol. 1935, pp. 391—398.
20. J. Neyman, "Fiducial argument and the theory of confidence intervals." *Biometrika*, Vol. 32, 1941, pp. 128—150.
21. Jerzy Neyman, "Lectures and conferences on mathematical statistics anb probability." Publisher: Graduate School of the U. S. Department of Agriculture, Washington D. C., 1952.
22. E. L. Lehmann, "Testing Statistical Hypotheses." Wiley, New York, 1959.
23. A. Kolmogoroff, "Sur l'estimation statistique des paramètres de la loi de Gauss." (Russian) *Bull. Acad. des Sci. de l'URSS*, Vol. 6, 1942, pp. 3—32.