OPTIMALITY AND SYMPOSIA: SOME HISTORY

ERICH L. LEHMANN

1. INTRODUCTION

As a contribution to a symposium focusing on optimality, this paper will consider the origin of the concept of optimality in statistics and the history of the earliest statistical symposia. These two very disparate subjects are linked by the fact that they originated in the imagination of the same person, Jerzy Neyman. A brief account of his life will therefore be useful as a framework for the two main subjects.

Neyman was born in 1894 to Polish parents, but he grew up in Russia and saw Poland for the first time when he moved there in 1921 at age 27. In Warsaw he became associated with the Polish School of Mathematics and published a paper in its journal, Fundamenta Mathematicae. However, mathematical positions were in short supply and his knowledge of statistics (which he had learned from Sergei Berstein) was more marketable. So he worked as a statistician in agriculture and meterology until he found university positions in Warsaw and Krakow, where he lectured in both mathematics and statistics.

The Polish authorities were interested in Neyman's statistical work but felt unable to evaluate it. To obtain an assessment of its worth, they provided him with a fellowship for the academic year 1925/26 at Karl Pearson's Laboratory in London, which was then the center of the statistical world. His success or failure there, they hoped, would furnish them with a basis for judgment.

His year in London was a great disappointment for Neyman, since the mathematical level at Pearson's Laboratory was much lower than he had expected. To salvage something from his stay in the West, he obtained a fellowship for a second year of leave which he spent in Paris. There the lectures of Borel and Lebesgue and Hadamard's seminar drew him back into pure mathematics.

At this crucial moment Neyman's life was changed by a letter he received in late 1926 from Karl Person's son Egon. Continuing a conversation they had started in
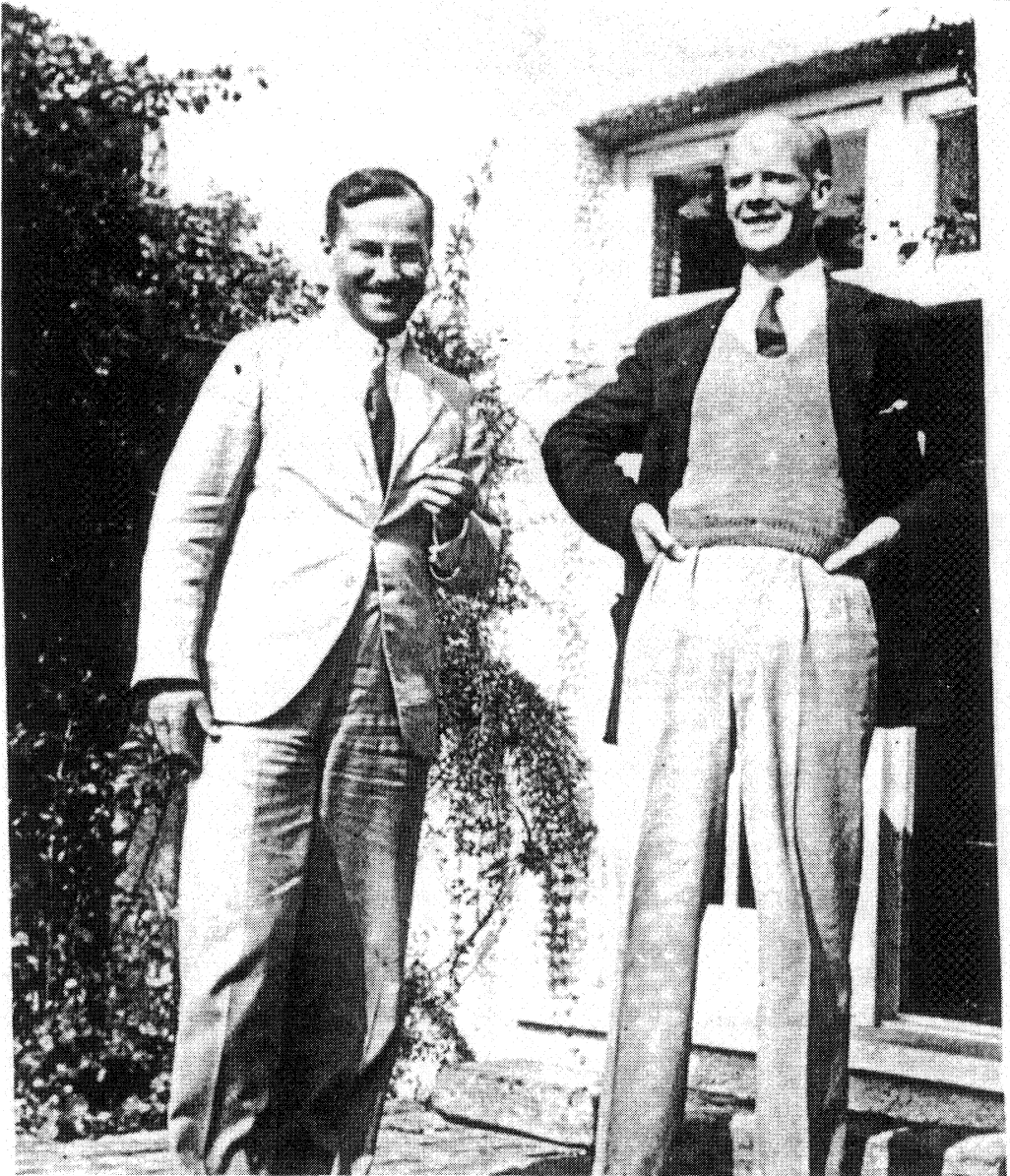
Jerzy Neyman and Egon Pearson

London, Egon Pearson's letter proposed that they work jointly on a problem that had been occupying him for some time.

The problem had its origin in Fisher's work during the preceding decade, which culminated in his enormously successful book of 1925, "Statistical Methods for Research Workers". The book presented Fisher's new tests based on the t, F and χ^2 distributions and his analysis of variance. However, these methods were put forth without any justification, backed only by Fisher's authority.

The question Pearson now asked himself, and then raised with Neyman, was: Why these tests rather than any of the many others that could be proposed? The problem captured Neyman's imagination and resulted in a collaboration carried out mainly through correspondence (with Neyman in Poland and Pearson in England) supplemented by occasional visits.

Neyman's letters to Pearson have been preserved and, even without Pearson's letters which are lost, make it possible to trace the progress of the work step by step. Their study shows two clearly distinguishable phases. During the early years Pearson is the leader and originator. He explains what he is doing to Neyman to whom all this is quite new and who often misunderstands. Neyman's main contribution at this stage is to force Pearson to greater clarity and to help working out the more complicated examples.

This first part of the work resulted in a long 1928 paper on the likelihood ratio method of hypothesis testing. This approach had been proposed by Pearson at the start of their work. In a large number of examples it was now seen to give the answers previously suggested by intuition. In particular, under the assumption of normality it led to Fisher's tests. Pearson was confirmed in his belief that it provided the general answer he had been looking for.

2. THE BIRTH OF OPTIMALITY THEORY

However, Neyman was not satisfied. He agreed that likelihood ratio was an appealing principle but felt that it was lacking a logically convincing justification. The tables were now turned. Neyman took the lead and gradually arrived at a radically new formulation which he first mentions to Pearson in a letter of February 20, 1930. To understand this and the following letter some explanations are necessary.

Neyman and Pearson define a test by a contour $\phi(X_1, \ldots, X_m) = c$. The hypothesis is rejected if the sample point falls inside this contour and accepted when it falls outside. (Today, the reverse of this definition would seem more natural). The contours given by the likelihood ratio criterion the authors call the λ- contours. It should also be mentioned that Neyman's English at the time was still far from perfect. In particular, there are many spelling errors which in the following quotations have been retained.

From the letter of Feb. 20, 1930:

"At present I am working on a variation calculus problem connected [with the] likelihood method. The results already obtained are a vigorous argument in favour of the likelihood method. I considerably forget the variation calculus and until the present time I have only results for samples of two. But in all cases considered I have found the following: We test a simple hypothesis concerning the value of some caracter $a = a_0$, and wish to find contours $\phi(X_1, \ldots, X_n) = c$ such that 1) the probability P(a) of a sample lying inside the contour (which probability is determined by the hypothesis H) is equal

$$P(a_0) = \epsilon$$

where ϵ is a certain fixed value, say .01. (This is for controlling the errors in rejecting a true hypothesis) and 2) that the probability $P(a_1)$ determined by some other hypothesis H' that $a = a_1 \neq a_0$ of sample lying inside the same contour be maximum. Using such contours and rejecting H when Σ (the sample point) is inside $\phi = const.$ we are sure that a true hypothesis is rejected with a frequency less than ϵ , and that if H (the hypothesis tested) is false and the true hypothesis is say H', then most often the observed sample will be inside $\phi = const$ and hence the hypothesis H will be rejected.

I feel you start to be angry as you think I am attacking the likelihood method! Be quiet! In all cases I have considered the $\phi = const.$ contours are the λ contours!" (A facsimile of the first part of this passage is shown on the next page.)

Letter of February 20, 1930

In this letter Neyman is still struggling with the problem of testing a simple hypothesis against a simple alternative. Two weeks later, he essentially has the solution. In this letter he outlines a number of problems on which the authors want to work. One of these problems is:

From the letter of March 8, 1930

"f) To finish what I started to do with the variation calculus. You will understand it in a moment. To reduce to a given level the errors of rejecting a true hypothesis, we may use any test. [Today we might say, any test statistic]. Now we want to find a test which would 1) reduce the probability of rejecting a true hypothesis to the level $\leq \epsilon$ and 2) such that the probability of accepting a false hypothesis should be minimum. – We find that if such a best test exists, then it is the λ-test. I am now shure [sic] that in a few days I shall be ready. This will show that the "λ-principle" is not only a principle, but that there are arguments to prove that it is really "the best test"."

The result which Neyman announces here is of course what is now known as the Neyman-Pearson lemma. The formulation he sets forth in these two letters of 1930 constitute the birth of exact optimality theory.

Neyman realizes that this solution only settles the case of a simple alternative and begins to worry about the more practical situation in which there are many alternatives. In a letter of June 3, 1930 he considers testing $H : \sigma = \sigma_0$ against $\sigma \neq \sigma_0$ in the normal case with zero mean. He states that "In this case also the [best test] is independent of the alternative, 'but not intirely [sic]'"; it depends on whether σ is greater or less than σ_0.

He returns to this problem a year later and in a letter of August 17, 1931 suggests that the test should reject $H : \sigma = \sigma_0$ when $\Sigma X_i{}^2$ is $< c_1$ or $> c_2$. But how to choose c_1 and c_2? He mentions equal tails as a possibility but then interrupts himself and exclaims:

> "I think I have got the point!"

And he goes on to argue that the probability of rejection under an alternative should never be less than under H; what we now call unbiasedness.

Their work on optimum tests of hypotheses was published by Neyman and Pearson in 1933. It was followed in 1934 by Neyman's paper establishing a rigorous theory of survey sampling and his 1937 paper on confidence intervals. In these three papers Neyman established a new point of view and in the process achieved a transformation that is sometimes referred to, not always favorably, as the mathematization of statistics.

3. THE MOVE TO AMERICA

Neyman's work now was becoming known, and he received various invitations to lecture. Of particular importance was a 1937 American lecture tour ending in Washington, D.C. where he gave a series of lectures and conferences which were a huge success.

Shortly after returning from this tour to London, he received a letter from a person of whom he had never heard, Griffith Evans, chairman of the mathematics department of the University of California at Berkeley. Completely unexpectedly, the letter offered him a position in Evans' department. After considerable hesitation and some negotiations Neyman decided to accept the offer and in 1938 moved to Berkeley. He was 44, at the halfway point of his life. The great achievement of the first (European) half had

been his entirely new approach to inference. The principal achievement of the second (American) half was to be the creation of an institutional base for the dissemination of this new approach.

Although Evans held Neyman in great esteem and gave him strong support, the two men had very different visions of Neyman's role which inevitably led to conflicts. Evans considered statistics a branch of mathematics such as algebra or differential equations, and thought of Neyman as a Professor of Mathematics whose specialty was statistics. For Neyman, on the other hand, statistics was a separate discipline, which should be represented by a separate department. He fought unrelentingly for independence and gradually obtained ever increasing autonomy. However, it took nearly 20 years before he reached his goal of a separate department of statistics, a department that was to grow to a faculty of over 20 and by the end of the century had graduated more than 400 Ph.D.'s.

But this development started slowly. Between 1938 and 1941, Neyman established a skeleton course program and a small statistical laboratory with a small temporary staff. After the outbreak of war in December 1941 growth of the academic program came to a halt as Neyman's efforts and those of his fledgling laboratory were concentrated on war work. It did not resume until the end of the European war in the spring of 1945.

4. THE BERKELEY SYMPOSIA

In June 1945, with the end of the war in sight Neyman returned to science and the development of his laboratory. As a first step, he proposed a symposium on mathematical statistics and probability, "to mark the end of the war and stimulate the return to theoretical research." The meeting took place in August 1945, with a sequel in January 1946 for those who had been unable to come at such short notice. Evans reported to the Provost that the symposium

"was an outstanding success and the success was due to Mr. Neyman's foresight in seeing its possibility at this time, and to his initiative and resourcefulness in planning it. It constitutes a significant page in the history of the University of California.

But what is of even more importance, it was convincing evidence of the growing importance of the relation of statistics to experimental work in many branches of science and of the service which the statistical laboratory is rendering to the university and the public."

The Proceedings of the Symposium, a volume of 500 pages consisting of 30 papers, appeared after some delays in 1949.

It seems likely that this first symposium was not part of a long range plan, but was conceived on the spur of the moment. However it was so successful, and Neyman enjoyed so much acting as host on his own turf, that by 1948 he began planning for a second symposium, which took place in 1950. By the third, in 1955, he was clearly viewing the symposia as an ongoing series.

The third symposium was on a much grander scale than the earlier two. The proceedings took up five volumes totaling over 1000 pages. The titles of the volumes:

(1) Theory of Statistics
(2) Probability Theory
(3) Astronomy and Physics
(4) Biology and Problems of Health
(5) Econometrics, Industrial Research, and Psychometry

reflect Neyman's broad range of interests.

At 5-year intervals Neyman organized three more symposia of increasing size and complexity. The proceedings of the sixth came to 6 volumes with a total of close to 3000 pages. By 1975, time for the seventh, Neyman was tired and the enterprise had become a burden to the department. And so the sixth symposium turned out to be the last.

During their lifetime the symposia constituted the most important international statistical meetings. All the leading probabilists and statisticians participated in one or more of them, as did many of the greatest physicists, astronomers, mathematicians, economists, ... The proceedings contained many seminal papers; for example Stein estimation, Stein's idea of adaptive inference, and Robbins' Empirical Bayes approach all were introduced in the Proceedings of the 3rd Symposium. Thus the symposia played an important role in establishing Berkeley as an internationally known center of statistics.

In an extensive review of the second symposium in the British journal Nature, M.G. Kendall predicted that "the Berkeley symposia will take a high place among the statistical literature of the world." The third symposium received an even more interesting review. It appeared in Science under the title: "Statistics – Servant of all Sciences", and its author–surprisingly–was Jerzy Neyman.

What gives this review its special interest is that it contains a succinct account of Neyman's statistical philosophy. He stresses the role of indeterminism–and hence of probability and statistics–in modern science. "As a result", he writes, "the field of application is literally limitless. Briefly, their role in scientific research consists
i) in providing tentative stochastic models of given classes of phenomena,
and
ii) in developing optimal methods of dealing with the observations in order either to supplement the original models with certain details such as the various constants they involve [i.e. estimation] or to decide for or against the further retention of these models [hypothesis testing]."

Point (ii), optimal methods of analysis, was of course his program of the 1930's. To this he now adds that statistical considerations must be preceded by construction of appropriate probabilistic models, and that this is a task for the statistician (in collaboration with a subject matter expert).

Modeling as an explicit programmatic requirement was new at the time and added an important component to Neyman's statistical philosophy. It also became a central aspect of his Berkeley research, particularly his work on some large-scale projects in several fields: competition of species, accident proneness, distribution of galaxies, effectiveness of cloud seeding, and the mechanisms of carcinogenesis. He was working on the proceedings of a conference on the last of these subjects during the last days in the hospital before his death on August 5, 1981.

We may consider Neyman to be the godparent of this symposium, both through his founding of optimality theory and the inspiration provided by his symposia. A detailed account of his life is given in Constance Reid's book "Neyman–From Life".

I should like to thank Mrs. Reid for sharing with me Neyman's letters to Pearson and other documents she collected while working on her book.

REFERENCES

[1] Le Cam, L. and Neyman, J. (Eds) (1982). Probability Models and Cancer. North Holland. Amsterdam.

[2] Neyman, J. (1934). On two different aspects of the representative method: The method of stratified sampling and the method of purposive selection. J. Roy. Statist. Soc. 97, 558-625.

[3] Neyman, J. (1938). Lectures and Conferences on Mathematical Statistics. Graduate School of the U.S. Department of Agriculture. Washington. [2nd expanded Ed. 1952].

[4] Neyman, J. and Pearson. E.S. (1928). On the use and interpretation of certain test criteria. Biometrika 20A, 175-240, 263-294.

[5] Neyman, J. and Pearson. E.S. (1933). On the problem of the most efficient tests of statistical hypotheses. Phil. Trans. Roy. Soc. Ser. A. 231, 289-337.

[6] Pearson, E.S. (1966). The Neyman-Pearson story. In Research Papers in Statistics (F.N.David, Ed). John Wiley. New York.

[7] Reid, C. (1982). Neyman - from Life. Springer, New York.

ERICH L. LEHMANN
DEPARTMENT OF STATISTICS
367 EVANS HALL
UNIVERSITY OF CALIFORNIA
BERKELEY, CA 94720