(b. Little Elm, Texas, 17 June 1906; d. Princeton, New Jersey, 7 March 1964)

mathematical statistics.

Wilks was the eldest of the three children of Chance C. and Bertha May Gammon Wilks. His father trained for a career in banking but after a few years chose to operate a 250-acre farm near Little Elm. His mother had a talent for music and art and instilled her own lively curiosity in her three sons. Wilks obtained his grade-school education in a one-room schoolhouse and attended high school in Denton, where during his final year he skipped study hall regularly in order to take a mathematics course at North Texas State Teachers College, where he received an A.B. in architecture in 1926.

Believing his eyesight inadequate for architecture, Wilks embarked on a career in mathematics. During the school year 1926–1927, he taught mathematics and manual training in a public school in Austin, Texas, and began graduate study of mathematics at the University of Texas. He continued his studies as a part-time instructor in 1927–1928, received an M.A. in mathematics in 1928, and remained as an instructor during the academic year 1928–1929.

Granted a two-year fellowship by the University of Iowa, in the summer of 1929 Wilks began a program of study and research leading to receipt, in June 1931, of a Ph.D. in mathematics. National research fellowships enabled him to continue research and training in mathematical statistics at Columbia University (1931–1932), University College, London (1932), and Cambridge University (1933). Wilks's scientific career was subsequently centered at Princeton, where he rose from instructor in mathematics (1933) to professor of mathematical statistics (1944).

Wilks married Gena Orr, of Denton, in September 1931; they had one son, Stanley Neal Wilks. He was a member of the American Philosophical Society, the International Statistical Institute, and the American Academy of Arts and Sciences, and a fellow of the American Academy of Arts and Sciences, and fellow of the American Association for the Advancement of Science. He also belonged to most major societies in his field.

Wilks’s education was extraordinary for the number of prominent people involved in it. At the University of Texas, his first course in advanced mathematics was set theory, taught by R. L. Moore, noted for his researches in topology, his unusual methods of teaching, and his contempt for applied mathematics. Having a strong practical bent, however, Wilks was more interested in probability and statistics, taught by Edward, L. Dodd; known for his researches on mathematical and statistical properties of various types of means, Dodd encouraged Wilks to pursue further study of these subjects at the University of Iowa (now the State University of Iowa).

At Iowa, Wilks was introduced by Henry L. Rietz to “the theory of small samples” pioneered by “Student” (W. S. Gossett) and fully developed by R. A. Fisher, and to statistical methods employed in experimental psychology and educational testing by E. F. Lindquist.

Wilks chose Columbia University for his first year of postdoctoral study and research because Harold Hotelling, a pioneer in multivariate analysis and the person in the United States most versed in the “Student”-Fisher theory of small samples, had just been appointed professor there in the economics department. At Columbia, Wilks attended the lectures at Teachers College of Charles E. Spearman, considered the father of factor analysis, and became acquainted with the work at Bell Telephone Laboratories of Walter A. Shewhart, originator of statistical quality control of manufacturing processes.

Wilks spent the first part of his second year writing a joint paper with Egon S. Pearson in the department of Karl Person at University College, London. At Cambridge University he worked with John Wishart who had been a research assistant to both Karl Pearson and Fisher, and whose work in multivariate analysis was close to Wilk’s main interest.

Wilks’s first ten published papers were contributions to the branch of statistical theory and methodology known as multivariate analysis, and it was to this area that he made his greatest contributions. His doctoral dissertation, written under Henry L. Rietz, provided the small-sample theory for answering a number of questions arising in use of the technique of “matched” groups in experimental work in educational psychology. It was preceded by a short note, “The Standard Error of the Means of ‘Matched’ Samples” (1931). This note and dissertation are the first in a long series of papers on topics in multivariate analysis suggested to Wilks by problems in experimental psychology and educational testing.
It was, however, his paper, “Certain Generalizations in the Analysis of Variance,” that immediately established Wilks’s stature. In this paper he defined the “generalized variance” of a sample of \( n \) individuals from a multivariate population, constructed multivariate generalizations of the correlation ratio and coefficient of multiple correlation, deduced the moments of the sampling distributions of these and other related functions in random samples from a normal multivariate population from Wishart’s generalized product moment distribution (1928), constructed the likelihood ratio criterion for testing the null hypothesis that \( k \) multivariate samples of sizes \( n_1, n_2, \ldots, n_k \) are random samples from a common multivariate normal population (now called Wilk’s \( \Lambda \) criterion) and derived its sampling distribution under the null hypothesis, and similarly explored various other multivariate likelihood ratio criteria.

Three other papers written in 1931–1932 concerned derivation of the sampling distributions of estimates of the parameters of a bivariate normal distribution from “fragmentary samples” - that is, when some of the individuals in a sample yield observations on both variables, \( x \) and \( y \), and some only on \( x \), or on \( y \), alone; derivation of the distribution of the multiple correlation coefficient in samples from a normal population with a nonzero multiple correlation coefficient directly from Wishart’s generalized product moment distribution (1928) without using the geometrical notions and an invariance property utilized by Fisher in his derivation (1928); and derivation of an exact expression for the standard error of an observed “tetradius difference,” an outgrowth of attending Spearman’s lectures.

“Methods of Statistical Analysis . . . for \( k \) Samples of Two Variables” (1933), written with E. S. Pearson, and “Moment-Generating Operators for Determinants of Product Moments . . .” (1934) are the products of Wilks’s year in England. The first consists of elaboration in greater detail for the bivariate normal case of the techniques developed for the multivariate normal in his “Certain Generalizations . . .” and reflects his and Pearson’s growing interest in industrial applications including a worked example based on data from W. A. Shewhart (1931). The second may be regarded as an extension of the work of J. Wishart and M. S. Bartlett, who had just completed an “independent” derivation of Wishart’s product moment distribution “by purely algebraic methods” when Wilks arrived in Cambridge. His next important contribution to multivariate analysis, “On the Independence of \( k \) Sets of Normally Distributed . . . Variables” (1935), appears to have been written to meet a need encountered in his work with the College Entrance Examination Board, as do many of his later contributions of multivariate analysis.

In addition to his extensive and penetrating studies of likelihood ratio tests for various hypotheses relating to multivariate normal distributions, Wilks made similar investigations (1935) relating to multinomial distributions and to independence in two-, three-, and higher-dimensional contingency tables. He also provided (1938) a compact proof of the basic theorem on the large-sample distribution of the likelihood ratio criterion for testing “composite” statistical hypotheses—that is, when the “null hypothesis” tests specifies the values of, say only \( m \) out of the \( h \) parameters of the probability distribution concerned. Jerzy Neyman’s basic paper on the theory confidence-interval estimation appeared in 1937. The following year Wilks showed that under fairly general conditions confidence intervals for a parameter of a probability distribution based upon its maximum-likelihood estimator are, on the average, the shortest obtainable in large samples.

In response to a need expressed by Shewhart, Wilks in 1941 laid the foundations of the theory of statistical “tolerance limits,” which actually are confidence limits, in the sense of Neyman’s theory—not, however, for the value of some parameter of the distribution sampled, as in Neyman’s development but, rather, for the location of a specified fraction of the distribution sampled. He showed that a suitably selected pair of ordered observations (“order statistics”) in a sample of sufficient size from an arbitrary continuous distribution provides a pair of limits (statistical “tolerance limits”) to which there corresponds a stated chance that at least a specified fraction of the underlying distribution is contained between these limits, thus providing the “distribution-free” solution needed when the assumption of an underlying normal distribution of industrial production is unwarranted. Wilks also derived the corresponding parametric solution of maximum efficiency in the case of sampling from a normal distribution (based on the sample mean and standard deviation) and an expression for the relative efficiency of the distribution-free solution in this case.

In 1942 Wilks developed formulas for the probabilities that at least a fraction \( N_o/N \) of a second random sample of \( N \) observations from an arbitrary continuous distribution (a) would lie above the \( r \) th “order statistic” (the observation in increasing order of size), \( 1 \leq r \leq n \), in a first random sample of size \( n \) from the same distribution, or (b) would be included between the \( r \) th and \( s \) th order statistics, \( 1 \leq r < s \leq n \), of the first sample; and illustrated the application of these results to the setting of one- and two-sided statistical tolerance limits. This work was Wilks’s earliest contribution to “nonparametric” or “distribution-free” methods of statistical inference, an area of research of which he provided an extensive review in depth in “Order Statistics” (1948).

Wilks was a founder of the Institute of Mathematical Statistics (1935) and remained an active member. The Institute took full responsibility for the Annals of Mathematical Statistics, and Wilks became editor, with the June 1938 issue. He served through the December 1949 issue, guiding the development of the Annals from a marginal journal, with a small subscription list, to theforemost publication in its field.

Although Wilks became an instructor in the department of mathematics at Princeton University at the beginning of the academic year 1933–1934, he did not give a formal course in statistics at Princeton until 1936, owing to a prior commitment that the university had made with an instructor in the department of economics and social institutions who had been sent off at university expense to develop a course on “modern statistical theory” two years before; and owing to the need for resolution by the university’s administration of an equitable division of responsibility for the teaching of statistics between that department.
Wilks was deeply interested in the whole spectrum of mathematical education. In “Personnel and Training Problems in Statistics” (1947) he outlined the growing use of statistical methods, the demand for personnel, and problems of training, and made recommendations that served as a guide in the rapid growth of university centers of training in statistics after World War II. Drawing on his experience at Princeton, he urged, in “Teaching Statistical Inference in Elementary Mathematics Courses” (1958), teaching the principle of statistical inference to freshmen and sophomores, and further proposed revamping high school curricula in mathematics and the sciences to provide instruction in probability, statistics, logic, and other modern mathematical subjects. During his last few years he worked with an experimental program in a school at Princeton that introduced mathematics at the elementary level, down to kindergarten.

NOTES

1. An unfortunate consequence of the father’s predilection for alliteration in naming his sons is that publications of Samuel Stanley and Syrrel Singleton Wilks (a physiologist and expert in aerospace medicine) are sometimes lumped together under “S.S. Wilks” in bibliographic works, such as Science Citation Index.


3. The background of this delay and its ultimate resolution are discussed in detail by Churchill Eisenhart. in “Samuel S. Wilks” in bibliographic works, such as Science Citation Index.

BIBLIOGRAPHY


Following Wilks’s death, his “working papers on subjects requiring statistical analysis; letters, reports and papers relating to professional organizations,” were donated by his widow and Princeton University to the *American Philosophical Society*; for further details see *Guide to the Archives and Manuscript Collections of the American Philosophical Society* (Philadelphia, 1966), 146. Another dozen item of correspondence (1946, Papers (MS group 695), Sterling Memorial Library, *Yale University*. Wilks’s professional books and journals have been placed in the S. S. Wilks Room in New Fine Hall, Princeton University.


Churchill Eisenhart

[Contribution of the National Bureau of Standards, not subject to copyright.]