

Wilks, Samuel Stanley | Encyclopedia.com

Complete Dictionary of Scientific Biography COPYRIGHT 2008 Charles Scribner's Sons
23-29 minutes

(*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.¹ 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 Pearson 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.

Wilk'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, \dots, n_k are random samples from a common multivariate normal population (now called Wilk’s Λ 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 “tetrad 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 by 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” tested 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_0/N of a second [random sample](#) of N observations from an arbitrary continuous distribution (a) would lie above the r th “order statistic” (r 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.² He served through the December 1949 issue, guiding the development of the *Annals* from a marginal journal, with a small subscription list, to the foremost 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

(which theretofore had been solely responsible for all teaching of statistics) and the department of mathematics.³ Wilks was promoted to assistant professor in 1936. In the fall term he taught a graduate course, the substance of which he published as his *Lectures . . . on . . . Statistical Inference, 1936–37 . . .*; and in the spring of 1937 he gave an undergraduate course, quite possibly the first carefully formulated college undergraduate course in mathematical statistics based on one term of calculus.

Wilks's service to the [federal government](#) began with his appointment in 1936 as a collaborator in the [Soil Conservation Program](#) of the U.S. Department of Agriculture. He continued to serve the government as a member of the Applied Mathematics Panel, National Defense Research Committee, Office of Scientific Research and Development: chairman of the mathematics panel, Research and Development Board, Defense Department: adviser to the Selective Service System and the Bureau of the Budget; a member of various committees of the [National Science Foundation](#), the [National Academy of Sciences](#), and NASA; and an academic member of the Army Mathematics Advisory Panel. In 1947 he was awarded the Presidential Certificate of Merit for his contributions to antisubmarine warfare and the solution of convoy problems.

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 Syrrrel 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*.
2. For a fuller account of the founding and early years of the *Annals of Mathematical Statistics*, see the letter from Harry C. Carver, dated 14 Apr. 1972, to professor [W.J.] Hall reproduced in *Bulletin of the Institute of Mathematical Statistics*. **2**, no. 1 (Jan. 1973), 11–14; and Allen T. Craig, "Our Silver Anniversary," in *Annals of Mathematical Statistics*. **31**, no. 4 (Dec. 1960), 835–837.
3. The background of this delay and its ultimate resolution are discussed in detail by Churchill Eisenhart, in "Samuel S. Wilks and the Army Experiment Design Conference Series," an address at the twentieth Conference on the Design of Experiments in Army Research, Development and Testing, held at Fort Belvoir, Va., 23–25 Oct. 1974, published in the *Proceedings of*, this conference (U.S. Army Research Office Report 75–2 June 1975), 1–47. This account also contains material unavailable elsewhere on Wilks's family and early career, together with extensive notes on the American institutions and personages that played important roles in it.

BIBLIOGRAPHY

I. Original Works. "The Publications of S. S. Wilks." prepared by T. W. Anderson, in *Annals of Mathematical Statistics*, **36**, no. 1 (Feb. 1965), 24–27, which gives bibliographic details for five books, forty-eight articles, and twelve "other writings," appears to be complete with respect to the first two categories but not to the last. All forty-eight articles are repr. in T. W. Anderson, *S. S. Wilks: Collected Papers—Contributions to Mathematical Statistics* ([New York](#), 1967), as are Anderson's lists of Wilks's publications, in rearranged form (xxvii–xxxiii). Particulars on thirty-one additional "other writings" are given by Churchill Eisenhart, "A Supplementary List of Publications of S.S. Wilks," in *American Statistician*, **29**, no. 1 (Feb. 1975), 25–27.

Among the more important of Wilks's publications are three holograph books: *Lectures by S. S. Wilks on the Theory of Statistical Inference 1936–1937*, [Princeton University](#) ([Ann Arbor](#), Mich., 1937); *Elementars' Statistical Analysis* (Princeton, 1948), quite conceivably the first carefully developed undergraduate course in mathematical statistics based on one term of calculus; and *Mathematical Statistics* ([New York](#), 1962), a far more advanced, comprehensive treatment—*Mathematical Statistics* (Princeton, 1943) was an early version of some of the same material, prepared partly with the help of his students. He also wrote *Introductory Probability and Statistical Inference: An Experimental Course* (New York, 1957; rev ed., Princeton, 1959; Spanish trans., Rosario, Argentina, 1961), with E.C. Douglas F. Mosteller, R.S. Pieters, D.E. Richmond, R. E. K. Rourke, and G.B. Thomas; and *Introductory Engineering Statistics* (New York, 1965; 2nd ed., 1971), with Irwin Guttman (2nd ed. with Guttman and J.S. Hunter).

Of his research papers, the most notable are "The Standard Error of the Means of 'Matched' Samples," in *Journal of Educational Psychology*, **22**, no. 3 (Mar. 1931), 205–208, repr. as paper 1 in *Collected Papers*; "On the Distributions of Statistics in Samples From a Normal Population of Two Variables With Matched Sampling of One Variable," in *Metron*, **9**, nos. 3–4 (Mar. 1932), 87–126, repr. as paper 2 in *Collected Papers*; his doctoral dissertation; "Certain Generalizations in the

Analysis of Variance,” in *Biometrika*, **24**, pts. 3–4 (Nov. 1932), 471–494, repr. as paper 6 in *Collected Papers*; “Methods of Statistical Analysis Appropriate for k Samples of Two Variables,” *ibid.*, **25**, pts. 3–4 (Dec. 1933), 353–378, repr. as paper 7 in *Collected Papers*, written with E. S. Pearson; “Moment-Generating Operators for Determinants of Product Moments in Samples From a Normal System,” in *Annals of Mathematics*, 2nd ser., **35** no. 2 (Apr. 1934), 312–340, repr. as paper 8 in *Collected Papers*; “On the Independence of k Sets of Normally Distributed Statistical Variables,” in *Econometrica*, **3**, no. 3 (July 1935), 309–326, repr. as paper 9 in *Collected Papers*; “The Likelihood Test of Independence in Contingency Tables,” in *Annals of Mathematical Statistics*, **6**, no. 4 (Dec. 1935), 190–196, repr. as paper 11 in *Collected Papers*; “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses,” *ibid.*, **9**, no. 1 (Mar. 1938), 60–62, repr. as paper 14 in *Collected Papers*; and “Weighting Systems for Linear Functions of Correlated Variables When There Is No [Dependent Variable](#),” in *Psychometrika*, **3** no. 1 (Mar. 1938), 23–40, repr. as paper 16 in *Collected Papers*.

See also “Shortest Average Confidence Intervals From Large Samples,” in *Annals of Mathematical Statistics*, **9**, no. 3 (Sept. 1938), 166–175, repr. as paper 17 in *Collected Papers*; “An Optimum Property of Confidence Regions Associated With the Likelihood Function,” *ibid.*, **10**, no. 4 (Dec 1939), 225–235, repr. as paper 20 in *Collected Papers*, written with J. F. Daly; “Determination of Sample Sizes for Setting Tolerance Limits,” *ibid.*, **12**, no. 1 (Mar. 1941), 91–96, repr. as paper 23 in *Collected Papers*; “Statistical Prediction With Special Reference to the Problem of Tolerance Limits,” *ibid.*, **13**, no. 4 (Dec. 1942), 400–409, repr. as paper 26 in *Collected Papers*; “Sample Criteria for Testing Equality of Means, Equality of Variances, and Equality of Covariances in a Normal Multivariate Population,” *ibid.*, **17**, no. 3 (Sept. 1946), 257–281, repr. as paper 28 in *Collected Papers*; “Order Statistics,” in *Bulletin of the American Mathematical Society*, **54**, no. 1 (Jan. 1948), 6–50, repr. as paper 32 in *Collected Papers*; and “Multivariate Statistical Outliers,” in *Sankhya*, **25A**, pt. 4 (Dec. 1963), 407–426, repr. as paper 48 in *Collected Papers*.

Two important papers on teaching and training in statistics are “Personnel and Training Problems in Statistics,” in *American Mathematical Monthly*, **54**, no. 9 (Nov. 1947), 525–528; and “Teaching Statistical Inference in Elementary Mathematics Courses,” *ibid.*, **65**, no. 3 (Mar. 1958), 143–152.

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](#). Wilk’s professional books and journals have been placed in the S. S. Wilks Room in New Fine Hall, Princeton University.

II. Secondary Literature. The biography of Wilks by Frederick Mosteller in *International Encyclopedia of the Social Sciences*, XVI (New York, 1968), 550–553, provides an informative summary of the highlights of Wilks’s life, work, and impact in diverse professional roles. Wilks’s research contributions and other writings are reviewed in the comprehensive obituary by T. W. Anderson in *Annals of Mathematical Statistics*, **36**, no. 1 (Feb. 1965), 1–23 (repr. in *S. S. Wilks: Collected Papers*), which is preceded by a photograph—not in *Collected Papers*, which is preceded by a photograph—not in *Collected Papers—of Wilks at his desk. A less technical but equally full account of Wilks’s life and work is Frederick Mosteller, “Samuel S. Wilks: Statesman of Statistics,” in American Statistician*, **18**, no. 2 (Apr. 1964), 11–17; there is some additional illuminating information in the obituaries by W.G. Cochran, in *Review of the International Statistical Institute*, **32**, nos. 1–2 (June 1964), 189–191; and John W. Tukey, in *Yearbook, American Philosophical Society for 1964* (1965), 147–154. The obituary in *Estadística* (Washington, D.C.), **22**, no. 83 (June 1964), 338–340, tells of his activities in connection with the inter-American Statistical Institute.

The eight articles that constitute “Memorial to Samuel S. Wilks” in *Journal of the American Statistical Association*, **60**, no. 312 (Dec. 1965), 938–966, are rich sources of further information, insight, and perspective: Frederick F. Stephan and John W. Tukey, “Sam Wilks in Princeton,” 939–944; Frederick Mosteller, “His Writings in Applied Statistics,” 944–953; Alex M. Mood, “His Philosophy About His Work,” 953–955; Morris H. Hansen, “His Contributions to Government,” 955–957; Leslie E. Simon, “His Stimulus to Army Statistics,” 957–962; Morris H. Hansen, “His Contributions to the American Statistical Association,” 962–964; W. J. Dixon, “His Editorship of the *Annals of Mathematical Statistics*,” 964–965; and the unsigned “The Wilks Award,” 965–966.

Other publications cited or mentioned in the text are: R. A. Fisher. “On the Mathematical Foundations of Theoretical Statistics,” in *Philosophical Transactions of the Royal Society*, **222A**, no. 602 (19 Apr. 1922), 309–368; and “The General Sampling Distribution of the Multiple [Correlation Coefficient](#),” in *Proceedings of the Royal Society*, **121A**, no. A788 (1 Dec. 1928), 654–673; E. F. Lindquist. “The Significance of a Difference Between ‘Matched’ Groups, in *Journal of Educational Psychology*, **22** (Mar. 1931), 197–204; J. Neyman. “Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability,” in *Philosophical Transactions of the Royal Society*, **236A**, no. 767 (30 Aug. 1937), 333–380, repr. as paper no. 20 in *A Selection of Early Statistical Papers of J. Neyman* (Cambridge–Berkeley–[Los Angeles](#), 1967); J. Neyman and E. S. Pearson, “On the Use and Interpretation of Certain Test Criteria. Part I,” in *Biometrika*, **20A**, pts. 1–2 (July 1928), 175–240, repr. as paper no. 1 in *Joint Statistical Papers of J. Neyman and E. S. Pearson* (Cambridge–Berkeley–[Los Angeles](#), 1967); On the Use and Interpretation of Certain Test Criteria. Part II. “*ibid.*,” pts. 3–4 (Dec. 1928), 263–294, repr. as paper no. 2 in *Joint . . . Papers*; and “On the Problem of k Sample,” in *Bulletin international de l’Académie polonaise des sciences et des lettres*, no. 6A (June 1931), 460–481, repr. as paper no. 4 in *Joint . . . Papers*; Walter A. Shewhart. *Economic Control of Quality of Manufactured Product* (New York, 1931), 42; J. Wishart, “The Generalized Product Moment Distribution in Samples From a Normal Multivariate Population,” in *Biometrika*, **20A**, pts. 1–2 (July 1928), 32–52; and J. Wishart and M. S. Bartlett, “The

Generalized Product Moment Distribution in a Normal System,” in *Proceedings of the Cambridge Philosophical Society. Mathematical and Physical Sciences*.**29** , pt 2 (10 May 1933), 260–270.

Churchill Eisenhart

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