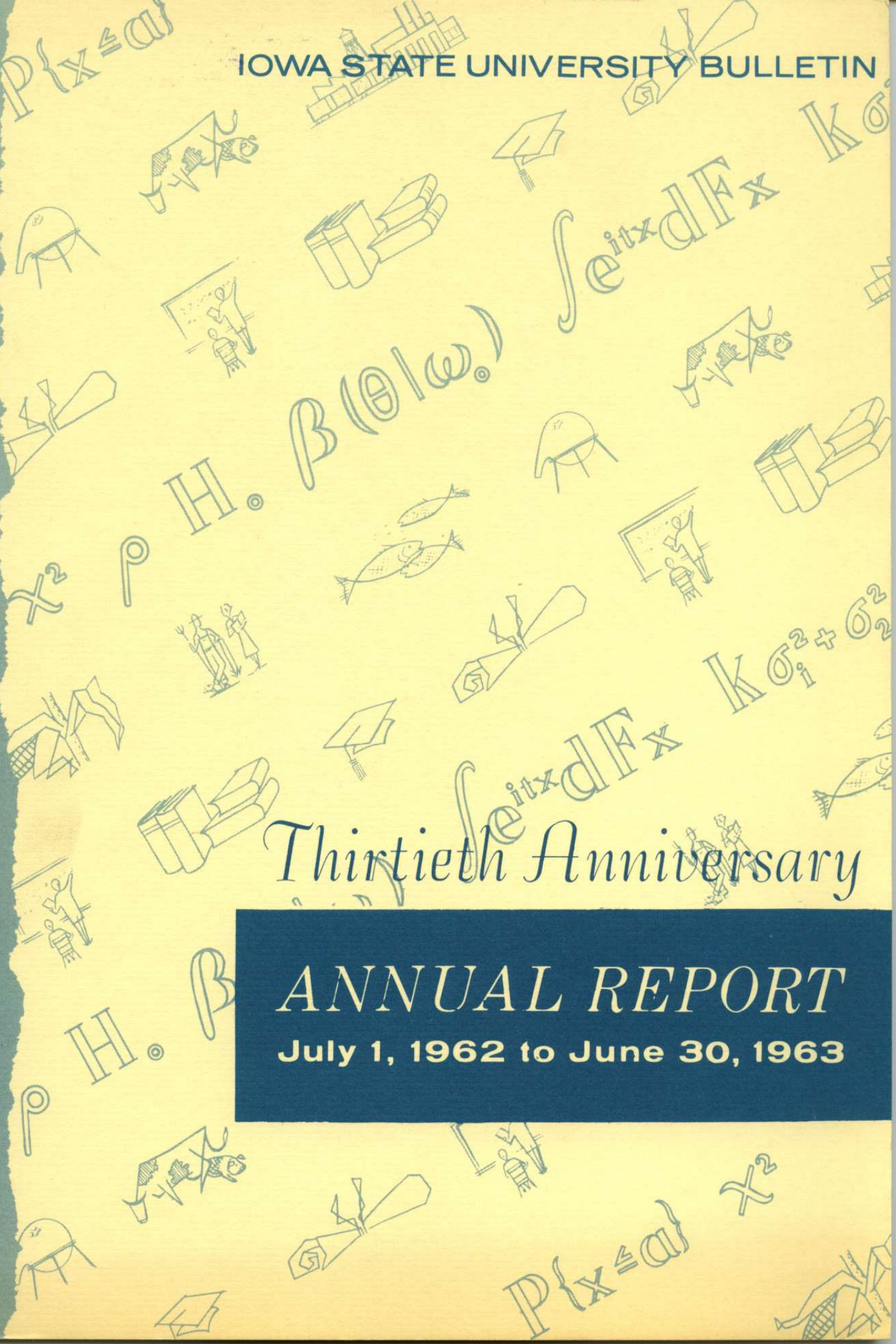
The Statistical Laboratory Established 1933 Iowa State University

IOWA STATE UNIVERSITY BULLETIN



Thirtieth Anniversary

ANNUAL REPORT

July 1, 1962 to June 30, 1963

President James H. Hilton
Iowa State University of
Science and Technology
Ames, Iowa

Dear President Hilton:

This is the Annual Report of the Statistical Laboratory staff at Iowa State for the period July 1, 1962 through June 30, 1963. It reports on the research, consulting, teaching and operational work of members of the university statistical center as complementary activities of the center's five components: the Statistical Laboratory, established as a research and service institute under the President's Office; the teaching Department of Statistics of the College of Sciences and Humanities; the statistics department of the Agricultural and Home Economics Experiment Station; and the statistics participant unit of the Industrial Science Research Institute; together with the research field office of the Statistical Standards Division, Statistical Reporting Service, United States Department of Agriculture, which is located in the Statistical Laboratory. Inasmuch as the statistical endeavors of individuals are carried on through institutions and departments of Iowa State University, this constitutes an annual report from those organizations.

Respectfully submitted on behalf of
the Statistical Laboratory staff,

T. A. Bancroft

Director, Statistical Laboratory and
Head, Department of Statistics; head
statistics department, Agricultural and
Home Economics Experiment Station

THIRTIETH ANNIVERSARY

THE STATISTICAL LABORATORY

Iowa State University

ANNUAL REPORT 1962-1963

IOWA STATE UNIVERSITY BULLETIN
Ames, Iowa

Vol. LXII December 4, 1963 No. 27

Published weekly by Iowa State University of Science and Technology, Ames, Iowa. Entered as second-class matter at the Post Office at Ames, Iowa, under the Act of August 24, 1912.

ORGANIZATION CHART

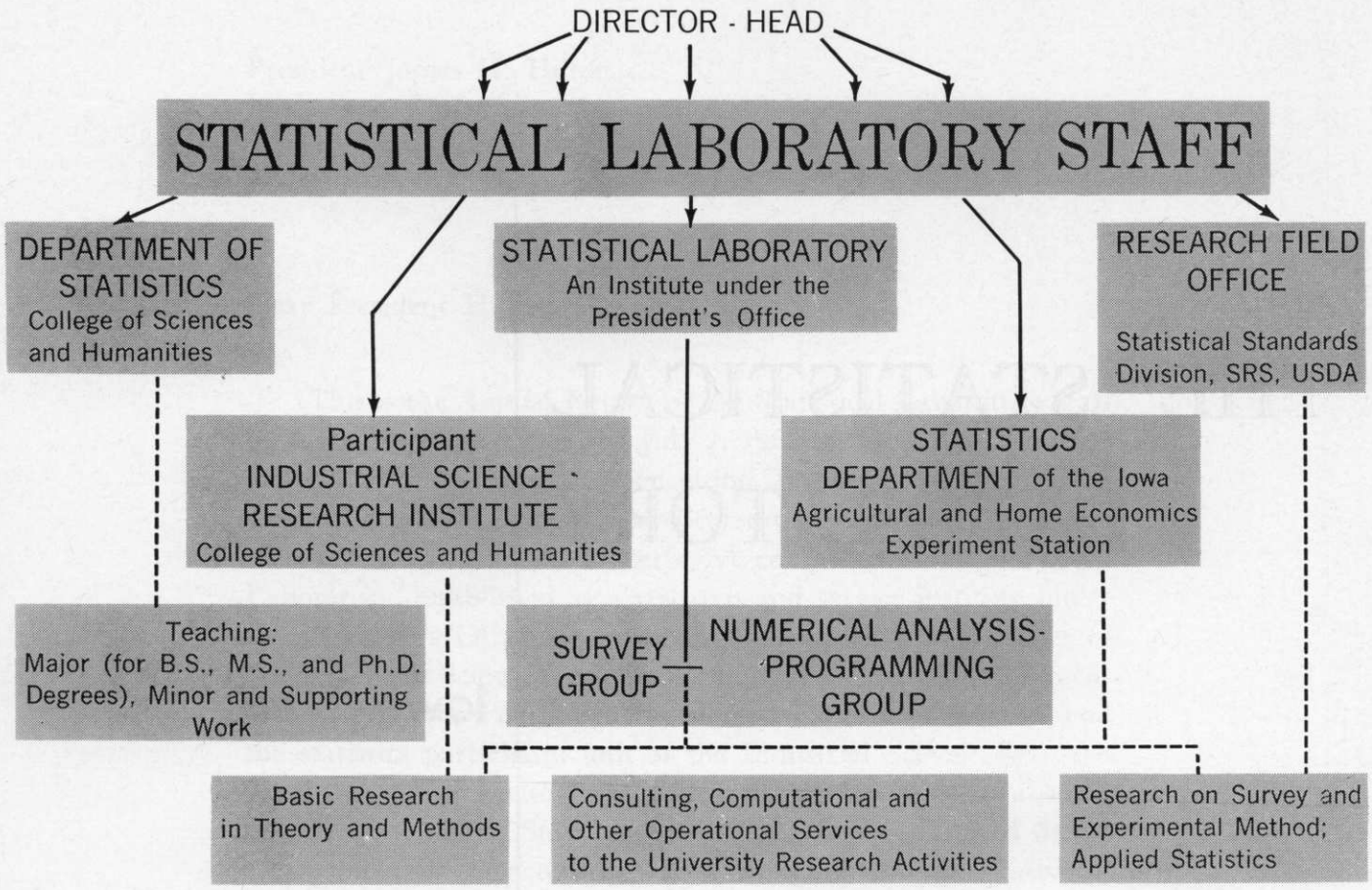
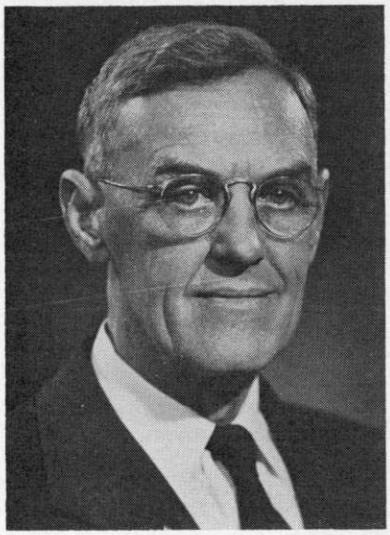


TABLE OF CONTENTS

Thirtieth Anniversary of the Statistical Laboratory 1933-1963

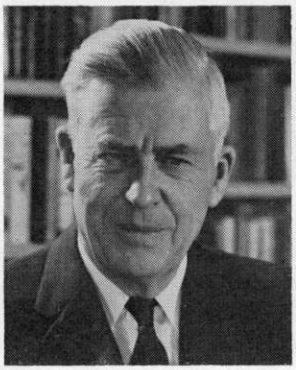


George Waddel Snedecor Director of the Statistical Laboratory, 1933-1947.

This year marks the Thirtieth Anniversary of the founding of the Statistical Laboratory. It is fitting that we honor the founding director, George Waddel Snedecor, and his co-workers in the early days of the Laboratory, and that we summarize its current strength and look to its future.

Snedecor came to Iowa State in 1913 to teach in the Department of Mathematics. He soon developed an interest in statistical applications and taught basic courses in statistics under the Department of Mathematics as early as 1915. In 1924 Henry Wallace, then assistant editor of *Wallace's Farmer* and later Secretary of Agriculture of the United States, met with twenty faculty research workers on Saturday afternoons to study multiple regression. Assisting Wallace were Snedecor and Charles F. Sarle, who was in Des Moines then as USDA Crop Estimator of Iowa. As a result of Wallace's lectures, the demand for professional help in statistics grew so rapidly that in 1927 a Mathematics Statistical Service was instituted with Prof. Snedecor and A. E. Brandt, assistant professor of mathematics, in charge. Calculating and punched card tabulating equipment were installed in the Physics Building where direct current was available. The following year due to increased interest in statistical services on the campus, larger quarters had to be provided in Beardshear Hall. Mary Clem, who at that time was employed by the Department of Economics, learned from Snedecor and Brandt how to operate the machines and started her long career with the computing services of Iowa State. She is still a valued member of the staff of the University Computation Center.

In 1933 the Statistical Laboratory was organized as a research institute directly under the President's Office, with Snedecor as its first director. It was the first statistical center of its kind in the United States. The staff consisted of Snedecor, A. E. Brandt, Mary Clem, and Gertrude Cox, who had received the first master's degree in statistics given by the Department of Mathematics at Iowa State two years earlier. Interest in the work of the Laboratory and demands for its services grew so rapidly that larger quarters were soon needed, and the Laboratory was moved to the Old Office Building where it was housed until 1939.



Henry A. Wallace

A. E. Brandt

The Statistics Section of the Iowa Agricultural Experiment Station was organized in 1935 and from this time on the AES has given financial support to the Laboratory in return for its services in consulting and joint research in all areas of agriculture. Under the terms of the Project Agreement of 1938 between Iowa State College and the United States Department of Agriculture, the staff and facilities of the Laboratory were greatly increased. Seven resident collaborators were stationed in Ames by the USDA, six new Laboratory positions were created, and the computing force was correspondingly enlarged. At this time A. J. King and Raymond J. Jessen, USDA collaborators joined the staff. Bernice Brown, who had been the second student to earn a master's degree in statistics (1932) and had been a valued tutor and teacher of statistics, became a member of the statistics section of the Agricultural Experiment Station and the Bureau of Agricultural Economics. Gerhard Tintner, coming from the University of Vienna and the Cowles Institute, joined the staff of Iowa State in 1937 in the department of Economics and Mathematics and introduced specialized research in economic time series. (In September 1953 he was given a joint appointment

Left to right: William G. Cochran, Gertrude Cox, George W. Snedecor, C. P. Winsor. April 12, 1940.

Statistical Laboratory Group, 1941

in statistics.) The next year Prof. W. G. Cochran came to the staff from the Rothamsted Experimental Station, and C. P. Winsor came from Harvard. The earlier preoccupation of the Laboratory with statistical methods was now superceded by a promising union of theory with practice. In the fall of 1939 the Statistical Laboratory was installed in the new Service Building where it was centrally located on the campus and equipped for its specialized work of consulting and research with the various departments of the College that required its services.

During these early years men who had developed the various areas of statistics visited the Laboratory as visiting professors and short-term members of the staff. In 1931 and again in 1936 R. A. Fisher was a visiting lecturer. Graduate students from foreign countries and from many sections of the United States came to the campus to study under this great leader. Prof. John Wishart came from Cambridge in 1934 to serve as summer lecturer on covariance problems. In 1937-38 Dr. F. Yates of the Rothamsted Experimental Station, England, was in residence as lecturer on the design of replicated experiments. Dr. Jerzy Neyman, then of the Galton Laboratory, University of London, visited the Laboratory in the spring of 1938 and counseled on the general problems of sampling.

Concerning the growth of the Statistical Laboratory during its first decade, Dr. Sarle wrote in 1958:

> The rapid development of the Statistical Laboratory really dates from the 1936 summer visit of Prof. R. A. Fisher. Directly and indirectly much of the credit for this rapid expansion during the succeeding decade goes to Henry A. Wallace, Arnold J. King as USDA research project leader, and the vision and cooperation of Prof. Snedecor and the ISC administration . . . The Bureau of the Census also deserves a great deal of credit as it financed a large part of the work involved in the development of the nation-wide Master Sample of Agriculture, which was used in connection with the 1945 Agriculture Census. This gave area sampling a sound basis for further development and use . . . The remarkable development of statistics at Iowa State, both in research and in teaching, came about in response to the basic statistical needs of US agricultural research in both experimental and sampling design.

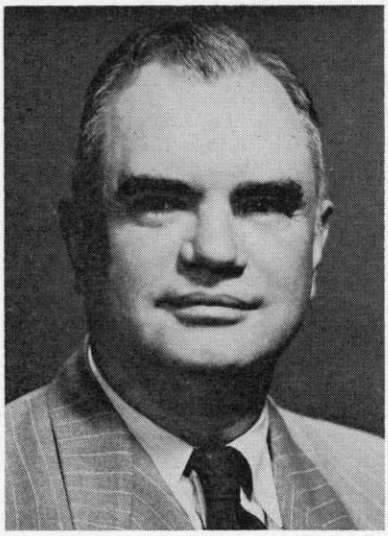
The activities of the Statistical Laboratory were affected by the war as students and staff became involved in the war effort. However the increased need for survey data by the government, which was in part due to the war, led to expansion of the service work of the Laboratory. Research was extended to include problems related to the federal government through the cooperative projects with the Bureau of Agricultural Economics, the Census Bureau, and the Weather Bureau. The Master Sample Project, begun in 1943 with the Bureau of the Census and the Bureau of Agricultural Economics, became an important part of the work of the Laboratory. Its function was two-fold: (1) to prepare the essential basic data and materials for samplings of the United States, not only of its agriculture but also of its population; and (2) to prepare, using the basic material in (1), specific samples for field use. In 1945 Jessen and King participated in a State Department mission in Greece making sampling investigations to determine the conditions affecting the Greek elections. In 1948 the Rockefeller Foundation supported an economic and social survey of Crete, and Jessen served as chief statistical advisor, with N. V. Strand, who had been connected with Iowa State intermittently since 1935, as participant.

Left to right: O. Kempthorne, Mary Clem, Vincent Lundquist, Raymond J. Jessen, Carl Marshall, Arnold J. King, Joseph Dodson. About 1947.

In 1947 the Department of Statistics was organized as an independent department in the Division of Science, and courses were offered not only for students majoring in statistics but also as service courses for students majoring in other fields in which statistical methods are important research tools. The granting of the degrees of bachelor of science, master of science, and doctor of philosophy in statistics was authorized.

It was also in 1947 that George W. Snedecor retired after 14 years as director of the Statistical Laboratory and 34 years after he had joined the Iowa State faculty as an instructor in mathematics. It is particularly appropriate that the establishment of a department for training young statisticians should culminate the years of work of this early statistician. Though Snedecor retired from his administrative duties, he continued to lecture and to teach at Iowa State and other statistical centers in the United States and in South America. To quote from the 1947-48 Annual Report, "Snedecor is recognized throughout the world for his pioneer work in the development of statistical methods; his publications and his students have given the College the reputation of being one of the few leading institutions in this country in statistics." In 1958, in conjunction with the 100th anniversary of the institution whose name was now to be Iowa State University, Dr. Frank Spedding presented Snedecor to President James H. Hilton for conferring of the honorary degree of Doctor of Science. Spedding said, "Dr. Snedecor was primarily responsible for the development of the Iowa State Statistical Laboratory. His text on statistical methods is widely used and is now in its 5th edition with translations in at least three languages. As teacher, lecturer and author, Dr. Snedecor has served Iowa State well for nearly half a century."

From 1947 to 1949 a committee administered the activities of the Statistical Laboratory: Jessen was Acting Director of the Laboratory and Acting Head of the Department of Statistics and in charge of sampling; King administered the off-campus work with the USDA and other agencies; and A. M. Mood, who had returned to Iowa State following the war, had charge of teaching and theory.



T. A. Bancroft, Director of the Statistical Laboratory since 1950.

In 1950 T. A. Bancroft became Director of the Laboratory and Head of the Department of Statistics. Under his leadership the growth of the Laboratory has continued, both in the volume of work done and in the range of assistance given to many departments of the University and to off-campus agencies. Sampling, the design of surveys, and the performance of operational work for research workers in agriculture continued. Among the early off-campus-supported projects were the Atomic Energy Commission project, of which Ray Mickey was chief investigator and Chalmer J. Roy was associate leader; the Camp Dietrick project in biological and chemical warfare for the Army Chemical Corps, with Stanley Issacson as chief investigator and Bernard Vinograde and Carl Lagenhop of the Department of Mathematics as valuable consultants; a study for the federal Bureau of Human Nutrition and Home Economics, on which Paul Homeyer consulted; and the Air Force project, started in 1953, which was the Laboratory's first Wright Air Systems Development contract. The work on the Air Force project was in two areas: on one phase T. A. Bancroft was chief investigator and H. O. Hartley, D. V. Huntsberger, and Helen Bozivich also worked on the project; on the other phase O. Kempthorne was chief investigator and Martin B. Wilk was associate. Research for the Air Force has continued to the present time.

In recent years the number of such off-campus contracts and grants has greatly increased. In the beginning the work of the Laboratory was largely in applied statistics in the field of agriculture. Later new theory and methodology were developed in many areas of mathematical statistics, sampling, experimental design, genetic statistics, economic statistics, psychological statistics, industrial and engineering statistics, scientific computing and biological statistics.

Many studies of a fundamental nature are now supported by grants and contracts with off-campus agencies and administered by the Industrial Science Research Institute: National Science Foundation, Office of Ordnance Research, Office of Naval Research, Army Research Office, Wright Air Systems Division, National Aeronautical and Space Administration, National Institutes of Health, and others. These grants have provided stipends for graduate students and associates and have been an invaluable aid both to the students and to the research which they support. Research has also continued for the Bureau of the Census, USDC, and for USDA. An area which has had great growth

Statistical Laboratory Group, 1958

in recent years is the National Institutes of Health Training Grant Program which began in 1958 with a five-year grant. Training and research under a new grant for a second five years are now getting under way. This program is described in detail on pages 20-23 of this report. A Statistical Laboratory Annual Report describing the work of the staff has been published each year since 1944.

Routine computing equipment had been available in the Statistical Laboratory from its beginning and soon some small electrical computers were purchased for research computation and teaching. The Laboratory gave assistance with numerical and statistical analysis, computer programming, data processing and scientific computing. In 1955 an Interdepartmental "Working Committee for Improvement of Computational Facilities at Iowa State" was appointed with Bancroft as chairman. The committee recommended (1) that an IBM 650 be installed on a rental basis at a 60% educational discount and (2) that a digital computer, the Iowa State Cyclone, be constructed. Construction of the Cyclone began in 1956, was completed in 1959, and was greatly enlarged in 1961, when an addition to the Electrical Engineering Building was contructed to house it. In 1957 the IBM 650 was installed in the northwest corner of the Service Building. It was moved to the first floor of the new addition to the Service Building in 1961. Later an IBM 7074-1401 was installed, and in April 1962 all electronic computing equipment was transferred from the Statistical Laboratory to the new University Computation Center.

The Numerical-Analysis Programming Group was organized at this time. It provides consultation in statistically-oriented computer techniques, computer programming, and processing of data requiring the services of an analyst-programmer. It functions in a liaison capacity between the researcher and the computer. The objective of the group is to assist staff and graduate students with their computational problems in substantive areas which make use of statistics as a primary research tool.

The Laboratory has worked with the College of Engineering and recently three staff members have held joint appointments in Statistics and Industrial Engineering. With the Engineering Extension Service and the Department of Industrial Engineering, the Laboratory has cooperated for several years in arranging an annual conference for an off-campus group on Statistical Quality Control. The Laboratory anticipates increased cooperation with the College of Engineering through the Engineering Experiment Station and through the Engineering Extension Service and its con-

nection with the Center for Industrial Research and Service (CIRAS), which is a functional part of the University. CIRAS was organized because of the increasing interest in promoting existing industries and establishing new ones in Iowa.

The Statistical Laboratory is primarily a research and consulting institute of the University and its research and consulting in cooperation with the Iowa Agricultural Experiment Station, the Engineering Experiment Station and the Industrial Science Research Institute are still the primary interest of the staff. With the expanded facilities offered by the new addition to the Service Building and the increased interest that has developed in the use of statistics in all phases of scientific and industrial research, continued growth of the Iowa State Statistical Laboratory seems evident.

Doctorates in Statistics at Iowa State

As we honor those who have built the reputation of Iowa State for its work in statistics, it seems appropriate that we list the names of those who have earned their doctorates here. They are working in their specialized fields of interest in almost every state of the Union and in many foreign countries. To quote Dr. Bancroft, "By continuing to provide able staff to assist in establishing new statistical centers and to strengthen existing ones, the Statistical Laboratory paradoxically renews its own strength and finds its influence ever expanding."

Iowa State University has awarded the honorary degrees of Doctor of Science to three statisticians for their outstanding contributions in the field of statistics:

1936—R. A. Fisher
1958—George W. Snedecor
Gertrude Cox

The degree of Doctor of Philosophy has been awarded to the following:

1940—Holly C. Fryer
1941—Richard L. Anderson
1943—T. A. Bancroft
Jerome Ching Li
1947—Samuel Lee Crump*
David B. Duncan
1948—Osmer S. Carpenter
Walter T. Federer
Clifford J. Maloney
1949—Bernard Ostle
Robert George D. Steel
1950—Garnet E. McCreary*
1951—Arthur M. Dutton
1952—Franklin A. Graybill
M. Ray Mickey
Donovan Thompson
1953—Daniel G. Horvitz
Virgil L. Anderson
1954—David V. Huntsberger
Edgar M. Jacobs
1955—Helen Bozivich
Edward C. Bryant
John F. Hofmann
Martin B. Wilk
Donald L. Brakensiek
(Jt. w/Agr. Engr.)
1956—G. Ronald Herd
Carl E. Marshall
1958—John LeRoy Folks
Basilio A. Rojas
William H. Williams
George Zyskind
Marion R. Bryson
Howard L. Taylor
(joint w/Econ.)
1960—S. K. Katti
Burton T. Onate
Alan Ross
Martin R. Dorff
Harold J. Larson
Sidney Addelman
Neeti R. Bohidar
(Jt. w/Animal Breed.)
1961—Roger S. McCullough
Jose Nieto de Pascual
J. N. K. Rao
T. N. Throckmorton
1962—Donald Ray Jensen
(Jt. w/Agronomy)
Edwin J. Hughes
Carol Bates Edwards
Ronald R. Hocking
Ilbok Lee
John E. Graham
W. J. Hemmerle
Mauritz van Aarde
Robert F. White
William T. Lewish
Dale Grosvenor
Paul Leaverton
Dale O. Richards

In some cases a major was earned in another department, with a strong minor in statistics. For example, Judd McGuire, William Kwolek, John Gill, and Donald J. Sisson had minors in statistics and are now working almost entirely in the field. Wayne A. Fuller, Dewey L. Harris, and Donald K. Hotchkiss, all of whom earned their doctorates in the College of Agriculture with minors in statistics, are now full time members of the Statistics staff. They devote a large share of their time to consulting and carrying on joint research with staff and graduate students in agricultural economics, animal breeding and genetics, and dairy science.

The George W. Snedecor Award was established in 1954 by the Statistical Laboratory to honor Prof. Snedecor in a permanent way. It is given annually to the most outstanding candidate for the Ph.D. degree in statistics at Iowa State and consists of a year's membership in the Institute of Mathematical Statistics, a subscription to the **Annals of Mathematical Statistics**, and a small cash award.

Winners of the George W. Snedecor Award have been the following:

Helen Bozivich, 1954
Martin B. Wilk, 1955
John F. Pauls, 1956
William H. Williams, 1957
Robert F. White, 1958
Martin Dorff, 1959
Scott Krane, 1960
J. N. K. Rao, 1961
Edwin J. Hughes, 1962
Ronald R. Hocking, 1963

The opportunity is now given for contributions to be made to increase the size of this award which was made possible by former students and friends of Prof. Snedecor. Contributions should be mailed to

The Statistical Laboratory
Iowa State University
Ames, Iowa

---

* Garnet E. McCreary b. 1923 - d. 1955

Samuel Lee Crump b. 1919 - d. 1963
A memorial sketch by Arthur M. Dutton appears in The American Statistician, 17: 3. June 1963.

Personnel

The Statistical Laboratory is one of several independent research institutes on the campus of Iowa State University. The five components of the center have a single director and share personnel. In fiscal terms this means that a staff member's salary may be provided from several sources—e.g., Department of Statistics, the Statistical Laboratory budget set up by the President's Office, a statistical project of the Iowa Agricultural and Home Economics Experiment Station, or a research contract arranged by the Statistical Laboratory and the Industrial Science Research Institute. The Laboratory also has a number of research grants and contracts with federal agencies which provide funds for research for both staff and graduate students at all levels.

The members of the statistical center, including its affiliated fellows and graduate students, are listed below.

THE STATISTICAL LABORATORY STAFF FOR THE YEAR 1962-63

Under the administrative direction of

James H. Hilton, D.Sc. .....President

Chalmer J. Roy, Ph.D. .....Dean, College of Sciences and Humanities; Director, Industrial Science Research Institute (ISRI)

Floyd Andre, Ph.D. .....Dean, College of Agriculture; Director, Iowa Agricultural and Home Economics Experiment Station

Theodore A. Bancroft, Ph.D. .....Director, Statistical Laboratory, Head, Department of Statistics; Head, Statistics Department, Iowa Agricultural and Home Economics Experiment Station

Professors

T. A. Bancroft
Herman O. Hartley
David V. Huntsberger
V. S. Huzurbazar (Visiting)
Oscar Kempthorne
Norman V. Strand
Gerhard Tintner (Joint appointment with the Departments of Economics and Sociology and Mathematics until Dec. 1, 1962)

Professor Emeritus

George W. Snedecor (in absentia)

Associate Professors

Robert J. Buehler
K. C. Chanda (Visiting)
C. Philip Cox
Herbert T. David
Wayne A. Fuller
Campbell C. Mosier
Leroy Wolins (Joint appointment with the Department of Psychology)
George Zyskind

Assistant Professors

Foster B. Cady
Dewey L. Harris
William J. Hemmerle
Donald K. Hotchkiss
Howard W. Jespersen
J. N. K. Rao
Dale O. Richards (Joint appointment with Department of Industrial Engineering)

Post Doctoral Associates

A. W. Qureshi
B. V. Shah (until August 1962)

Instructors and Associates

Harold Baker
Edward J. Carney
Robert Cochran
Eugene Dayhoff
John Gill
Francis Giesbrecht
John E. Graham
Dale Grosvenor
Klaus Hinkelmann
Ronald R. Rocking
Thomas C. Jetton
Ilbok Lee
Harold Rosenberg
Florence Tetreault
Mauritz van Aarde

Graduate Assistants

Rodney Basson
Thomas D. Burnett
Russell J. Burns
Charles Cress
Thomas E. Doerfler
Ahmed El Mawaziny
Alan Feddersen
Edmund Fuller
James R. Heard
William G. Hill
Louis W. Johnson
William D. Lawing
John Lesem
Richard W. Mensing
Philip M. Mills
Martin S. Rosenzweig
Chang-sheng Shih
Gary Spencer
Richard Swanson
David R. Thomas
Gordon D. Wakefield
Henry Walter

National Institutes of Health Trainees

Ross Adams
Michael Billings
Patricia Conn
Paul Leaverton
Thomas Roseberry
Darrol Heggen (in absentia)

General Electric Fellows

Fred Ramsey
Lee H. Smith

Iowa State U. Research Foundation Fellow

Donna Ruhl

Affiliated Fellows

Carlos Cavallini	AID Argentina
Ahmed El Mawaziny	Gov't UAR
Lawrence Gould	NIH Western Reserve U.
Jose Gutierrez	FAO Philippines
Mohammed Ishaque	FAO Pakistan
Rudolfo Mengido	AID Argentina
Douglas Murfield	USDA
Esmat Nouri	Gov't UAR
Nangnoi Suwanphant	Asian Foundation Thailand
Phrensi Svasti-Salee	FAO Thailand
George Weiss	Gov't Canada
Herman Wiedenhofer	Gov't Venezuela
Stephan Williams	USDA

National Science Foundation Undergraduate Science Education Participants

Sharon Fie (until March 1, 1963)
Edmund Fuller (until March 1, 1963)
Alan Johnson
Nancy Preston
Susan Reynolds
James R. Veale

Student Assistants (Undergraduate)

Larry Conger
Alan L. Jaycox
Leroy Snyder
James R. Veale

General Office Staff

Administrative Assistant, Margaret G. Kirwin
Mabel M. Peterson, Accountant
Eleanor F. Bolton, Technical Writer-Editor
Secretarial Staff:
Audrey Graham, Experimental Design, until May 30
Iveta Zeliadt, after May 30
Barbara Mason Downey
Anne Leicht, Survey Group
Glenda Sampson, General Office
Shirley Saveraid, Numerical Analysis-Programming Group

Numerical Analysis—Programming Group

Mary Ann Carney, Technician
June F. Smith, Technician

Survey Group

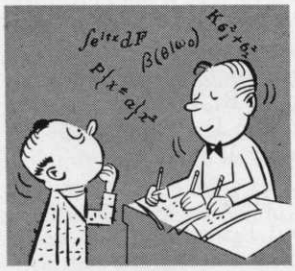
Survey Supervisor, Helen Ayres
Clerical Supervisor, Marjorie Mason

Clerks:

Lucile White	Hazel Cook (Interviewer)
Anna B. Woodrow	Jeanne Wagaman
Mabel Matthews	Ava Klopf
Marie Osterman	Alice Eldridge
Evelyn Howe	

In addition numerous field interviewers are temporarily employed by the Survey Group during the time various sampling projects are carried out.

Consulting and Joint Research



The members of the staff of the Statistical Laboratory through their consulting activities come in contact with research work being carried on in agriculture, the sciences, engineering, home economics, and veterinary medicine at Iowa State University. The Statistical Laboratory budget, provided by the President's Office, supports consultation in mathematical statistics and computational methodology, as well as general statistical consultation in areas where no other specific funds have been provided for advisory work. Recent problems are exemplified in the material which follows.

In general statistical consulting

T. A. Bancroft consulted with the staff of the Department of Entomology, including T. A. Brindley and George R. Pesho, on the design and analysis of experiments with the European corn borer. Problems relating to adequate model specifications and scaling were considered. Some of the work on the model specifications led to a joint paper by Bancroft, Chanda and Brindley, which was presented at the meetings of the Institute of Mathematical Statistics and the Biometric Society at Harvard University in May 1963. (See page 45.)

Bancroft and N. V. Strand held consultations with the University of Iowa Medical School regarding a nutrition survey.

Bancroft and Wayne Fuller of the Department of Statistics and W. G. Murray of the Department of Economics consulted with members of the State Tax Commission in Des Moines regarding the possibility of a study of equalization of property taxes in Iowa.

In April Bancroft consulted in Ames with Dr. Basilio Rojas, Director of the National College of Agriculture at Chapingo, Mexico, on the establishing of a statistical computing center at Chapingo. Conferences were arranged for Dr. Rojas with Dr. George M. Browning of the Agricultural Experiment Station and with Dean Roy and Dean Page. Bancroft's assignment by the Ford and Rockefeller foundations in Mexico in February is reported on page 16.

In April Bancroft acted on a project of the Site Committee of the National Institutes of Health at the University of Pittsburgh. In May he met with the NIH committee in Cincinnati, Ohio, to review research proposals.

Bancroft consulted in Ames on April 9 with Morris Hansen of the USDA Bureau of the Census on the matter of the establishment of a long term joint program between the Census Bureau and the Statistical Laboratory.

Bancroft and members of the sampling group consulted with Earl Houseman, Head of the Standards and Research Division of the Statistical Reporting Service, USDA, on May 1, regarding continuing joint work with the USDA.

In June Bancroft, Strand, Fuller and Harold Baker consulted with a group of twelve foreign technicians who were interested in the area of sampling and census work.

Bancroft and a group of staff members in the social sciences at Iowa State consulted with the Rector of the U. of Costa Rica and Mario Romero, Iowa State graduate in statistics, regarding the establishment of a joint program, including statistics, between the University of Costa Rica and Iowa State.

H. O. Hartley consulted with agricultural economists on the application of nonlinear and integral programming to the optimization of projects involving mutually exclusive subprojects and/or a diminishing vector to scale economy. He also consulted with industrial engineers on the application of the system of Pearson Type Frequency Functions to problems in industrial retirement. He worked with geneticists on the deviation of exact distributions for measures of heritability in the analyses of progeny trials. Numerous visitors on the campus consulted Hartley on the planning of Monte Carlo computations on high-speed computers.

Hartley consulted with a mechanical engineer on the solution of a partial differential equation of an elliptical type arising in turbo design. This equation presented a difficulty in that the boundary conditions were mixed. The function was prescribed over a subset of the bound-

ary, was periodic over another subset, and had a prescribed normal differential over the remainder of the boundary. The relaxation (ordinary or overrelaxed) was found to diverge at first but was forced to convergence by the device of keeping certain critical boundary values at fixed levels, solving the problem for each combination of levels and then meeting the boundary condition (normal deviation) by interpolation.

D. V. Huntsberger reviewed with a USDA biologist the statistical aspects of a paper, "Sharp-Tailed Grouse Habitat and Populations." He advised on the analysis of the data and made suggestions for corrections and additions. Huntsberger advised an agronomist from the Near East on the analysis of data for an experiment concerning varieties, levels of nitrogen and levels of irrigation. He advised a biologist on the selection and evaluation of various measures involved in the exploratory study of ovaries and corpus luteua of deer. He continued his consultation with a biologist on the analysis and interpretation of bacteria counts and morphology of milk in the various stages in the production of nonfat dried milk. For an agricultural engineer he gave advice concerning some extreme value theory as applied to meterological data and its application in design loads of structures.

C. P. Cox consulted with a research worker from Simpson College on the analysis of genetic experiments investigating possible differences in mean litter sizes of various types of congenitally blind mice. He consulted with a graduate student in Fish and Wildlife Management on the analysis of quantal bioassay type experiments to determine the toxicity (LD50) of insecticides on caddis and mayfly. It was also found that the data arising in this study were particularly suitable for examination by assay methods considered by Paul Leaverton in his Ph.D. thesis. With an associate in the Department of Animal Science Cox discussed the planning and analysis of experiments testing progesterone compounds for their ability to control the onset of estrus in cattle without impairment of fertility.

Dewey L. Harris and Wayne A. Fuller helped a veterinary hygiene staff member design a survey for the assessment of degree of incidence of a respiratory disease of beef cattle in certain counties in Iowa.

Francis Giesbrecht worked with a graduate student in Child Development on a project involving the analysis and interpretation of a semantic differential test administered to crippled children.

H. T. David assisted in preparing the final report on the findings of a Home Economics Experiment Station study, "The Work Pants Wear Study," the purpose of which was to determine which of five types of garments would give the best wear under work conditions. David set up the plan for analysis of the final data in the form of a regression analysis of an unbalanced incomplete block design. The results of this analysis led to point of interval estimates of the durabilities of the five types of garments.

David consulted with a member of the staff in Industrial Administration on the establishing or prediction equations for sales and other business indicators. A prediction method was devised depending both on the marginal statistical significance of the predictors individually and also on their significance in conjunction with the entire set of possible predictors.

With a research economist David worked on the application of Markov chain theory to the study of the changes, in successive time periods, of the representation of various firm sizes in an industry. He consulted with a graduate student in economics on the insertion of a prior estimate for near-redundant regression coefficients. With a graduate student in biochemistry, David consulted on a problem in discrimination between two hypotheses concerning the shape of a large molecule involved in reproduction by using relative counts of certain characteristics of off-spring cells.

In engineering

Herbert T. David consulted with engineering staff and students on a number of projects.

With an agricultural engineer, David consulted on the design and layout of an experiment to evaluate the effect on disposal efficiency of various attributes of sewage disposal lagoons. With another agricultural engineering staff member and graduate student, he discussed the format of data collection and the analysis of an experiment investigating the effect of several methods of plowing and related soil-distribution methods on certain physical characteristics of soil. He consulted with another agricultural engineering staff member and a graduate student on a project to study the effect on cultivation-loss of terrain curvature and slope.

David also consulted with a civil engineer on two problems: One was a study intended to evaluate, by means of variable lag regression, the effective time lag between ambient temperature and subsurface pavement temperature and related effects of ambient temperature. Another was a problem developing a relation between cement and UCS index on soil-cement strengths. David consulted with a civil engineer on the problem of analyzing the records of an asphalt-producing operation in an attempt to judge the efficiency of the operator in charge of adjusting the process to changing environmental conditions. David and Huntsberger consulted with a civil engineer about subject matter content of an experimental course in statistics for civil engineering students.

David consulted with a number of chemical engineering research workers. With a graduate student, he worked on the design of an experiment to test the hypothesis that effective cake resistance, in compression-permiability testing, depends on the ratio of cake-volume to the cake surface area in contact with cell walls. He worked with another graduate student on setting up a Monte Carlo investigation to study momentum-loss in filter flow. For another phase of this research, an experiment was designed to measure the effect of filter cake weight and concentration in a filter cell. This experiment was set up in two blocks, adding to a full 4x4 factorial in such a way that only the QxL and QxC effects were confounded with blocks. With another graduate student he worked on two problems concerning the development of a resin-type glue: (1) the design of an experiment to test the effect of the aging of the two main ingredients of the glue on the strength of the glue; and (2) the design of an experiment to assess the effects of four different factors on the strength of the same glue. The design used here was a completely orthogonal 5 x 5 Latin square. Four of the available six alphabets were used for each of the four factors and the remaining two alphabets were used to balance out a suspected linear time trend in the master batch.

David worked with a graduate student in electrical engineering on the application of regression techniques to the handling of coded information in computer design. He consulted with a staff member in nuclear engineering in the setting-up of a Monte Carlo study of the energy distribution in a pipe resulting from a particle source at its centroid. A similar experiment was concerned with ascertaining whether a complicated theoretically derived formula for the gradient of mass diffusion in a cylinder could be replaced by a simple linear function for an experiment measuring the phenomenon in a particular experimental situation. The adequacy of the simpler relation was tested in the form of the null hypothesis that certain regression coefficients were zero.

Dale Richards consulted with various members of the staff of the College of Engineering on problems of a statistical nature. One such problem involved trying to fit a linear regression when both variables are subject of error. Possible procedures were outlined and recommendations were made.

Engineering Extension and the Department of Industrial Engineering, with the cooperation of the Statistical Laboratory, held a third annual conference for an off-campus group on "Statistical Quality Control at Work" on May 9-10, 1963. Dale O. Richards and E. J. Carney, who hold joint appointments in Industrial Engineering and Statistics, were co-chairman of the conference. Harry D. Greiner, who is Manager of Quality Control of Radio Corporation of America, Moorestown, N. J., was the featured speaker. He spoke on "Dynamic Quality Control—More Management than Statistics." T. A. Bancroft welcomed the group to the campus. E. J. Carney spoke on "The Language of Sampling"; Dennis Lawing on "What Makes a Good Quality Control Procedure—Statistical Problems"; and D. O. Richards on "Continuous Sampling and Related Topics". Herbert T. David addressed the group on "Typical Quality Control Problems—Theoretical Approaches."

In the Behavioral Sciences

Wolins is involved in much of the research on the campus in the behavorial sciences. He helped with various aspects of psychological testing, with problems of the formulation and design of investigations, and with the analysis of results.

In consulting with the people who do research in the behavioral sciences, one problem seems to recur in a large number of experiments: how to obtain response from people so that the numbers associated with the responses are additive. This problem occurs whenever stimuli are presented to subjects and the subject is directed to indicate his judgments about the stimuli. In this type of problem the experimenter is frequently interested in reducing his data by combining or "adding" the numbers associated with stimuli which are positively intercorrelated.

Several graduate students in psychology have worked on this problem. The evidence for additivity comes from a variation of paired comparison. Each subject is directed first to choose the member of the pair he "prefers", or judges to be more favorable, or agrees with most fully. After he makes this judgment, he then indicates how certain he feels about his choice. That is, he might be asked to indicate how certain he is that he would make the same choice if he were given this choice again at some future time. The subject responds in terms of a "proportion" varying form .5, indicating almost complete uncertainty, to .99, indicating almost complete certainty. These proportions are transformed to normal deviates.

The evidence suggests this procedure results in linearity or additivity. It also seems that the procedure can be adapted to all types of psychological measurement where the experimenter is interested in obtaining the results of the process of judgment.

The Iowa Agricultural and Home Economics Experiment Station

The IAHEES provides financial support to permit certain members of the Statistical Laboratory Staff to consult on statistical problems in the biological and agricultural sciences. Foster Cady and Donald K. Hotchkiss, under the Agricultural Experiment Station Project 101, consulted with a large number of research workers and faculty members of designs, techniques, and analyses of investigations in the areas of soils, field crops, animal nutrition and management, botany, entomology, horticulture, forestry, food technology and zoology. Hotchkiss also worked with students and staff members in Home Economics in areas of human nutrition and in household equipment. Oscar Kempthorne and Dewey L. Harris under AES project 1448 assisted with problems involving genetics. N. V. Strand and Wayne Fuller performed statistical consulting services for sampling investigations under AES Project 113. The examples which follow are representative of the consulting services of the staff in these areas.

Statistical services in the plant sciences

Under Project 101, Foster Cady continued to provide statistical consultation services in the plant sciences, primarily in agronomy. botany, horticulture and forestry.

Several studies were carried out under different environments (locations and years). The ramifications of considering years and locations at random, approximate F tests, homogeneity of error, optimum number of replications per experiment, years and locations, and combining experiments where the treatments are not the same were among the problems encountered. Examples were a study investigating the relationship between second-ear development of corn and genotype x environment interaction, a horticulture experiment studying the effect of soil type on onion production, and a corn breeding experiment comparing the relative effectiveness of recurrent selection.

Essentially many of the consulting problems involved fairly standard designs and analyses, but some phases of the problems were different from the usual textbook solutions. Assistance was given in the analysis of variance calculations of a potato fertility split plot experiment where each whole plot consisted of a 3^3 factorial confounded in blocks of size 9. In a cytogenetic study comparing the number of two types of chromosomes in different families, the observations were variances. An onion bulb shape study involved consideration of criteria for measuring bulb shape. A modification of the usual plot analysis was necessitated in a corn breeding experiment when the subplots were not the same for all the

whole plots. Problems arose in an entomology experiment investigating the feeding habits of potato leafhopper when the sampling error was much larger than experimental error. An explanation for the large sampling error was hypothesized by a possible systematic gradient when the samples of each experimental unit were taken.

Investigations relating yield to several independent factors were handled by multiple regression techniques. The following were examples: relating yield of forage to soil fertility and grass composition factors; predicting depth of topsoil in a soil conservation study where several dummy variables were used; and evaluating the influence of temperature and precipitation on corn yields under different fertility levels and cropping systems. In the later study an attempt was made to express rainfall and temperature in the overall equation by Fisher's regression integral concept.

Practical considerations in carrying out an experiment were weighted in the design of experiments. With field experimentation, equipment considerations are usually important, as was the case in a discussion of the relative merits of a split plot experiment comparing tillage equipment treatments where large turn-around area is needed; or the example of different levels of auxin causing widely varying plant height in a soybean physiology experiment.

Inverse regression was used in two investigations. In a forestry experiment, estimates were made of different seedling compensation points (value of X when $Y = 0$) given the regression of absorbed carbon dioxide on light intensity. The other use was in a soil moisture study. In both of these problems, a recent publication, "Calibration" by H. T. David and F. B. Cady, was utilized.

Some of the other activities included planning orthogonal sets of treatment comparisons in a soil microbiology study; aid in the analysis and interpretation of a correlation study evaluating insect predators on corn; design of an experiment to study soil water stress with soybeans; use of sampling and experimental errors along with cost factors for the determination of optimum number of trees per plot in a forestry seedling establishment study; analysis of a composite design in the study of the response of turf to fertilizer; evaluation of mathematical models studying soil loss; optimum plot size for weed control experiments when less than the experimental plot was harvested; and design for a laboratory study investigating seed germination.

Statistical services in animal sciences

D. K. Hotchkiss has consulted with numerous research scientists in his work on AES Project 101. Interesting consulting problems with the animal scientists have included analysis of unbalanced experimental data, resolving complex regression equations, and interpretation of experimental results. Numerous consultees requested assistance in design of their experiments or in the analysis of a set of data which did not conform to standard textbook examples.

Four experiments were analyzed utilizing least squares solution of equations and the multiple regression program. In one of these the influence of season of the year and stage of lactation on several fatty acid constituents in milk was examined. With the assistance of Wayne A. Fuller, a least squares analysis was used to separate the influence of the factors. A response curve for each factor was fitted to the data after an adjustment was made for cow differences and the influence of the other factor. The data also provided information on two procedures for determining the total polyene content of the milk fat. One procedure incorporated a set of constant multipliers. The appropriateness of these constants was reevaluated, using the chemical determinations from this study. Some discrepancy between the established formula and the results in the current study was noted.

A type of incomplete multiple classification model was analyzed in a fish growth study. The data were in the form of a diagonal half replicate with three major classifications; year of birth, year of observation, and age of fish at time of recording. Sex as well as location of fish at time of catch were included in the study. Fuller assisted in setting up a least square analysis of these data. As reported by K. Hinkelmann, the linear response of the major classification could not be determined with this type of incomplete half replicate.

In a calf nutrition study, the extent of carbohydrate digestion was evaluated as it was affected by different sources of carbohydrate. The observations were recorded over time after feeding. Because of the expense and the mortality rate connected with the establishment of an ileal fistula (one of the locations of sampling), data was not always obtainable and occasionally results for an entire treatment were unavailable. The resulting disproportionality throughout the data complicated the analysis, and the least squares solution of the normal equations was required to isolate the influence of the treatment factors. The complexity of the resulting interaction terms necessitated a detailed breakdown of each calf's response in order to resolve these data completely.

A triangular partially balanced incomplete block design was utilized in a study to detect blood lipid changes as various fatty acids were incorporated into the basic calf ration. The data obtained were in the form of a split plot design, the split being observations made over time. The balance of this study was destroyed when one complete treatment was not consumed by the calves. Thus a least squares analysis was required. The limitations of the regression program required that the main plot analysis be run first, followed by the subplot analysis. Interaction between the main and subplot factors was evident, thus necessitating the plotting of a response curve for each treatment.

A multiple regression approach was used to develop a prediction equation for chloresterol content of blood as it was influenced by age and weight of growing chickens. The high correlation between these two factors essentially eliminated the consideration of both in the equation. Consequently a single variable equation for each was published, and a paper will be presented at a meeting of the Poultry Science Association. A similar regression approach was used for research workers in dairy bacteriology who requested an equation for describing the influence of salt, moisture and environmental variables on cheese ripening processes when bacterial populations were used as the measure of response.

Mice were used in a study to determine the influence of selected dietary food emulsifiers on numerous blood and biological factors. The ration effects were a part

of a factorial set of treatments which included sex of mice and growth periods. Unequal numbers of mice resulted, due to death losses not caused by the treatment. An analysis of unweighted means was used successfully to detect factor differences and large interactions of the factors. In some instances a detailed examination of the trend response was undertaken.

Rabbits were used to evaluate the relative acceptability of three forages. In an acceptability study the consumption of any one ration from a set of three is highly dependent upon which rations are offered simultaneously. Consequently each rabbit was offered less of each forage than would satisfy his appetite, forcing some consumption of at least two of the three offered. A balanced incomplete block design was used. The preference rank (1, 2, or 3) was recorded and these values analyzed.

For a staff member in Veterinary Medicine, the volume of readily available dust from a medical aerosol bomb was evaluated by using the approximately normal distribution properties of the "log particle diameter". The distribution statistics ($\bar{x}$, s) of five fractions of the dust were determined. These fractions were then segmented into portions readily absorbable and those too large for air suspension or too small for precipitation. The accumulated proportion of each fraction within the specified limits of absorbable particles was determined, using properties of the normal distribution.

Statistical services in Food Technology and Home Economics

Hotchkiss has consulted with research workers in Food Technology and Home Economics.

Assistance was requested by members of the staff in Food Technology in the design of balanced incomplete block experiments which would not exceed the tasting capacity of panel members. These studies dealt with different processes for preparing pork roasts, the influence of meat quality on steak acceptance, processing techniques as they affected stored poultry, and the influence of storage on cooked turkey acceptance. Correlations of various chemical measures were made with the mechanical and organoleptic properties of the food products.

Assistance was given in combining two independent studies for the comparison of similar treatments where methods of handling turkey in the institutional type kitchen were being studied.

Continued assistance was given in the design and analysis of an incomplete block study to determine the optimum conditions for washing "wash and wear" fabrics.

A modification of a triangular test was used to evaluate flavor differences as they were affected by starter cultures in stored cheddar cheese. Each judge was offered three cheese samples to be scored for flavor, using a 1 to 10 scale. Unknown to the judges, two samples were from the same starter culture, and the third was from a different starter culture. The variability of the scores of the identical cheese samples was used to estimate the within judge variance component, which in turn was used to evaluate statistically any starter culture effects. Cheese differences were evaluated using an F statistic.

Statistical services in genetics

The Genetic Statistics Group continued to consult with the Agricultural Experiment Station personnel in the departments of animal science, poultry science, agronomy, entomology and horticulture on the statistical aspects of their genetic research. Both Kempthorne and Harris consulted extensively with the members of the corn breeding group on various design and analysis aspects of their studies on corn improvement through breeding. They consulted on possible modes of interpretation of multi-way cross data, under the assumption that the lines entering the crosses are fixed. Harris also consulted on a study concerned with the development of selection indexes for soybeans and an assessment of the relative effectiveness of various indexes. This study offered some challenging features due to the fact that soybeans is a self-pollinating species, whereas most previous applications of index selection procedures have been for random mating species or for varietal selection.

Harris also consulted with an animal science staff member concerning the theoretical aspects of crosses between populations with differing gene frequencies and the interpretation of covariances between relatives under such a situation. A scheme for estimating components of variance from a study of identical twins on two levels of nutrition was developed for a graduate student in Animal Breeding. In this study an orthogonal set of mean squares was suggested by the nature of the data, but the number of mean squares was greater than the number of components of variance involved in the expectation of the mean squares. Thus, a weighted least squares procedure, with the mean squares as observed values, was suggested. Another study in Animal Breeding involved variance component estimation from a voluminous amount of data with quite disproportionate subclass numbers. A scheme was developed by Harris and the graduate student for choosing segments of the data in which equality of subclass numbers existed. The results of the subsequent analysis were examined for the suggested implications for testing of dairy sires.

Harris assisted a Genetics Department staff member in the development of a procedure for analyzing a cytogenetic study of the influence of irradiation on oats.

Francis Giesbrecht completed a study of an experimental design suitable for separating genetic and environmental changes in populations under selection. He worked with A. W. Nordskog on a project involving the analysis of some data available from a relaxed selection program in chickens.

Harris and the other members of the Genetic Statistics Group consulted on numerous other studies, but the above studies should serve as examples of the challenging and interesting problems encountered.

Statistical services for sampling investigations

Statistical Services for Sampling Investigations, Project 113, is a consulting project with the Iowa Agricultural and Home Economics Experiment Station under N. V. Strand and Wayne A. Fuller. Statistical services were provided within the areas of economics and sociology, agricultural engineering, animal and poultry science, rural sociology, forestry, agronomy, textiles and clothing and education. Consultation and advice have been given on design of surveys and experiments, collection of data, outlining appropriate analyses of data col-

lected, supervising required computations and assisting in the interpretation of results.

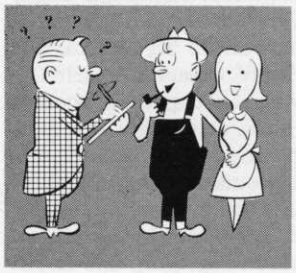
Much of Strand's work on Project 113 was done in connection with the work of the Survey Group on such projects as those discussed in the next section of this report. Other projects representative of Strand's consulting activities follow.

With Earl O. Heady of the Department of Agricultural Economics and James Vermeer of USDA, Strand worked on a joint research project to evaluate certain aspects of the 1962 feed grain program. For the Department of Food and Nutrition he advised on appropriate sampling designs for the measurement of nutrition qualities in the diet of young children in the North Central states. He also assisted a graduate student in Home Economics with a study of the choices of home furnishings by young homemakers. Strand helped a member of the Department of History, Government and Philosophy on the sampling aspects of a study of farmer reactions to the 1964 wheat program. He also helped a graduate student in the Department of Technical Journalism with a study to measure the reactions of persons in a 10-county area in southern Iowa toward publicity concerning area programs. With John R. Timmons and Donald Kaldor of the Department of Economics and Sociology he drew a sample for use in a study of land tenure in the North Central Region of the United States.

Wayne Fuller also consulted on a number of studies under AES Project 113. He worked with the agricultural engineers on the frequency of storms and runoff. The project was concerned with summarizing weather data on severe storms and run-off to obtain information on frequency, with the ultimate purpose of relating this to gully-growth. He also worked with a member of the staff of the Department of Economics on the analysis of a survey study of land contracts. The effects of purchasing land through land contracts on the purchaser's capital and income position were studied. These farmers, contacted in 1956, were interviewed again in 1962 to measure their financial progress. Fuller also consulted with members of the economics staff on a study of the reporting of land values by real estate brokers. Possible break-downs of the variations in reported land values were considered. Fuller and Foster Cady studied the effects of weather on corn production.

With members of the staff of the Department of Education Fuller consulted on the design of surveys of high school graduates to evaluate their education in the light of their occupational experiences and accomplishments. The purpose of this study was to gain information about the educational needs of students in the future. With members of the Department of History and Government, he studied the relationship between population and the costs of county government from 1920-1960. The results of the study show that the cost per capita for the sparcely-settled counties is high, and for the counties with greater population the cost per capita is relatively lower. Fuller also consulted with an economist on the relationship between a person's goal attainment and his financial situation and personal characteristics.

The Survey Group



The Survey Group of the Statistical Laboratory, with its associated workers and interviewers, provides direct operational services to the rest of the campus on all aspects of sampling, surveys and census-type studies. Research in survey methodology and consulting services are combined with operational work through the financing of the Statistical Laboratory, the Statistics Department, the Iowa Agricultural and Home Economics Experiment Station and the Industrial Science Research Institute. Professional staff members of the Survey Group for 1962-63 were H. O. Hartley, Norman V. Strand, Wayne A. Fuller, Leroy Wolins and J. N. K. Rao, a group of associates, Harold Baker, Robert Cochran and Thomas C. Jetton, and graduate assistants Louis W. Johnson, John Lesem, Philip Mills and Martin Rosenzweig. Helen Ayres and Marjorie Mason were supervisors of field interviewers and clerical personnel.

During the year, the Agency for International Development supported one student and USDA supported two in-service trainees who took courses in the department and worked with the Survey Group personnel in a special training program involving agricultural statistics. Two National Science Foundation Undergraduate Science Education students worked with the Survey Group.

Data obtained through the applied research and operational services of the Survey Group were used in a master's degree thesis in the Department of Agricultural Economics. ("The Conditions of Entry into Farming in Iowa 1959-60" by Thomas C. Jetton, Iowa State University Library, May 1963.)

The Survey Group participated in a number of studies during the year which were administered jointly by the Statistical Laboratory and some other University institute or state or federal agency, such as the Iowa Agricultural and Home Economics Experiment Station, the Agricultural Adjustment Center, the Industrial Science Research Institute, the United States Department of Agriculture, the Census Bureau, and the Iowa Conservation Commission.

Some projects, which were initially described in the 1961-62 Annual Report, were completed; others described then were continued; and still others were started in 1962-63. Progress is reported here on representative studies carried on by the Survey Group.

The Farming Opportunities in Iowa project, with D. R. Kaldor of the Department of Economics and Sociology and N. V. Strand as co-leaders, was continued. The purpose of this study was to estimate (1) how many farmers entered farming in 1958-60; (2) how many left farming; and (3) the size, quality and other characteristics of the farms given up. An analysis was completed of the conditions of entry of new farmers during the 1959-60 period. Thomas Jetton performed the operational and analytical work on this project and wrote his master's thesis on it. Findings show that approximately

2500 persons began farming per year in the two-year study period in the universe sampled. Although beginning operators were found to vary widely in age from 16 to 64, more than two-thirds were under 30 years of age. More than 82% of the operators observed started as sole proprietors. As might have been expected, most of the operators were tenant farmers. Although the average net worth of the beginning operator was found to be slightly less than $9000, the distribution was quite skewed with more than 70% of the operators having less than the mean net worth for all operators. Gifts were found to play an important part in helping the beginning farmer get started.

An Agricultural Experiment Station bulletin is being prepared to report the findings of the study on the conditions of entry of beginning farmers and on their initial year's financial experience. Another bulletin is planned to give an analysis of the components of change in the number of farm operators in Iowa.

Another study, a follow-up of the study of senior high school boys' career plans, which was reported in the 1961-62 Annual Report, was started. A survey was made to learn what the young men in the original study actually did do between June 1959 and February 1961, what their education and work experiences were during that period, and what their plans were for the future. Thomas Jetton wrote the mail questionnaire and supervised the coding and punching of the data. The results will be analyzed by D. R. Kaldor. Data collected in the 1959 survey were used for a study made to test the relationship between grade point, amount of time spent on extracurricular activities, intelligent quotient and the use of a car in school.

Again one of the main tasks of the Survey Group has been the contract with the Bureau of the Census. The study of Response Errors for Agricultural Data Collected by Mail was completed. Wayne A. Fuller, as leader, and Harold Baker worked on the project in 1962-63. The analysis of data showed very little difference among the four types of questionnaires with respect to bias or mean square errors. The inclusion of 1959 census data to provide a "benchmark" against which the respondent could relate changes did not appreciably reduce reporting errors. It is expected that the questionnaries used in the evaluation of mail responses will be incorporated in the forms to be used in the evaluation of the 1964 Census of Agriculture.

Edmund Fuller and Nancy Preston, National Science Foundation undergraduate science education participants, developed computer procedures and techniques for machine edit of the questionnaires. (See Undergraduate research report on page 30.) Research on another study for the Census Bureau is reported on page 29.)

Harold Baker prepared a report for the Bureau of the Census on the livestock portion of the evaluation of the 1959 census, which was reported in the 1961-62 Annual Report. This report, which discussed the procedures used and the results obtained, completed the 1959 study and will be incorporated in a report to be published by the Bureau of the Census.

Data were collected during the summer of 1962 for the Iowa Park System Study, a study which was carried out in cooperation with the Iowa Conservation Commission under the direction of Strand and Fuller. Questionnaires were distributed on a random basis to occupants of cars at the exits of 28 Iowa parks. Persons sampled in this manner were asked to fill out the questionnaire at their convenience and to mail them to the Statistical Laboratory for processing. In addition a sample was drawn from the camper registration lists at the parks; these persons were sent questionnaires through the mail. Personal interviews were obtained for a subsample of the non-respondents to the mail phase. A subsample of the original sample, including respondents and non-respondents (out of state as well as in state park visitors), was interviewed to obtain information on their outdoor recreational activities for the entire season, including the parks they visited and their expenditures in various categories. For the first phase of the study, 9000 questionnaires were obtained and tabulated. Tabulations included number of visitors by age and sex, distance travelled to the park, length of stay, source of information about the park, facilities used, activities most popular with park visitors, and suggestions for park improvements. A partial summary of the data has been made and presented to the Iowa Conservation Commission, and a report will be prepared early in 1963-64.

The Business Impact project, supported by the Iowa College-Community Research Center, the Department of Economics at Iowa State, and the State University of Iowa, was continued. The purpose of this study is to analyze the effect on towns and cities of expenditures by households and farms. The work of the Survey Group was directed by N. V. Strand. Coding was completed and tables were run on percentages on various categories of expenditures in towns of specific sizes. Dale Grosvenor prepared the IBM 7074 programs to make the tabulations. Other tabulations and analytical material is scheduled to be completed in 1963-64.

A three-year study, the Effects of Product Development, Merchandizing, and Promotion, AES Project 1404, was begun in collaboration with research workers in agricultural economics, food technology, animal and poultry sciences. One of the main purposes of this study is to learn the effect of meat purchasing habits of householders on the flow of livestock from the farmers to the buyers, packers and retailers. A sample was drawn in Webster County, Iowa, from which approximately 800 householders were to be interviewed over an eight-week period in June-July 1963. N. V. Strand set up the sample design, a feature of which was a rotation panel scheme wherein each householder would remain in the sample only four weeks. Members of the project committee specified objectives and prepared lists of items on which data were to be gathered.

N. V. Strand worked with Richard Forsythe of the Department of Poultry Science on a study to determine the relative preference of housewives for fresh or frozen chickens. A questionnaire was prepared, the sample was drawn, and data was collected. The questionnaire and IBM codes were also used by scientists in California and Minnesota, who conducted surveys in their own states. In Iowa the study centered in Marshalltown. Preliminary tabular material summarizing the results was prepared and submitted to Forsythe for analysis and report writing on the project.

A survey was made for Bob R. Holdren of the Department of Economics and Sociology to learn the shop-

ping habits of householders in northwest Des Moines, particularly their food-purchasing habits. Interviews were conducted with a random sample of 300 residents to learn where residents shopped and to obtain information about expenditures patterns. Following the interview, at which data was taken for purchases made in the week preceding the interview, a questionnaire for the coming week was left to be filled out and mailed to the Laboratory. Data processing phases of the survey were commenced upon completion of the field work.

The Survey Group assisted a graduate student working under Geoffrey Shepherd in the Department of Economics on a study in Hand County, South Dakota. The purpose of the study was to evaluate the effects of the Soil Bank and Feed Grain programs on the local businesses in a rural area. Hand County, a typical plains county, was chosen because it had a high degree of participation in these government programs. The Survey Group assisted in construction of the questionnaires, supervised the field work and coded the data for IBM use.

Operational assistance was given by the Survey Group to other departments on campus on a number of studies. Such assistance was given W. F. Kenkel of the Department of Economics and Sociology on a study of the effects of whiplash injuries in automobile accidents. The Survey Group coded and tabulated the data for this project. For George M. Beale and Joe A. Bohlen, professors of sociology, the Survey Group drew a sample for a Farm Machinery Study to learn farmers' attitudes toward and preferences for various types of machinery. Assistance was given in setting up and in training interviewers for a survey of the reading habits of south Iowans, a study which is part of a ten-county program. Samples were drawn for a Soybean Seed Quality Study for the Iowa Crop Improvement Association and the Department of Agronomy, for a Foot Health Survey for the American Podiatry Association of Topeka, Kansas, and for an Iowa City Nutrition Study for the Department of Internal Medicine at the State University of Iowa.

Work on a Feed-Grain Study began in the winter of 1963, a cooperative study of the USDA and the Iowa Agricultural Experiment Station. One of the main purposes of the investigation was to compare the relative advantages to farmers participating in or not participating in the Feed-Grain Program. E. O. Heady is project leader for the AES and James Vermeer for the USDA. The Survey Group recruited and trained the interviewers and collected data in Iowa, South Dakota, Minnesota and Indiana. The Group also coded data for Missouri, Nebraska, Idaho, Washington and Texas. This work will be completed in 1963-64.

Attitudes and Values Associated with Agricultural Land Tenure, a study in which the Survey Group is collaborating with the State University of Iowa, is part of a research program of the north Central regional tenure committee of the Land-Grant colleges. Marshall Harris, USDA representative at the Agricultural Law Center, SUI, and Harold Mark, Ph.D. candidate and a member of the Sociology Department at SUI, originated and supervised the project. The Survey Group drew the sample, worked on the final drafts of the questionnaires used in the study, recruited, trained and supervised the interviewers, and performed the coding for the IBM. A set of cards was sent to SUI for use by Mark in making analyses of the data for a preliminary report. He also used the data for his Ph.D. thesis. The Survey Group will make the final tabulations on the result of the survey.

Tabulations for the Soil Conservation Service, USDA, were continued under the direction of N. V. Strand. Much of the work done in 1962-63 was in placing land resource area and watershed codes on IBM cards. Some states originally used soil code symbols which were not consistent from county to county within a state. In order to make the data useful for purposes other than the soil and water conservation inventory, which was the primary purpose of this study, it was necessary to revise the soil coding system to a state rather than a county basis and to repunch the soil fields on the IBM cards.

Watershed tabulations were made for specific areas in Mississippi and Nebraska. Tabulations were also completed for Idaho.

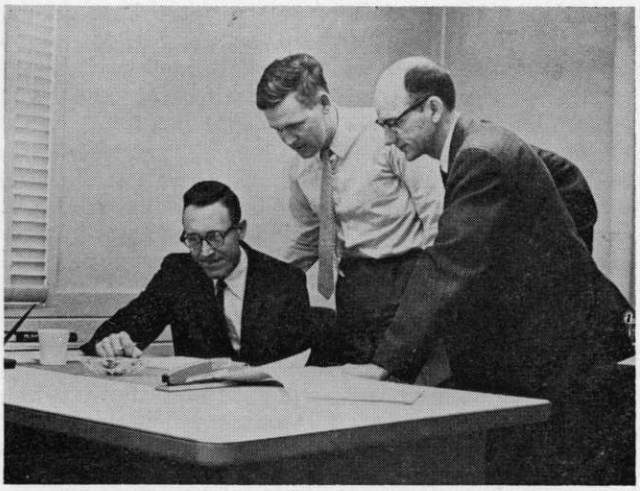
States involved in a special study of the Ohio River Basin included Ohio, Indiana and specific counties in Tennessee, Illinois, Kentucky and North Carolina. Special sets of tabulations were made for use of economists in ERS, USDA, of data for these areas. Similar studies are under way on the Upper Mississippi Basin and the Missouri River Basin.

A study of soil types is being made for the Tennessee Valley Authority using SCS data. Tabulations will be made for specific counties in Alabama, Georgia, Kentucky, Mississippi and Tennessee on land use acreages by watershed, land resource area, by soil type and land capability units.

The Numerical Analysis-Programming Group

This group concentrated much of its effort during the year to provide the capability for processing on the IMB 7074 an extensive range of problems in analysis of variance and linear programming. In addition to its work in these two areas, the group did consulting work on a diversity of research problems. Staff and graduate students from the departments of Agronomy, Animal Science, Bacteriology, Economics and Sociology, the Engineering Experiment Station, Child Development, Household Equipment, Industrial Engineering, Poultry Science, and Psychology requested computing assistance. Consulting services were also given to WOI and the National Animal Disease Laboratory.

In addition to specialized programs relating to analysis of variance problems, the AARDVARK system described on page 19 was designed, programmed, and utilized for a wide variety of production problems. Data analyzed ranged from judgments of creativity and competence of art school pupils to genotypic and phenotypic data simulated by Monte Carlo procedures. The problems also varied in the degree of analytical assistance required in defining the problem as opposed to computational assistance in calculation of specified analyses. W. J. Hemmerle consulted with a staff member of the National Animal Disease Laboratory in the analysis of data on blood glucose and creatine in a study to determine the cause of acute Swine Erysipelas.



William J. Hemmerle, E. J. Carney and Dale Grosvenor in one of the new offices of the Numerical Analysis-Programming Group in Bldg. N.

E. J. Carney conducted an experiment in conjunction with the Industrial Engineering department in an egg products plant to determine the effects of case re-use on egg solid yield and labor cost. More than 150 analyses were processed by AARDVARK or prior versions of this system for a number of different departments.

In the area of linear programming, the IBM Library Program S2 was adapted for I.S.U. use by D. Grosvenor, and the Penn State Activity Analysis Program was modified, extended, and made operational by D. Grosvenor and D. Thomas (page 19). Several research problems in linear programming were subsequently processed with the assistance of Grosvenor and Thomas. One of these problems involved the use of the complete description method of computing the vertices in the feasible solution space to develop and extend some theorems about stochastic linear programming. Another problem was designed to test alternative methods of defining labor and capital restrictions and coefficients to treat more realistically the problem of differential rates of capital turnover.

Consulting and programming assistance was given on many other problems in substantive areas utilizing statistics as a primary research tool. Programs were written for use in Factor Analysis. Mary Ann Carney prepared a general program to compute sets of correlation coefficients within and between blocks of variables for test score data. Programs were prepared for the department of Bacteriology by June Smith to find basic similarity levels between strains of bacteria for classification purposes.

W. J. Hemmerle consulted on a Poultry Science problem involving a least squares analysis of data. He devised a method whereby characteristic roots and vectors could be calculated to determine suitable restrictions for solving a system not of full rank. E. L. Fuller assisted in obtaining the computer solution. W. J. Hemmerle and E. J. Carney assisted A. W. Qureshi in designing and coding computer subroutines for Monte Carlo generation of genotypes and genotypic and phenotypic values in the simulation of genetic processes.

Some special purpose programs were prepared for regression analysis problems in sampling and experimental design. A graduate student in the Engineering Experiment Station was assisted with transformations of initial variables and subsequent regression analysis on the transformed data. Distance and population functions were evaluated for a graduate student in psychology and studied in a regression analysis. Optimum values were determined for formula coefficients of a composite estimator in rotation sampling. A procedure was designed and implemented to edit farm census questionnaires by the computer. Efficiency factors were determined for different experimental designs, and variances of general combining ability comparisons were computed.

Dale Grosvenor and a graduate student in the Department of Economics and Sociology visited at the ESSO Computing Center of ESSO Research and Engineering Co. at Plainfield, N. J. in August, 1962. They discussed linear programming techniques with the numerical analysts, programmers and economists of ESSO Oil Co. and became familiar with the use of their LP/90 Linear Programming program so that a programming model being formulated could subsequently be run on their IBM 7090 computer. Grosvenor again visited the ESSO Computing Center in November with another graduate student from the Department of Economics and Sociology to consult with numerical analysts regarding the formulation and computation of linear programming models concerning an interregional analysis of agricultural crops among the states of India. Both these trips were made at the request of Earl O. Heady of the Center for Agricultural and Economic Adjustment at Iowa State.

E. J. Carney assisted T. A. Bancroft with tabulations for a paper presented at the Minneapolis meeting of the American Statistical Association. Fortran programs were written for computing biases and mean square errors for various time values of the regression coefficients and significance levels of preliminary tests.

Mary Ann Carney assisted Bancroft in the preparation of data for a paper given at Harvard University in May 1963 by computing a series of analyses of variances and related tests on transformed data and preparing the graphs and tables relating to these computations.

Off-campus advisory assignments

Staff members, in addition to their regular consulting duties, are often called upon to act as consultants to off-campus groups. T. A. Bancroft spent three weeks in Mexico in February on a special assignment for the Rockefeller and Ford foundations. He served as Special Temporary Aide to the Graduate College, National College of Agriculture in Chapingo, and to the National University and the Technological Institute in Mexico City. At Chapingo he evaluated a proposal for establishing a statistical laboratory, including a high-speed computing service. In Mexico City he was consulted on plans for an expanded program in high speed computing at the National University and at the new National Polytechnic Institute Computer Center and the coordination of these programs.

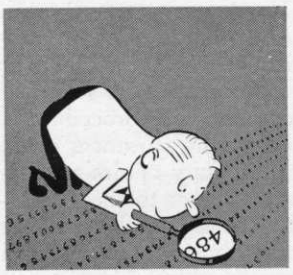
O. Kempthorne participated in a conference on undergraduate courses in Statistics held at the Center

for Advanced Study in the Behavorial Sciences, Stanford University on Nov. 17, 1963. The conference was organized by the Panel on Mathematics for the Biological, Management and Social Sciences of the Committee on the Undergraduate Program in Mathematics of the Mathematical Association of America. The Conference drafted a report giving recommendations for a one year course in probability and statistics for students with a solid back-ground in calculus; a recommendation for a mathematical course in probability and statistics that does not have a calculus prerequisite; and a recommendation for a course at an intermediate level based on about a semester of calculus and some previous preparation in finite mathematics.

Kempthorne spent the month of June 1963 in Argentina under the sponsorship of the National Institute of Technical Agriculture (INTA), the Institute of Calculus of the Faculty of Exact Sciences of the University of Buenos Aires, and the Rockefeller Foundation. He gave a course of about 25 lectures on the principles of experimental design and inference to a group consisting of statisticians of INTA, of the School of Statistics of the University of Rosario and of the University of Montevideo. He also gave general talks on the basic ideas of experimental statistics to the staff of INTA at Castelar and to the staff and students in statistics at the University of Rosario. He gave a lecture on some statistical problems in medical research to the School of Clinical Science of the Medical School of the University of Buenos Aires. He also consulted with a group working on prevention of hailstorms by iodide seeding, and a group concerned with the interpretation of beef herd data. He visited three of the larger agricultural experiment stations of INTA, Castelar (the Beltsville of Argentina), Pergamino and Balcarce.

H. O. Hartley consulted with the research staff of I. E. Dupont de Nemours Co., Wilmington, Delaware, during a period of three weeks in July, 1962. Research workers of the company presented a great variety of problems, such as textile technology, chemical research, operational research, simulation of production processes by analogue computers, and solutions of scientific problems by numerical analysis. Special seminars comprised nonlinear least squares estimation and introduction to sample methodology. A special talk of general interest to research staff of the textile technology section on design of experiments was arranged. Two reports concerning problems in administration of scientific personnel and high speed computer policies were presented during the period.

Current Research



In the research program of the Laboratory both the development and extension of basic theory and its application to new statistical methods and techniques are emphasized. The Laboratory cooperates with other research institutes and experiment stations on the campus in arranging programs of statistical research and in obtaining support through joint grants and projects. Many studies of a fundamental nature are supported by the increasing number of such grants and contracts with off-campus agencies. In most instances they provide for supervised graduate research on topics proposed by the staff. The Statistical Laboratory budget, however, supports only projects which are of specific interest to regular university research programs.

Statistical Laboratory Research Projects

In general theory

T. A. Bancroft and Harold J. Larson, formerly of I. S. U. and now of Stanford Research Institute and the U. S. Naval Post Graduate School, conducted joint research on two problems in the area of incompletely specified linear regression models. One problem, based on Larson's Ph.D. thesis, is described in a publication summary on page 31.

The other problem, based on Larson's M.S. thesis, was concerned with consequences of an experimenter separating the independent variables in a multiple regression model into two classes: (1) those the experimenter is sure are necessary for accurate predictions, and (2) those 'doubtful' variables which will be included in the model if the joint preliminary hypothesis (that the 'doubtful' coefficients are simultaneously zero) is accepted. Designating the predictand finally adopted y*, then y* is dependent on either k predictors (keep all independent variables) or on m predictors (accept the hypothesis that k-m of the coefficients are jointly zero, and delete the corresponding k-m independent variables from the model). The expected value (thus exhibiting an explicit evaluation of the bias) and the mean square error of y* were derived. Some tabular values were obtained to illustrate the numerical size of the bias in certain situations. An example was constructed to illustrate the use of those tabular values and the mean square error was numerically evaluated and discussed for the example. The inference procedures used above, incorporating a preliminary test of significance, belong to a general class of such procedures relating to the analysis of incompletely specified models. A review of earlier work in this area was undertaken.

Bancroft extended some of the earlier work with Larson by considering the problem of adding versus deleting predictor variables in the analysis of incompletely specified regression models. Considering the standardized bias to be the only, or decidedly the most important, criterion of comparison, it was shown that Procedure A (deleting non-significant predictor variables by starting with the last of the ranked 'doubtful' predictor variables) would in general be preferred to Procedure B (adding significant predictor variables by starting with the first of the ranked 'doubtful' predictor variables.) This appears to be in contradiction to the 'practical' recommendation to use a decision rule based on testing sequentially for possible addition of predictor variables of a 'doubtful' ranked set (Procedure B) since it provides positive information or terminates

at any step beginning with the first. However, different results may be obtained from a comparison of the respective mean square errors. A preliminary report on this research was presented at the annual meetings of the statistical societies in Minneapolis, September 7-10, 1962.

T. A. Bancroft, K. C. Chanda, and T. A. Brindley conducted research and presented an invited paper at the joint Biometric Society (ENAR) and IMS meetings at Harvard in May 1963 on "Some Problems of Specification Arising in Entomological Data Analysis." It is suggested that an appropriate "biologically significant" model relating observed larvae or damage counts through time or space to population growth or spread should represent a family of distributions, including as a special member one that also would be an appropriate specification to be used in the analysis of the usual comparative experiment using counts made at a single point in time. In Part I some attention is given to effective ways of obtaining an appropriate transformation as a means of interpreting counts made at a single point in time so that treatment effects in a comparative experiment may be judged. In Part II an attempt is made to formulate a dynamic stochastic model for the distribution of larvae at any given point in the time period of the entire life cycle of the insects. It is assumed here that the distributions of the number of egg masses over the particular geographical area and the number of eggs per egg mass are both Poisson and independently so. The eggs produce larvae which are assumed to be subject to a continuous death on immigration rate $\mu_n(t) = n\mu(t)$ and $\lambda_n(t) = \lambda(t)$ respectively through the entire cycle.

In another study Chanda investigated the asymptotic sampling properties of the sample distribution functions of random variables, $X_1, X_2, \ldots, X_n$ being realization from a linear normal process $\{X(t)\}$. If $F_n(x)$ denotes the sample distribution function associated with these random variables, then it has been shown that the joint distribution of $\sqrt{n}\{F_n(x_i)-F(x_i)\}(1\leqq i\leqq k)$ where $F(x)$ is the univariate distribution function associated with X_1 and $x_1, \ldots, x_k$ are arbitrarily chosen real numbers, is, asymptotically, normal with mean zero. More generally, if $\{Z(t)\}$ is stochastic process defined by $Z(t) = G\{X(t)\}$ where $G(x)$ is integrable with respect to the normal measure with zero mean, then the property of asymptotic normality can be extended to the class of statistics $\sum_{t=1}^{n} Z(t)/n$. The particular case considered before assumes that $G(x)$ is a characteristic function.

Chanda also initiated investigation of the sampling properties of tests for categorical data. The present report concerns the portion of investigation relating to binomial data. It is the usual practice to employ large sample tests for testing equality of binomial probabilities for several samples even when the sizes of the samples are only moderately large. The corresponding loss of validity of the test criteria can be quite large. An attempt was made to reduce this loss by working out the small sample moments to order $\frac{1}{n}$ (It is assumed that the sizes of the samples are the same and equal to n) and modifying the classical χ^2 tests. Three different tests have been considered in this context, viz, (1) the standard χ^2 test, (2) the likelihood ratio test, and (3) the test based on the classical arc sine transformation.

H. T. David extended his work on the asymptotic independence of the order statistics and statistics of structure-d to order statistics other than the extremes. The main tool is a multivariate probability integral transformation recently discussed by M. Rosenblatt in the Annals of Mathematical Statistics.

David and Rudolfo Mengido investigated the construction of a symptotically exact truncation rules for certain types of binomial sequential probability ratio tests. Such probability ratio tests constitute random walks with two absorbing barriers. This research attacks the problem of truncation through the convergence properties, under iteration, of the transition matrices associated with these random walks.

In Survey Methodology

H. O. Hartley and Robert Cochran continued their research on multiple frame surveys as described in the 1961-62 Annual Report. The theory of using multiple frames in survey designs is extended to new complex situations of two-stage designs. Formulas were obtained for the optimum allocation of primary and secondary sample sizes to the various frames, as well as for the optimum weight coefficients. Hartley commenced work on the application of nonlinear programming to problems in survey design optimization. In particular, the problem of multiple purpose optimization in stratified sampling was considered in detail. The procedure is one in which tolerances are set for the variances of all the estimators arising in the multiplicity of design purposes, and the cost is minimized subject to these tolerances. If the reciprocals of the strata sample sizes are introduced as new variables, this approach results in a non-linear programming problem in which the objective function is nonlinear, separable and convex in the variables and all restraints are linear.

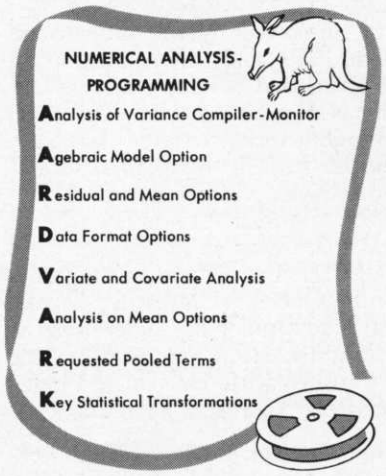
In engineering research

Herbert T. David's joint research with engineering staff and graduate students centered on two main areas. In the first area, a fairly general study was undertaken to determine the applicability of various statistical techniques to Civil Engineering experimentation. Greatest promise in this regard was shown by control chart procedures, use of control runs, balanced fractional experimental plans, systematic designs, probability-paper plots, and extreme-value analysis. Several pilot experiments were carried out to illustrate the application of these techniques. In the second area, multiple regression was applied to the problem of finding good correlates of soil-cement strength and the ensuing problem of soil-cement strength prediction.

Study was made into the area of analysis of life test data by Dale O. Richards, whose Ph.D. thesis is being written on "Incompletely Specified Models in Life Testing." Consideration was given to the effect preliminary tests of hypotheses have on subsequent inferences under various distributional assumptions. Under the assumption of a parent two parameter exponential distribution,

quite extensive results were obtained when studying the effect of preliminary tests of the location parameter on subsequent inferences about the scale parameter. Inferences both in estimation and testing of hypotheses were considered. For comparative purposes in estimation, the bias and the mean square error of the resultant estimators were calculated. For comparative purposes in testing the hypotheses, the size and power of upper and lower one-sided 5 percent tests were examined. Recommendations were made for use of specific levels of significance of the preliminary test for certain known a priori information about the location parameter for each of the situations of estimation and testing hypotheses. Assuming no known a priori information, tables and figures were developed to indicate how the characteristics examined vary for different values of the location parameter. One might refer to the problem considered as the incompletely specified exponential distribution.

In Computational Methodology



Aardvark—A Compiler-Monitor System for Analysis of Variance.

AARDVARK has been designed to handle an extensive number of problems in the area of analysis of variance utilizing a language or notational scheme consistent with the usual symbolic description of the statistical problem. A problem is defined by a set of problem statements which may include covariates, variate and covariate transformations, selected means, residuals, or estimates to be computed, as well as other factors relevant to the analysis. The statistical model may be stated algebraically, with further specifications employing the symbols which appear in the model. In some common cases, algebraic specification of the model is optional. Data is described by specifications that relate to the problem statements.

An aim of AARDVARK is to process problems in a routine one pass manner with convenient data preparation prior to processing. Annotated output consists of analysis of variance tables, crossed product terms when covariates are specified, and selected means and residuals (i.e. estimates of effects) when desired. Algorithms used in computing the analysis of variance minimize loss of significance. Various error provisions allow for salvaging of partially completed jobs in case of data card errors. AARDVARK takes advantage of the large memory and high speed of the IBM 7074 computer to enable large problems to be handled swiftly and economically. At the same time, it is versatile enough to make the processing of small problems compare very favorably in cost with desk calculator processing.

A mimeographed reference manual which explains the system has been prepared and is available from the Statistical Laboratory.

Dale Grosvenor and David Thomas modified and extended the Pennsylvania State University Activity Analysis Linear Programming Systems for use at Iowa State. This program, written in FORTRAN, uses the ordinary simplex method with single precision floating point calculations for solving linear programming problems. The size of problem is limited to 60 rows including the functional row. The number of columns is essentially unlimited (approximately 30,000). Data input must include the entire matrix with all zero elements included. If desired, the final solution will be substituted into the original equations as a check on the accuracy of the calculations. Provision is made to revise elements of the matrix after the original data have been input. This is useful either in correcting data or in solving several very closely related problems using the same basic matrix.

An IBM Library program, S2, a linear programming system, was adapted for Iowa State University use by Dale Grosvenor. The S2 program, written in a special linear programming compiler language, uses the revised simplex method with some single precision and some double precision floating point calculations. The size of the problem is limited to approximately 200 rows and 2000 columns. The input format is essentially the Share standard format as is used in LP/90 on the IBM 7090. This program permits multiple resource columns in the input. The effective resource column for optimization purposes is a given linear combination of those which have been input. Multiple functional rows may be input and they also may be used separately or in combination, as with the resource columns.

Write-ups and program decks for both the S_2 linear and the activity analysis programs are available from the Numerical Analysis-Programming Group of the Statistical Laboratory.

The work of Howard Jespersen, who holds a joint appointment in the Statistical Laboratory and the Computation Center, resulted in a general purpose program in multivariate linear regression. This system of programs will (1) form the correlation matrices on up to 130 variables with no more than 99,999 observations (and no missing data); (2) give the totals, means, raw and corrected sum of squares, mean squares and standard deviations for each variable; (3) allow for any number of variables to be selected as independent and any of the remaining to be dependent; (4) solve the resulting normal equations formed in (3) and give the analysis of variance of the regression, the solution, their standard errors and t-tests; (5) will plot any one variable against another by use of a "bivariate plotter;" and (6) after obtaining the solution in (4), will calculate the predicted values from the normal equation.

Hemmerle engaged in research concerning the automation of analysis of variance procedures. Because of the difficulty involved in the man-machine interface, particular emphasis was placed on the use of algebraic specifications, rather than absolute machine type, in order to express the statistical analysis. Aardvark is an outgrowth of his research for his doctorate.

The research of Dale Grosvenor was concerned with those models of linear programming in which some or all of the variables are restricted to be integer-valued in the final solution.

The computational aspects of integer programming were considered from the the standpoint of computing efficiency and size-of-model limitations. A new computational algorithm was presented for the full-integer and partial-integer algorithms. This new method permits the use of the product form of the inverse in its calculations. Available computer programs used an in-core type of procedure, which limited the size of the input model to those which could be stored in the immediate-access storage of the computer. The use of the so-called "product-form-of-inverse" method relaxes the size limitation considerably.

Consideration was given to the problem of conditioning the matrix of linear programming coefficients. A specific conditioning technique was presented which used an analysis of variance procedure on the logarithms of the coefficients.

H. O. Hartley carried on research on the application of integer programming to project optimization in Agricultural Economics. These projects are often faced with a choice between mutually exclusive subprojects. Integer programming attaches to the i-th of these subprojects an integer variable V_i bounded to lie within the range $0 \leq v_i \leq 1$ so that $v_i = 0$ (i-th subproject out) or $v_i > 1$ (i-th subproject in). By adding the restriction $\Sigma v_i = 1$ precisely one subproject is selected and by adding the restrictions $\Sigma_j x_{ij} \leq v_i M$ the activities x_{ij} entering into the i-th subproject are all forced out when $v_i = 0$. This method appears to have been originally suggested by George Dantzig but has here been more thoroughly explored and developed.

An experiment in election forecasting at the local level, utilizing limited resources, was conducted to determine the outcome the 1962 Iowa gubernatorial race. Robert Clyde of WOI and William J. Hemmerle conferred with Dr. Donald Boles of the Department of History and Government and J. N. K. Rao and T. A. Bancroft of Statistics in establishing a statistical method for estimation commensurate with available data sources. Hemmerle designed the operational system, utilizing telephone lines to the Associated Press in Des Moines, the IBM 650, and other pertinent computing equipment. He managed the activities on election evening as well as two mock election trial runs. A staff of 25 persons from the Numerical Analysis-Programming Group, WOI, and the Computation Center were involved in this procedure. The IBM 650 program was prepared by Dale Grosvenor. Three estimators were presented to the viewing public over WOI-TV, a Combined Stratified Ratio Estimator, a Separate Stratified Ratio Estimator and an uncorrected observed percentage. These estimates were graphed as the evening progressed. However, only one prediction was made. With 13.95% of the total vote in, at approximately 10 P.M., a prediction was made, using the combined ratio estimator, of 52.20% for Hughes. The final percentage was 52.44.% A paper written by Clyde, Hemmerle and Bancroft has been accepted for publication in the **Public Opinion Quarterly.**

National Institutes of Health Training Grant Program

Since 1958 the Statistical Laboratory has had a grant from the National Institutes of Health. The purpose of the grant is to train graduate students in statistics with a view to stimulating their interest in biometry or medical statistics or public health as a career; to raise the quality and level of competence of these students; and to contribute to biometry generally by including basic research in their programs. Support for the program is assigned to students who show promise in basic, as well as applied, research. A five-year grant from NIH terminated June 30, 1963, and a second five-year grant has now been made for the 1963-68 period. The original program was designed to support three graduate trainees during any one fiscal year. In 1960 the contract was rewritten to support up to six trainees studying concurrently. The 1963-68 contract also gives support to six trainees and provides certain additional funds to staff these programs more adequately.

Three students were supported by the program in 1958, a fourth started in 1959, and three more began work in 1960. John Gurland had several graduate students in addition to the NIH trainees who worked on related topics, and 18 publications resulted from the research of Gurland and his students between 1958 and 1961. In 1960-61 Stanley Nash, visiting associate professor from the University of British Columbia, worked with the trainees while Gurland was on leave from Iowa State. Special additional NIH support was given James Kilpatrick, on leave from Queen's University, Belfast, Ireland, for postdoctoral study under the supervision of T. A. Bancroft during 1960-61. He is now lecturer and consultant in medical statistics with Dr. D. J. Finney at the University of Aberdeen, Aberdeen, Scotland.

T. A. Bancroft became director of the NIH-sponsored program at Iowa State in 1961, and C. P. Cox joined the staff to serve part time on the program. Cox has graduate degrees from Oxford University, England, has worked under Dr. F. Yates at Rothamsted, and has numerous publications in fields related to the NIH program. He conducts research in the area of biological statistics, guides NIH trainees and teaches an expanded course in bioassay (Stat. 535). In 1961-62 Bancroft, Cox and H. T. David directed the graduate programs of six trainees supported by the NIH grant. Other members of the statistics staff have also assisted these students.

In 1962-63 four master's degrees and two Ph.D.'s were awarded to NIH trainees by Iowa State University. One trainee continued his work and three new ones entered the program. Bancroft, Cox, David and K. C. Chanda were responsible for their research. Chanda joined the staff as visiting associate professor and added new courses in Stochastic Processes, Nonparametric Theory and Methods, and Advanced Time Series, which are available to NIH trainees and other students.

A balanced program of statistical theory and application based on modern statistical methodology is required of all graduate students in statistics. In addition to these core courses, NIH trainees are encouraged to take courses in the biological sciences and to participate in

National Institutes of Health Trainees, 1959-1963.

Top: Eugene Cohen, Carol Edwards, Roger Connelly, Paul Leaverton, James Zweifel, Fred Ramsey.
Bottom: Thomas Roseberry, Darrol Heggen, Ilbok Lee, Ross Adams, Michael Billings, Patricia Conn.

NIH-sponsored summer institutes. At such institutes students receive training in epidemiology and other specialized areas of public health and medical applications of statistics. Paul Leaverton, for example, attended the Graduate Summer Session in Statistics in the Health Sciences at the University of California, Berkeley, in 1962. The NIH program also provides travel funds for assignments of students to public health and medical centers, such as the NIH Center in Bethesda and other university and public health centers.

The experiences of a few typical students may be of interest in an appraisal of the NIH program. One girl, an NIH trainee, received both her master's and doctorate degrees under the program. She presented a paper at the 1961 annual meeting of the American Statistical Association and a publication resulting from her research has appeared in the Journal of the American Statistical Association. Upon completing her studies at Iowa State she became assistant professor in the Department of Behavioral Science in the Medical Center of a state university.

Another student worked on a number of statistical techniques which are commonly employed to analyze quantal response bioassay data. The results were reported in his doctor's degree thesis. He is now employed as a biostatistician in the Industrial Statistics Department of a large pharmaceutical firm.

A third NIH trainee worked on the problem of the potential toxicity of chemicals used in modern agriculture on human beings, as a result of which he earned his M.S. degree. He is now a statistical consultant at a state university Institute of Agricultural Medicine.

Four graduates became commissioned officers assigned to the U. S. Public Health Service. Three of them worked as statisticians with the National Cancer Institute in Bethesda, Maryland, and the fourth with the Communicable Disease Center in Atlanta, Georgia.

A cooperative arrangement has been made between the Statistical Laboratory and Western Reserve University Medical School. Iowa State has provided course work in Statistics for four Western Reserve NIH trainees, Betty Eberle and Howard Marsh in 1961-62, Lawrence Gould in 1962-63, and Lois Hendges, who began her work in June 1963.

The twelve trainees who have been supported by the NIH Training Program since its inception in 1958, the degrees they have earned, and the positions to which they have gone are listed here:

Trainee	ISU degrees	Present position
Roger R. Connelly June 1958-April 1960	B.S. 1958; Work completed for M.S., except thesis	Statistician, USPHS, National Cancer Institute Silver Springs, Md. On leave at Harvard School of Public Health, Cambridge, Massachusetts.
Carol B. Edwards June 1958-Sept. 1962	M.S. 1959; Ph.D. 1962	Ass't. Prof. of Behavioral Science, Medical Center, U. of Kentucky, Lexington.
Thomas D. Roseberry Sept. 1958-June 1960 Sept. -1962-	B.S. 1958; M.S. 1960	Statistician, USPHS, Nat. Cancer Institute, Silver Spring, Md. Returned to ISU to work on Ph.D., Sept. 1962.

Trainee	ISU degrees	Present position
Paul E. Leaverton Nov. 1959-	B.S. 1956; M.S. 1961	1957-59 — Commissioned officer USPHS, Communicable Disease Center, Atlanta, Georgia. Returned to ISU to work on M.S. 1961. Remained at ISU to work on doctorate.
James R. Zweifel June 1960-Aug. 1962	B.S. 1960; M.S. 1962	Commissioned officer, Biometry Branch, National Cancer Inst. Bethesda, Md.
Eugene B. Cohen Sept. 1960-Aug. 1962	M.S., 1962	Biostatistician, E. J. Squibb and Sons, Pharmaceuticals, New Brunswick, N. J.
Darrol W. Heggen Sept. 1960-May 1961	B.S. 1960; M.S. 1962	Statistician, Institute of Agricultural Medicine, State Univ. of Iowa, Iowa City
Ilbok Lee July 1961-Feb. 1963	B.S. 1956; M.S. 1959 Ph.D. 1963	Biostatistician, Parke, Davis and Co. Detroit, Mich.
Ross Adams April 1962-	M.S. 1963	To continue work toward Ph.D. at ISU
Michael Billings Sept. 1962-		Starting work toward Ph.D. degree
Patricia Conn Sept. 1962-		Starting work toward Ph.D. degree
Fred Ramsey June 1963-	M.S. 1963	Starting work toward Ph.D. degree

Twenty-nine Iowa State University graduates in Statistics are now employed in the general areas of biometry, medical statistics or public health. Many of these are former NIH trainees, and their positions show that the statistics program at Iowa State is producing graduates competently trained in these areas.

NIH research in 1962-63

During the 1962-63 academic year the NIH program has provided traineeships for four master's degree and seven Ph.D. degree candidates. The research of Eugene B. Cohen, Darrol Heggen and James R. Zweifel for the master's degree and Carol Bates Edwards and Ilbok Lee for their doctorates was reported in the 1961-62 Annual Report, although their degrees were granted in 1962-63. Their thesis abstracts, which are to be found on pages 39 to 42 of this report, summarize their research Ross Adams received his M.S. in May; Paul Leaverton has almost completed his work for the Ph.D. Thomas Roseberry returned to Iowa State to continue his research and to work for his doctorate after two years with the United States Public Health Service in Bethesda, Maryland. Michael Billings, Patricia Conn and Fred L. Ramsey have only begun their work toward their doctor's degrees.

Ross Adams, working under the direction of T. A. Bancroft, carried on research with a linear regression procedure for obtaining point estimates of missing values in time-trend investigations,—that is, continuous treatment, longitudinal experiments analyzed at a single predetermined terminal point in time. The method obtains an estimate directly from the linear regression model, after a series of preliminary tests. The method takes into account the relationship between response and time. K. C. Chanda provided advice in the serial correlation aspects of this research. A paper will be presented on this new technique at the annual meeting of the American Statistical Association in Cleveland in September 1963.

Thomas Roseberry continued work in drug screening which was begun while he was with the USPHS in Bethesda. "The Construction of Two-and Three-Stage Screening Procedures" by Roseberry and Edmund A. Gehan, Head, Biometrics Section, NIH, in Bethesda, appeared in **Cancer Chemotherapy Reports** in May 1962. This paper describes a method for constructing screening procedures which allow rejection at any stage but acceptance at only the final stage. A set of tables is given which greatly simplifies the calculations necessary for obtaining accept-reject rules and operating characteristic curves.

Since his return to Iowa State University, Roseberry has worked on an extension of this paper, investigating the optimality of the procedures. This work, which includes much more of the mathematical development of the method, has been accepted for publication in **Biometrics,** with Edmund A. Gehan as co-author. Also the tables for constructing accept-reject rules and operating characteristic curves are to be included in a book now in preparation.

Investigation of statistical procedures applicable in the analysis of bioassays when some of the usual assumptions are not fulfilled was undertaken by Paul Leaverton, under the direction of C. P. Cox. The standard statistical methods for the analysis and interpretation of bioassays are based on a number of assumptions concerning the relevance of the mathematical model to the behavior of the experimental material. The usefulness of such methods depends upon the appropriateness of these assumptions in a particular bioassay. Situations arise in bioassay studies where some of the usual assumptions are not justified so that the standard methods are inapplicable.

One basic assumption, which must be considered in the statistical analysis of all bioassays, is that the dose-response function, $F(z)$, is monotonic. General tests for montonicity, that is that

$F(z_1) \leq F(z_2) \leq \ldots \leq F(z_k)$,

where the parametric form of $F(z)$ is not specified, where developed for both the quantal and quantitative response cases.

Another assumption which is frequently made is that the response function is or can be made essentially linear, with respect to dose or log dose, by an appropriate transformation of the response variable. Standard tests of this assumption are available in the literature. The problem of estimating relative potency was examined when no exact parametric form, such as a linear model, for $F(z)$ is assumed.

For the relative potency of one preparation with respect to another to be a constant value, the condition of similarity must be fulfilled. This condition requires that the test preparation must behave as though it were a dilution (or a concentration) of the standard preparation in an inert diluent. Tests for the hypothesis of similarity based on an assumed parametric form for $F(z)$ are available for the usual models employed in bioassays. General methods for investigating similarity, without assuming an exact form for $F(z)$, were considered.

Although all of the standard statistical procedures available for the analysis of bioassays are based on the assumption of similarity, many cases arise in practice in which this convenient assumption is violated. By the usual statistical procedures an assay is considered "invalid" if the hypothesis of similarity is rejected. Efficient alternative procedures for analysis in such cases are not generally available. Methods have also been developed for the analysis and presentation of bioassay results when the assumption of similarity is not necessarily appropriate.

Ilbok Lee, who did his research under the direction of C. P. Cox, was also assisted by K. C. Chanda in the final stages of his research. In the method previously available for the analysis of multinomially grouped polychotomous quantal time response study, which was the subject of Lee's research, the assumption was made that the response is a polynomial function of time and dosage level. The corresponding analysis becomes complicated even if the polynomial is assumed to be of a somewhat simple form (the simplest nonlinear). Based on the usual Markovian structure of the time series data, the new method introduced considerable simplification on the assumption that the transitional probabilities associated with the Markov process were themselves simple functions of the dosage level. This enabled the effects due to time and dosage level on response to be separated, and the resulting estimation procedures were accordingly simpler than those based on the polynomial model.

WASD project on analysis of variance procedures

Research under the contract with the Wright Air Systems Development, AF 33(616) 8269, and the Industrial Science Research Institute was continued with Oscar Kempthorne as chief investigator and George Zyskind as leader. Eugene Dayhoff, associate, and Rodney Basson and Gary S. Spencer, graduate assistants, worked on the project. Robert F. White completed his thesis, which covers a phase of the project.

In the subject of linear models, a general account of the connection among the various augmentation and deletion procedures of parameters and/or observations is being worked out. Thus, the connection between covariance analysis and general elimination procedures is made explicit, as are also the formal relations between covariance models, general tests of subhypotheses and problems of model misspecification. The formal equivalence of covariance analysis and analysis of data with missing observations is also exhibited.

Further work has also been done on complex error structures. In particular, conditions under which best linear estimation of estimable parametric functions are identical with simple least square estimation have been further elaborated upon, with attention given to specialized subsets of estimable functions.

A paper by George Zyskind which explores some properties of the analysis of models by a method termed "residual analysis" will be published in the Journal of the American Statistical Association in December. The paper establishes a precise connection between covariance analysis and residual analysis results. Another paper by Zyskind on the randomization properties and consequences in the generalized balanced incomplete block design will be published in the December issue of the Annals of Mathematical Statistics.

R. F. White, working under Kempthorne in connection with his thesis, an abstract of which appears on page 43 has integrated much of the previously existing theory under the notion of balance and completeness of classification structures. A necessary and sufficient condition has been obtained for a sample mean to have a standard cap-sigma expansion previously introduced by Zyskind. A general theorem has also been derived on expectation of products of sample observations. The examination of single degrees of freedom contrasts has received special emphasis and is thought to hold promise. A variety of specific examples has been worked out including a generalization of the partially balanced incomplete block design, the n-dimentional lattice designs, the Finney partitions, the Modified Latin square, and the incomplete Latin square. Throughout, the study maintains an intimate relation between the method of sampling of the population and the inference to be made by way of a linear model.

An extension of Hooke's bipolykays of degree two was made by Eugene Dayhoff and modifications were introduced to encompass all balanced structures for any number of factors. These newly-defined functions were termed generalized polykays. A proof was established that the generalized polykay corresponding to an admissible mean is identical with the cap-sigma associated with that mean. The problem of obtaining variances and covariances of mean squares is essentially that of finding variances and covariances of the generalized polykays. Results have been obtained for the one-fold nested, two-fold nested, and the two-way crossed classifications. Partial results have been obtained for the three-way crossed classification. Explicit expressions for the generalized polykays, i.e., expressed in terms of generalized symmetric means, have not been obtained,

although the general procedure for obtaining these has been outlined. Additional work is being done relating the crossed and nested structures in order to work out the variances and covariances of the remaining threefold structures.

A report by Sutter, Zyskind and Kempthorne has been published by WASD, "Some Aspects of Constrained Randomizations," and is summarized on page 36.

In addition an extensive survey was made of the literature on nonparametric tests based upon the permutations of the observation. This was done in order to investigate known properties of the randomization test and to find possible alternative tests to the randomization test.

At present investigations are being made by Gary Spencer, directed by George Zyskind, with the aid of the 7074 IBM computer, involving power curves of the possible alternative tests mentioned above by an empirical study based upon samples from known distributions.

A monograph by Rodney Basson and G. Zyskind on missing plot techniques is nearing completion. The essential equivalence of the various methods in use has been demonstrated and a number of theoretical results have been derived. Practical emphasis is placed on the computational merit of the method of covariance which can be uniformly used to obtain exact tests of treatment differences in all the well known designs. Attention has also been given to designs with two and more errors.

In the matter of yield surface exploration, a technical report by Thomas E. Doerfler and Kempthorne, "The Compounding of Gradient Error in the Method of Parallel Tangents," has been revised. The report considers the problem of determining by successive experiments the input conditions which optimize a response. An investigation is made of the role of observational error on the performance of two competing iterative methods. Two path methods are considered, steepest descent and steepest descent partan, in the presence of a gradient error. Certain restrictive assumptions are made, and equations are developed and then averaged over a collection of starting points and errors for each competing method. Numerical results for representative inputs are considered. The results favor the partan method when the gradient error is relatively small. As the error increases, partan is less efficient, but steepest descent exhibits peculiar irregularities. At certain starting values steepest descent actually performs more to the experimenter's advantage. The overall pattern of the results exhibits the superiority of partan in the range of input values considered. Both methods appear to be satisfactory for nearly circular contours regardless of the magnitude of the errors.

Some additional work has been done on the use of randomly chosen fractions of factorial systems, initiated by Shah and Kempthorne in the previous year. In multifactorial systems a practical problem is to study the effects of the several factors with a limited number of experimental runs or trials. The work done examines the use of randomly chosen fractions of the total of possible combinations. This method allows tests of main effects and some low order interactions and may also allow an approximate test for the presence of other interactions.

NASA study of a statistical correction model

A study for the National Aeronautical and Space Administration, under the direction of H. O. Hartley, and administered by the Statistical Laboratory and the Industrial Science Research Institute, was initiated in June 1961, reported on in the 1961-62 Annual Report, and completed during the current year. The problem was to test by numerical evaluation a new statistical regression technique developed by the Statistical Laboratory for reducing the error field of the Joint Numerical Prediction Unit's current baratropic short-range weather forecast.

The final analysis revealed that there was indeed a considerable reduction in the variances of the predicted error field of height by using the predictors X_1, X_2, X_3, X_4 under study. However, it was found that unfortunately the fitted regression coefficients $b_1, \ldots b_4$ did not behave in a stable manner. This precludes the use of the error prediction equation for operational purposes. The stratified analysis revealed that, as expected, the prediction equation had the largest R^2 for the grid points, even those parts of the northern hemisphere in which the meteorological measurements would be expected to be most precise.

ARO project in survey methodology and sampling theory

This study, sponsored by the Army Research Office and administered by the Statistical Laboratory and the Industrial Science Research Institute, on "Unequal Probability Sampling Theory and the Time Series Aspects of Panel Surveys" was completed in June 1963. H. O. Hartley directed the project, and J. N. K. Rao was chief investigator. Carl F. Schach and John E. Graham worked on the study. A complete account of the first phase of the work "Unequal Probability Sampling Theory in Finite Populations" was given in the 1961-62 Annual Report, and the abstract of Schach's thesis for the master's degree which reports his work on the project is given on page 38 of this report.

H. O. Hartley, J. N. K. Rao and John E. Graham continued research on the "Time Series Aspects of Panel Surveys" in 1962-63. The results of this study were incorporated into Graham's Ph.D. thesis. (See thesis abstract, page 42.) When sampling on a number of successive occasions, there are advantages associated with using a partial replacement plan or rotation design for sampling units. Composite estimators which utilize the information collected on prior sampling occasions as well as the current occasion can introduce important efficiency gains over the use of estimators utilizing current records only. The variance of the composite estimators of the current occasion mean, and of the change in means between the previous and current occasions in more complex rotation designs, have been developed. Under the assumption of an exponentially decaying lag correlation between observations on the same unit over time, tables have been prepared which indicate the optimum values of the weight coefficients in the composite estimators, the optimum number of consecutive visits by a unit in the sample, and the optimum number of occasions at which a unit would be dropped out of the sample before returning it for

another cycle, for various values of the correlation coefficient. Comparison with an arithmetically decaying lag correlation indicated that the exponential correlogram was robust to deviations from the assumed model. These results are all summarized in Technical Report 2.1 "Rotation Designs for Sampling on Repeated Occasions" by J. N. K. Rao and John E. Graham.

Generalized composite estimators were introduced to exploit, e.g., the possibly high year-to-year correlations which are encountered for many of the characteristics measured in the Census Bureau's Current Population Survey. Such estimators are a weighted average of two composite estimators, one of which utilizes only the sample data collected on the current occasion, one year ago, two years ago, etc. . . . Calculations have shown that worthwhile efficiency gains were resultant when both the monthly and yearly correlations were high relative to the use of either of the two component parts of the generalized composite estimator.

Multi-component estimators which resemble the generalized composite estimators save for the fact that two weighting coefficients are required for the former compared with three for the latter, were also considered. There is some difficulty however in determining the variance of the multi-component estimator in general rotation designs in that the solution of a higher order difference equation with variable coefficients is required. Hence the study was restricted to the study of a rotation design in which a unit remains in the sample for three consecutive enumeration periods before dropping out of the sample completely. The derivation of the variance formula appropriate to the multi-component estimator of the current occasion mean in such a design situation involved the solution of a second order difference equation.

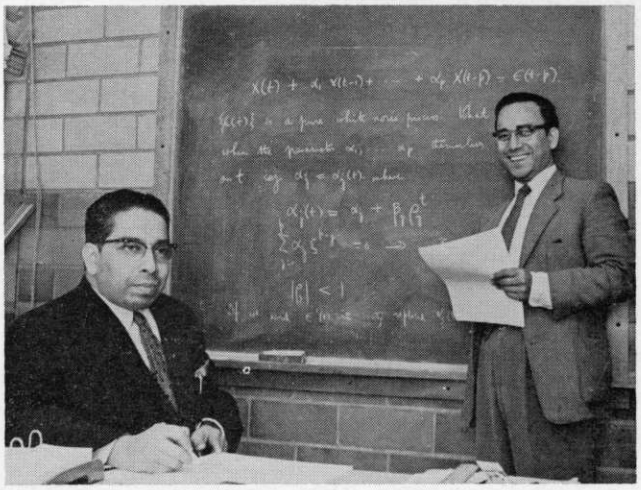
Comparison with a generalized composite estimator in the same design situation indicated that both were about equally efficient as estimators of the current occasion mean. Technical Report 2.2, "Generalized Composite and Multi-Component Estimators in Rotation Sample Designs," by John E. Graham gives an account of these latter studies.

Research in sufficient statistics

The National Science Foundation has awarded a grant (G.P.—1155) to Iowa State for a year's research on "Properties of Sufficient Statistics," and V. S. Huzurbazar, who has been investigating the problem, has extended his stay as Fulbright Visiting Professor. Huzurbazar has found some serious errors and omissions in Pitman's derivation of the general form of distributions admitting sufficient statistics when the range depends on the parameter, and he has been working on constructing a valid proof of the case in question. He has also obtained a simple and elegant proof of Pitman's conjecture in this case that a sufficient statistic exists only when one of the extremities of the range is a monotone decreasing function of the other extremity.

Huzurbazar has also been working on the application of the likelihood ratio criterion for testing of statistical hypotheses to the general class of distributions admitting sufficient statistics, with the object of unifying isolated results found in the current literature for particular members of the class in question.

He also proposes to investigate in general terms some interesting properties of sampling distributions of sufficient statistics.



Dr. V. S. Huzurbazar and Dr. K. C. Chanda, visiting professors, consult on a problem in statistics.

NSF project of the foundation of statistical inference

Work under National Science Foundation Grant G-14831, with Robert J. Buehler as chief investigator, was terminated in June 1963 with the publication of two technical reports which summarize the research on this project. Technical Report 1, "A New Consistency Test for Fiducial Distributions and Some Related Fiducial Theory," by Buehler, begins with a discussion of Lindley's test of consistency of the fiducial method in which the fiducial distribution is treated as a prior distribution. After some other preliminaries, Pitman's fiducial distributions of location and scale parameters are derived in a slightly generalized form. A few new results are given on the uniqueness of fiducial distributions of a single parameter and on the consistency tests of Lindley and Sprott.

The problem of "prediction" of future observations is also considered. Fisher has indicated that when the fiducial distribution of a parameter is known, it can be used, as in the Bayesian theory, to obtain the fiducial distribution of one or more future observations. The resulting "fiducial prediction distribution" will depend on the previous observations but not on the parameter. Requiring that the solution thus obtained shall have a valid frequency interpretation is shown to lead to a new consistency criterion, different from Lindley's, but in the same spirit. It is shown that in case of a single parameter, a single "past" observation, and a single "future" observation, the present consistency test is satisfied whenever Lindley's is. It is not known whether the converse is true. The result is then extended to the case of n previous observations either from a location parameter family or from a location and scale parameter family, and to families obtainable from these by separate transformations of the variates and parameters. In these cases there automatically exist "conditionally sufficient" statistics, and consequently there is no need for an assumption of sufficiency at any point.

Technical Report 2, "Prediction in Location and Scale Parameter Families," by Fred L. Ramsey and R. J. Buehler, is based in part on Ramsey's M. S. thesis. (See page 44.) The problem is considered of predicting a future observation y given previous observations $\Theta_1, \ldots, \Theta_n$, where all observations depend on a location parameter Θ. The product of the likelihood of y by the Pitman fiducial distribution of Θ gives the joint fiducial distribution of y and Θ. An expectation identity is established in which fiducial expectations of certain invariant functions are shown to equal expectations over portions of the sample space in which ancillary statistics take fixed values. It is thereby shown that the mean of the fiducial distribution of y is the minimum mean square error invariant predictor of y. An analogous result for the joint location and scale parameter case involves the joint fiducial distribution of the two parameters and the future observation.

NSF project on methods of finding optimum operating conditions

Oscar Kempthorne, chief investigator, and Robert J. Buehler, B. V. Shah, associate, and Thomas E. Doerfler, graduate assistant, carried on research on a project sponsored by the National Science Foundation (G-19940) and administered by the Statistical Laboratory and the Industrial Science Research Institute.

The general problem under consideration is the comparison of strategies for finding optimum conditions, that is the values of $x_1, x_2, \ldots, x_n$ which optimize (maximize or minimize) an unknown function $f(x_1, x_2, \ldots, x_n)$, which is a constantly recurring problem in science and technology. The present work has been restricted to "path methods", defined as follows: The experimenter proceeds from an initial point in the factor space along a straight line whose direction is determined by the derivatives at the initial point. When the optimum on the line is reached, the first derivatives again are used to determine a new direction. Examples are the "one factor at a time" method [Friedman and Savage (1947)], the method of "steepest ascent" [Box and Wilson (1951)] and the method of parallel tangents (PARTAN) [Shah, Buehler and Kempthorne (1961), Shah, Buehler and Kempthorne (1962)].

Previous work (Buehler, Shah and Kempthorne, Technical Report No. 1, Technical Report No. 3), compared these methods in the absence of error of observation or computing for ellipsoidal contours.

In order to investigate the performance of the partan method when the contours are not ellipsoidal, a two-dimensional example due to Beale has been investigated. In this example

$$\phi(x, y) = \sum_{i=1}^{3} u_i^2$$

$$u_i = c_i - x(1-y^i)$$

$$(c_1, c_2, c_3) = (1.5, 2.25, 2.625).$$

The example is thought to be a particularly awkward one. It was examined on a computer with no errors involved except those due to rounding. It was found necessary to adopt somewhat arbitrary rules on choice of scale units and on a rule for searching for the optimum on a line. Three strategies for indefinitely continued search were examined: (a) continued partan in which one uses the steepest ascent partan algorithm at all times, (b) iterated steepest ascent partan in which one "starts again" after using up the dimensions of the problem and (c) iterated scale invariant partan similarly. It appeared that the third strategy was somewhat inferior.

The application of the partan algorithms to nonlinear least squares fitting problems has also been considered.

NSF project on crosses of genetic material

A study jointly supported by the National Science Foundation (G-16491) and the Agricultural Experiment Station (Project 1478) was completed in June 1963 by O. Kempthorne as principal investigator and Klaus Hinkelmann, associate.

The project is directed to the construction of designs by which the totality of possible single (two-way) crosses and of possible three-way crosses among a set of n lines may be sampled to enable estimation of genetic effects of the lines. The designs are called partial diallel crosses (PDC) and partial triallel crosses (PTC), respectively.

For constructing PDC's a correspondence has been set up between genetic lines and treatments, and between a single cross ixj of the lines i and j and a block containing the two treatments i and j. Since a complete diallel cross corresponds to a BIB design it was to be expected that PDC's correspond to certain types of PBIB designs. PBIB designs with blocks of size two and λ-parameters equal to one or zero have then been used to construct PDC's. In this connection a new class of m-associate class PBIB designs was developed. For the present purpose the designs are more flexible and more useful than most 2-associate class PBIB's.

Using the model $y_{ij} = \mu + g_i + g_j + e_{ij}$, in which g_i is the effect of the i-th line and e_{ij} is the error, and an m-associate class PDC, the solution of the normal equations can be reduced to the solution of $(m+1)$ equations in $(m+1)$ unknowns. The general form of this system is given, together with formulae for the average variance of the comparison $\hat{g}_i - \hat{g}_j$ and the efficiency of a design.

A three-way cross is denoted by (i, j)k, where i and j are the half parents and k is the full-parent. For analyzing a PTC the model $y_{(ij)k} = \mu + h_i + h_j + g_k + e_{(ij)k}$ has been adopted. The aim was to construct PTC's which allow the estimation of all linear contrasts $\hat{h}_i - \hat{h}_{i'}$ and $\hat{g}_k - \hat{g}_{k'}$. Such designs are called connected PTC's. The method used to achieve this was to set up a correspondence between PTC's and generalized PBIB's with blocks of size three. Conceptually, the lines are divided into two groups G_1 and G_2, G_1 containing the lines in their function as half-parents, G_2 containing the lines in their function as full-parents. Using this artificial device and the property of generalized PBIB's that treatment in different groups may occur at a different number of times, it is possible to construct PTC's such that every line occurs r times as full-parent and 2r times as half-parent. PTC's which have a circulant structure have been examined in some detail.

The analysis of connected PTC's is given in general,

reducing the solution of the normal equations to essentially the solution of a system of m+1 equations, where m is the number of associate classes in a PBIB plan used in the construction of the generalized PBIB and hence the PTC.

NSF project in genetic statistics

Linear statistical models where some of the elements of the model may be realistically considered to be random variables from a normal distribution are the models usually termed "components of variance" models. Dewey L. Harris considered the problem of estimation of the elements of these models as a portion of his work under a project sponsored jointly by the National Science Foundation (G-18093) and the Agricultural Experiment Station (Project 1505). In this study it has been shown that three relevant approaches to this type of estimation—(1) conditional expectation of the elements given the observed values, (2) minimum mean squared error estimation of these elements from linear functions of the observed values, and (3) maximum likelihood estimation where the join distribution of the random elements and the observed values is considered—all lead to the same estimators. The seeming inapplicability of two other approaches—(1) the method of fitting constants and (2) maximum likelihood estimation considering the conditional distribution of the observed values given the random elements — was also considered. Experience led to the conclusion that the maximum likelihood approach involving the joint distribution when applied to a specific form gives estimation formulae rather directly in a tractable form. A paper has been written which discusses the theoretical aspects of this problem and gives the derivation of the estimation formulae for several statistical models relevant to a wide class of experimental situations, especially to various genetics situations. The equivalence of these estimation approaches to the considerations involved in the genetic selection index approach is also discussed.

Also considerable time has been spent in preparing for publication a comprehensive paper on the work over a two-year period concerned with the assessment of the influence of errors of parameter estimation upon index selection procedures. Part of this work was reported in the 1961-62 Annual Report.

In addition, other calculations have been carried out on the IBM 7074 for the numerical assessment of the influence of errors of parameter estimation for combinations of population parameters not previously considered. These calculations will probably form the basis for an additional paper on this subject.

NSF project on Monte Carlo studies in genetic selection

The National Science Foundation (G-19218) and the Agricultural Experiment Station (Project 1508) jointly supported studies in genetic selection by O. Kempthorne and L. N. Hazel, who is a member of the Department of Animal Science, and Dewey L. Harris. John L. Gill, Klaus Hinkelmann, and W. G. Hill, associates, and A. W. Qureshi and N. R. Bohidar, postdoctoral associates have worked on the project.

The aim of the present project is to obtain deeper understanding of selection in the polygenic case with particular attention to the role of (a) finiteness of population, (b) linkages, (c) genetic model particularly involving epistacy and (d) natural selection in addition to artificial selection. This will be supplemented by some mathematical investigations aimed at understanding the results of the simulation studies.

A pilot study by J. L. Gill was completed in December. Results are reported in his Ph.D. thesis in Animal Breeding, "Effect of Population Size, Selection Intensity, Linkage and Non-additive Variability upon Genetic Change in Simulated Populations." Populations were generated on the Cyclone computer as determined by the following 4-level factors:

(1) parent population size: (8, 12, 16, 32)
(2) Selection intensity: (1/2, 1/4, 1/6, 1/8)
(3) Environmental variation: $(0, \sigma^2_G/3, \sigma^2_G, 3\sigma^2_G)$
(4) Linkage: (.005, .05, .2, .5)

σ^2_G being the initial genotypic variance and the linkage values being recombination frequencies between the five adjacent loci on each of eight chromosomes.

Nine different types of gene action were considered: (1) additive effects, (2) complete dominance, (3) over dominance, (4) optimum gene number, (5) duplicate factors, (6) complementary factors, (7) only additive-by-additive genetic variation initially, (8) only additive-by-dominance genetic variation initially, (9) only dominance-by-dominance genetic varation initially.

One full replicate of the possible combinations in a 4^4 factorial plan with nine models requires 2,304 trials, and a 1/16 fractional replicate of the basic plan was used. The results of this investigation are being prepared for submission to scientific journals. It is felt that this study has given valuable preliminary information and has suggested that a scientifically reasonable and useful picture of the role of some population parameters in polygenic selection can be developed by the Monte Carlo approach.

Another study is underway, conducted by Esmat Nouri using the program of Gill, to assess the influence of gene model, linkage, selection intensity, environmental variation and the interactions among any of these upon the genetic progress from selection and the amount of fixation when all reproduction is from matings of selected full-sibs. The data generated for this study represents one complete replicate of a 9×4^3 experiment where the factors are models, selection intensity, linkage, and environmental variation, as in the pilot study. The complete factorial allows examination of all effects and interactions.

One of the most difficult aspects of selection theory is the joint role of linkage and epistacy. One should try as far as possible to obtain theoretical results based, of course, on approximations by which to evaluate Monte Carlo results. The theory of the covariance of relatives in infinite random mating populations has been used, for example by Griffing, to develop some theory of polygenic selection. It therefore seemed appropriate to attempt to develop the theory of the covariances of relatives in infinite populations with linkage and epistacy. M. Van Aarde's Ph.D. thesis "Covariances of Relatives in Random Mating Populations with Linkage," (see page 43) dealt with his research in this area.

Investigations have been initiated by K. Hinkelmann to study the effects of truncation selection under somewhat more relaxed conditions than those made by Griffing. Any change in gene frequency is associated with changes in the gene effects and, consequently, with changes in the variance components.

N. R. Bohidar, who earned his doctorate at Iowa State, returned for three month's work on the interaction of artificial and natural selection. The aim of this research is to obtain some knowledge of the interaction of artificial selection for phenotypically superior individuals with natural selection against homozygosity. Individuals are first subjected to natural selection in which the fitness (f) of an individual is given by functions of h, where h is the proportion of heterozygous loci. Artificial selection was then applied to these individuals, the genotypic value being determined by one of four models: additive, dominance, overdominance, and optimum number. Three different values for recombination between successive loci on the 5 chromosomes of 8 loci were incorporated. A constant amount of environmental variance was included. The program was written and the computations were made, but the analysis of the data has not yet been completed.

A program for simulating genetic process on the IBM 7074 has been developed and tested by A. W. Qureshi. The program is generalized in the sense that it can be adapted for a variety of experiments. The input variables consist of the number of generations, identification of "treatments" or "treatment combinations," number of replications, number of male and female parents and offspring allowing for unequal numbers and unequal selection intensities in the two sexes, and the levels of linkage and environmental variance. Basic subroutines for generation of genotypes in the initial population, formation of gamete, and for the evaluation of a genotype according to a certain model of gene action were written in autocoder language.

The procedure followed for simulation is here described briefly. Initial genotypes are generated by turning the electronic switches in the computer "on" or "off" at random. Complete heterozygosity at each locus is easily simulated in this way. A genotypic array is defined as a three-dimensional variable, the subscripts indicating the individual number, parental origin of gamete and chromosome number in gamete. At present the program allows for 4 chromosomes each of 10 loci. Random mating is simulated by sampling parents with replacement. Sampling without replacement can be done by changes in the object program. A separate program to choose parents according to relationship and to control the number of offspring from each individual has been written.

The number of $++$, $+-$, and $--$ phases in the genotype is counted for genotypic evaluation, and a record of fixation at each locus in the population is maintained. Selection of individuals on the basis of their basis of their high phenotype is programmed, but selection on the basis of low deviation from optimum can be made by making a few changes in the object program. The average machine time per generation required to execute the above processes and calculations for an individual with two offspring is .0728 seconds.

W. G. Hill and K. Hinkelmann have been studying models for potential use in simulation studies. Attention has been restricted to gene models in which the gene effects are the same at all loci, and the genotypic value Y can be expressed as a function of the number y_2, of duplex $(++)$ loci, the number, y_1, of simplex loci $(+-)$ and the number, y_0, of nulliplex loci. Gene frequency was taken to be the same at all loci. Studies have been initiated on the inheritance of threshold characters, such as life and death, where the expression of the effect is discrete but the underlying genetic effects are continuous.

AES Research on design of experiments and analysis of data

Project 890 is a continuing project supported by on-campus funds of the Agricultural Experiment Station and directed by O. Kempthorne. Research has been conducted in three areas: (1) the design of general comparative experiments; (2) the logic of experimental inference; and (3) the theory and design of investigations on quantitative inheritance. The general principles of experimentation and inference are useful in most empirical science, and the investigations on quantative inheritance will be useful in breeding research.

With D. F. Cox of the Animal Science Department the application of randomization test procedures to survival data has been examined and illustrated in the paper published in Biometrics, "Randomization Tests for Comparing Survival Curves." This paper is summarized on page 34.

Considerable attention has been devoted to classification models of a particular type arising in studies to evaluate environmental and genetic trends in biological material. A simple case arises where one has year of origin of genetic stock and number of years of relaxation of selection, and one wishes to fit an additive incomplete classification model in which the observation is equal to a year of origin effect plus an effect due to number of years of relaxation of selection plus an effect due to year of testing of material. The structure of the data results in an inability to separate the effects one wishes to estimate. (See abstract of paper by K. Hinkelmann p. 34.) Any such type of experiment requires the introduction in one way or another of some "control" measure. A possible way of introducing a control in genetic studies is by the use of repeat matings, which was advocated by Goodwin, Dickerson and Lamoureux. This type of control has been examined by Giesbrecht and Kempthorne. It is reasonable in such a situation to suppose that there are maternal-year effects and this complicates the issue. A detailed examination of the repeat mating design has been made, and it appears that a reasonable statistical analysis is complex. Basically, the formulation of a control for evaluation of any treatment is a subjective substantive question and underlies the whole of scientific method. Statistical procedures can be applied to assess the numerical efficiency of a control measure but not its scientific validity.

Considerable effort of Kempthorne on Project 890 has been in the direction and work of grant research reported elsewhere.

AES research on mathematical genetics and genetic statistics

Project 1448, which is also a continuing project supported by on-campus funds of the Agricultural Experiment Station, provides for consultation (See page 12) and research in mathematical genetics and genetics statistics under Dewey L. Harris. Harris arranged the Quantitative Genetic Seminars during the year. These are listed on page 49 of this report.

As a major portion of his research activities, Harris has spent considerable time on the development of a set of basic parameters for populations in which all reproduction is by self-fertilization. This mating system is the natural mating system in several monoecious plant species. The parameters are involved in the generation means and in the covariances between relatives for a quantitative genetic trait. These statistics are the major statistical considerations in the analysis of data from such populations.

This parameter system was developed by first developing a rather complicated system of notation for a mathematical description of the probabilistic aspects of such populations. This descriptive system suggested a system of parametrization that seems to have considerable utility relative to previously presented parametrizations. Major weaknesses of the present system involve the assumption of no genetic linkage and the assumption that 100% of the reproduction is by self fertilization. A paper on this work will be presented by Harris at the XIth International Congress of Genetics in The Hague, The Netherlands, September 2-10, 1963.

AES Research on design of surveys and analysis of data

Wayne A. Fuller continued research under Project 1005, a continuing project supported by on-campus funds of the Agricultural Experiment Station. Work on the estimation of growth curves for pork muscles continued in cooperation with the Department of Animal Science, and a manuscript, prepared on the results of the study, has been accepted for publication in the Journal of Animal Science. Work also continued on estimation when errors are correlated, and a paper, "Estimation of Models Containing Correlated Errors," was presented at the meeting of the American Statistical Association in Minneapolis in September 1962. Estimates for optimum rates of nitrogen application over time were derived, using the data from the continuous corn trials of the Iowa State University agronomy farm at Ames. A manuscript has been prepared on the analysis of year-to-year variability in fertilizer response. Research was conducted in areas of double sampling and regression estimation.

Fuller continued his research with C. Philip Baumel of the Department of Economics and Sociology on a study of the management of cooperative elevators. Mathematical models were formulated relating management practices and characteristics of managers of small agricultural business firms to the output of the firm and its profits. Indices of management practices were constructed from information collected in a survey of country elevators, and these data were used in estimation of the management models.

USDC, Bureau of the Census, Research in sample census methods in agriculture

The 1962-63 contractual research project in statistics supported by the USDC, Bureau of the Census, and administered jointly by the Agricultural Experiment Station, has as its objectives (1) the development of questionnaires and procedures which are usable for the purpose of obtaining information quarterly on topics which are included in the agricultural census, and (2) the securing of information on quality of responses and response rates, both for personal and mail interviews repeated quarterly. Wayne A. Fuller served as project leader and Harold Baker, Thomas Jetton and Helen Ayres worked on the study.

A panel operation was conducted in eight states at three periods during the year. The first run was conducted in January when 240 farmers were interviewed in four states and 480 farmers were sent mail questionnaires in eight states. The Survey Group trained and supervised interviewers in Alabama, California, Texas and Iowa. Four other states, Pennsylvania, Montana, Wisconsin and Washington, were included in the study, but in these states only mail and telephone procedures were used. Information was collected on farm transactions for the three preceding months, particularly to get a complete record of livestock and feed grain transactions. An effort was also directed toward the collection of information on farm labor.

On April 1 the second run was conducted. All participants in the survey were sent mail questionnaires. Non-respondents in the four states where interviews had been conducted in January were interviewed again. Telephone interviews were conducted with non-respondents in states where personal interviews were not conducted.

A third run is planned for July 1, 1963.

It is expected that the results of this survey will be used by the USDC, Bureau of the Census, in auxiliary surveys and evaluation procedures associated with the 1964 agricultural census.

SRS, USDA, research

The Laboratory's Production Studies for Corn, Grain Sorghum, and Pasture were continued by agreement with the Research and Development Branch, Statistical Reporting Service, USDA. A joint project of the Statistical Laboratory and the Agricultural Experiment Station, the work of this project 1207 is supported mainly by funds from the Statistical Reporting Service. Principle personnel were Norman V. Strand and Wayne A. Fuller, Harold Baker, Robert Cochran and Kenneth Simons. Improving the accuracy of yield estimates and forecasts serves to improve the efficiency of administration of government programs and to aid farmers and feeders in planning their operations.

Corn samples for 1962 were weighed, dried and reweighed, length and circumference of ears were measured, moisture content was assessed, and a report was prepared. Harold Baker supervised the processing of corn samples, and the resulting measurements were sent to the Agricultural Estimates Division, USDA, Washington, D.C.

Cochran worked on the grain sorghum study. Observations on plant and head characteristics and number of plants were obtained from plots on a sample of 50 farms in nine southwest Iowa counties. Each plot was visited three times between August 7 and October 15. These data were used to provide information on optimum plot size and in the development of methods of making early season yield forecasts.

Work was continued by Kenneth Simons and Harold Baker on the development of objective methods for determining pasture production. Plots were laid out in six of the university pastures to furnish preliminary information of the practicability of the cage technique for estimating pasture production and harvest through grazing. Plots were relocated and clipped at one and two week intervals during the pasture season. Eye estimation and subsampling techniques were introduced into the trials. The 1963 study started in May, with similar goals as last year. Techniques used were also similar, in that the cage technique was continued but with some changes in the methods used.

Carlos Cavallini worked on methods for handling extreme observations (outliers) obtained in the continuing surveys conducted by USDA. He found that in each individual cell the sample may contain extreme observations to the right of the mean. Therefore the available data were divided into two groups by an arbitrary boundary. One group contains all those values less than or equal to the arbitrary boundary. The other group consists of all values greater than the arbitrary boundary. Three possible estimators of the individual cell totals were considered; (i) a ratio estimator; (ii) an estimator based on an additive model; and (iii) an estimator based on a multiplicative model. The results showed that the sum of the estimated mean square errors over states and years for each of the estimators were lower than the sum of the estimated variances of the unbiased estimator for all items.

SCS, USDA Special survey projects involving applied research

Sampling in Soils Surveys, Project 1312, under the cooperation of the Statistical Laboratory and the Soil Conservation Service, USDA, and the Agricultural Experiment Station, was continued under the direction of N. V. Strand.

USDA Branch Field Office

A research field office of the United States Department of Agriculture is located within the Statistical Laboratory. This field office is a part of the Statistical Standards Division of the Statistical Reporting Service established by the federal government. Its general function is to conduct research in connection with sampling problems, survey procedures and theory of mutual interest to the Statistical Standards Division and the Statistical Laboratory. The field office is staffed on a part-time basis by N. V. Strand and M. Kirwin. Its work during 1962-63 was again connected mainly with the Corn, Sorghum and Pasture Study already described.

NSF Grant for expansion, modernization and renovation of facilities

A grant was made in January 1962 by the National Science Foundation, matched by the University, for the expansion, modernization and renovation of the Mathematical Statistics Research Facilities at Iowa State. The area for the Experimental Design-Genetic Statistics Group was completed in March 1962 and reported in the 1961-62 Annual Report. The work assigned to the computer-oriented Numerical Analysis-Programming Group was completed in January 1963, and consists of a conference room and five offices for staff and a room for the secretary for the group. Consultation in statistically-oriented computer techniques, computer programming and processing of data requiring the services of an analyst-programmer have been carried on under the supervision of H. O. Hartley and W. J. Hemmerle. Dale Grosvenor and E. J. Carney, with two graduate assistants and two technicians completed the personnel in the area.

Undergraduate research

Under the National Science Foundation — Undergraduate Science Education Program, the Department of Statistics offered six undergraduates the opportunity to participate in research during the summer of 1962 and the academic year 1962-63. T. A. Bancroft was overall director of the program.

Sharon Fie worked under the supervision of N. V. Strand on the 1962 Park User Survey for the Iowa State Conservation Commission. She worked on estimates for park gate percentages and wrote parts of the report on the survey procedure. Susan Reynolds worked under Strand on the sorghum production study. She made statistical analyses and prepared a part of the report. She also wrote instructions for interviewers and processed data relating to weight and size of sorghum heads.

Edmund Fuller and Nancy Preston worked under W. J. Hemmerle on automatic editing of Agricultural Census questionnaires. They developed computer procedures and techniques to duplicate and expand hand editing methods used in a pilot study relative to retrieving information by mail. Some of the problems encountered in this study were the establishment and application of internal and external consistency checks, tolerance levels, reasonableness criterion and questioning weighting factors. They developed all the necessary logic and completed the computer programming required for the project. The questionnaires were ultimately processed on the IBM 7074. A subsequent study of the machine-edited results showed that the project was highly successful. Alan Johnson assisted Hemmerle in the design and implementation of the Analysis of Variance Compiler-Monitor System. He programmed and tested several complex algorithmic subroutines for array manipulation. James R. Veale, under Hemmerle, programmed and tested double-precision floating point Autocoder subroutines for use with FORTRAN. He also worked on a program to determine a "best" index of percent Republican for the pre-election study relative to predicting the outcome of the 1962 Iowa gubernatorial election, which is reported on page 20.

Postdoctoral Fellows

David Jowett visited the Laboratory from June 1962 to May 1963 as a postdoctoral fellow under the sponsorship of the Rockefeller Foundation. He studied theoretical statistics, design of experiments and genetic statistics at Iowa State under the direction of Oscar Kempthorne.

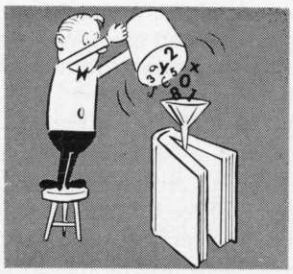
A. W. Qureshi came to Iowa State in October 1962 from Texas A and M where he had earned his doctorate while in the United States on a student's visa from India. He has studied partially as a postdoctoral fellow in design of experiments and genetic statistics and partially as an associate on the Monte Carlo genetic selection project under O. Kempthorne.

Leroy Wolins directed the research of two NSF postdoctoral research participants during the summer of 1963. Dr. Milton D. Jacobson teaches at University of Minnesota, Morris, Minnesota. He is working on the development of statistical techniques applicable to multivariate psychological data where the scales are not comparable.

Dr. Dean Kliewer teaches at Tabor College, Hillsboro, Kansas. Dr. Kliewer's Research is being jointly supervised by Dr. A. C. MacKinney of the Psychology Department and Dr. Wolins. He is applying some of the experimental techniques which Dr. Jacobson is developing, thus evaluating these techniques in terms of how succintly they summarize empirical results.

Wilhelm Seyffert of the Max Planck Institute for plant breeding research at Cologne, Germany, spent the month of April 1962 at the Statistical Laboratory acquainting himself with recent work in biometry and biometrical statistics.

Publications



The members of the Statistical Laboratory have written numerous articles describing their research, many of which have been published in professional journals. Staff members have served in editorial positions or as referees for these journals. Robert J. Buehler has been in charge of abstracting articles for the Journal of the American Statistical Association for the International Journal of Abstracts, Statistical Theory and Methods. He was also a cooperating member of the editorial staff of the Annals of Mathematical Statistics. Buehler, Herbert T. David, H. O. Hartley, B. V. Shah and George Zyskind have served as referees for the Annals of Mathematical Statistics during 1962-63.

Record of Published Research

Publications dealing with the research of the Laboratory are listed in the following pages, with a brief summary of each. Some of these publications are included in the Reprint Series of the Laboratory and are available upon request. These are indicated by an asterisk.

***Harold J. Larson** (U.S. Naval Postgraduate School, Monterey, California)† and T. A. Bancroft, "Sequential Model Building for Prediction in Regression Analysis, I," in Annals of Mathematical Statistics, 34:2, 462-479. June 1963. (Jour. Paper J-4567, Iowa Agri. and Home Ec. Exp. Sta., Project 169.) Reprint Series No. 126 Statistical Laboratory, Iowa State University.

Two different sequential procedures for arriving at a regression prediction equation are examined. In the first procedure the experimenter fits a linear regression model including k independent variables (previously ranked in order of importance). He then tests the coefficient of the "least important" independent variable to be zero. If he accepts this hypothesis he tests that the coefficient of the second least important variable is zero; if he accepts this second hypothesis he tests that the coefficient of the third least important variable is zero, etc., continuing in this manner until he rejects such a hypothesis. He then deletes from his prediction equation all independent variables whose coefficients did not differ significantly from zero and retains the one whose coefficient did differ significantly from zero, as well as all independent variables whose coefficients were not tested. In the second procedure the experimenter starts with a group of r independent variables ($0 \leq r < k$) which he knows are necessary to his prediction equation. He has measurements available on an additional set of k-r "doubtful" independent variables (again previously ranked in order of importance). The experimenter then tests the hypothesis that the coefficient of the "most important" doubtful independent variable is zero. If he rejects this hypothesis he then tests that the coefficient of the second most important independent variable is zero. If he rejects this second hypothesis he tests the coefficient of the third most important independent variable to be zero, etc., continuing in this manner until he accepts such a hypothesis. He then includes in his prediction equation all those doubtful independent variables whose coefficients differed significantly from zero, along with the original set of r he knows to be necessary. The bias and mean square error of the resulting prediction equation is evaluated for both procedures, assuming the population variance to be known. A few tabular values are presented allowing numerical evaluation of the bias function for both procedures, for certain parameter values. A discussion of incompletely specified models is presented and the relationship of this work to other papers in this field is indicated. A comparative discussion of the two procedures and some numerical examples are included.

* Indicates reprints are available.

† When research was carried on by a staff member who has gone on to another position, the institution with which he is now connected is given in parentheses.

*L. E. Everson, **C. S. Shih,** and **Foster B. Cady,** "A Comparison of the Hand and Uniform Methods for the Purity Analysis of Kentucky Bluegrass (Poa Pratensis) Seed," in Proceedings of the International Seed Testing Congress, Lisbon, Spain, 27:2, 476-488. 1962. (Jour. Paper No. J-4279, Iowa Agri. and Home Ec. Exp. Sta. Project No. 1083) Reprint Series No. 118, Statistical Laboratory, Iowa State University.

An inter-laboratory study was conducted to determine the variation of Kentucky bluegrass (Poa Pratensis) test results among laboratories of the International Seed Testing Association (ISTA). The ISTA results, using the hand method, were compared with results using the uniform blowing procedure adopted by the Association of Official Seed Analysts (AOSA). Calculated tolerance limits showed that differences among ISTA test results were excessive and could result in difficult buying and selling among countries. The variation of test results among laboratories was much smaller when the uniform blowing procedure was employed. Information from a previous study was utilized in the design of the present experiment to eliminate the need of a within-laboratory error component for the uniform blowing procedure.

Foster B. Cady, with Herbert T. David (See David, 1963)

K. C. Chanda, "Sampling Properties of Tests of Goodness-of-Fit for Linear Autoregressive Schemes," in Journal of the Royal Statistical Society, Series B, 24:2, 492-510. 1962.

Sampling properties of tests for trend-free linear autoregressive schemes due to Bartlett and Diananda, Quenouille and Whittle have been investigated. The actual procedure of investigation consists of the calculation of the first two moments of the various test-criteria which are χ^2's with k degrees of freedom where k is the number of asymptotically independent linear compounds on which the χ^2's are based. The purpose is to examine the effect of the combination of individual constituents of χ^2 on the validity of the test-criteria when the sample is not large. It has been demonstrated that when the sample is not very large, the contributions due to the terms involving k/n are considerable. In view of the apparent loss of validity thus entailed, a modified test has been suggested to be used when the sample size is small.

K. C. Chanda, "On Bounds of Serial Correlations," in Annals of Mathematical Statistics, 33:4, 1457-1460. December 1962.

The author has proved that the serial correlations r_s of order $s(0 \leq s < n)$ based on n observations from a process $\{X_t\}$ $(t=1, 2, \ldots)$ have attainable upper and lower bounds which are greater than 1 and smaller than -1 respectively. This, in a way, improves the popular convention among mathematical statisticians that $|r_s| \leq 1$. The note ends with two tables exhibiting the extreme values of r_s for various selected values of S and n.

***K. C. Chanda,** "On the Efficiency of Two Sample Mann-Whitney Test for Discrete Populations," in Annals of Mathematical Statistics 34:2 612-617. June 1963. Reprint Series No. 125, Statistical Laboratory, Iowa State University.

Most of the nonparametric tests available in the literature use the assumption that the distributions concerned are absolutely continuous. Under this assumption the power efficiences of these tests for parametric families of distribution are relatively well-known. One particular aspect of the nonparametric test, however, has to the best of the author's knowledge not been discussed so far. This concerns the efficiency of nonparametric tests when the distributions are discrete relative to the most powerful tests available for such distributions when the latter belong to parametric families. The special case of the two-sample Wilcoxon or Mann-Whitney test has been discussed. The power efficiency of this test has been worked out for the class of discrete distributions which are of the exponential type, and finally three examples have been discussed.

Herbert T. David and **Foster B. Cady,** "Calibration," a section in **Methods of Soil Analysis,** a monograph sponsored by the American Society of Agronomy and the American Society for Testing Materials. 1963.

This paper discusses the use of a calibration curve in estimating the concentration of an unknown. Both confidence intervals and point estimates are derived, based on inverse regression. Three examples, typifying situations frequently met in practice, are given. The first two examples involve cases where the form of the calibrating curve is linear, but differ in the assumed structure of the experimental error. The third illustrates an approximate graphical method for the case where the assumption of linearity is not valid.

***Herbert T. David:** "The Sample Mean Among the Moderate Order Statistics," in Annals of Mathematical Statistics, 33:3, 1160-1166. September 1962. Reprint Series No. 111, Statistical Laboratory, Iowa State University.

The first part of this paper is devoted to the derivation of the asymptotic distribution of the sample proportion above a normal sample mean. The second part of the paper, leaning on the first, treats the asymptotic joint distribution of runs of various lengths above the sample mean. The approach here proves general enough to cover asymptotic run distributions arising when the dichotomy criterion is magnitude relative to a sample function other than the sample mean and the population is other than normal.

***Herbert T. David,** "The Sample Mean Among the Extreme Normal Order Statistics," in Annals of Mathematical Statistics 34:1, 33-35. March 1963. Reprint Series No. 122, Statistical Laboratory, Iowa State University.

This paper begins with a discussion of convex spherical polyhedra. The discussion touches on conditions for degeneracy, sub-polyhedra, lunes, and simplices, and terminates with Schlafli's fundamental differential re-

lation for the measure of such a polyhedron. The paper then proceeds to the computation of bounds for the measure of an equilateral spherical simplex. The asymptotic measure of an equilateral spherical simplex then is computed by means of these bounds. (As was brought out in the course of the referring, this asymptotic measure has recently been computed elsewhere, but by a different method.) These results are applied to the computation of the asymptotic value and bounds of the probability that a normal sample mean falls between successive order statistics of fixed order. These computations constitute the asymptotic solution of Youden's "Demon Problem", and yield probabilities of order substantially lower than had previously been hypothesized.

*C. A. O'Flaherty, **Herbert T. David** and D. T. Davidson, "Controlling Data Quality in Soil Stabilization Studies." Proceedings of the American Society for Testing Materials, 62:1297-1312. 1962. Reprint Series No. 120, Statistical Laboratory, Iowa State University.

The purpose of this investigation is to present some methods whereby data quality in soil stabilization investigations can be evaluated and controlled. These include the following: (1) A recommendation regarding the preparation of the soil samples prior to the actual investigation and a method whereby any remaining material nonuniformity can be detected. (2a) A method for evaluating the reliability of a typical soil-additive curing chamber which is most useful for determining the presence of any effects associated with temperature or humidity differentials in various parts of the curing room. Although not considered in this paper, this method may be extended to yield estimates of positional effects which could serve to adjust the data. (2b) A method, which may be applied simultaneously with that just described, for determining if there is any significant operator variability—in the form of a time trend—in a large series of soil-additive strength determinations. (3) A graphical method for detecting outliers in a large series of strength determinations involving four specimens per test condition. This procedure involves the establishment of a constant coefficient of variation, CV, for the entire series of determinations. (4) A method involving the use of control specimens for assessing the hypothesis that all the data in an investigation are subject to the same CV. If such a CV constancy is not indicated, it is shown how the data can be partitioned into constant-CV sets.

A common thread in this study is the extensive use of control specimens. The effective tasks performed by these uniformly spaced specimens seem to be well worth the extra investment involved in their preparation.

W. L. Circeo, D. T. Davidson and **Herbert T. David,** "Strength Maturity Relations of Soil-Cement Mixtures." Highway Research Bulletin 353, 84-97, 1962.

Straight-line relationships can be used to predict the compressive strength of soil-cement at a future time of curing. The accurate prediction of future compressive strengths can be useful in decreasing the cement content of certain soil-cements, and as a basis for periodically increasing maximum soil-cement road and airfield load capacities. The strength-age relationship of a soil-cement can be determined from data obtained from standard laboratory tests. Analysis of 120 sets of data indicates that a semi-logarithmic plot tends, on the whole, to predict future compressive strengths best for all soil types. Compressive strength can be evaluated both graphically and by equation.

Wayne A. Fuller, Ray E. Wakeley, Walter A. Lunden, Pearl Swanson and Elizabeth Willis, "Characteristics of Persons 60 Years of Age and Older in Linn County, Iowa." Special Report No. 33, Iowa Agricultural and Home Ec. Exp. Sta., June 1963. Project No. 1345.

Early in 1957, 695 persons 60 years of age or older were interviewed in Linn County, Iowa. This group was randomly selected from the population of such persons residing in the county. The overall picture for the older persons in Linn County was one of general satisfaction and desire to remain a part of the community. The majority did not picture themselves as facing any insurmountable problems. On the other hand, persons with insufficient income, feelings of loneliness and isolation, and other problems, while a definite minority, form a group sufficiently large in absolute numbers to merit special programs and institutions.

***Dewey L. Harris:** "The Influence of Errors of Parameter Estimation upon Index Selection." A chapter in **Statistical Genetics and Plant Breeding.** Ed. by W. D. Hanson and H. F. Robinson, National Academy of Sciences, National Research Council, Publication 982. 623 pp. + iv. Washington, D.C. 1963. (Jour. Paper No. J-4133, Iowa Agr. & H. Ec. Exp. Sta., Project 1055). Reprint Series No. 128, Statistical Laboratory, Iowa State University.

Index selection is a procedure used in various species for the simultaneous selection for several quantitative genetic traits in a population. The basic theory provides that if certain assumptions are satisfied and if population parameters are known, a linear index can be developed that will lead to maximum attainable genetic improvements when this index is used as the selection criterion in the population. In practice, these population parameters are not known exactly and estimates are used in their place.

This paper describes a study of the influence of the errors in these estimates upon the resulting progress from selection when the calculated index is used as a selection criterion. The estimation situation that is considered is based upon the analysis of variance and covariance of the relevant traits when the sample data are classified into paternal-half-sib groups.

These results indicate that a rather sizeable amount of data is necessary for the development of a calculated index that is likely to lead to progress near the maximum attainable progress.

H. O. Hartley, with J. N. K. Rao and W. G. Cochran (See Rao, December 1962)

*H. O. Hartley, "On Multiple Frame Surveys." Proceedings of the Social Statistics Section of the American Statistical Society, 203-206. 1962 Reprint Series No. 115, Statistical Laboratory, Iowa State University.

The situations encompassed by the term "multiple frame surveys" may be described as follows: In sample survey methodology one often finds that a frame known to cover approximately all units in the population is one in which sampling is costly, while other frames (e.g., special lists of units) are available for cheaper sampling methods. However, the latter usually cover only an unknown or approximately known fraction of the population. The paper develops a general methodology of utilizing any number of such frames without requiring any prior knowledge of the extent of their mutual overlap.

The technique of multiple frame surveys has been used in the past occasionally and under special circumstances. For example, the 1960 Survey of Agriculture of the Bureau of the Census uses two frames: a frame based on the conventional area sampling approach, and a frame of farms conceptually and operationally associated with the A-1 listings of the last (1959) Census of Agriculture. Earlier, the Statistical Laboratory at Iowa State University had used a two frame approach in a small study of "Effects of Industrialization on Farming" which was carried out for the Department of Economics and Sociology at Iowa State University.

*H. O. Hartley: "In Dr. Bayes' Consulting Room." American Statistician, 17:1, 22-24. Feb. 1963. Reprint Series No. 119, Statistical Laboratory, Iowa State University.

This is a feature article dealing with the concept of the subjective belief of the Neo-Baysesian School.

H. O. Hartley, "The Analysis and Design of Experiments with the help of Digital Computers." A section in **Applications of Digital Computers,** edited by Freiberger and Prager, 179-194. Ginn and Co. 1963.

This is predominantly an expository article confined to certain specialized aspects of the analysis of data arising from designed experiments or in operational research of an experimental nature. It is mainly concerned with aspects of numerical analysis arising in the wellknown computations for data from an open linear model. New results present a theory of estimating the parameters in a linear model when they, and their estimate, are constrained to a convex space.

*Klaus Hinkelmann: "A Commonly Occurring Incomplete Multiple Classification Model." Biometrics 19:1, 105-117. March 1963. (Jour. Paper J-4528, Iowa Agri. and Home Ec. Exp. Sta. Project 1478 and NSF G 16491.) Reprint Series No. 123, Statistical Laboratory, Iowa State University.

The linear three-way classification model $y_{ij(k)} = \mu + \Theta_{1i} + \Theta_{2j} + \Theta_{3k} + e_{ijk}$ with the effect-number relationship $k = i + j - l$ has been investigated from the point of view of estimating parametric functions. It is shown that because of the effect-number relationship the rank of the model is reduced by one as compared to the usual three-way classification model. As a consequence no linear effects of the "treatments" are estimable but quadratic effects are estimable. General estimability conditions are given and from these a basis for the space of estimable functions has been derived. The scientific implications with regard to some biological experiments are discussed.

H. H. Van Horn, Jr., S. M. Kassir, N. L. Jacobson, P. R. Shellenberger, **D. K. Hotchkiss,** and R. S. Allen: "Relative Effectiveness of Various Anti-Foaming Agents for Pasture Bloat Therapy," Journal of Animal Science 22:1, 86-92. February 1962. (Jour. Paper No. J-4372, IAHEES. Project No. 1267.)

The experiments reported in this paper were designed to evaluate the effectiveness of certain oils, emulsifiers, and combinations of the two in treatment of pasture bloat and in reducing foaming properties of rumen fluid. The observations, required to reduce the bloat severity by the one bloat score, were transformed using logs to normalize the distribution of the response. The analysis of variance was the principle statistical technique employed. A nonorthogonal comparison of treatment was necessary because of unequal numbers of observations in each treatment group.

*Oscar Kempthorne: "The Role of System of Mating in the Determination of Means, Variances, and Covariances in Genetic Populations." A chapter in **Statistical Genetics and Plant Breeding,** published by the National Academy of Sciences, National Research Council. Publication 982. 623 pp.+ iv. Washington, D. C. 1963. (Journal Paper No. J-4148, Iowa Agri. and Home Ec. Exp. Stat. Project 890) Reprint Series No. 127. Statistical Laboratory, Iowa State University.

The basic idea is that some situations of relevance arise in studies of quantitative genetics by means of certain definite genetic operations such as mating each individual to a random member of the population, or with a full sib or to itself. Each operation is denoted by a symbol such as R for mating to a random individual, F for mating to a full sib, and S for selfing.

It is shown that consideration of the genetic properties of the particular operation leads to a model (a) for the genotypic value of the starting individual or individuals, (b) the genotypic value of individuals arising by repeating the operation an arbitrary number, n, of times and hence to variances and covariances of relatives. An example worked out in some detail is the selfing of autotetraploids.

*D. F. Cox and **Oscar Kempthorne:** "Randomization Tests for Comparing Survival Curves." Biometrics 19:2, 307-317. June 1963. (Jour. Paper No. J-4379, Iowa Agr. and H. Ec. Exp. Sta., Projects 1424 and 890.) Reprint Series No. 130, Statistical Laboratory, Iowa State University.

Randomization tests provide an approach to the problem involved in the comparison of survival curves from biological material. The procedure avoids many of the difficulties that arise in other analyses based on

parametric models. The test is based on all or part of possible random partitions of the experimental units into groups of comparable size as those in the experiment. A criterion of interest is evaluated for each partition and the proportion of the partitions where this value exceeds that in the actual experiment is used to assess the significance of the observed result. The validity of such a test is derived from the physical randomization used to assign the experimental units to the various treatments in the original experiment.

The procedure is taken as the basis for the evaluation of the results from work with swine concerning the genetic effects of paternal irradiation on early survival. The large number of experimental units involved necessitated the use of empirical sampling among all possible partitions of the data. Different procedures that can be applied are described and related to the conduct of the actual experiment.

O. Kempthorne, with Gary Sutter and George Zyskind (See Sutter, January 1963)

Joseph J. Moder (Postdoctoral Fellow, Iowa State University; Georgia Institute of Technology): "A Sequential Search Procedure for Locating a Response Jump." Technometrics 4:4, 610-614. November 1962.

This paper considers the problem of locating the point on some scale at which the probability of occurrence of some attribute goes from zero to one. A "trial-wise" optimal sequential search procedure is given; conditions under which this search procedure is "experiment-wise" optimal are discussed. It is shown that in most cases the intuitive rule of testing the mid-range of the interval of uncertainty is near optimal. A simple improvement of this intuitive procedure can be made by using the average of the mid-range and the median of the a priori distribution of the jump point, defined on the interval of uncertainty. An extension to the multivariate case is also given.

J. N. K. Rao, assistant professor, joined the staff of the Statistical Laboratory and the Department of Statistics in 1960. He conducts research in survey sampling, mathematical statistics and economics, and teaches courses in Design of Surveys.

Born in India, Rao was educated in Indian schools and received his M. A. degree in statistics from the University of Bombay in 1956. He was a research scholar in the Forest Research Institute, Dedra Dun, India, from 1956-58, where he devoted much of his time to research in sampling techniques and published five papers in this field. He came to the United States in 1958, enrolled as a graduate student at Iowa State University, worked as a graduate assistant and later as a research associate, and received his Ph.D. degree in February 1961. He was the recipient of the George W. Snedecor Award in 1961.

Rao was appointed instructor in statistics in March and became assistant professor in July 1961, a post which he held until May 1963 when he returned to India to work with the National Council of Applied Economic Research in New Delhi.

Rao is the author of some 18 publications in the field of sampling and mathematical statistics. He is a member of Phi Kappa Phi, Sigma Xi, the American Statistical Association, the Institute of Mathematical Statistics, and was elected in 1962 as a Fellow of the Royal Statistical Society of England.

***J. N. K. Rao:** "On Sampling with Varying Probabilities and with Replacement in Subsampling Designs". Journal of the Indian Society of Agricultural Statistics, XIII, 1 and 2, 211-217. 1961. Reprint Series No. 114, Statistical Laboratory, Iowa State University.

In sub-sampling, it is usual practice to select the primaries with replacement and with varying probabilities, due to difficulties in the theory of sampling with varying probabilities and without replacement. This leads to three different methods of selecting the secondaries. In method l, if the i^{th} primary is selected λ_i times, $m_i\lambda_i$ secondaries are selected without replacement and with equal probabilities from the i^{th} primary. In method 2, if the i^{th} primary is selected λ_i times, λ_i subsamples each of size m_i are independently drawn of each other from the i^{th} primary with equal probability and without replacement, each subsample being replaced after it is drawn. In method 3, when the i^{th} primary is selected λ_i times, a fixed size of m_i is drawn from the i^{th} primary with equal probability and without replacement and the estimate from the i^{th} primary is weighted by λ_i. It is known that method l has smaller variance than method 2, and method 2 has smaller variance than method 3. But, the three methods have different expected costs, assuming that expected cost in a primary is proportional to expected sample size from the primary. Therefore it would appear more reasonable to compare the efficiency of the three methods for the same expected sample size. Here a comparison of the variances has been made for the same expected sample size but interestingly the conclusions remain the same regarding efficiency.

*J. N. K. Rao, H. O. Hartley and W. G. Cochran: "On a Simple Procedure of Unequal Probability Sampling Without Replacement." Journal of the Royal Statistical Society, Series B, 24:2, 482-492. December 1962. Reprint Series No. 113, Statistical Laboratory, Iowa State University.

To draw a sample of size $n(\geq 2)$, the following sampling procedure is considered: Split the population of N units at random into n groups. Select one unit independently from each group with probabilities proportional to the initial probabilities p_i. It is shown that the variance of the estimator of the population total for this procedure is smaller than the variance of the usual estimator of the population total in sampling with replacement with probabilities p_i. Also the estimated variance is always positive. The optimum choice group size is N/n, so that if N is not a multiple of n(say $N=nr+k$) we choose k groups each of size r+1 and n-k groups each of size r. The results are extended to multistage sampling.

*J. N. K Rao: "On the Estimation of the Relative Efficiency of Sampling Procedures." Annals of the Institute of Statistical Mathematics (Tokyo), XIV, 2, 143-150. 1962. Reprint Series No. 117, Statistical Laboratory, Iowa State University.

There are generally several ways of utilizing supplementary information in sample survey designs. For example, if the variate $\underline{x}$, which is correlated with the characteristic of interest $\underline{y}$, is known for all the units in the population, one can use any one of the following three methods: (1) Select a simple random sample of size $\underline{n}$ and use a ratio estimator utilizing $\underline{x}$. (2) Stratify the population into homogeneous strata according to $\underline{x}$ and draw random samples n_h from the shape such that $\Sigma n_h = n$. (3) Select a sample of size $\underline{n}$ with probabilities proportional to sizes $\underline{x}$ (p. p. s.) with or without replacement.

After choosing one of these methods, from whatever knowledge one has regarding the relative accuracy of these methods (which is rather limited), and drawing the sample accordingly, it would be of interest to find how much gain or loss in efficiency one would obtain if one of the remaining two methods was to be used. In this paper we have considered the three methods in all possible ways and estimated the variance of the estimator in one method from a sample drawn by a different method. The criterion of unbiassedness is used in finding these estimated variances. The precision of these estimated variances is not investigated.

*J. N. K. Rao: "On Three Procedures of Unequal Probability Sampling without Replacement." Journal of the American Statistical Association, 58:301, 202-215. March 1963. Reprint Series No. 124, Statistical Laboratory, Iowa State University.

In this paper three well-known procedures of unequal probability sampling without replacement are considered. Using an asymptotic approach, Hartley and Rao (Annals of Mathematical Statistics 33:350-74, 1962) have obtained compact expressions for variance and estimated variance for one of these procedures. Here are developed formulae for variance and estimated variance for the other two procedures, using their asymptotic approach. It is shown that for large and moderate size populations the three procedures have almost the same variance so that there is little to choose between the three procedures on the basis of variance alone. Therefore, some discussion of other factors for choosing between the three precedures is given.

*J. N. K. Rao and G. Tintner: "The Distribution of the Ratio of the Variances of Variate Differences in the Circular Case." Sankhya, the Indian Journal of Statistics, Series A, 24: 4, 385-394. 1962. (Jour. Paper No. J-4093, Iowa Agr. & H. Ec. Exp. Sta., Project 1200). Reprint Series No. 121, Statistical Laboratory, Iowa State University.

In time series analysis, the variate difference method is used to test the order of the finite difference at which the trend or the systematic part in the time series is approximately eliminated. There is no exact test available in the literature except for the one proposed by Tintner (The Variate Difference Method, Bloomington, Indiana, 1940) based on a method of selection which uses only a portion of the observations. In this paper, the statistic V_{k+1}/V_k is proposed to test that the trend is approximately eliminated at the kth finite differencing of the series where V_k is the variance of the series of the kth differences. Its exact distribution assuming that the observations are $NI(\Theta, \sigma^2)$ is derived under a circular definition of the universe. The lower 5% and 1% points of the statistics V_2/V_1 and V_3/V_2 are tabulated for various values of N, the size of the sample. In practice, one uses the non-circular statistic with these percentage points for the circular statistic as an approximation, especially with long time series.

Gary Sutter, George Zyskind, and O. Kempthorne: "Some Aspects of Constrained Randomization." Interim Technical Report on Air Force Project 33 (616)-8269 with Wright Air Systems Division, published by Aeronautical Research Laboratories, Office of Aerospace Research, U. S. Air Force, Wright-Patterson Air Force Base, Ohio. January, 1963. iv+ 30 pp.

In Chapter I the definition of constrained sets of plans and the problems to be considered are given. In Chapter II a correspondence between constrained sets of plans and RT×RT square arrays and a correspondence between constrained sets of plans and resolvable balanced incomplete block designs are given. These correspondences led to the enumeration of many constrained designs. In chapter III some relationships between the variance of the error sum of squares under complete randomization and under constrained randomization are given. In Chapter IV the power of the randomization test under constrained randomization was considered for nearly linearly, randomly, and semi-randomly distributed basal yields. The empirical results exhibit a tendency of the sensitivity of the test for the detection of treatment differences to increase with decreases in the variance of the error sums of squares of the constrained sets.

Balder von Hohenbalken and **G. Tintner:** "Econometric Models of the OEEC Member Countries, the United States and Canada, and Their Application to Economic Policy," in Weltwirtschaftliches Archiv, Bd. 89, 1, 29-86. September 1962. (Journal Paper J-4145, Iowa Agri. and Home Ec. Exp. Sta., Project 1200.)

This paper has a two-stage objective. Section 1 is rather conservative. The authors estimated some simple macro-economic models on the basis of national account data of a wide variety of economies, including almost all member countries of the OEEC (since called OEDC —Organization for Economic Co-operation and Development), the United States and Canada. Although the statistical methods employed are standard, their assumptions and limitations are pointed out in some detail. Section 2 gives the sources and the tables used, the results of the estimation, the tests for the parameters estimated and their economic interpretation.

Section 3 includes the second objective which was to derive an analytical framework which could make use of parts of the previous findings for purposes of economic policy. By converting one of the original static (and nonlinear) models into a system of simultaneous elasticities, quantitative policy proposals for several countries (where immediate results allowed these calculations) were produced. As a conclusion, possible generalizations of the elasticity approach, using some ideas of Tinbergen's theory of economics policy, were shown.

J. K. Sengupta and **G. Tintner:** "On Some Economic Models of Development Planning," Economia Internazionale, 16:1, 34-50. February 1963. (Journal Paper J-4468, Iowa Agri. and Home Ec. Exp. Sta., Project 1200.)

Recent applications of aggregative growth models to problems of national planning and development have led to interesting generalizations by introducing intersectoral parameters and constraints and optimizing considerations involved in the operational decision-making of a policy-maker. In this article the authors have attempted a comparative evaluation of the aspects, from the viewpoint of operations, of the Indian planning model and the Dutch planning model with a view to suggesting some extensions of the Indian model in investment planning. The additions suggested include the interpretation of the planning model for India in terms of growth development, and the derivations of the costs from this growth development in terms of a nonlinear production function and an effective interpretation of the stochastic program. In the latter case a convex objective function is introduced and it is shown how the planner has to choose the optimum policy variables. One can easily indicate the rules from an alternative optimal decision involving the statistical distribution of the objective function. The method of distribution, based on an interpretation of the stochastic program is illustrated by an empirical application.

J. K. Sengupta and **G. Tintner:** "On Some Aspects of Trend in the Aggregative Models of Economic Growth," in Kylos, Fasc., 1, 47-61. 1963. (Journal Paper J-4415, Iowa Agricultural and Home Economics Experiment Station Project 1200.)

In the framework of an aggregative Domar-type growth model, an attempt has been made here at empirical specification of trends more general than the exponential. The income trend is analyzed in its two aspects, e.g., in terms of overall real income and its components in the form of output-mix measured by the ratio of investment to consumption. On the basis of long-run data on income and its components for the United States and the United Kingdom, the preliminary statistical results show that the hypothesis of a mixed-logistic trend is not refuted by the available data. Since the second-order autoregressive equations did not fare better than the first-order ones in terms of goodness of fit, the latter results alone are reported.

A theoretical analysis has also been made of the conditions under which a mixed-logistic trend can be derived from the aggregative growth models. These conditions are specified as a set of constraints on the production function and the investment demand function underlying the growth model. This has the implication that the overall growth process can be conceived as a succession of stages of development or a change of regimes. The role of stochastic processes in the analysis of long-run economic development needs however a more detailed investigation, before any definite conclusion can be derived about a specific type of trend.

G. Tintner, Charles Millham and J. K. Sengupta: "A Weak Duality Theorem for Stochastic Linear Programming," in Unternehmensforschung Band 7; 1, 1-8. 1963. (Journal Paper J-4470, Iowa Agricultural and Home Economics Experiment Station Project 1200.)

A linear programming problem is said to be stochastic if one or more of the coefficients in the objective function or the system of constraints or resource availabilities is known only by its probability distribution. A distinction is usually made between two related approaches to stochastic linear programming, the active and passive approach respectively. An extension of the duality theorem of non-stochastic or deterministic programming problem has been attempted in this paper in the area of stochastic linear programming in its two approaches. The method of proof is based on the idea that since the parameter space defined by a stochastic linear programme is the topological product of the real line with itself, it forms a first countable topological space. Using a set of distinct and selected points in the parameter space the concepts of feasibility, optimality and duality are extended to stochastic linear programming problems of arbitrary dimensionality. Based on the nonsingular regions of the parameter space of a stochastic linear programming problem the theorem utilizes the conditions of convergence of the sequence of distinct and selected points in the parameter space to a limit point and thereby generalizes the duality theorem in the stochastic case. Furthermore is it shown that the regions of feasibility of the active and passive approaches of stochastic linear programming may be different, so that on this basis it may be possible to establish some inequality relations for the optimal solutions defined for the respective feasible regions.

Tintner, G. with J. N. K. Rao (see Rao, 1962)

Maxine A. Hinton, Hester Chadderdon, Ercel Eppright and **Leroy Wolins:** "Influence on Girl's Eating Behavior." Jour. of Home Econ. 54:10, 842-846, December 1962. (Jour. Paper No. J-4392, Iowa Agri. and Home Ec. Exp. Sta., Project 1428)

Previous research suggests a relationship between diet adequacy and menarche; teen-age girls have poorer diets than other groups.

A sample of 140 girls enrolled in the Boone, Iowa, school system were selected for this study. Approximately equal numbers of 12, 13 and 14 year old girls were selected, and in each age group an attempt was made to obtain as many girls who have not yet menstruated as those who have.

The dependent variables were derived from diaries of food consumption which the girls kept for one week during the summer and for three days during the winter. Many other psychological, physiological, and demographic "measures" were obtained. The only clear relationships with the dependent variables occurred with the psychological variables and indices of parental conflict. The results suggest that girls who have poor social relationships also have poor relationships with their parents and have relatively poor eating habits and diets.

R. M. Pitkin, M.D., and W. B. Goddard, M.D., F. A.C.O.G., with statistical analysis by **James R. Zweifel:** "Paracervical and Uterosacral Block in Obstetrics—A Controlled, Double-Blind Study." Obstetrics-Gynecology 21:6, 737-744. June 1963.

For a report of Zweifel's work, see the abstract of his thesis on page 40 of this report.

George Zyskind, with Gary Sutter and O. Kempthorne (See Sutter, January 1963.)

Book Reviews

Guide to Tables in Mathematical Statistics, J. A. Greenwood and H. O. Hartley. Reviewed in Technometrics 4:2, 276, June 1963, by D. B. Owen, Sandia Corporation.
Also reviewed in Annals of Mathematical Statistics 34:1, 353, March 1963, by H. A. David, Virginia Poly. Inst.

Some Papers Concerning the Teaching of Statistics, T. A. Bancroft and D. V. Huntsberger. Reviewed in Biometrics 19:1, 192, March 1963, by B. G. Greenberg, University of N. Carolina.

Mathematical Statistics, John E. Freund, Reviewed in Journal of the American Statistical Association 57:-299, 717, September 1962, by T. A. Bancroft, Iowa State University.

Methods of Least Squares and Principles of the Theory of Observation, Yu V. Linnik, Reviewed in Journal of the American Statistical Association 57:299, 719, September 1962, by O. Kempthorne, Iowa State University.

Small Particle Statistics, An Account of Statistical Methods for the Investigation of Finely Divided Materials. G. Herdan. Reviewed in Food Technology 16:9, 94, September 1962, by O. Kempthorne, Iowa State University.

ABSTRACTS OF THESES

Edwin Joseph Hughes: "Maximum Likelihood Estimation of Distribution Parameters from Incomplete Data." Ph.D. thesis. Iowa State University Library. July 1962.

If variate values occurring in a population are not observable in a sample, or if individual variate values occurring in a population are observable only as grouped together in a sample, then the data of the sample are called, respectively, incomplete truncated data or incomplete censored data. In this work, two procedures are described for the estimation of distribution parameters of a wide range of distributions, both discrete and continuous, from incomplete data of a very general type, requiring in most cases only tables of distribution areas and ordinates.

Both procedures described are for maximum likelihood estimation. The first procedure arises from the observation that the only complicating factor introduced in maximum likelihood estimating equations by the incompleteness of the data—the derivative with respect to the parameter of the cumulative distribution function—can, for many common distributions, both discrete and continuous, be expressed in terms of the distribution area and ordinate functions. The second estimation procedure is that described earlier by Hartley for discrete distributions and extended here to continuous distributions. This procedure arises from the observation that, by introducing "pseudo-frequencies" for variate values in the subsets of truncation and censorship (under a system of "proportional allocation"), the maximum likelihood estimating equations from incomplete data can be put into the form of maximum likelihood equations from complete data. Pseudo-frequencies are calculated from proportional allocation equations requiring the use only of tables of distribution areas and ordinates.

Incomplete data from discrete distributions may involve an arbitrary subset of truncation and a finite number of arbitrary subsets of censorship, all subsets non-overlapping. Incomplete data from continuous distributions may be described similarly for subsets consisting of finite sums of non-degenerate, non-overlapping intervals.

The two estimation procedures are applied to the Poisson, binomial, negative binomial, geometric, normal, gamma, exponential, and uniform distributions.

A theorem is proved establishing sufficient conditions for the convergence of Hartley's iterative estimation procedure.

Carl F. Schach: "Sampling Problems Involving Rotational Survey Designs and Unequal Probability." M. S. thesis. Iowa State University Library. July 1962.

This study is concerned with two particular aspects of sample survey methodology. The first aspect is concerned with estimation of the relative efficiency of three sampling procedures—simple random sampling, stratified random sampling, and unequal probability sampling. Some of the methods available are applied to data obtained from an Iowa high school survey and

from an Iowa farm boy survey. The use of these methods may be of aid to a research worker who has selected a sample by one procedure and wishes to know, for future use, what the estimated variance of the estimator would have been if he had employed a different sampling method.

The second method is concerned with sampling of a population on repeated occasions in time by rotational survey design. The mean of a characteristic was estimated by a composite estimator for the units in the population at one particular time. A problem of particular importance is the determination of optimum weights for the two parts of the estimator. Optimum values for these weights and length of time a unit remains in the sample are determined for certain cases.

Gary J. Sutter: "Some Aspects of Constrained Randomization." M.S. thesis Iowa State University Library. July 1962.

The idea of constrained randomization is due to W. J. Youden. Some of the plans possible under complete randomization for the arrangement of T treatments on RT experimental units have a systematic ordering of the treatments and therefore may be considered undesirable. The basic idea is therefore to consider the existence and utility of subsets of the possible plans which have particular properties. The requirement is imposed that the subset of plans is to be unbiased, and any such subset is called a constrained design.

Correspondences are established of constrained sets of plans to RT×RT square arrays and to resolvable balanced incomplete block designs. An extensive literature survey of methods of constructing resolvable balanced incomplete block designs is summarized, leading to knowledge of the existence of many constrained designs.

To evaluate some of the effects of constrained randomization, three constrained sets for a number of given sets of design parameters were constructed with widely differing variances of the error sum of squares under linear basal yields. The power of the randomization test was determined for each of these three constrained sets under several circumstances. The results indicate that with nearly linear basal yields a constrained set, for which the variance of the error sum of squares is less than that under complete randomization, has greater power than complete randomization.

Bert Reed Webster: "The Distribution of Generalized Mean Square Ratios in Nested Classifications of Data." M.S. thesis. Iowa State University Library. July 1962.

In this thesis the exact probability distribution was derived for certain statistics derived from the three mean squares in a three stage hierarchical analysis of variance. The statistic defined consisted of the ratio of linear functions of the mean squares. The well known facts involving the distribution of the mean squares were used in obtaining the probability distribution function. This consisted of defining the probability function as a triple integral and then evaluating this integral. This was accomplished by setting limits on the integral corresponding to the primary stage of sampling. A second equation for the probability function was obtained by interchanging the order of integration by setting limits on the integral corresponding to the secondary stage of sampling. In both cases, the probability distribution function was obtained as a double binomial sum.

The next section of this thesis consisted of the adaption of the general theory to measures of heritability. Six estimates for heritability were defined. Using four of these estimates and the two equations obtained for the probability distribution functions, several combinations of data were run on the IMB 650 electronic computer and the values for probability were obtained and recorded.

A comparison of the results obtained showed that an increase in the number of primary units had a greater effect on increasing the precision of the distribution than did an increase of either the secondary or tertiary units. In terms of the genetical illustration, this would mean that an increase in the number of sires, rather than an increase in either the number of dams per sires, or the number of progeny per dam, would be more effective in increasing the precision of these particular measures.

By use of the formulas used in this thesis, the means of the remaining distributions could be obtained and further comparisons could be made relative to these means.

Eugene Bertrand Cohen: "Statistical Studies in the Epidemiology of Toxoplasmosis." M.S. thesis. Iowa State University Library. August 1962.

The need for interrelating the concepts and methods of biology with concepts and methods of statistics and mathematics is continually increasing and, for example, many problems in epidemiology have strongly biometrical aspects.

In this thesis the theory of 2×2×2 contingency tables has been applied to an epidemiological study on Toxoplasmosis carried out by Dr. F. H. Top and Dr. W. F. McCulloch at the State University of Iowa. Epidemiology is defined and the general areas of the study are described. Toxoplasmosis is discussed and the skin- and dye-test methods of detecting either active or previous infection with Toxoplasmosis are explained. In the present study, statistical methods for examining the relationships between skin-test or dye-test positives and the epidemiological factors (i.e., degree of animal contact, residence background, and age group) are given.

The objectives of the study were to examine:

1. the hypothesis that the test result (positive or negative) was independent of the other factors under study;
2. the conditional independence between the factors under study and the test result for each student;
3. the hypothesis of no three-factor interaction.

Likelihood-ratio methods are used to verify the appropriate chi-square computations. Bartlett's cubic equation for obtaining expected numbers is shown to result in only one admissible real root since any other gives rise to negative values for some of the expected frequencies.

The results and interpretations of the analysis were discussed, the main result being that moderate and marked contact with swine and horses was significantly associated with test positivity among medical (and total) students but not with veterinary students.

Ahmed El Mawaziny: "Some Fiducial Probabiliy Problems Involving Multiple Ancillary Statistics." M. S. Thesis. Iowa State University Library. August 1962.

This thesis is concerned with the uniqueness of the fiducial distribution of a single parameter Θ. When t is a sufficient statistic with cumulative distribution $F(t;\Theta)$, then the cumulative fiducial distribution of Θ given t is $1-F(t;\Theta)$, provided this quantity has the properties of a distribution ("Condition B" of Lindley). If the statistic (t, a) is sufficient and a is ancillary (that is, the distribution of a is independent of Θ), the cumulative fiducial distribution of Θ given (t, a) is $1-F(t \mid a; \Theta)$. A problem of uniqueness arises when a and a′ are nonequivalent ancillaries for which (t, a) and (t', a') are both sufficient. Several examples of this kind are considered. In each case it is found that "Condition B" is violated, so that no fiducial distribution can be obtained. If the requirement of sufficiency is ignored, then nonunique results can be obtained. This is illustrated by the case of (x, y) bivariate normal with zero means, unit variances, and covariance ρ. The nonsufficient pairs (t, x) and (t, y) (where t is the maximum likelihood estimator of ρ) are shown to give two distinct confidence interval solutions.

James Roger Zweifel: "Statistical Analysis of Experiments on the Efficacy of Anesthetic Drugs Used in Childbirth." M.S. thesis. Iowa State University Library. August 1962.

The main statistical problem considered in this thesis was the estimation of the parameters of the normal distribution from a multi-censored sample. The principal goal was to provide estimates when the censoring points occur at random and are in nowise determined by the experimenter. The maximum likelihood estimators which were derived may, however, be used for a wide range of situations in which censoring occurs. They are valid for all cases of single or double censoring in which the censoring point (or points) is determined by the investigator. They are valid in all instances where censoring occurs at random and is due to some cause acting independently. They are equally valid when the sampling units appear in groups for which only a single observation can be made, i.e. when the occurrence of some event within the group precludes any further observation of the remaining units.

The estimates can be obtained very easily on a desk calculator with the aid of a single table which has been included. An alternative solution is described which may be used when a high speed computer is available.

The asymptotic variances have been found and an empirical study was completed in order to study the behavior of the estimators in small samples.

The methods have been applied using data obtained in an investigation of anesthetic drugs used in childbirth, carried out by Dr. Roy M. Pitkin at the State University of Iowa.

Carol Bates Edwards: "Multivariate and Multiple Poisson Distributions." Ph.D. thesis. Iowa State University Library. November 1962.

Multivariate and multiple Poisson distributions are derived as limiting distributions in several situations. Supplementary results which either characterize these distributions or are required by the derivation procedures are also presented.

The first situation pertains to general sequences of distributions converging to the multivariate Poisson and investigates their properties. It is shown by mean of a special case of a theorem by Chernoff, proved here independently, that (1) sub-sums of random variables whose joint distribution converges to a multivariate independent Poisson converge in distribution to a multivariate Poisson, and (2) sub-sums of random variables whose joint distribution converges to a multivariate Poisson converge in distribution to a multivariate multiple Poisson.

It is also shown that the multivariate Poisson distribution satisfies the multivariate Carleman uniqueness criterion, and thus, upon application of a theorem by Haviland, that a sequence of random variables whose moments converge to those of a multivariate Poisson, converges in distribution to that same multivariate Poisson. An example is constructed of a non-independent bivariate distribution with Poisson marginals such that $E[X_1X_2]=E[X_1]E[X_2]$, $E[X^2_1X_2]=E[X^2_1]\cdot E[X_2]$, $E[X_1X^2_2]=E[X_1]E[X^2_2]$.

The second situation pertains to the derivation of the multivariate and multivariate multiple Poisson distributions as limit distributions of numbers of runs (as usually defined) and other configurations on a circle. Under the condition of constant expectations, it is shown, by means of a generalization of a theorem of Fréchet, that the number of black and white ball runs on a circle is asymptotically multivariate independent Poisson. A corollary to this shows that the sum of numbers of runs of the same length is asymptotically univariate Poisson, a result analogous to that of von Mines. A second corollary states that the bivariate distribution of number of runs of black balls of length $k(n)$ and number of runs of both black and white balls of length $k(n)$ is a bivariate Poisson which assigns positive mass to only one-half the positive quadrant.

When more general configurations are considered in the bivariate case, a correlated Poisson results.

When configuration lengths are independent of n, the asymptotic multivariate distribution of these configurations involves independent, equivalent or constant (zero) Poisson random variables.

A theorem is also proved, which, using the property of infinite divisibility, states that all configurations whose counts are essentially asymptotically marginally Poisson, must be multivariate Poisson. This shows that there cannot essentially exist a set of configurations whose counts are marginally Poisson but not multivariate Poisson.

The third situation considered is that of the derivation of a multiple Poisson from suitable Markov process postulates.

Raymond R. Hocking: "Mathematical Programming in Statistical Estimation Theory." Ph.D. thesis. Iowa State University Library. November 1962.

In this study there are two main problems. The first of these is to find a computationally simple solution to the convex programming problem: that is, the problem of determining the minimum of a convex objective function subject to the restriction that the solution is constrained to lie in a convex region defined as the intersection of regions bounded by convex functions. The existing methods for solving this problem are computationally quite difficult, primarily because they require that the trial solutions, at each stage of an iterative procedure, remain inside the convex region. The algorithm proposed here is quite different in that the trial solutions approach the minimum from outside the convex region, hence avoiding the cumbersome techniques for correcting back to the convex region. In addition, this procedure has the advantage that, with a few modifications, the existing computer codes for the modified simplex method may be used.

The second problem with which this study is concerned is that of estimating the parameters in a linear regression model when these parameters are constrained to lie in a convex region of the parameter space. Computationally, this problem can be formulated as a problem in convex programming, in fact, the constrained minimization of a quadratic function. Certain properties of this quadratic programming problem suggest an alternate algorithm which, although computationally more difficult than the general convex programming algorithm, has certain desirable features, and in addition it can be applied to a slightly larger class of restraining functions. Some of the properties of the estimators are investigated with particular attention given to asymptotic properties. For certain special cases, the exact probability distribution of the estimators is derived.

Darrol W. Heggen: "Evaluation of the Toxic Effects of Pesticides on Human Cells in Vitro." M.S. thesis. Iowa State University Library. November 1962.

This thesis is concerned with the examination of methods of analysis which are being developed at the Institute of Agricultural Medicine, State University of Iowa, for evaluating the toxicities of insecticides commonly used by Iowa farmers in their spraying operations. The laboratory studies were conducted using human cells cultivated in vitro, which were exposed to varying doses of experimental insecticides. At specified intervals, the responses of the cells to the treatments were measured by either spectrophotometric or cell counting techniques. For the former of these, it was then desired to estimate, and place limits upon, the dose necessary to inhibit 50% of the growth increase from the beginning of the experiment to the end. This involved certain linearizing transformations, the application of Fieller's theorem, and in a few cases the utilization of the Behrens-Fisher distribution to set approximate limits. The data obtained from experiments involving cell counts were examined using analysis of variance techniques. This also involved investigating the mechanisms of growth of variously treated cells by the fitting of orthogonal polynomials.

Philip Mason Mills: "Response Errors Associated with Reported Farm Acreages." M.S. thesis. Iowa State University Library. November 1962.

Response errors in farm acreage associated with four mail questionnaries and a screening type of interview questionnaire were estimated and compared. The effect of such response errors on the estimated population mean and variances was examined. Estimates of mean difference and mean square error were obtained for the mail questionnaires and the screening survey. In one region, Texas, 289 usuable mail questionnaires were returned, from which 129 were interviewed. The interview and subsequent review were used to establish a standard which was used as the "true value." The difference between the mail and the standard information constituted the response error. Information from people not interviewed was utilized through a regression estimation procedure. The analysis did not indicate that one mail questionnaire was significantly better than another.

For the regression estimates the mean difference between the mail and standard information was 1.70 acres and for the screening survey -27.20 acres. Although the latter difference was large, it did not exceed twice the standard error. The mean square error for the mail questionnaire was estimated to be 6,293 acres squared, and for the screening survey it was 182,-361 acres squared. The mail questionnaire did a significantly better job of collecting the information than the screening survey, which was attributed to the form of questionnaires used and not to the type of survey.

The mean square errors were significantly smaller in the groups which had exhibited little change since the previous Census. Differences in response errors were associated with certain types of farm operations.

Brenda Jane Morrison: "Stochastic Linear Programming with Reference to Economic Models." M.S. thesis. Iowa State University Library. November 1962.

The empirical probability approach to stochastic linear programming was used to obtain the solution to a dynamic linear programming problem in economic planning. The object of the problem was to find what was the optimum allocation of investment in a two sector model of the economy of India, in order to obtain the highest national income. The two sectors were industries which produced capital goods and industries which produced goods for consumption. The coefficients of the left-hand side of the constraint equations were considered to be independent stochastic variables each with a Gamma distribution.

The results indicated that although an allocation of one third of the available capital to capital goods industries and two thirds to consumer goods industries produced the highest mathematical expectation of national income, the reverse allocation of two thirds and one third gives a lower mathematical expectation but the lower 5% level of the distribution is appreciably higher. Hence, in the former case, the returns are greater but the risk is higher; and in the latter, the returns are lower but the risk is lower too.

It must be concluded that the probability approach to stochastic linear programming is involved and laborious if the problem is of any size at all. Deriving the distribution of the objective function, if it is more than the simple sum of the activity levels, is almost impossible.

Kenneth Kay Simons: "Statistical Procedures for Estimating Dry Matter Consumed by Grazing Animals." M.S. thesis. Iowa State University Library. November 1962.

This study investigated statistical methods of estimating the quantity of dry matter consumed by grazing animals. An associated determination of interest was the growth rate of pastures. The data from an experimental survey were utilized to estimate the size, combination and allocation of plots yielding minimum variance for a fixed cost. Observational time periods of one and two weeks were considered.

There was some evidence to suggest that growth and grazing regress on the amount of dry matter outfield at the beginning of the period, and on the difference between dry matter outfield at the beginning and at the close of the time period.

Estimators were constructed for grazing and growth when the estimation interval covers several observational time periods. These estimators utilized the grazing estimates, growth estimates, and yields from unprotected plots at the beginning and end of the estimation interval. An optimum sampling plan for the growth estimator was determined for the special case in which the number of unprotected plots taken at the beginning of the interval is equal to the number of unprotected plots taken at the end of the interval.

Ilbok Lee: "Bioassay with Quantal Response Observed at Different Times." Ph.D. thesis. Iowa State University Library. February 1963.

Suppose that doses $z_1, \ldots, z_m$ are administered to numbers $n_1, \ldots, n_m$ subjects respectively and that the times of observation $t_1 \ldots t_k$ are pre-specified for all dose levels. At the end of each one of these time intervals every individual subject is examined and classified, according to its response into one of the mutually exclusive classes 'dead', 'moribund', and 'alive'. At the conclusion of an experiment, there will be for each subject, an individual history of its survival time, the time for which it stayed in the state of being moribund, and the time at which it was observed as dead.

The present study introduces a procedure for obtaining efficient estimators of assay parameters by utilizing all the information on each individual subject. This technique is developed by using a time homogeneous Markov process to formulate a relevant mathematical model. The interrelations between the stochastic process and the logistic function models are obtained.

For simplicity the trichotomous quantal response is primarily considered; the principles, however, can be readily extended to the case of polychotomous quantal response. It is considered that the procedures here described and, in particular, the introduction of the Markovian principles, should be applicable to a wide variety of bioassay situations.

John Elwood Graham: "Rotation Designs for Sampling on Successive Occasions. Ph.D. thesis. Iowa State University Library. May 1963.

If precisely n units are selected from a finite population of N units on each of a number of successive occasions, a survey design will (a) number of the N units at random and (b) specify in advance which of the unit numbers will be in the sample on each of the occasions. Such a design can therefore be regarded as a *fixed* specification of (b) (the "rotation plan") which is applied to the actual units in the population by arranging them in one out of the N′ possible permutations in (a).

Specifically, the rotation plan is formulated here as follows: n_2 units remain in the sample for $r \geqq 2$ consecutive occasions, where $n = rn_2$ and n is the constant sample size on any occasion. These units drop out of the sample for the next $m \geqq r$ consecutive occasions, return to the sample for another r consecutive occasions, and so on without limit. The unbiased composite estimator of the current occasion mean, $\bar{X}_0$, is

$$\bar{x}'_0 = Q(\bar{x}'_{-1} + \bar{x}_{0,-1} - x_{-1,0} + (1-Q)\bar{x}_0,$$

where $\bar{x}_0$ is the mean of all n sample units on the current occasion, $\bar{x}_{0,-1}$ and $\bar{x}_{-1,0}$ are respectively, the means on the current and previous occasions of the $(r-1)n_2$ units common to both of these occasions, $\bar{x}'_{-1}$ is the composite estimator for the previous occasion and $0 \leqq Q \leqq 1$. The composite estimator of the change between the previous and current occasions is $d'_0 = \bar{x}'_0 - \bar{x}'_{-1}$. The general variance formulae are derived for $\bar{x}'_0$ and d'_0, and the optimum values of r and Q are determined under specified correlation models for the character of interest over time. The theory of sampling on successive occasions is then extended to include two-stage sampling designs where either primary or secondary sampling units are rotated.

The variance of the composite estimator $\bar{x}'_0$ is derived in the Current Population Survey design situation wherein a unit remains in the sample for four consecutive months, drops out of the sample for the next eight months, returns for another four months, and then drops out once more but does not return again. It is assumed that the correlation between observations on the same unit separated by $12j+i$ months is of the form $\rho^i{}_1 \rho_2{}^j$ (i being a month index and j a year index). A "generalized composite estimator" which explicitly exploits the possibly high year-to-year correlations under the foregoing correlation model is introduced and its variances given.

A so-called "multi-component estimator" is employed in the special case of a rotation design with $r=3$ and $m=\infty$. The solution of a second order difference equation is involved in deriving its variance. This estimator is compared in efficiency with both a composite estimator and a generalized composite estimator in the same design situation.

William J. Hemmerle: "The Logical Structure of Analysis of Variance and Its Implementation on Digital Computers." Ph.D. thesis. Iowa State University Library. May 1963.

A language or notational scheme is formulated to provide a basis for the logic of analyzing statistical statements and specifications to determine computations to be performed. Care has been taken to make this notation reasonably consistent with the usual symbolic description of an analysis of variance problem. The analysis of variance model may be written utilizing the alphabet in a relatively natural manner in the assignment of effect symbols and subscripts. Further specifications would then employ the symbols appearing in the model. Algorithmic computational procedures are developed tailored to the notation used. Relative to these procedures, techniques for incrementation in n-dimensional array summation are presented, as well as automatic determination of factorial terms contained in a nonfactorial term.

The capability of including covariates and statistical transformations of variate or covariates in the problem specifications and the logic of their implementations is discussed. Methodology of handling the case of unequal cell frequencies, both by exact and approximate methods, is also described. A consequence of the notation developed is a means of denoting structures that is useful in their enumeration. Considering some ordering restrictions of effect symbols imposed by this notation, some interesting relationships between structures and associated models are derived.

The concepts and methodology presented here formed the foundation for the design of the analysis of variance compiler-monitor system currently in use on the IBM 7074 at Iowa State University. With the exception of certain extensions and modifications for the purpose of presentation, the logic described has been verified in the course of making the system operational. A few of the pertinent logical diagrams are presented in this thesis to illustrate some of the procedures employed.

Izak Mauritz van Aarde: "Covariances of Relatives in Random Mating Populations with Linkage." Ph.D. thesis. Iowa State University Library. May 1963.

This study is directed to developing mathematical expressions for covariances of relatives in a Mendelian population of diploid organisms when loci are linked. The population is assumed to be at equilibrium under random mating.

With n loci, an individual may produce 2^n different kinds of gametes. The 2^n kinds consist of 2^{n-1} pairs, such that the two members of any pair occur with the same frequency. It follows that the relative frequencies of the 2^n kinds of gametes may be expressed in terms of $(2^{n-1}-1)$ parameters. A useful parametrization is obtained by relating the relative frequencies to the yields of a two-level factorial system, and by using its conventional parametrization.

A system of operators is developed and applied to obtain a breakdown of the genotypic value of a random individual in terms of the population mean and (4^n-1) uncorrelated random variables. It is found that the covariance of any pair of non-inbred relatives can be expressed as a linear combination of the variances of these (4^n-1) random variables.

An initial part of the study is confined to consideration of the covariances of relatives which derive from one common ancestor through single lines of descent. It is shown that the system of operators leads to the required covariances. This is accomplished by identifying a particular class of operators which exist within the general system, and which were termed "random mating operators."

A subsequent part of the study is directed toward finding the covariance of the genotypic values of any pair of non-inbred relatives in general. The results of this part of the study are more general than the results which are obtained by application of the random mating operators.

The study is concluded with an investigation of a class of problems which arise from the simultaneous consideration of quantitative and qualitative traits. When metric traits are studied in the presence of so-called "marker genes," the classical and biometrical methods may be integrated. Statistics are developed which may be useful in testing for linkage association between marker genes and genes which control the variation of some quantitative characteristic. These statistics enable one to obtain information on the extent to which the genes of a particular chromosome segment control the variation of a quantitative characteristic both with regard to additive gene effects and dominance deviations.

Robert F. White: "Randomization Analysis of the General Experiment." Ph.D. thesis. Iowa State University Library. May 1963.

From rigorous definitions concerning the nature of population and samples structures, general results are obtained which apply to obtaining expected values of sample quadratic forms in a wide variety of sampling situations. Starting with the concept of a "factor" of a structure, the study proceeds, through the notion of the "nesting" relation, to define structures which are "balanced," "complete," and "balanced complete." The study shows that the nesting relation is the only fundamental property of population structures.

The populations studied are all those with a balanced complete structure. The method prescribed by the study for constructing random samples encompasses a general procedure for randomizing experiments. Hence the results of the study apparently apply to all experiments.

The two underlying assumptions in the results obtained are that the population is balanced complete and that the sampling method is of the prescribed form, called "symmetric uniform probability." A Boolean algebra of point sets is used, rather than the indicator variable and index algebra of prior workers.

The primary result, called the "fundamental second moment theorem," makes no assumption about the completeness or balance of the sample structure. It gives the variance-covariance matrix of the sample, in very simple form, in terms of quantities called "cap sigmas," which are functions of the population structure and are independent of the sample structure.

A second result concerns the expected values of those quadratic forms which are the squares of sample "partial means." This class is important in the analysis of variance. It is shown that in order for such an expected value to have the simple form called the "standard sigma expansion," it is necessary and sufficient that the sample structure be balanced. In particular, it is neither necessary nor sufficient that the sample structure be complete.

These results are applied to obtain the expected mean squares in some rather complex experimental designs, of which the most generally interesting is the "general incomplete block" design. It is to be noted that these expected values are obtained without the use of an assumed linear model. Their interesting correspondence to the results which would be obtained from such a model is noted.

The expectation which is involved in the above stated results is "unconditional," and corresponds to the ordinary notion of a "fully random" linear model. This is sufficient for the general "expected mean square" problem in the analysis of variance. It does not apply to the problem of obtaining specific "single degree of freedom" expected mean squares. An introduction to this latter problem is taken up in this study and an approach is outlined which corresponds closely to the unconditional approach.

Ross Willard Adams: "Estimating Missing Values by a Regression Method in Time-Trend Experiments." M.S. thesis. Iowa State University Library. May 1963.

This thesis presents a regression method for obtaining point estimates of missing observations, in time-trend experiments, for both completely randomized and randomized block designs. Time-trend experiments are defined as single experiments with continuous treatment application and periodic response measurement through time with one analysis at a predetermined terminal point in time. Four distinct relations between response and time are considered. These are: (1) homoscedasticity of responses at each measurement period, without autocorrelated errors between the measurement periods, (2) homoscedasticity with autocorrelated errors, (3) heteroscedasticity without autocorrelated errors, and (4) reteroscedasticity with autocorrelated errors.

For each possible relation a point estimate is obtained by least square procedures after transformations suitable to each of the four situations were applied. The estimated values of the regression parameters are determined such that the point estimate of the missing observation at the terminal point has minimum variance. This is achieved through preliminary testing procedures involving the comparison of statistics obtained from both the complete and incomplete records. The estimated missing value, its variance and confidence interval are determined in each situation. These procedures are illustrated by numerical examples.

The estimation technique described appears to be a superior method for the point estimation of missing values, in the time-trend experiments considered, as compared to current empirical missing value methods.

Fred Lawrence Ramsey: "Some Fiducial Prediction Problems." M.S. thesis. Iowa State University Library. May 1963.

Following the work of Fisher, Pitman and Buehler, this thesis is concerned with the prediction of future observations from the fiducial distributions of parameters. Buehler has proposed and discussed an explicit extension of the fiducial distribution of a parameter to a "fiducial prediction" distribution of a future observation. Suppose y is a future observation whose distribution $F(y;\Theta)$ depends on the parameter Θ, and suppose that $G_f(\Theta \mid x_1, \ldots, x_n)$ is the fiducial distribution of the parameter based on an observed sample of x-values. Then the proposed fiducial prediction distribution of y, given the sample, is

$$G_f(y \mid x_1, \ldots, x_n) = \int F(y;\Theta)\, dG_f(\Theta \mid x_1, \ldots, x_n).$$

In the case where the parameters are location and scale parameters, the fiducial distributions of the parameters have been given by Pitman. The fiducial prediction distributions obtained from these have been investigated. Identities have been obtained relating expectation with respect to the fiducial prediction distribution to conditional expectation over part of the sample space. Using these identities it has been shown that a number of predictors having desirable properties can be defined in terms of the fiducial prediction density. For example. it is shown that the mean of the fiducial prediction distribution is the minimum mean square error "invariant" predictor.

The formal definition of a fiducial prediction distribution extends immediately to the case of joint prediction of a number of observations. Fisher considered the normal case in which the mean and variance of a future sample were to be predicted, and showed that as the size of the predicted sample tends to infinity, the fiducial prediction density approaches the fiducial density of the population mean and variance. This result has been included in a general formulation in which it is shown that convergence in probability of sample functions to the parameters implies that the joint prediction distribution of the sample functions tends to the joint fiducial distribution of the parameters.

Papers presented at scientific and professional meetings

Papers presented at professional meetings reflect the work that has been done in the Statistical Laboratory. Many of these papers will appear later in publications. Abstracts are often published in professional journals. In such cases, references are given.

T. A. Bancroft and **Harold J. Larson:** "Adding Versus Deleting Predictor Variables in the Analysis of Incompletely Specified Regression Models," a paper presented at the 122nd Annual Meeting of the American Statistical Association in Minneapolis, September 7-10, 1962. Annals of Mathematical Statistics 33: 4, 1486. Abstract No. 22. December 1962.

T. A. Bancroft, K. C. Chanda and T. A. Brindley: "Some Problems of Specification Arising in Entomological Data Analysis," a paper presented at the meetings of the Institute of Mathematical Statistics and Biometric Society (ENAR) at Harvard University, May 5-7, 1963.

Neeti H. Bohidar: "The Probability Approach to the Construction of the Recurrent Relation of the Panmictic Indices under the Sex-linked Transmission," a paper presented at the meetings of the American Statistical Association in Minneapolis, September 7-10, 1962. Biometrics 18:4, 619. Abstract No. 836.

Neeti H. Bohidar: "The Mathematics of the Simulation of Genetic Systems," a paper presented at the Minneapolis meetings of the American Statistical Association. Biometrics 18: 4, 619. Abstract No. 835.

Robert J. Buehler: "The Method of Parallel Tangents for Finding an Optimum," a paper presented at the Gordon Conferences for 1962, New Hampton School, New Hampton, N.H., August 1, 1962.

Robert J. Buehler: "The Method of Parallel Tangents," an invited paper presented at the annual meeting of the American Institute of Chemical Engineers in New Orleans, March 12, 1963.

Robert J. Buehler: "A New Test of Fiducial Consistency," a paper presented at the sectional meeting of the Institute of Mathematical Statistics at the University of Wisconsin, June 14-15, 1963.

D. W. Staniforth, W. G. Lovely and **Foster B. Cady:** "Sampling Procedures for Estimating Weed Yields." a paper presented at the North Central Weed Control Conference in St. Paul, Minnesota, December 3-5, 1962.

K. C. Chanda: "Asymptotic Expansion for a Class of Distribution Functions," a paper presented at the annual meetings of the American Statistical Association in Minneapolis in September 1962. Annals of Mathematical Statistics 33:4, 1488. Abstract No. 29.

K. C. Chanda: "On Asymptotic Normality of a Class of Statistics Related to Linear Stochastic Processes," a paper presented at the sectional meetings of the Institute of Mathematical Statistics at the University of Wisconsin, June 14-15, 1963.

Herbert T. David: "Asymptotic Independence of (d)-Structured Statistics of Kolmogoroff-Smirnoff Type," a paper presented at the annual meetings of the American Statistical Association in Minneapolis, September 7-10, 1962. Annals of Mathematical Statistics 33:4, 1489. Abstract No. 31.

Herbert T. David, D. T. Davidson and L. J. Circeo: "The Relationship Between Cement Content and Freeze-Thaw Loss of Soil-Cement Mixtures," a paper given at a meeting of the Highway Research Board in Washington, D.C., January 7-11, 1963.

Carol B. Edwards and **Herbert T. David:** "Poisson Limits of Bivariate Run Statistics," a paper given at the Minneapolis meetings of the American Statistical Association September 7-10, 1962. Annals of Mathematical Statistics 33:4, 1489. Abstract No. 33, December 1962.

Wayne A. Fuller: "Estimation of Models Using Correlation Errors," a paper presented at the annual meetings of the American Statistical Association in Minneapolis, September 7-10, 1962. Journal of the American Statistical Association, 58: 302, 551. June 1963.

Dale Grosvenor: "Digital Computer Applications in Agricultural Engineering," at a forum for the Mid-Central Section of the American Society of Agricultural Engineers at Kansas State University, March 29, 1963.

Dewey L. Harris: "Biometrical Genetics in Man," a paper for the Conference on Methods and Goals in Human Behavior Genetics Research at the University of Louisville, April 16, 1963.

H. O. Hartley: "Matching Frame Surveys," a paper presented at the annual meeting of the American Statistical Association in Minneapolis, September 7-10, 1962.

H. O. Hartley: "Some Large-and Small-Sample Properties of Nonlinear Least Square Estimators." a paper presented at the Minneapolis meeting of the American Statistical Association. Journal of the American Statistical Association 58: 302, 552. June 1963.

H. O. Hartley: "The Impact of High Speed Computers on Research in the Life Sciences," the opening address at the Biomathematics Conference at the Texas Medical Center, Houston, Texas, March 28, 1963.

H. O. Hartley: "Multiple Frame Surveys," and "Monte Carlo Methods," two sectional addresses at the Houston meetings. March 28, 1963.

H. O. Hartley: "Solutions of Statistical Distribution Problems by Monte Carlo Methods," at the annual meeting of the American Association for the Advancement of Science at Philadelphia, December 26-30, 1962.

H. O. Hartley: "Nonlinear Regression Estimation," a talk given at a meeting of the local chapter of the American Statistical Association at Minneapolis on April 25, 1963.

H. O. Hartley: "Applications of Nonlinear Programming to Problems in Survey Design," a paper presented at the sectional meeting of the Institute of Mathematical Statistics at the University of Wisconsin, June 14-15, 1963.

O. Kempthorne and D. F. Cox: "The Use of Randomization Tests in Comparing Survival Curves," at a meeting of the Biometric Society (WNAR-ENAR) at Corvallis, Oregon, August 29-30, 1962. Biometrics 18:4, 629. December 1962. Abstract No. 864.

O. Kempthorne: "Can Multivariate Data Be Analyzed by Univariate Methods?", a paper presented at the meetings of the American Statistical Association in Minneapolis September 7-10, 1962. Journal of the American Statistical Association 58: 302, 555. June 1963.

O. Kempthorne: "Fisher's Contribution to Mathematical and Statistical Genetics," a paper presented at the meetings of the American Association for the Advancement of Science in Philadelphia in December 1962.

Marvin M. Lentner and **Robert J. Buehler:** "Some Tests for Gamma Parameters with an Application to a Reliability Problem," a paper presented at the meetings of the American Statistical Association in Minneapolis in September 1962. Annals of Mathematical Statistics 33:4, 1496. Abstract No. 54, December 1962.

Fred L. Ramsey and **Robert J. Buehler:** "Prediction in Location and Scale Parameter Families," a paper presented at the annual meeting of the Institute of Mathematical Statistics at the University of Wisconsin, June 14-15, 1963.

J. N. K. Rao and **John E. Graham:** "On Rotation Sampling," a paper presented at the American Statistical Association meetings in Minneapolis, September 7-10, 1962. Annals of Mathematical Statistics 33:4, 1498. Abstract No. 61.

J. N. K. Rao: "Unbiased Ratio Regression Estimators in Multistage Sampling," a paper presented at the Institute of Mathematical Statistics and Biometric Society (ENAR) at Harvard University, May 5-7, 1963.

J. K. Sengupta, E. J. Thomas, and **Gerhard Tintner:** "Stochastic Theory of Economic Development," a paper presented at the American Meeting of the Econometric Society in New York, December 27-29, 1961. Econometrica 30:3, 586. July 1962.

Leroy Wolins and E. B. Hutchins: "Factor Analysis of Statements Describing Student Environment in American Medical Colleges," and

Leroy Wolins, K. E. Johnson and A. C. MacKinney: "Direct Magnitude Estimation of Scale Values of Attitude Statements Compared with a Normal Transformation of Scale Values Derived from the Method of Equal Appearing Intervals," both papers presented at the annual meetings of the Midwestern Psychological Association, at Chicago, May 2-4, 1963.

George Zyskind: "On Conditions for Equality of Best and Simple Linear Least Squares Estimators," a paper presented at the annual meetings of the American Statistical Association at Minneapolis, September 7-10, 1962. Annals of Mathematical Statistics 33: 4, 1502. Abstract No. 75. December 1962.

Seminar talks

T. A. Bancroft: "The Fundamentals of Statistical Inference," a seminar talk given at National University, Mexico City, Mexico, February 18, 1963.

Robert J. Buehler: "Fiducial Prediction and Fiducial Consistency," a seminar talk given at the University of Minnesota, February 21, 1963.

Robert J. Buehler: "Fiducial Prediction and Fiducial Consistency," a talk given at a seminar for the Department of Statistics, University of Chicago, April 19, 1963.

H. O. Hartley: "Elements and Applications of a Method for Mathematical Optimizations," for the U. S. Steel Corp., Pittsburgh, Pa., October 23, 1962.

H. O. Hartley: "Elements and Applications of Linear Programming," a seminar talk given at Texas A and M College on April 1, 1963.

H. O. Hartley: "Sufficient Estimators and Their Relation to the Calculus of Variations," at a seminar for the Department of Statistics, University of Minnesota, April 26, 1963.

V. S. Huzurbazar: "Some Variants Associated with Probability Distributions," a paper presented at a seminar at the University of Minnesota, Minneapolis, November 2, 1962.

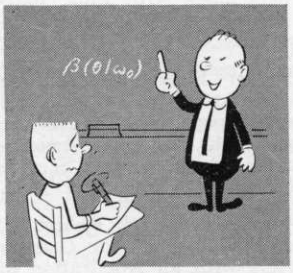
O. Kempthorne: "The Present Status of the Theory of Quantitative Inheritance," at the Research Branch of the Canada Department of Agriculture, Central Experimental Farm, Ottawa, Ontario, Canada, November 26, 1962.

O. Kempthorne: "The Design of Investigations" and "The Determination of Optimum Operating Conditions for a Production Process," two papers presented at St. Mary's College, Winona, Minnesota, on November 10, 1962.

O. Kempthorne: "Experimentation in Industry," at the School of Engineering University of Buenos Aires, on June 10, 1963;
"Sampling in Enumeration and Experimental Design," at Castelar Agricultural Experiment Station on June 14;
"Inference in Cooperative Experiments," at the University of Rosario on June 21;
"Some Statistical Problems in Medical Research," at the School of Clinical Science of the Medical School, University of Buenos Aires, on June 25;
and two lectures on "The Study of Yield-input Relationships," at the General University on June 25 and 26.

J. N. K. Rao: "Some Contributions to Unequal Probability Sampling Without Replacement," a talk to the Statistics Research Group, Bell Telephone Laboratories, Murray Hill, N. J. on April 11, 1963.

Teaching



The Department of Statistics in the College of Sciences and Humanities at Iowa State offers work leading to the degrees of Bachelor of Science, Master of Science, and Doctor of Philosophy with majors in statistics. Each major is built around a common core of courses in theory and methods, with other courses being chosen to fit the individual student's background and interests. At all levels, the program in statistics is designed to emphasize the close relationship between sound application and modern statistical theory. The Department also offers minor and supporting work in statistics. When desired, a joint-major program may be arranged between statistics and a substantive field of application for the Ph.D. degree.

Training facilities include two student laboratories equipped with 75 electric calculating machines for course instruction and for student use during nonclass hours. Use is also made of the high speed computing equipment and electrical equipment for punch card analysis of data which are located in the Computation Center of the University. A new IBM 7074-1401 high speed computer was installed in the fall of 1962. The Statistical Laboratory Library includes reprints of articles considered important in theoretical and applied statistics, reprints of staff publications, graduate theses, various statistical journals, reference books and other publications obtained on an exchange, gift or loan basis.

COURSE OFFERINGS IN STATISTICS

The Department of Statistics continues to grow as the use of statistics increases in all branches of science and industry, and changes are made in the curriculum as necessary to keep pace with the needs of the students. Foster Cady continued to teach an experimental course in elementary biological statistics which has now been approved as a new service course in statistics (Stat. 201A) to be required in several curricula in agriculture. H. T. David has helped organize the curriculum for a similar experimental course for students in civil and chemical engineering (Stat. 201B). The course in Psychological Statistics, which is listed jointly with Psych. 505, has been expanded into a two-quarter sequence (Stat. 505-506). Leroy Wolins teaches experimental design in phychological research (Stat. 505) and factor analysis (Stat. 506). To prepare students adequately for the application of statistical procedures to genetic studies and the interpretation of genetic

data, the course in genetic statistics has been expanded into a two-quarter sequence (Stat. 536-537) taught by Dewey L. Harris. Stat. 537 was previously taught as a special topics course. K. C. Chanda has taught special courses in stochastic processes and nonparametric methods in statistics (Stat. 599) and a course in time series analysis (Stat. 646) which is listed jointly with Economics 646. H. O. Hartley and W. J. Hemmerle taught Scientific Applications of Digital Computers (Stat. 580-581 and 680, which are listed jointly with Math. 580-581 and 680). Both courses include selected research topics and emphasize the applications rather than the mechanics of programming, thus broadening the curriculum in the scientific application of high speed computers.

The courses offered by the Department of Statistics during the year 1962-63 follow:

Courses for Undergraduate Students Only

		Cr.		
201	Principles of Statistics	5	FWS*	Billings, Cochran, Conn, Feddersen, Huntsberger, Lawing, Tetreault, Walter
327	Elementary Business Statistics	3	F	Tetreault
341, 342	Intro. to Theory of Statistics (Math. 341, 342)	3	FW	Huntsberger, Tetreault
380	Introduction to High Speed Computers (Math. 380)	3	FWS	Jespersen

Courses for Graduate Minors and Undergraduates

401, 402	Statistical Methods for Research Workers	4	FWS	
	A. Animal Sciences			Cox, Hotchkiss
	B. Plant Sciences			Cady
	C. Physical Sciences, Mathematics and Engineering			David
	D. Economics			Cochran
	E. Undergraduate			Huntsberger, Tetreault
	F. Psychology, Sociology Child Development and Vocational Education			Wolins
411	Experimental Designs for Research Workers	3	S	Cady
421	Survey Designs for Research Workers	3	S	Fuller
431	Elementary Statistical Quality Control	3	S	Richards
446, 447, 448	Statistical Theory for Research Workers	3	FWS	Huntsberger
480, 481, 482	Processing of Data	2	FWS	Mosier
499	Special Problems as arr.		FWS	Cady, Cox, Hemmerle, Huntsberger, Mosier, Strand
	Principles of Statistics for Biological Sciences (499B)	3	WS	Cady, Hotchkiss

*Taught Fall, Winter and Spring quarters.

Courses Primarily for Graduate Students, Major and Minor

501	Intermediate Statistical Methods	3	F	Bancroft
505	Psychological Statistics (Psych. 505)	3	S	Wolins
511, 512	Design of Experiments	3	WS	Kempthorne, Zyskind
521, 522	Design of Surveys	3	WS	Rao
531	Industrial Statistics: Sampling Inspection	3	F	David
532	Industrial Statistics: Design of Experiments	3		(Taught Alt. with 539)
535	Biological Statistics	3	S	Cox
536	Genetic Statistics (Genetics 536)	3	F	Harris
538	Elementary Econometric Statistics (Econ. 538)	3	F	Tintner
539	Operations Research Methods (Econ. 539, Math. 539)	3	W	David
541, 542, 543	Theory of Statistics (Math. 541, 542, 543)	3	FWS	Buehler, Huzurbazar
554, 555	Probability (Math. 554, 555)	3	WS	
580	Scientific Machine Computing	3	WS	Hartley, Hemmerle
599	Special Topics as arr.		FW	
	A. Theory			Chanda
	B. Methods General Scientific Computing			Hartley Hemmerle
	C. Design of Experiments			Harris
	D. Design of Surveys			

Courses for Graduate Students, Major and Minor

601	Advanced Statistical Methods	3	F	Hartley
611, 612	Advanced Design of Experiments	3	WS	Kempthorne, Zyskind
621	Advanced Design of Surveys	3		Hartley
622	Seminar on Design of Surveys as arr.			Hartley
638	Advanced Econometric Statistics (Econ. 638)	3		
641	General Theory of Linear Hypothesis (Math. 641)	3	F	Zyskind
642	Distribution Theory (Math. 642)	3	F	Buehler
643	Theory of Estimation and Testing of Hypotheses (Math. 643)	3	W	Huzurbazar
646	Time Series (Econ. 646)	3	S	Chanda
647	Multivariate Analysis (Math. 647)	3	S	Huzurbazar
649	Recent Developments in Statistics and Probability (Math. 649)			
688	Seminar of Theory of Statistics and Probability (Math. 688)	3	W	Buehler
699	Research 1, 3, 5, 6, 11		FWS	Bancroft, Buehler, Cox, David, Fuller, Harris, Hartley, Kempthorne

Summer Quarter

The summer quarter is divided into two 5½-week periods. Since the fiscal year began on July 1, 1962, and ended on June 30, 1963, only the courses offered during the second summer session of 1962 and the first summer session of 1963 are given in this report.

Second Summer Session—1962

402	Statistical Methods for Research Workers	3	Hotchkiss, Tetreault
421	Survey Designs for Research Workers	3	Fuller
448	Statistical Theory for Research Workers	3	Huntsberger
449	Special Problems	as arr.	
539	Operations Research Methods	3	Tintner
599	Special Topics	as arr.	
699	Research	as arr.	

First Summer Session—1963

401	Statistical Methods	4	Cox, Cochran
411	Experimental Design	3	Huntsberger
447	Stat. Theory for Research	3	Tetreault
499	Special Problems	as arr.	
599A	Intermediate Applied Decision Theory	3	David
599	Special Topics	as arr.	
699	Research	as arr.	

Roster of Graduate Majors and Special Students

Post Doctoral Visiting Fellows

(Mrs.) Indira Bhanot
David Jowett

Ph.D. Candidates

Rodney Basson
Marshall Brunden (Joint Stat-Forestry major)
Edward J. Carney (Joint Stat-Ind. Engr. major)
Robert Cochran
Charles E. Cress
Eugene Dayhoff
Thomas E. Doerfler
Carold Bates Edwards
Ahmed El Mawaziny
Francis G. Giesbrecht
John E. Graham
Dale Grosvenor
Jose S. Gutierrez
William J. Hemmerle
Klaus Hinkelmann
Ronald R. Hocking
Edwin J. Hughes
William D. Lawing
Paul Leaverton
Ilbok Lee
Dale O. Richards (Joint Stat.-Ind. Engr. major)
Lee H. Smith
Florence Tetreault
Mauritz van Aarde
Henry Walter
Robert F. White

M.S. Candidates

Ross Adams
Harold Baker
Michael Billings
Thomas Burnett
Russell Burns
Carlos Cavallini
Eugene B. Cohen
Patricia Conn
Alan Feddersen
Edmund Fuller
Burwell C. Gooch
Lawrence Gould
James E. Heard
Darrol Heggen
William G. Hill
Sheikh Ishaque
Louis W. Johnson
John Lesem
Douglas E. Murfield
Esmat Nouri
Fred Ramsey
Thomas D. Roseberry
Harold Rosenberg
Martin S. Rosenzweig
Donna Jean Ruhl
Carl F. Schach
Chang-Sheng Shih
Kenneth K. Simons
Gary Spencer
Gary J. Sutter
Nangnoi Suwanphant
Phrensi Svasti-Salee
Richard W. Swanson
David R. Thomas
Gordon D. Wakefield
Rudolfo Mengido
Richard W. Mensing
Philip M. Mills
Brenda Jane Morrison
Bert Webster
George Weiss
Stephan R. Williams
James R. Zweifel

Special Students

Carlos Gadpaille—Panama
Fazil Momand Rahim—FAS-USDA Afghanistan
Beryl Reckford—Jamaica
Bekele Teguegne—Tech. Asst. Program Ethiopia
(Mrs.) Dougchai Weekasul—AID Thailand

The George W. Snedecor Award in Statistics was granted by the Statistical Laboratory to Edwin L. Hughes in 1962 and to Ronald R. Hocking in 1963. This award, established in 1954, is given annually to the graduate student judged by the senior staff to be the most outstanding among those candidates for the Ph.D. degree in statistics at Iowa State who have successfully completed doctoral preliminary examinations in statistics within the preceding calendar year.

Degrees granted and positions taken

During the 1962-63 academic year 8 candidates were awarded doctorate degrees, 12 master's degrees, and 7 bachelor's degrees in statistics. Titles and abstracts of theses written as part of the requirements for advanced degrees are given on pages 38 to 44 of this report.

Recipients of the Ph.D. degree

Edwin J. Hughes (July 1962, under H. O. Hartley) in September 1962 became assistant professor of statistics at Oregon State University, Corvallis, Oregon.

Carol Bates Edwards (November 1962, under Herbert T. David) in September 1962 became assistant professor, Department of Behavioral Science, Medical Center, University of Kentucky, Lexington, Kentucky.

Ronald R. Hocking (November 1962, joint major in statistics and mathematics, under H. O. Hartley and Clair G. Maple) returned to Michigan State College of Mining and Technology, Houghton, Michigan, where he is assistant professor of mathematics.

Ilbok Lee (February 1963, under C. Philip Cox) took a position as biostatistician in the Statistical Methods Department of Parke, Davis and Co., Detroit, Michigan.

John E. Graham (May 1963, under H. O. Hartley) returned to Ottawa, Ontario, Canada, where he is with the Agriculture Division, Dominion Bureau of Statistics.

William J. Hemmerle (May 1963, joint major in statistics and mathematics under H. O. Hartley and Clair G. Maple) will continue to serve as assistant professor of statistics at Iowa State University where he is in charge of the Numerical Analysis-Programming Group of the Statistical Laboratory.

Mauritz van Aarde (May 1963, under O. Kempthorne) returned to the Union of South Africa to serve as teacher and senior lecturer in statistics at the University of Stellenbosch.

Robert F. White (May 1963, under O. Kempthorne) will continue his work with French, Kline and Smith Laboratories, Philadelphia.

Recipients of the M.S. degree

Carl F. Schach (July 1962, under H. T. David) took a position with Minnesota Mining and Manufacturing Co., St. Paul, Minnesota.

Gary J. Sutter (July 1962, under O. Kempthorne) went to CEIR, Inc., Dugway, Utah.

Bert Webster (July 1962, under H. O. Hartley) entered the Navy.

Eugene B. Cohen (August 1962, under C. P. Cox) became a biometrician with E. J. Squibb and Sons, Pharmaceuticals, New Brunswick, New Jersey.

Ahmed El Mawaziny (August 1962, under R. J. Buehler) is continuing study toward a doctorate in Statistics at Iowa State.

James R. Zweifel (August 1962, under C. P. Cox) became a commissioned officer working in the Biometry Branch, National Cancer Institute, Bethesda, Maryland.

Darrol Heggen (November 1962, under C. P. Cox) is statistician for the Institute of Agricultural Medicine, State University of Iowa Medical College, Iowa City.

Philip M. Mills (November 1962, under Wayne A. Fuller) is employed by Jet Propulsion Laboratory, Pasadena, California.

Brenda Jane Morrison (November 1962, under G. Tintner) took a position with the British Columbia Insurance Co., Victoria, B.C.

Kenneth K. Simons (November 1962, under Wayne A. Fuller) is employed by Martin Aircraft, Denver.

Ross Adams (May 1963, under T. A. Bancroft) will continue graduate work at Iowa State.

Fred L. Ramsey (May 1963, under R. J. Buehler) will continue his study for a doctor's degree in statistics.

Recipients of the B.S. degree

Thomas R. Douglas (August 1962) is an associate engineer at Nortronics, a division of Northrop Corp., Redondo Beach, California.

Craig K. Mitchell (August 1962) is 2nd Lieutenant in the Army Engineer Corps.

Ronald R. Devitt (August 1962) is 2nd Lieutenant in the U. S. Army serving in Germany.

Ronald R. Anderson (May 1963) is in Ames, Iowa.

Larry A. Conger (May 1963) is in Davenport, Iowa.

Paulette Sander (May 1963) is living in Ames, Iowa.

James R. Veale (May 1963) will do graduate study in statistics at Iowa State University.

Seminars

Statistical Laboratory— Department of Statistics series

Each year the Statistical Laboratory and the Department of Statistics sponsor weekly seminars which are offered on a non-credit basis and are open to students and faculty from other departments on the campus as well as those of the statistical center. Talks on current staff and graduate research projects and on more general developments in particular areas of statistics precede informal discusssion. The 1962-63 program included the following topics and speakers:

Fall Quarter 1962

September 12: Outlook for Statistics at Iowa State. T. A. Bancroft

September 26: On a class of birth and death processes. Gerhard Tintner

October 3: Toward a common psychological scale. Leroy Wolins

October 10: Straight regression lines deserve straight confidence lines. Herbert T. David

October 17: Estimation of normally distributed random elements. Dewey L. Harris

October 24: Fiducial argument and distribution involving multiple ancillary statistics. Ahmed El Mawaziny

October 31: Nonparametric statistical inference. K. C. Chanda

November 7: Joint meeting of Central Iowa Chapter ASA at Iowa City. Crossing phenomena for cumulative distribution functions. Prof. Meyer Dwass, Northwestern University

November 14: The 7074 Computer System. Ralph Klopfenstein, Director, Iowa State Computation Center.

November 21: A solution of the nonlinear programming problem of minimizing a convex function subject to convex restrictions. Ronald Hocking, Michigan College of Mining and Technology

Winter Quarter 1962-63

November 28: Statistical prediction of 1962 Iowa gubernatorial election results. William J. Hemmerle

December 4: Joint meeting of Central Iowa Chapter of ASA at Ames. Some aspects of biological assay. Prof. D. J. Finney, F.R.S., University of Aberdeen, Scotland; now Visiting Professor at Harvard.

December 5: The comparison of regression coefficients. Prof. D. J. Finney, F.R.S.

December 12: Linear forms in the order statistics from an exponential distribution. Elliot A. Tanis, State University of Iowa.

December 19: Bioassay with quantal response observed at different times. Ilbok Lee

January 9: Fiducial prediction, the problem of the Nile, and Fisher's Underworld. Robert J. Buehler

January 16: Some large and small sample theory in nonlinear least squares estimation. H. O. Hartley

January 23: Non-negative estimates of variance components. W. A. Thompson, Jr., University of Delaware

January 30: Sample size determination in simple analytical surveys. Joe Sedransk, Harvard University.

February 6: On simple random sampling with replacement. P. K. Pathak, Michigan State University

February 13: Operators in random mating populations. Mauritz van Aarde

February 20: Statistical problems in the measurement of chronic morbidity in an interview survey. Dr. Theodore D. Woolsey, National Center for Health Statistics, Washington, D.C.

Spring Quarter 1963

March 6: On the relation between covariance and a certain other residual analysis two stage procedure. George Zyskind

March 13: Unbiased ratio and regression type estimators in multistage design. J. N. K. Rao

March 20: An estimation procedure for the parameters of a multi-censored normal distribution. C. P. Cox and James R. Zweifel

March 27: Calibration, Foster B. Cady

April 3: The general form of distributions admitting sufficient statistics when the range depends on the parameter. V. S. Huzurbazar

April 10: Evaluation of the 1960 censuses. Dr. Morris H. Hansen, Assistant Director, Department of Commerce, Bureau of the Census, Washington, D.C.

April 17: Inference problems concerned with incompletely specified models in life-testing. Dale O. Richards

April 24: Invariants of probability distributions and their applications to inverse probability. V. S. Huzurbazar

May 1: Analysis of the Iowa Continuous Corn Experiment. Wayne A. Fuller

May 8: Joint meeting of the Central Iowa Chapter of ASA at Ames Dynamic quality control—more management than statistics. Mr. Harry Greiner, Manager Quality Control Division, Radio Corp. of America, Moorestown, New Jersey

May 15: Rotation designs for sampling on successive occasions, John E. Graham

Quantitative Genetics series

Seminars are held regularly on topics in quantitative genetics, participated in by staff and graduate students in Statistics, Animal and Poultry Science, Agronomy, Horticulture and Genetics. Dewey Harris and Oscar Kempthorne have served as chairmen of an interdepartmental committee in charge of the series. The following seminars were held:

September 27: A new approach to two old estimation problems: I. Sire evaluation in D.H.I.A. data; II. Sire and cross evaluation in diallel cross data. Dewey L Harris

October 9: On a relationship between partial diallel crosses and partially balanced incomplete block designs Klaus Hinkelmann

October 23: Covariation of relatives when relationship by pedigree is modified by knowledge of the transmission of marker genes. Mauritz van Aarde

November 6: Indirect estimates of genetic gains under selection. A. W. Qureshi

December 4: Selection in simulated genetic populations. John L. Gill

December 18: The effects of paternal irradiation on quantitative genetic parameters. Richard Willham and David Cox

January 15: A five-year study of genotype environment interaction. A. W. Nordskog

January 29: A missing link in bull sampling programs—the selection of bull dams. Dietmar Flock

February 12: Genetic variation in nipple numbers in swine. Richard L. Willham

March 5: The evolution of the linked gene complex. Dr. Walter Bodmer, lecturer in Genetics, Cambridge University on leave at the Department of Genetics, School of Medicine, Leland Stanford University.

March 19: Control systems in maize. Peter A. Peterson

April 2: A general selection index approach to hybrid soybean populations. Billy E. Caldwell

April 16: a review of the article "Prediction Formulae for General Combining Ability Selection Methods Utilizing One or Two Random Mating Populations," (Aust. J. of Biol. Sci. 15: 4, 1962. by Bruce Griffing.) Charles Cress

April 30: Diallel crosses in inbred lines of poultry repeated over locations and years. Dr. Laurence H. Baker, By-Line Poultry Farms

Computer Science series

The Computer Science series is an interdisciplinary seminar series which encompasses those areas of interest generally referred to as computer science. Seminars were cosponsored by the academic department with which the speaker was associated and the Computation Center. H. O. Hartley organized the seminars for the Fall Quarter, after which they were arranged by the Computation Center.

Members of the Department of Statistics who participated are listed here:

October 4: *Mathematics and Statistics*—A new method of convex programming. H. O. Hartley

October 11: *Statistics*—Automatic applications programming-analysis of variance. W. J. Hemmerle

October 25: *Statistics and Economics*—The capabilities of the S_2 linear programming code for the 7074. Dale Grosvenor

November 29: *Statistics*—Prediction of 1962 Iowa gubernatorial election results. William J. Hemmerle

December 20: *Statistics and Computation Center*—Regression analysis on the IBM 7074-1401 Computer system. Howard Jespersen

January 17: *Statistics*—Use of Monte Carlo methods in scientific computing. H. O. Hartley

March 4: *Statistical Laboratory*—Selection and grouping of bacterial strains utilizing the Cyclone Computer. June Smith

May 9: *Statistics and Computation Center*—The national machine record conference for colleges and universities. C. C. Mosier and Wayne Ostendorf

May 16: *Industrial Engineering and Statistical Laboratory*—Industrial Engineering and digital computers. Edward J. Carney

Miscellaneous seminars:

Members of the staff of the Statistical Laboratory occasionally participate in seminar programs sponsored by other departments of the university, and special seminars are held occasionally when off-campus speakers are available. The following list is representative of such seminars.

In the fall three seminars were held in Econometrics at which members of the Statistical Laboratory spoke:

October 16: Some Comments on simultaneous equation estimates. Wayne A. Fuller

November 6: Estimation in a Markov Scheme. K. C. Chanda with reference to K-class estimators. J. N. K. Rao

November 13: Estimation of simultaneous economic relations

December 7: *Joint Mathematics and Statistics Seminar*—Markov branching process and semigroups of operators. A. T. Bharucha-Reid, Wayne State University

January 11: *Forestry Seminar.*—Comparing regression lines. Foster B. Cady

March 4: *Joint Bacteriology and Genetics Seminar*—An experimental and theoretical analysis of the effects of deoxyribonucleases on the transformation of linked markers in B. subtilis. Dr. Walter Bodmer, lecturer in Genetics, Cambridge University, currently on leave at Leland Stanford University School of Medicine.

April 18: *Special Seminar in Statistics*—Randomization analysis of the general experiment. Robert F. White, Smith, Kline and French Laboratories, Philadelphia

May 2: *Special Seminar in Statistics*—On fractional replication. Dr. Walter Federer, professor of Statistics, Cornell University, now on leave at the Army Mathematics Research Center, University of Wisconsin

A publication of the Statistical Laboratory
Iowa State University
Eleanor F. Bolton, editor