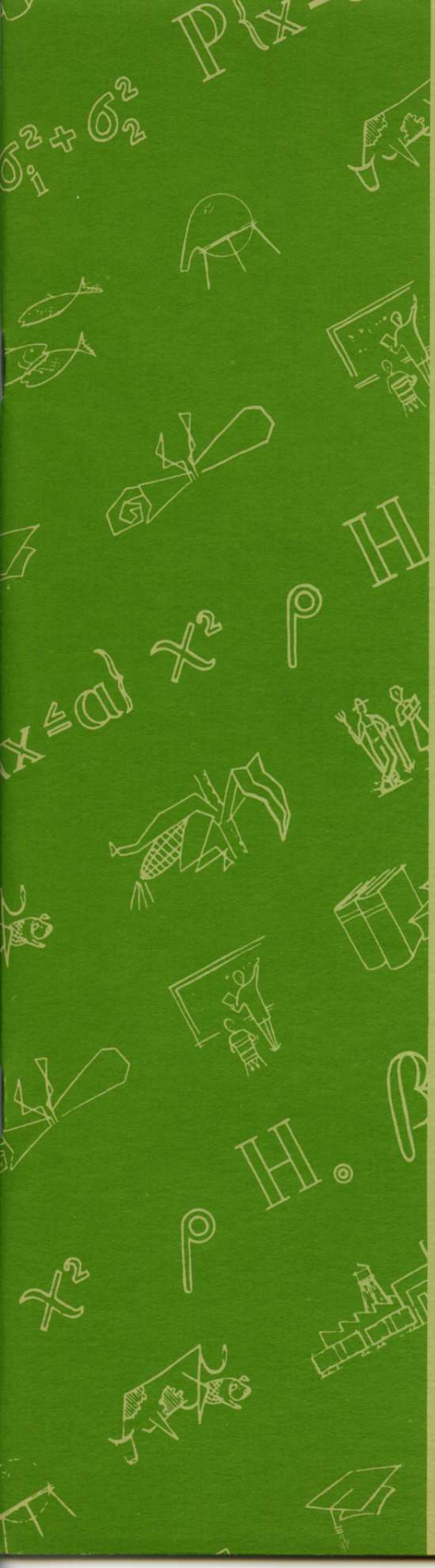


STATISTICAL LABORATORY

ESTABLISHED 1933

ANNUAL REPORT

July 1, 1969 to June 30, 1970

IOWA STATE UNIVERSITY

Ames, Iowa

President W. R. Parks
Iowa State University of
Science and Technology
Ames, Iowa

Dear President Parks:

This is the annual report of the Statistical Laboratory at Iowa State University for the fiscal year July 1, 1969, through June 30, 1970. It reports on the research, consulting, teaching and operational work of the staff of the university statistical center.

The statistical center is composed of:

(1) The Statistical Laboratory, a research and service institute under the president's office;

(2) The teaching Department of Statistics in the College of Sciences and Humanities;

(3) The statistics department of the Agriculture and Home Economics Experiment Station;

(4) The statistics participant in the Sciences and Humanities Research Institute and the Engineering Research Institute; and

(5) The research field office of the Statistical Standards Division, Statistical Reporting Service, United States Department of Agriculture, which is located in the Statistical Laboratory.

The staff members of the Statistical Laboratory work in cooperation with many institutions and departments of Iowa State University. This annual report is a review of these activities as well as a record of the activities carried on solely by the Statistical Laboratory.

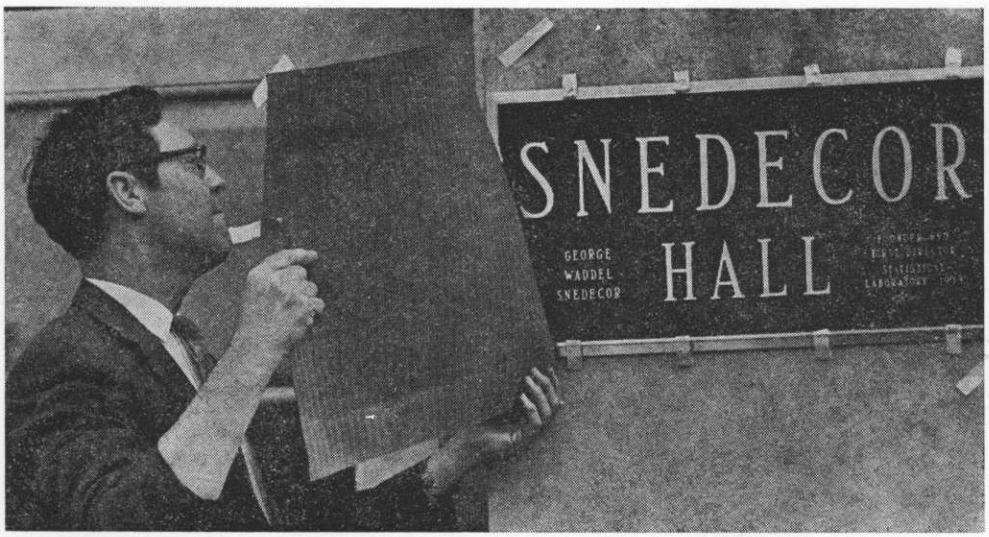
Respectfully submitted on behalf
of the Statistical Laboratory staff,

T. A. Bancroft

T. A. Bancroft
Director, Statistical Laboratory;
Head, Department of Statistics;
Head, Statistics Department,
Agriculture and Home Economics
Experiment Station

THE STATISTICAL LABORATORY

ANNUAL REPORT 1969-1970



Left: James Snedecor unveils the plaque.

Right, top: T. A. Bancroft presided; President W. Robert Parks spoke; middle: Gertrude Cox and Mrs. T. A. Bancroft poured at the reception; bottom: the Snedecor family (son James, grandniece Karen, nephew Dick with his wife) admires the portrait sketch of George Snedecor.

Snedecor Hall

"Iowa State has honored George W. Snedecor before, but no greater honor can it bestow than the recognition that is represented in the action we take here today," President W. Robert Parks said at special ceremonies May 18 when the Service Building was named Snedecor Hall. According to President Parks, "The name Snedecor is synonymous with statistics at Iowa State University . . . for the building that houses the statistics department and the Statistical Laboratory, there could have been no other name than Snedecor Hall."

The May ceremony gave formal recognition to action taken by the Board of Regents November 14 to honor professor emeritus George W. Snedecor. This culminated efforts of the Statistical Laboratory faculty, and friends of Professor Snedecor, which began in July 1963.

Professor Snedecor's health made it impossible for him to travel from California. He was represented at the ceremonies by his son, James, who unveiled a "Snedecor Hall" plaque on the front of the building, and a portrait sketch of Professor Snedecor in the entrance hallway. The ceremonies were concluded with an evening seminar by Professor W. G. Cochran of Harvard University, co-author of the sixth edition of Snedecor's book, *Statistical Methods,* and a member of the Statistical Laboratory staff from 1939 to 1946.

Personnel

Iowa State's statistical center includes the Statistical Laboratory, an institute under the president's office; the Department of Statistics in the College of Sciences and Humanities; the Statistics Department of the Agriculture and Home Economics Experiment Station; and the statistics participants in the Sciences and Humanities Research Institute and the Engineering Research Institute. The director of the Statistical Laboratory reports to the president of the university through the vice president for research. He also serves as head of the Department of Statistics and Statistics Department, and in these capacities reports to the respective dean and director.

Personnel are shared by the various components of the statistical center, which means that a staff member's salary may be budgeted from several sources, depending on the emphasis of his work. In addition, a number of cooperative agreements with federal agencies provide research funds for both staff and graduate students.

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

THE STATISTICAL LABORATORY STAFF FOR THE FISCAL YEAR 1969-70

Under the administrative direction of

W. Robert Parks, Ph.D. . . President of the University
J. Boyd Page, Ph.D. . . Vice President for Research; Dean of the Graduate College
Chalmer J. Roy, Ph.D. . . Dean, College of Sciences and Humanities; Director, Sciences and Humanities Research Institute
Floyd Andre, Ph.D. . . Dean, College of Agriculture; Director, Iowa Agriculture and Home Economics Experiment Station
Theodore A. Bancroft, Ph.D. . . . Director, Statistical Laboratory; Head, Department of Statistics; Head, Statistics Department, Iowa Agriculture and Home Economics Experiment Station

Professors

T. A. Bancroft
C. Philip Cox
David Cox - on faculty improvement leave
Herbert T. David
Wayne A. Fuller
Donald K. Hotchkiss
David V. Huntsberger
Oscar Kempthorne - Distinguished Professor, College of Sciences and Humanities
Basilio Rojas — visiting, winter and spring quarters
J. K. Sengupta — joint appointment with Department of Economics

George W. Snedecor – Professor Emeritus – in absentia
Norman V. Strand
B. V. Sukhatme
P. V. Sukhatme – visiting, spring quarter
James Walsh – joint appointment with Department of Psychology
Leroy Wolins – on leave
George Zyskind

Associate Professors

Barry Arnold – joint appointment with Department of Mathematics, in Mexico through summer 1969
David Jowett
C. C. Mosier – joint appointment with Computation Center
Edward Pollak – joint appointment with Department of Genetics
Richard D. Warren – joint appointment with Department of Sociology and Anthropology

Assistant Professors

Harold Baker
Terry Dickinson – joint appointment with Department of Psychology
Chien-Pai Han
Roy Hickman
Paul Hinz
James L. Hutter – joint appointment with Department of Political Science
Dean Isaacson – joint appointment with Department of Mathematics
William Kennedy
Glen Meeden
Richard Mensing
Abel Mexas (promoted as of spring quarter) – on assignment in Mexico
Shashikala Sukhatme

Instructors and Associates

Leon Burmeister
G. L. Ghai
Louis Jensen
John Kinney
John Lin
Jeff Meeker
P. Papaioannou
Vasantha Solomon
Vincent Sposito
Richard Stein
Victor Tang
James Veale
Eric West

Postdoctoral Associate

Jerome J. Pella

Visiting Scholars

Isabelle Gravett – through winter quarter
Sundardi Wirjosudirdjo

Statistical Data Processing Service

Bud J. Meador, Supervisor
Charlotte Bentley
Shirley Cheatum – fall, winter and spring quarters

Graduate Assistants

(The status of graduate students often changes. Students who have held the title of graduate assistant during the year are listed here.)

Forrest Aspengren
George Battese
Richard Chamberlain
Eliahou Cohen
Jean-Denis Desrosiers
Her Tzai Huang
Cary Isaki
J. D. Jobson
Henry Kelker
Dennis Mar
James Mellon
Kenneth Merritt
Martin J. O'Connell
Kenneth Offord
C. M. Patel
David Pyne
Sue Ritchie
J. R. Schmid
Wendell Smith
Terry Svejda
William Warde
Irene Yung

NIH Trainees:

Pamela Doctor
Thomas Fears
Ronald Jacobson
Roland Loup
Peter O'Brien
Lonnie Vance

NDEA Fellows:

Richard K. Dorsch
A. Ronald Gallant
John Goebel
Nancy Heath
G. Nick Lauer

NSF Fellows:

Thomas Keefe
Richard Madsen

Supported Graduate Students:

Guillermo Adames-Suari, Government of Mexico
Yahia Ahmed, Institute of International Education
Pathra Chatkeo, AID, Census Bureau, Thailand
Isidoro David, Rockefeller Foundation, Philippines
Amiri Gamshadzahi, USDA, Iran
Soner Gonen, Hacettepe University, Turkey
Joseph E. Grimes, Department of Mathematics
Norman Hoesly, Department of Mathematics
Jairoj Jayavadhanangkur, AID, Census Bureau, Thailand
Lyle Lautenschlager, USDA
Brian D. Macpherson, University of Manitoba, Canada
Muhammad A. Malik, USDA, Pakistan
Angel Martinez, Ford Foundation, Mexico
Nimmagadda Murthy, AID, India
Ahmed Salem, Government of UAR
Marcela Salinas, Institute of International Education
Ivan Sampaio, IRI
Malte Sund, Volkswagen Foundation
George Wang, Department of Economics
James P. Whipple, Western Reserve University
Franklin Wolf (joint statistics-industrial engineering), Department of Industrial Engineering
Robert Young, USDA

Unsupported Graduate Students:

Ray Bailey
Kuo Ming Chien
Peampan Davivongs
Modesto A. Freytes
Elizabeth Hsu
Geung Ho Kim
Hae Ja Chung Kim
Charles MacIsaac
Yupha Onthuam
Dennis Tsai
Milton Winger
Chartsee Xumsai

General Office Staff

Margaret G. Kirwin, Administrative Assistant
Kathleen Ringgenberg, Accountant
Susan Alice Brown, Technical Writer-Editor
Avonelle Jacobson, Supervisor, Teaching Section
Shirley Saveraid, Secretary, Statistical Numerical Analysis and Data Processing Section
Iveta Zeliadt, Secretary, Experimental Design-Genetic Statistics Section — through fall quarter; beginning spring quarter
Neta Perry, Secretary, Experimental Design-Genetic Statistics Section — winter and spring quarters
Jan Bates, Secretary
Marilyn Black, Secretary — winter and spring quarters
Norma Christian, Secretary
Donna Cooper, Secretary
Carol Shafer, Secretary — through fall quarter

Survey Section

Sandi Partlow, Secretary
Marjorie Mason, Office Manager
Hazel Cook, Survey Supervisor
Betty Fell, Programmer
Clerks:
Helen Carney
Evelyn Green
Ava Klopf
Anna B. Woodrow

Consulting and Joint Research

When researchers from substantive areas who use statistics as a research tool wish to determine scientific conclusions from data, they frequently need the assistance of a skilled statistician who is up-to-date on the latest advances in statistical methodology. All members of the Statistical Laboratory staff are available for consulting, although certain personnel are budgeted to devote a major portion of their time to these services.

Consulting services are provided for the statistical analysis and interpretation of data, statistical design of experiments and surveys, and the development and extension of new statistical methods and techniques. There is much variation in the nature of the statistical approach, subject matter involved, complexity of the statistical treatment and the degree of involvement of the consultant in a given study.

One key to constructive and useful consulting is the ability to produce, easily and efficiently, analyses of data based on the proposed models of the research. The best results are obtained when the research workers themselves become involved in the analysis procedures. Often one important service performed by the consulting statistician is to encourage a research worker to examine his objectives carefully in view of his proposed experiment to ensure that it will answer his questions.

Statistical consulting services are made available to many staff members and graduate students under the support of the Iowa Agriculture and Home Economics Experiment Station. The Experiment Station budget also provides some funds for statistics research and the development of new statistical methods.

During the year Agriculture Experiment Station Project 101, Statistical Services in the Animal Sciences and Plant Sciences, has supported the consulting services of Dr. D. K. Hotchkiss, Dr. David Jowett, Dr. Paul Hinz and several graduate students, including Leon Burmeister, Richard Chamberlain, Eliahou Cohen and James Mellon. Dr. Basilio Rojas assisted with this project while he was a visiting professor. Students and faculty members from all areas of agriculture and related fields have submitted data for analysis. Frequent consultations have taken place on the design and analysis of experiments, and progress has continued in the implementation of general computer programs carrying out statistical analysis.

Project 113, Statistical Services for Sampling Investigations, made it possible for Professor N. V. Strand and Harold D. Baker of the Survey Section to provide consultative advice on sample surveys for departments of the Agriculture and Home Economics Experiment Station.

Dr. Edward Pollak receives some support for consulting activities from Agriculture Experiment Station Project 1448, Consultation and Research in Mathematical and Statistical Genetics.

Arrangements with the Engineering Research Institute, similar to that with IAHEES, result in the support of some of the consulting done by Dr. H. T. David.

Statistical Laboratory staff members may spend as much as 5 hours per week in consulting and advisory services. The average time spent in these activities totals more than 8½ hours per week for each of the 17 senior staff members who regularly spend at least one hour per week consulting.

Nearly half of the research workers who seek consulting advice are affiliated with the College of Sciences and Humanities. Slightly more than a fourth are from the College of Agriculture. The rest are from all of the other colleges of the university: Home Economics, Education, Veterinary Medicine and Engineering. The departments which have sought advice on the greatest number of projects (although this does not necessarily indicate that the most time was devoted to these departments) are economics, sociology, psychology, agronomy and family environment. Some projects require only a few minutes of a consultant's time, while others require many hours over a period of many months.

About half of this consulting involves assistance with analysis of data. Two or more staff members often will be consulted, independently or jointly, on a single project. The project may involve individual students, staff members or a group of researchers. The results may appear in a thesis, or be presented as a paper at a meeting, or be published. Sometimes the consultant's contribution will be such that he is recognized as a paper's co-author.

It is neither possible nor practical to enumerate in detail all of the consulting work done by each staff member. However the following very brief condensation will give an indication of the types of projects about which staff members are consulted, and the statistical processes used:

A time series was used on a study of the biochemical and physiological responses of dogs to tobacco smoke. Regression with confounded factors was selected for a study of changes in the dietary intake of adult women with age. Analysis of variance was combined with trend analysis to study the fat characteristics resulting from the storage of poultry products. A chi-square evaluation of grouped data was used to evaluate analysis techniques for a study of the effectiveness of pasteurization. Fertilizer experiments were designed for greenhouse studies comparing the potash availability in Iowa soils. Regression was used to determine factors affecting the swelling of wood.

General least squares and correlation analysis were applied to the evaluation of experimental teaching methods developed for use in vocational agriculture programs in public high schools. General least squares (time series and cross section) were used in a study of interest rates and government securities. Path analysis was used to determine factors related to English achievement. Scaling and measurement procedures were applied to a study which compared scaling procedures. Factor analysis was used to determine the relationship between ACT scores and student records.

A new measure of correlation was developed for a study relating two measures of porosity. It was adapted to straight-line fitting by minimization of the sum of squared perpendicular distances. The method of moments was applied to estimating detector efficiency and the relative rates of reactions producing specified numbers of neutrons, for an analysis of neutron-yielding reactions in fluorine subjected to photon bursts.

Data analysis in a weighted nonlinear regression situation was used to estimate future daughters' milk yield for observed yields of dairy cows. Data on survival probabilities and fecundities were used in calculating the ultimate rate of growth of a population to predict the effects of current hunting policies.

Non-orthogonal analysis of covariance was used on a study of the effects of hormone and other factors on carcass quality. Data were analyzed and a split plot designed for a protein and vitamin study using rats. A theory was derived to predict genetic means in a population subjected to selection for a quantitative character, for a study of poultry hatchability.

Regression analysis with unusual interactions was used in a study of factors that influence stock market changes. Scaling and data analysis were used to compare the value of status as regarding fashion choices in clothing. Regression, analysis of variance, chi-square and factor analysis were employed in an investigation of usefulness and effectiveness of several statistical techniques using human subjects as an independent variable. Scales of attitudes towards consumer projects were constructed by factor analysis.

Two special sections of the Statistical Laboratory offer a great many consulting services to staff members, graduate students, and, when time permits, off-campus personnel. These sections are the Survey Section and the Statistical Numerical Analysis and Data Processing Section.

Survey Section

The Survey Section provides both consulting and direct operational services on all aspects of sampling, surveys and census-type studies. These services include design of surveys, drawing of samples, construction of questionnaires, training and supervising of field workers, collection of data by personal interview and mail questionnaires, coding and editing of data for IBM processing, analysis of data, inference procedures and maintaining liaison with programmers and the Computation Center on machine manipulation of data coded. The Survey Section is prepared to provide any or all of these services. This flexibility allows the section to serve the individual researcher better, by adapting its services to his needs and resources. A sizable file of basic data and materials for sampling urban and rural populations has been accumulated which is

useful for handling large-scale regional or national surveys.

In addition to working on research conducted through other departments of the university, the Survey Section annually participates in a number of studies administered jointly by the Statistical Laboratory and some other university, institute, state or federal agency. The United States Department of Agriculture and the Bureau of the Census have cooperated with the Statistical Laboratory for many years on a number of projects.

The listing which follows is only an indication of the type of work done by the Survey Section during the year, and is not a complete record of its services:

All of the operational work was performed (Survey Section personnel designed and drew the samples, helped with the questionnaires, wrote the manuals, conducted field work and did the coding, editing and machine runs) for a grass back slope study, a land ownership study, and studies of land purchase contracts, farm capital requirements and teacher mobility.

Interviewers were trained for the third national cancer survey and a reinterview study done on a subsample of members of the sample. The survey involves data collection, processsing and reporting of findings. Checking, editing, coding and punching were completed for a study of rural families conducted by a university in a neighboring state. Editing and testing of the questionnaire and field work were done for a study of political attitudes of residents of a selected city in Iowa.

Field work and data analysis were done on a study of the educational needs of farmers. Machine runs were done for tax work using SCS data. A sample was drawn, returns were edited and coded and machine runs were done for a study of the most prevalent hybrid corn varieties. Sampling was done for a mobility study of residents of Polk County, a civil defense survey, a smoking behavior study, a regional study of disadvantaged households, attitudes of alumni toward extension courses, low cost housing and a survey of attitudes toward television commercials.

Statistical Numerical Analysis and Data Processing Section

The Statistical Numerical Analysis and Data Processing Section is organized into two closely allied groups: the Statistical Data Processing group and the Statistical Numerical Analysis group.

Members of the Statistical Data Processing group devote their time to the actual processing of data through the computer and to consulting on problems in statistical data processing. During the past year, its second year of operation, the group handled 127 research projects, distributed as follows: College of Agriculture – 17, College of Education, – 11, College of Engineering – 7, College of Home Economics – 31, College of Sciences and Humanities – 49, College of Veterinary Medicine – 2, University Administration – 6, Iowa Crime Commission – 1, and USDA, ARS – 3.

The Statistical Numerical Analysis group is made up of persons within the section who conduct research in computer applications of statistics, consult on problems in computer application of statistics and develop general-purpose computer programs for statistical applications. This group provides the general programming support available within the section, allowing the data processing group to avoid programming which is not immediately applicable to a data processing problem on hand.

A computer program known as TARSIER was rewritten during the year. This program performs nonlinear regression analysis utilizing H. O. Hartley's modification of the Gauss-Newton technique for solving nonlinear systems. Up to 20 parameters and 10 independent variates are allowed in TARSIER.

A new program completed during the year was named HELARCTOS. It performs multiple linear regression computations and provides many user exercisable options such as residual plotting and testing of various hypotheses.

Personnel from the section provided programming and data processing support for 135 separate research projects during the year. In addition, assistance was provided to several courses in statistics and related fields. The following accounting indicates the type of consulting activities performed by the Statistical Numerical Analysis and Data Processing Section:

A computer program was developed for a study of the moments of a special distribution. Advice on data editing and coding were used for a study of farmer product mixes and the effect of changing mixes. Regression model building was used in a study to classify differences in growth rates of pine trees. Consulting on computer consolidation and analysis were provided for data supplied by the federal government on duck kills of birds banded in Iowa. Analysis of variance with unequal subclass numbers was applied to student scores on achievement tests given at various universities around the country.

Linear programming and sensitivity analysis were used in a study of the economic impacts with respect to variation in certain commodity prices. Quadratic programming was used to identify a partial competitive equilibrium for the crop and livestock sectors of U.S. agriculture. Parametric programming was used to find an optimal crop policy for a typical farm in Mexico. Convex programming was used to determine optimal sample size for sampling on successive occasions with multiple objectives.

A nonlinear recurrence formula was evaluated for a study of the dynamics of populations with overlapping generations. Iteration of matrices from Markov chains was used to study the genetic change in a finite population.

Current Research

Statistical research has produced many fundamental techniques which are of value to a great number of research workers. The Statistical Laboratory budget supports research projects which are of specific interest to regular university research programs. Other research is part of a continuing cooperative program with other campus research institutes and experiment stations. Grants and contracts with various off-campus agencies provide for supervised graduate research on statistical problems of common interest.

During the year Dr. H. T. David has been involved with research on sequential aspects of a less parametric life analysis. This research was supported by Project 655 of the Engineering Research Institute, Valuation Project Development, directed by Dr. Harold A. Cowles of the Department of Industrial Engineering.

Krane (1963) suggested polynomial graduation of observed failure rate curves. The research considered what the sequential point of view can bring to the estimation aspects of this problem. Unbiased estimators of good efficiency have been suggested both for estimating a single failure rate and for the simultaneous estimation of several. The latter case is the relevant one for the graduation problem.

Dr. Richard Warren continued as co-leader of Agriculture Experiment Station Project 1626 with Dr. George M. Beal and Dr. Joe M. Bohlen, Department of Sociology and Anthropology. This project is an attempt to determine the characteristics of managers and local agribusiness firms that will permit prediction of their success and ability to adjust to innovation and change. This year's activities included continued data analysis and the preparation of reports and papers.

While serving as a visiting professor during spring quarter, Dr. P. V. Sukhatme conducted research on a new appraisal of the protein problem. This material will appear in a book he is writing.

During the year Dr. Edward Pollak has served as co-investigator, with Dr. E. T. Hibbs, of Agriculture Experiment Station Project 1736. This project, Corn Rootworm Population Dynamics, is administered through the Department of Zoology and Entomology.

Field sampling corn rootworm eggs for study of distribution and density in specific fields was accomplished at two locations in Story County. Reference populations are isolated in greenhouses at Ames for genetic comparisons. Enhancement of natural infestation for experimental purposes in a nine acre field at Ankeny failed due to natural decimation of local numbers.

NIH Research

Four of the five trainees supported by the National Institutes of Health biometry training grant have participated in work experience programs during the year. These programs make it possible for the students to obtain experience under senior consultants on actual statistical problems under investigation at the cooperating institutions.

In addition to their work experience, the students, who all are doctoral candidates, complete a balanced program of statistical theory and application at Iowa State University, and take courses in biological statistics and in the biological sciences.

Pamela Doctor was at the Mayo Clinic in Rochester, Minnesota, during the summer, where she was assigned to a study of the effect of Jorpes secretion in intestinal action potential in the dog. Data were analyzed by a three-way classification design with sessions representing blocks. Mrs. Doctor participated as a subject in research on touch and pain thresholds, and prepared a seminar on the problems of the measurement of sensation. She also assisted with a study of biopsies of diseased and normal nerves, helping to obtain a distribution of the number of myelinated fibers of various diameters per nerve, and to investigate the relationship between the length and diameter of internodes in teased fiber studies. Mrs. Doctor's work at Iowa State is under the guidance of Dr. T. A. Bancroft.

Peter O'Brien has also been at the Mayo Clinic, where he has been involved with consulting during the year. Most recently he has been engaged in the development of a method for the diagnosis of mio-infarction in the human heart, based on analysis of vectorcadiograms. In his thesis research, under Dr. Wayne Fuller, he has obtained sequential procedures for selecting the best of several populations. A reduction over existing procedures in the average number of observations required to make a selection is the objective in each case. It has been demonstrated empirically that this objective has been accomplished when selection is from gamma, normal, multinomial and certain incompletely specified populations. A paper on this research has been submitted for publication.

Tom Fears worked during the summer at the Biometry Branch of the National Cancer Institute in Washington, D.C. He participated in a study conducted by the End Results Section to attempt to describe the extent of breast cancer in terms of the results of a clinical examination. The life table method and data from Memorial Hospital were used in the study, with the average annual mortality rate over ten years as an index. During the year Fears has completed his thesis research on robustness properties of certain nonparametric procedures used in estimating and testing, under the direction of Dr. K. L. Mehra.

Lonnie Vance spent nine weeks during the summer in the Department of Preventive Medicine and Environmental Health at the University of Iowa. He worked under former NIH trainee Paul Leaverton,

who received B.S., M.S. and Ph.D. degrees from Iowa State. Vance spent his summer getting experience and practice with biological and health statistical problems via consulting sessions, analysis of data, and review of literature relevant to the problems. He has been a full-time student during the year. His major professor is Dr. Chien-Pai Han.

Roland Loup has been working with Professor C. P. Cox on his doctoral thesis which is concerned with generalizing von Neumann's mean square successive difference statistic and developing applications of this generalized statistic. He has developed a test procedure employing the generalized statistic successively to test for the degree of polynomial which fits a series of normally distributed y-observations taken at equally spaced x-values. He has shown that when this test procedure is used in conjunction with the least squares test procedure for polynomial regression, additional information about the data, such as negative serial correlation among successive observations, may be detected. In addition to his thesis research, Loup taught a statistics laboratory course at Iowa State during winter quarter.

Research in Mathematical and Statistical Genetics

Research in mathematical and statistical genetics is supported jointly by the National Institutes of Health, Grant GM 13827, and Agriculture Experiment Station Project 1669. Dr. Oscar Kempthorne and Dr. Edward Pollak are principal investigators, with additional research being done by Gauri Ghai, W. G. Hill, Louis Jensen and B. D. H. Latter.

Theoretical analysis of various concepts of fitness has been made, with mathematical study of the dynamics of Mendelian populations under selection for viability and fecundity. A start has been made on the study of selection theory for large diploid populations with overlapping generations. Work on the fate of a line descended from a mutant gene resulted in the development of a method to obtain bounds on the probability that a line of mutant genes descended from one ancestral gene survives to generation n, where n =1,2, ... Bounds have also been obtained for the mean time to extinction, given extinction occurs. These methods apply to some offspring distributions regardless of whether the mean number of offspring is less than or greater than unity.

Research continued on the long-term theory of selection for a quantitative attribute in an infinite population determined by a single or by two loci. An algorithm has been developed for solving diffusion equations arising in genetic selection theory, and solutions have been compared to exact results derivable from probability transition matrices.

The dynamics of genetic populations under mixtures of mating systems has been studied, with particular reference to quantitative genetic theory when there is partial inbreeding. The theory of assortative mating has been examined with particular reference to similarities to and differences from inbreeding theory. Some work has also been done on the variance and covariance of relatives under mixed selfing and random mating.

Models for quantitative genetic theory have been explored by computer simulation with particular emphasis on centripetal selection for quantitative attributes in natural populations. The primary aim was to obtain a satisfactory model for data on electrophoretic variants in *Drosophila.* Statistical methods for treatment of data from selection experiments have been examined, with particular reference to the precision or realized heritability estimates. Some probability theory of selection with variation in population size and selective advantage has been developed.

Goodness of Fit Procedures

National Science Foundation Grant GP-5997, concerned with research on the general problem of testing goodness of fit of distributions in one or more dimensions, continued under the direction of Dr. Oscar Kempthorne.

Graduate student Eric West has been engaged in research on various aspects of the general problem of data analysis and evaluation, especially with regard to the inference problem of goodness of fit. Combining theoretical developments and advanced computational techniques, he has investigated the small sample properties of the standard Pearson chi-square test statistic for the multinomial distribution. For a variety of small sample cases he has computed the exact null distribution and exact power of the chi-square test. He has also made a large number of comparisons of some approximations for the power of the test. In addition, he has compared the chi-square and likelihood ratio tests for a simple null and composite alternative, with regard to null distribution and power. He is also carrying out research on the relationship among various tests of significance, as opposed to simple accept-reject rules.

Some Problems Involving Inference After a Preliminary Test of Significance

National Science Foundation Grant GP 9046, Some Problems Involving Inference After a Preliminary Test of Significance, was terminated June 30. Dr. T. A. Bancroft was principal investigator. The grant also supported research conducted by Dr. Chien-Pai Han, Dr. W. J. Kennedy and Dr. B. V. Sukhatme.

During the year Dr. Han has been considering the problem of testing equality of two means after a preliminary test of equality of variances. Depending on the results of the preliminary test, the test procedure is either the usual t-test or the situation is the same as that of the Behrens-Fisher problem, in which case a procedure given by Cochran (1964) is used.

After a number of transformations of the variables, the size and power of this testing procedure is being obtained and computer programs written to obtain numerical values of the size and power which make it possible to recommend the significance level of the preliminary test.

Dr. Kennedy's research, which resulted in his doctoral thesis, was concerned with model building for prediction in regression analysis based on repeated significance tests. An abstract of his thesis appears in the publications section of this report.

Dr. Sukhatme worked with graduate student Alfonso Carrillo on a study of the preliminary tests of significance in some sample survey inference problems. The results are contained in Dr. Carrillo's Ph.D. thesis, which also is summarized in the publications section of this report.

Design and Analysis of Sample Surveys

The contract with the U. S. Office of Education for research activities on the design and analysis of sample surveys terminated August 31. Dr. T. A. Bancroft assumed the responsibility of project director when Dr. J. H. Sedransk left the university.

During the two months the contract was in effect this year, graduate student K. T. deGraft-Johnson was awarded his doctorate for research supported by the contract. An abstract of his thesis, Some Contributions to the Theory of Two-Phase Sampling, appears in the publications section of this report.

Use of Preliminary Test(s) of Significance in Designing Surveys and Analysis of Survey Data

A new contract with the Office of Education of the Department of Health, Education and Welfare originated April 15 with Dr. B. V. Sukhatme as principal investigator. The title of the research is Use of Preliminary Test(s) of Significance in Designing Sample Surveys and Analysis of Survey Data.

The objective of this research is to extend and develop the theory and methodology of sampling involving the use of preliminary test(s) of significance to problems which are of considerable importance to statisticians and research workers. Although inference procedures incorporating preliminary test(s) of significance have been used extensively for some time by research workers and statisticians, only in recent years and for certain purposes have attempts been made to evaluate the properties of such inference procedures.

The research supported by this contract is concerned with three main topics:

1. Allocation of sample sizes to different strata,
2. Estimation of the mean in stratified sampling—combined ratio estimate or separate ratio estimate, and
3. Estimation of the mean in stratified sampling — combined regression estimate or separate regression estimate.

Research in Mathematical Statistics and Probability

A project concerned with research in mathematical statistics and probability began August 1. Dr. Barry Arnold, Dr. H. T. David, Dr. Dean Isaacson, Dr. Glen Meeden and Dr. Richard Mensing are co-investigators, with Dr. David acting as project coordinator.

Current research is underway in six areas:

1. Dr. Arnold is pursuing research on a characterization of the exponential distribution. He also is concerned with problems of estimation in truncated normal distributions.

2. Dr. David is extending research on game value distributions which originated in the Ph.D. thesis of former graduate student Donald J. Soults. Work so far has extended the 2xn analysis of Soults' thesis to the case of not necessarily identically distributed columns.

3. In queueing theory, one of the problems of interest is the limiting distribution of the waiting time of a customer when expected inter-arrival time exceeds expected service time. Dr. David and Dr. Mensing are investigating an alternative limiting distribution, pertaining to the case where expected service and inter-arrival times are almost identical.

4. Dr. Isaacson is extending research which resulted in a paper, "Stochastic Integrals and Derivatives." An abstract appears in the publications section of this report. Specifically he is investigating whether or not one gets convergence in the case of a continuous square integrable martingale.

5. Dr. Meeden is concerned with classification procedures. He hopes to find "good" classification procedures which minimize misclassification in some sense as the sample sizes tend to infinity.

6. Dr. Mensing is studying nonlinear estimation with special emphasis on the estimation of the variance of estimates.

Analysis of Variance and Experimental Design Research Procedures

Research supported by the Aerospace Research Laboratories, Office of Aerospace Research, United States Air Force, has continued under the direction of Dr. George Zyskind and Dr. Oscar Kempthorne.

Dr. Zyskind has continued work on linear model theory and randomization theory. In linear models some further results on conditional inverses and on their relevance to other aspects of the subject have been obtained. A definition is being developed appropriately generalizing the Moore-Penrose inverse in terms of non-othogonal projections. Attention is also being paid to certain geometrical aspects of linear estimation. Thus it is seen that best estimators can be obtained in a general case by finding the point of tangency of the column space of the matrix X in a linear model $y = X\beta + e$ with a system of ellipsoids whose axes are determined from the covariance matrix $\sigma^2 V$ of the model. Progress has been made in estab-

lishing connections between conditional inverses and the projections they induce on specified row and column spaces. For example, if X and U are matrices of the same rank such that UX is defined then the unique projection operator on the column space of X along the orthogonal complement of the column space of U′ is given by X(UX)°U, where (UX)° is any conditional inverse of UX. An interesting special case is the situation where Uy is a generator set of best linear unbiased estimators (b.l.u.e.'s) of linearly estimable parametric functions of the vector parameter β in the model $y = X\beta + e$, with $E(e) = 0$ and $E(ee') = \sigma^2 V$. It has been shown that a b.l.u.e. of the vector $X\beta$ is then given by $X\hat{\beta} = X(UX)^{\circ}Uy$ and that a b.l.u.e. of a parametric function whose expectation is a $'X\beta$ is a $'X(UX)^{\circ}Uy$. Here a $'X(UX)^{\circ}U$ is the projection of the vector a′ on the row space of the matrix U. Thus, the indempotent matrix $X(UX)^{\circ}U$ is a bilinear projection operator. A relation with the general Gauss-Markov theorem is obtained by choosing a suitable admissible conditional inverse V° of V such that the row space of $X'V^{\circ}$ is identical with the row space of U. Another application of this type of projection operator is in the solving of normal equations in which the solutions satisfy predetermined constraints.

Some new results in a study of Σ quantities, as generally defined by Zyskind in 1958, and their relationships to general design structures, can be reported by graduate student John Kinney, working under the direction of Dr. Zyskind.

Balanced experimental structures involving any number of factors A, B, C, . . . can be completely characterized by the various nesting and crossing relationships which obtain among these factors. Thus (S:P) (T) is one way of denoting the population structure of the usual randomized block. Far more general structures have been studied and it becomes important to specify the structures in a standard way. By modifying the Hasse diagrams used by Throckmorton (1961) one or more standard specifications of all structures involving as many as five factors can be made and work is proceeding on those with more than five factors.

Research is continuing on generalized polykays, a special effort being made to relate the occurrence of generalized polykays in design problems to that of polykays in sampling contexts as shown by Fisher (1928), Tukey (1950, 1956) and others. The emphasis of the current work is on unifying, if possible, the occurrence of polykays in statistically different contexts and providing a single exposition of all relevant work in the field.

A new result linking population cumulants with products of moments and cumulants of lower order has been obtained and its implications are currently under investigation.

Abel Mexas, working part-time on the project under the direction of Dr. Zyskind, has continued his investigations on the analysis of variance of balanced complete experimental data on a digital computer. He has also developed some formulae for counting the number of distinct balanced experimental structures with a given number of factors. In the case of classification data of less than maximal rank, progress was made in obtaining generating sets of functions for the set of linearly estimable functions within each factor of classification. A detailed description of the results is given in Mexas' doctoral thesis; an abstract appears in the publications section of this report.

Takis Papaioannou, working under the direction of Dr. Kempthorne, has continued his research on the statistical theory of information and related measures of information. The results of the research have shown that a theory of statistical information can be built up on the basis of three types of functional measures of information. These are the Fisherian, the modified Kullback-Leibler and the generalized Bhattacharyya-type measures of information. The results of this research appear in Papaioannou's doctoral thesis; an abstract is in the publications section of this report.

Another part of Papaioannou's research was devoted to the improvement and finalization of the PARTAN Steepest Descent Computer Optimization program. Certain improvements regarding the efficiency of the algorithms have been obtained. The program has been tested on several optimization problems and is now operational. It has performed successfully by finding satisfactory maximum likelihood estimates in a genetic model.

Dr. Kempthorne has continued investigations into and criticisms of the various attempted approaches to the foundations of statistical inference and inductive logic. He has presented discussion of the problems in seminars, remarks on the Cornfield paper (Biometrics, 1969), and at the Waterloo symposium on Statistical Inference.

Papers and publications which have resulted from research supported by this contract are reported in the publications section of this report, listed under the names of personnel involved with this research.

USDC, Bureau of the Census Research Project

Dr. Wayne Fuller continues as director of Project Cco-9165, a continuing cooperative program between the Statistical Laboratory and the U. S. Bureau of the Census.

The efficiency of the Winsorized mean as an estimator of the mean of Weibull populations has been investigated. It has been proven that the Winsorized mean is more efficient, equally efficient or less efficient than the sample mean if the shape parameter of the Weibull is greater than one, equal to one, or less than one, respectively.

The formalization and use of "anticipated" or "prior" information in survey design and estimation

also have been investigated. In one model the investigator is willing to express his prior knowledge as a distribution for each element of the population. In fact, useful results are possible if only the means and variances are specified. Unequal probability designs and ratio-type estimators have been developed for the model.

Alternative estimators of the variance of the estimated mean for a stratified sample with one observation per stratum have been compared. Approximate expressions for the variance of the estimated variances have been obtained assuming the response variable to be a linear function of an observable concomitant variable.

Design of Experiments and Analysis of the Data

Research has continued on the development of understanding of statistical methods and their application to data. Dr. Oscar Kempthorne has directed the work, supported by Agriculture Experiment Station Project 890. During the year emphasis has been focused on ideas of likelihood, ideas of information and tests of significance, including goodness of fit tests of significance.

Research to Formulate Techniques for Estimating Farm and Related Characteristics by Sample Surveys

Professor Norman V. Strand continued as director of Agriculture Experiment Station Project 1753, which included work on the soil and water conservation inventory, and estimation of farm characteristics by sample survey.

In the soil and water conservation inventory, county and state summary tables showing land use by soil capability, and soil and water treatment needs by capability unit, were completed for more than 2200 counties in 43 states. "Camera-ready" computer runs have been completed for use in a national summary publication of results of the conservation needs inventory. Presently information on soil, slope, erosion and land capabilities is being updated.

The estimation of farm characteristics included a study on procedures using a "snowballing" scheme to accumulate lists of names for four items of low frequency of occurrence on farms. The same four states used for this snowballing study: Illinois, New Mexico, Oklahoma and Tennessee, were used for a study of problems associated with the use of rural postal route numbers and names as a frame for drawing farm samples. A revised area frame designed for convenient and unbiased drawing of farm samples was completed for Iowa.

Analysis of Data and Design of Surveys

Research previously conducted under Agriculture Experiment Station Project 1005 is now being supported by Project 1806, with Dr. Wayne Fuller continuing as project director.

A procedure for selecting nonreplacement unequal probability samples has been developed. Joint probabilities are calculated directly from relatively simple explicit expressions.

For samples of size two the unbiased estimator is always more efficient than the unbiased estimator for replacement sampling. Modifications of the procedures to permit the selection of rotating samples of size greater than two have been developed.

Publications and Professional Activities

Statistical Laboratory staff members report their research activities at professional meetings, and publish their findings in professional journals. They also participate in the various professional societies and are involved with the publication of the journals. This is a record of these activities during the past fiscal year.

Editorial collaborators for the Journal of the American Statistical Association included C. P. Cox, H. T. David, Wayne Fuller, C. P. Han, Oscar Kempthorne, Glen Meeden, Abel Mexas, B. V. Sukhatme and George Zyskind. Dr. David, Dr. Han and Dr. Zyskind were referees for the Annals of Mathematical Statistics. Biometrics referees included Paul Hinz, Louis Jensen, Dr. Kempthorne, Edward Pollak and Dr. Zyskind. Staff members who refereed for Technometrics were Barry Arnold, Dr. David, Dr. Han and Dr. Zyskind.

Dr. Kempthorne and Dr. Pollak refereed for Genetics, and Dr. Kempthorne also was a referee for the American Journal of Human Genetics. Dr. Pollak was a referee for the Journal of Theoretical Population Biology. Dr. Arnold was a referee for the Journal of Mathematical Psychology, Dr. Fuller served on the editorial council for the American Journal of Agricultural Economics and James Walsh continued as associate editor of Perceptual and Motor Skills and served on the editorial board of Psychological Reports. Richard Warren was a reviewer for The Sociological Quarterly and Dr. Zyskind was a referee for Science.

RECORD OF PUBLISHED RESEARCH

This is a record of articles published by staff members and graduate students during the past fiscal year. When the research was conducted at Iowa State but the author has since accepted a new position, his current location is listed in parenthesis after his name. Some of these publications are included in the Sta-

tistical Laboratory's Reprint Series and copies are available upon request. These are indicated by an asterisk (*).

Earl O. Wright, George W. Reinbold, **Leon Burmeister** and **James Mellon:** "Prediction of Standard Plate Count of Manufacturing-Grade Raw Milk from the Plate Loop Count." Journal of Milk and Food Technology, 33:4, 168-170. April 1970. Journal Paper No. J-6379, Iowa Agriculture and Home Economics Experiment Station, Ames, Project 1050.

Plate Loop Counts and Standard Plate Counts of manufacturing-grade raw milk samples were compared. With bacterial counts of 200,000/ml and less, the Plate Loop Count approximated the Standard Plate Count very closely. With counts greater than 200,000, the agreement was not as close. Differences between counts from the two methods, however, were not statistically significant ($P < 0.05$) within either count range.

An adjustment factor was determined to enable the Plate Loop Count to more accurately predict the Standard Plate Count of milk with a Standard Plate Count > 200,000/ml. The Standard Plate Count can be reasonably predicted from the Plate Loop Count by the formula:

Standard Plate County (predicted) = (Plate Loop Count)$^{1.04}$.

C. P. Cox: "Introduction to the Design and Analysis of Continuous Response Bioassays." Chapter V in *Statistics in Endocrinology* by Janet W. McArthur and T. Cotton, Massachusetts Institute of Technology Press. 1970.

The design and analysis of parallel line and slope-ratio assays for point and interval estimation of relative potency are discussed. Numerical examples are presented.

E. L. Jeska, J. F. Williams and **D. F. Cox:** "*Ascaris suum*: Larval Returns in Rabbits, Guinea Pigs and Mice After Low-Dose Exposure to Eggs." Experimental Parasitology, 26:2, 187-192. October 1969.

Larval recoveries from the livers and lungs were highest from commercially available guinea pigs, then rabbits and mice after inoculation with 1000 embryonated *Ascaris suum* eggs and the differences were highly significant.

The larval recoveries from the livers and lungs of the mouse indicated that the number of larvae migrating between these organs is less than the number in the rabbit and guinea pig.

Studies in mice on factors that may cause variation in numbers of recoverable larvae showed that part of the inoculated egg dose was lost due to passage of feces and this loss involves a greater proportion of a small dose. Mice which had previous infections yielded more larvae than did parasite-free mice after inoculation with 1000 eggs.

Infections produced with 1000 embryonated eggs were found to be suitable for quantitative studies on host-parasite relationships employing *A. suum*. The relating of the magnitude of kinetics of larval migration between mice and animals of the other species is not possible without correction factors.

Frank L. Rawling, Jr., D. R. Boylan and **H. T. David:** "Effect of Wall Friction in Compression-Permeability Testing." I&EC Process Design and Development, 9:2, 161-164. April 1970.

Experimental work was undertaken to measure the effect of wall friction in compression permeability testing, to account for at least part of the difficulties in correlating data from compression-permeability cells with those from constant pressure filtrations. Two compression-permeability test cells differing only in diameter were constructed, and statistically controlled experiments were made with each cell using varying cake thicknesses and constant solids pressure. Variance analysis showed that specific resistance decreased as cake thickness increased. For cakes of the same thickness in the two cells, the specific resistance was greater for the test cell of the larger diameter. These results are those expected if friction exists between the cake and the wall of test cells.

***J. R. Gebert** (University of Arizona, Tucson) and **B. K. Kale** (University of Manitoba, Canada): "Goodness of Fit Tests Based on Discriminatory Information." Statistiche Heft, 10:3, 192-200. 1969. Reprint Series No. 254, Statistical Laboratory, Iowa State University.

In this paper the authors consider some tests based on sample spacings, for which the test statistics have an information theoretic derivation. Procedures applicable to a large class of sample spacings tests are employed to show the asymptotic normality of the test statistics and to study their power.

***G. L. Ghai:** "Structure of Populations under Mixed Random and Sib Mating." Theoretical and Applied Genetics, 39, 179-182. 1969. Journal Paper No. J-6127, Iowa Agriculture and Home Economics Experiment Station, Ames, Project 1669. Reprint Series No. 252, Statistical Laboratory, Iowa State University.

The investigation reported in this paper relates to various properties of population bred by a mixture of breeding systems, namely mixed random and sib mating. Expressions have been derived which give the genotypic frequencies in any given generation in terms of the initial values. Under the mating system considered, the population will eventually become stable, having a certain amount of heterozygosis depending upon the amounts of random and sib mating. The loss of heterozygosity in successive generations has been examined for varying amounts of sib mating in the population.

The formulae have been derived giving the mean and genotypic variance in any given generation of continued mixed mating. The effect of the mating system considered on mean and genotypic variance in successive generations has been discussed in detail in case of (i) absence of dominance and (ii) complete dominance.

°**Chien-Pai Han:** "Distribution of Discriminant Function in Circular Models." Annals of the Institute of Statistical Mathematics, 22, 61-69. 1970. Reprint Series, Statistical Laboratory, Iowa State University.

Suppose X, a p x 1 vector, is an observation which is known to have come from one of two multivariate normal populations. Denote the ith population by π_i which is $N(\mu_i, \Sigma_i)$ for $i = 1, 2$. Assuming that Σ_i is positive definite and circular, an appropriate transformation is made which reduces the circular matrices to canonical forms. The discriminant function is obtained and its distribution is found to be a weighted sum of non-central chi-square distributions when the parameters are all known. A similar distribution is given when the means are unknown and the covariance matrices are known. When all the parameters are unknown, an asymptotic expansion for the distribution is obtained.

Harpal S. Bal, **Roy D. Hickman** and Robert Getty: "Age Changes in the Weight of the Ovary in Swine." Journal of Gerontology, 24:2, 157-162. 1969.

Results of an investigation of qualitative and quantitative changes in the ovary of swine are reported in this paper. Data from 84 female swine, ranging in age from one day to eight years, were used to study the relationship between age and the weight of the ovary and to provide so-called normal values for ovarian weights. Age, ovary weight and body weight relationships were approximated by using "grafted" polynomial functions and then fitting such functions by usual least-squares regression methods. By equating segmented functions and their first derivatives at preselected "join" points, the empirical function over the entire range of age was assured of being both continuous and smooth at these "join" points.

°**W. G. Hill** (Institute of Animal Genetics, Edinburgh): "On the Theory of Artificial Selection in Finite Populations." Genetics Research, Cambridge, 13, 143-163. 1969. Journal Paper No. J-6036, Iowa Agriculture and Home Economics Experiment Station, Ames, Project 1669. Reprint Series No. 242, Statistical Laboratory, Iowa State University.

The bulk of the stochastic theory of genetic selection is based on a haploid model and a diffusion equation in time and gene frequency. In the present paper, exact mathematical theory of response to selection is given for a finite diploid population, in which the selection is upper truncation selection of the best N from a progeny population of N, on the basis of individual phenotype. A probability transition matrix is obtained and this permits obtaining results of repeated cycles of selection.

The numerical evaluation of the formula is discussed and a small example given. The results from the exact method and the approximate haploid diffusion method are compared with regard to response, chance of fixation of single genes, the rate of selection advance with repeated cycles of selection, and the optimum intensity of selection. The paper closes with a discussion of the status of the theory of genetic selection in diploid populations.

°**Dean Isaacson:** "Stochastic Integrals and Derivatives." Annals of Mathematical Statistics, 40:5, 1610-1616. October 1969. Reprint Series No. 251, Statistical Laboratory, Iowa State University.

This paper deals with the question of whether or not the stochastic integral has a property analogous to the Fundamental Theorem of Calculus. That is, if $Y(t, \omega) = \int_0^t \varphi(s, \omega) dM(s. \omega)$, does

$$\lim_{\Delta t \to 0} \frac{M(t + \Delta t, \omega) - M(t, \omega)}{Y(t + \Delta t, \omega) - Y(t, \omega)} = \varphi(t, \omega)$$ and in what sense does this limit exist? In the case where $M(t, \omega)$ is a Brownian motion, we get the desired convergence in probability. However, if $M(t, \omega)$ is a right continuous square integrable martingale, we have a counterexample to show convergence may fail. The question of whether convergence holds for any continuous square integrable martingale remains open.

Dean Isaacson: "Note on the Three Series Theorem." Annals of Mathematical Statistics, 40:5, 1884. October 1969.

The three series theorem as stated and proved by Feller in his second volume, can actually be reduced to a two series theorem. The condition

$$\sum_{n=1}^{\infty} P[|Xn| \geq c] < \infty$$ is unnecessary.

°**Louis Jensen** and **Edward Pollak:** "Random Selective Advantages of a Gene in a Finite Population." Journal of Applied Probability, 6, 19-37. 1969. Journal Paper No. J-5944, Iowa Agriculture and Home Economics Experiment Station, Ames, Project 1669. Reprint Series No. 248, Statistical Laboratory, Iowa State University.

In this paper the authors consider a haploid population having two genes. It is assumed that the frequency of one of the genes is a Markovian variate; that is, the individuals in each generation depend only on the previous generation of parents. It is further assumed that two independent random processes, occurring at different stages in the life cycle, act on the

population. These two processes are: (1) A selective advantage of either genotype, occurring with equal probabilities, acts on the many infants produced in each generation. This type of selection is defined as "selection being zero on the average." (2) A random sample of the infants is chosen to form the adult population for the next generation. If the adult population is large and selection is not intense, the distribution of gene frequencies can be approximated by a continuous distribution whose probability density is the solution to the forward Kolmogorov partial differential equation. This equation is solved by using consequences of the property of completeness for a complete set of orthonormal polynomials. The solution is compared to the true solution for special cases which are obtained by powering the relevant transition matrix. The differential equation solution approximates the true solution reasonably well. It is found that the fixation or loss rates of the gene are greater than the corresponding rates given by the random sampling process alone.

The authors also consider an alternative to processes (1). The alternative is (1 ′) A selective advantage or disadvantage of a genotype, occurring with equal probabilities, acts on the infants. Other writers have called this type of selection "selection being neutral on the average." It is shown that this type of selection is not neutral on the average. As before, the differential equation solution is found and is compared to the true solution. The differential equation solution approximates the true solution exceptionally well.

Oscar Kempthorne: *An Introduction to Genetic Statics.* 545 pp., Iowa State University Press, Ames. 1970.

This book presents the probability and statistical concepts essential to study and understand genetic data of both experimental and population type.

The first eleven chapters provide an introduction to the random nature of the Mendelian process and the consequences of this randomness to the theoretical dynamics of genetic populations and to the nature of data of basic experiments. The elementary theory of infinite populations in the absence and the presence of selection, and the beginning of the dynamic theory of finite populations are presented. The treatment of inbreeding is particularly extensive with examination of all the simple systems by elementary probability arguments. A detailed presentation of generation matrix methodology for study of theoretical inbreeding is given. Statistical procedures necessary for elementary qualitative genetic data are described; in particular, estimation and testing of genetic parameters. This material is totally relevant to experimental and population genetics, and comprises a course in mathematical genetics.

Chapters 12 to 23 give a complete introduction to the statistical theory of quantitative inheritance. The basic ideas presented include analysis of variance, statistical regression, and path analysis including the method of path coefficients. The text gives a definitive account ab initio of the analysis of variance. It applies these ideas to the partition of variance of a quantitative attribute in a genetic population, leading to the concepts of additive, dominance and epistatic variance. The theory of population observables, means, variances and covariances of relatives, is developed for random mating and selfing populations. Finally, the applications of the theory to specially prepared populations and to a population under equilibrium under assortative mating are given. The role of assumption in interpretation of data is stressed throughout.

This book was originally published by John Wiley and Sons, Inc. in 1957.

***Oscar Kempthorne** and **T. E. Doerfler** (Booz-Allen Applied Research, Inc., Chicago): "The Behaviour of Some Significance Tests Under Experimental Randomization." Biometrika, 56:2, 231-248. 1969. Reprint Series No. 250, Statistical Laboratory, Iowa State University.

For the past 40 years, there has been confusion about the differences between tests of significance and tests of hypotheses to the point where data interpretation is presented as an accept-reject process. In this paper the case of inference for a simple experiment, the paired design, is discussed. A rationale is given of the significance tester, and a comparison made of three tests of significance which can be viewed as competitors in the case of this design. These are the sign test, S, the Wilcoxon test, W, and the Fisher and randomization test, R. The conclusion is that there is indeed an ordering of the tests, which is R preferred to W preferred to S. The S test should never be used if the others are possible. An appendix gives a proof of monotonicity of the R test with respect to a shift alternative.

R. S. Kleber (St. Olaf College, Northfield, Minnesota): "The Probability of Duplication in Sampling with Replacement." American Mathematical Monthly. 75:10, 1084-1087.

While formulas exist for working out the probability of duplication when given the sample size and population size, the author has worked out highly accurate asymptotic formulas for determining the sample size when given the population size and probablity, and for determining the population size when given the sample size and probability. The accuracy of these formulas was established by extensive tests run on the computer.

Richard S. Kleber (St. Olaf College, Northfield, Minnesota): "A Classroom Illustration of a Nonintuitive Probability." The Mathematics Teacher, LXII:5, 361-362. 1969.

This paper describes a method of illustrating a nonintuitive probability in the classroom, using a

generalized application of the "matching-birthdays problem," the problem of determining the probability that two or more class members will share the same birthdate.

Class members are asked to write down integers between 1 and N, where N is determined by a table showing the probability of duplicate choices for classes of different sizes. The value of N which yields a probability of 0.50 is much larger than most students anticipate.

°**Glen Meeden:** "Best Tests for Testing Hypotheses about a Random Parameter with Unknown Distribution." Annals of Mathematical Statistics, 41:2, 585-591. April 1970. Reprint Series, Statistical Laboratory, Iowa State University.

Let X be a random variable with a family of possible distributions for X indexed by $\lambda \epsilon \Omega$. λ is the realization of a random variable Λ taking values in the space Ω. Let $\mathcal{G}$ be a family of possible a priori distributions G for Λ. After observing X, we wish to test H: $\lambda \epsilon \omega$ against K: $\lambda \epsilon \omega'$ where ω is a subset of Ω and ω' its complement. To determine good tests for this problem, we use an analysis similar to the one of the Neyman-Pearson theory of hypothesis testing. The existence of uniformly most powerful level α tests relative to $\mathcal{G}$ is proved for various choices of the family of a priori distributions.

°**Ahmed Z. Memon** (West Pakistan Agricultural University, Lyallpur) and **H. T. David:** "The Distribution of Lattice Join Counts." Bulletin of the Institute of Statistical Research and Training, 2:2, 75-83. Reprint Series No. 253, Statistical Laboratory, Iowa State University.

Consider m x n points of a two dimensional rectangular lattice with M(n-1) horizontal links and n(m-1) vertical links. At every point, B or W materializes with probabilities respectively p and q. Moran (1948) has found that the distribution of the number of BB joins tends to be normally distributed for large m and n. He has also remarked (1947) that, in problems of this type, when m and n increase but the probabilities of materialization are simultaneously adjusted so as to give a finite non-zero limit for the expected number of symbols of one type, distributions similar to the Poisson can be expected to arise.

It is pointed out in this paper that precisely the Poisson distribution is obtained as limit distribution for the number N of BB joins when m, n and p are varied in such a manner as to yield a finite non-zero limit for E(N); this in accordance with the general methodology suggested previously by David and Fuchs (1964 and 1965). On the way to this asymptotic result finite moments are derived, and a slight numerical discrepancy between the second moment derived here and that given by Moran (1947) is pointed out.

°**Charles J. Mode** (State University of New York at Buffalo): "A Multidimensional Age-Dependent Branching Process with Applications to Natural Selection. II" Mathematical Biosciences, 3, 231-247. Journal Paper No. J-5761, Iowa Agriculture and Home Economics Experiment Station, Ames. Reprint Series No. 247, Statistical Laboratory, Iowa State University.

This article is a continuation of a companion report (1968) on multidimensional age-dependent branching processes. It is divided into six sections: In Section 2 the covariance functions of the process are studied; in Sections 3 and 4 the convergence in quadratic mean and the convergence with probability one of the random function $N(t) \exp(-\alpha t)$ are investigated. Section 5 is devoted to a study of the limiting distribution of genotypic frequencies; in Section 6 some applications of the model to the theory of natural selection are suggested.

°**Edward Pollak:** "Bounds for Certain Branching Processes." Journal of Applied Probability, 6, 201-204. 1969. Journal Paper No. J-5758, Iowa Agriculture and Home Economics Experiment Station, Ames, Project 1669. Reprint Series No. 249, Statistical Laboratory, Iowa State University.

In this paper a branching process is considered for which the offspring distribution has the generating function f(t) and mean $f'(1) = m < 1$. The probability that a line descended from an individual still survives in generation n is asymptotically of the form cm^n. A method is derived whereby good bounds for c may be obtained. This method makes use of the first three moments of the distribution of offspring.

A. Sen and J. K. Sengupta: "Optimal Capacity Models Under Periodic Demand: A Survey of Peak Load Pricing Theory and Appraisal." Zeitschrift Fur Die Gesamte Staatswissenschaft, 125:3, 371-395. July 1969.

The objective of this paper is to analyze and appraise some of the basic models of optimal capacity expansion under the periodic demand case. One section presents alternative cost concepts which are then used in the following section for specifying optimal pricing and capacity expansion rules under constant-load conditions. Some of the basic models of capacity expansion under periodic demand in a somewhat general context are presented in another section, where the consequences of removing some of the simplifying assumptions are noted. A brief appraisal of the models is discussed at the end of each section.

J. K. Sengupta: "Some Observations on the Optimal Growth Path for an Underdeveloped Economy." Metroeconomica, 20:2, 149-172. August 1969.

An attempt has been made in this paper to generalize the model of optimal growth path for an underdeveloped economy, which has been formalized by

Goodwin (1961) in one of his highly fundamental contributions. The generalizations involve four basic types of analytic extensions, e.g., a comparison with the Harrod-Domar type growth path and a development planning model, a multi-sector generalization of the optimal growth path, an application of the optimality principle to a model of development of a dual economy, and lastly, the policy implications of Goodwin-type optimal paths.

The main object of the generalizations has been to show the usefulness of the optimality principle of the Goodwin model in different realistic types of economic framework of an underdeveloped country. It's believed some further extensions involving linear and nonlinear programming framework would prove operationally useful for development planning.

J. K. Sengupta: "Distribution Problems in Stochastic and Chance-constrained Programming." Essay No. 18, pp. 391-424 in *Economic Models, Estimation and Risk Programming: Essays in Honor of Gerhard Tintner,* edited by Karl A. Fox, Jati K. Sengupta and G. V. L. Narasimham. Springer-Verlag. 1969.

The problems of statistical distribution of the maximand are here analyzed under stochastic and chance-constrained linear programming. Uses of non-central chi-square, truncated normal, non-central F and other non-negative distributions of statistical reliability theory are indicated. This analysis would be applicable to economic models with production coefficients which are usually required to be non-negative.

J. K. Sengupta and J. H. Portillo-Campbell: "A Fractile Approach to Linear Programming Under Risk." Management Science, 16:5, 298-308. January 1970.

The implications of a fractile approach to linear programming under risk through maximizing a given fractile of the distribution of profits under linear programming restrictions are examined here theoretically, computationally and empirically.

Jati K. Sengupta and Gene Gruver: "A Linear Reliability Analysis in Programming with Chance Constraints." Swedish Journal of Economics, 221-246. December 1969.

A linear method of reliability analysis is developed for linear programming problems with chance- constraints. Interpreting each chance-constraint as a reliability measure, the approach develops a simple computable method of incorporating optimal system reliability for the linear programming system, under the assumption that the components of the resource vector follow either the exponential or the normal distribution. This method is useful in designing LP models with desired reliability levels.

Jati K. Sengupta and Amitabha Sen: "Models of Optimal Capacity Expansion for the Firm: An Appraisal." Metroeconomica, 21:1, 1-28. 1969.

An analytic appraisal of the theories of optimal capacity expansion policies of the firm is attempted. The objective is to examine the operational implications of some of the recent growth models of the firm and to suggest some lines of generalization. This analysis is not intended to be comprehensive in the historical sense. As a matter of fact, the probabilistic aspects of capacity expansion, the theories of inventory control and the models of production scheduling are not considered here, since it is felt that those aspects require a separate treatment.

The presentation starts with a general outline of various capacity expansion models followed by a technical appraisal of their structural aspects. The models are divided into three broad groups: (a) static capacity models, (b) capacity expansion models with scalar optimization over time, and (c) capacity expansion models with optimal intertemporal paths.

J. K. Sengupta: "On the Active Approach to Linear Programming." Metrika, 15, 59-70. 1970.

The problem of statistical distribution of the optimal objective function under the active approach of stochastic linear programming is investigated in this paper from two interrelated aspects. First, the active approach is viewed as a method of decomposition. Second, some results on the asymptotic form of distribution of extreme values are utilized to derive the limiting distribution of the maximand under the active approach.

J. K. Sengupta and B. C. Sanyal: "A Fractile Programming Approach with Extreme Sample Estimates of Parameters." Statistica Neerlandica, 24:1, 51-59. 1970.

The implications of extreme value estimates of net prices in the objective function of a linear programming problem are investigated. Analytically, the approach suggests interesting points of comparison between different measures of sensitivity of a linear programming problem, when a specified fractile of the distribution of profits is optimized.

J. K. Sengupta and Karl A. Fox: "A Computable Approach to Optimal Growth of an Academic Department." Zeitschrift fur die Gesamte Staatswissenschaft, 126:1, 97-125. January 1970.

The objective in this paper is to specify and develop alternative models of growth of an academic department in forms capable of being applied as decision rules operated by the departmental chairman (or any other suitable policy maker). These models involve different degrees of sophistication in dealing with different aspects of growth. Since the growth of a system is the subject of analysis, the various mathematical models presented have the common characteristic of being essentially dynamic and basically policy oriented; this implies that the models are defined over time (short, medium or long) and that the set of variables in any model should be divisible into

two subsets: one containing the so-called instrument variables (or control variables) of the policy maker and the other containing the so-called target variables. The instrument variables are at least potentially the controlling devices which can be set at different levels by the policy maker.

For analytical convenience the various mathematical models of educational planning for specifying consistent and optimal growth of an academic department may be classified into four interrelated groups as follows:

A. Equational Models (aggregative and disaggregative with or without explicit optimization criteria);

B. Models of Programming (linear and quadratic with or without decomposition aspects);

C. Probabilistic Models (linear and quadratic with or without transition probabilities specified for the system.)

It is to be noted that these groups of models are arranged in an order of increasing complexity, increasing information and data requirements and also increasing sensitivity to errors of specification.

J. K. Sengupta: "Optimal Stabilization Policy with a Quadratic Criterion Function." Review of Economic Studies, 37:109, 127-146. January 1970.

The stabilizing and optimizing characteristics of varying government expenditures under a macrodynamic model are examined in this paper in the light of adaptive and optimal control theory. Using Pontryagin's maximum principle and a quadratic objective function, it is shown how in linear and nonlinear models optimal controls may impart instability

*Ravindra Singh and **B. V. Sukhatme:** "Optimum Stratification." Annals of the Institute of Statistical Mathematics, 21:3, 515-529. 1969. Reprint Series No. 256, Statistical Laboratory, Iowa State University.

In this paper the authors consider the problem of optimum stratification on a concomitant variable x when the form of the regression of the estimation variable y on the concomitant variable x and also the form of the variance function $V(y|x)$ are known. Minimal equations giving optimum strata boundaries have been obtained for Neyman and proportional allocations. Since the minimal equations cannot be solved easily, various methods of finding approximate solutions have been given. A comparison of approximate solutions with the exact solutions is made for certain density functions.

P. V. Sukhatme and **B. V. Sukhatme:** "On Some Methodological Aspects of Sample Surveys of Agriculture in Developing Countries." Pp. 528-562 in *New Developments in Survey Sampling* - a symposium held at the University of North Carolina. N. Y. Johnson and H. Smith, editors. Wiley Interscience, New York. 1969.

This is a review of some of the difficulties encountered and methodological problems involved in using sampling in agricultural censuses and surveys under conditions in the developing countries.

P. V. Sukhatme and **B. V. Sukhatme:** *Sampling Theory of Surveys with Applications,* Second Edition. 452 pp., Asia Publishing House, London. 1970.

The book and its Spanish translation have been extensively used by students, research workers and professional statisticians all over the world. The revised edition retains essentially the same character as the first in giving a comprehensive and systematic account of sampling theory and its applications. This account has been updated in this second edition to include recent developments in theory and practice. Examples of recent surveys have been added to provide the reader with the methodological problems one faces in practice.

Special features of this edition are: (i) a separate chapter on sampling theory with varying probabilities, (ii) an expanded chapter on non-sampling errors giving an account of experiences of agricultural censuses conducted in developing countries to show the relative contributions of the different sources of error and the need to control them within available resources, and (iii) problems based on published work at the end of each chapter.

*Hanspeter Thöni (University of Bern, Switzerland): "A Table for Estimating the Mean of a Lognormal Distribution." Journal of the American Statistical Association, 64, 632-636. 1969. Reprint Series No. 246, Statistical Laboratory, Iowa State University.

If experimental data have been analyzed using a logarithmic transformation of the actual observations, taking the antilog of the mean of the transformed variables yields a biased estimate of the mean μ of the original variable. To obtain an unbiased estimate of this mean a correction for bias must be applied. The table included in this paper provides values of a function of the sample size and variance of data transformed to logarithms to base 10 which, when added to the mean of the transformed data, yields an unbiased estimate of the mean of the original variable after transforming back. This estimate is more efficient than the mean of the untransformed observations.

Charles Blake Keasey, **James A. Walsh** and Gary P. Moran: "The Effect of Labeling as an Informational Social Influence upon Color Perception." Journal of Social Psychology, 79, 195-202. 1969.

While normative social influence primarily affects response but not perceptual processes, there is reason to believe that informational social influence affects

perceptual processes also. Indeed, this follows from the definition of informational social influence. While these two sources of social influence typically occur together, it is possible to maximize the probability that one rather than the other has been manipulated.

The present study demonstrates that when informational social influence is emphasized by keeping the influence attempt covert, the influence source absent, and the stimulus situation ambiguous, (a) Ss' initial perception of a color stimulus anchors subsequent color determinations and (b) informational social influence interacts with the Ss' initial perceptions so as to modify them significantly.

Michael J. Kavanagh, Arthur C. MacKinney and **Leroy Wolins:** "Satisfaction and Morale of Foremen As a Function of Middle Managers' Performance." Journal of Applied Psychology, 54:2, 145-156. 1970.

The relationship between satisfaction and performance was investigated from a different viewpoint, middle managers' performance and their subordinates' satisfaction, in two samples. Analysis of the first sample revealed the existence of two satisfaction item referents, individual and group, which have not been discussed in the literature. Relating these two types of items to supervisor's performance revealed positive correlations with group items but little relation with items individualistic in reference. Results of the cross validation sample provide moderate support for the main results of the validation sample. Various explanations and data which all seem to indicate that cohesiveness of work groups varied across samples are presented. Implications of the results for the job satisfaction literature and industrial morale programs are discussed.

Leroy Wolins: "Technology and Technologists in Educational Research." Didakta Medica, 1, 7-10. 1970.

The interplay between technologists in education, statistics, computer science and psychometrics is described in the context of educational research, and the application of such technology is described in the context of economic constraints in an educational system.

°**George Zyskind** and **Frank B. Martin** (University of Minnesota, Minneapolis): "A General Gauss-Markoff Theorem for Linear Models with Arbitrary Non-Negative Covariance Structure." S.I.A.M. Journal of Applied Mathematics, 17, 1190-1202. November 1969. Reprint Series, Statistical Laboratory, Iowa State University.

Given the general linear model $y = X\beta + e$ having the covariance matrix, σ^2V, of the errors, with $\sigma^2 > 0$, V known and non-negative (possibly singular), the authors specify the complete non-empty class $\mathcal{V}$ of conditional inverses of V such that, for any estimable parametric function $\lambda'\beta$ and V° in $\mathcal{V}$ a best linear unbiased estimator of $\lambda'\beta$ is given by $\lambda'\hat{\beta}$ where $\hat{\beta}$ is any solution to the general normal equations $X'V^\circ X\beta = X'V^\circ y$. Properties of the solutions $\hat{\beta}$ are presented. It is further verified that if y is distributed as a multivariate normal variable, then $\lambda'\hat{\beta}$ is the maximum likelihood estimator of $\lambda'\beta$. A procedure for testing hypotheses, using solutions to the general equations, is also presented.

°**George Zyskind:** "Parametric Augmentations and Error Structures under which Certain Simple Least Squares and Analysis of Variance Procedures Are Also Best." Journal of the American Statistical Association, 64:328, 1353-1368. December 1969. Reprint Series, Statistical Laboratory, Iowa State University.

The development of this research exhibits conditions on the form of the covariance matrix so that simultaneously for all models with a common specified linear model part every simple least squares estimator (s.l.s.e.) is also a best linear unbiased estimator (b.l.u.e.). It is shown that models with such a covariance structure may be viewed as possessing just one error term. Linear parametric compounds in the testing of which such an error term is appropriate are explored. Linear models for many complex experiments, e.g., the split-plot design, with an induced covariance structure under which all linear s.l.s.e.'s are also b.l.u.e.'s, often possess several natural error terms. The usual analysis of variance for such situations is considered along with the behavior of pertinent ratios of mean squares. Under multivariate normality the usual noncentral and central F distributions are seen to result. Certain illustrations originating from finite randomized experiments are examined. Throughout the general development full rank and nonsingularity restrictions on linear model and covariance matrices respectively are not employed.

Discussion

Oscar Kempthorne: Discussion of "The Bayesian Outlook and Its Application" by Jerome Cornfield. Biometrics, 25:4, 647-654. December 1969.

Query Answer

"Maximum Likelihood Estimate in Intraclass Correlation Model." Chien-Pai Han. Technometrics, 11:4. November 1969.

Book Reviews

T. A. Bancroft: Review of *Research Papers in Statistics*, edited by F. N. David. Technometrics, 11:4, 838-839. November 1969.

James Hutter: Review of *Introduction to Statistical Data Processing* by Theodor D. Sterling and Seymour V. Pollack. Historical Methods Newsletter, II:3, 11-13. 1969.

Oscar Kempthorne: Review of *The Correlation between Relatives on the Supposition of Mendelian Inheritance* by P.A.P. Moran and C.A.B. Smith. The American Journal of Human Genetics, 20:4, 402-403. July.

Oscar Kempthorne: Review of *Statistics: Uncertainty and Behavior* by Richard Savage. Psychometrika, 35:1, 129-133. March 1970.

Oscar Kempthorne: Review of *A Selection of Early Statistical Papers of J. Neyman.* Journal of the American Statistical Association, 65:329, 455-456. March 1970.

James Walsh: Review of *Mathematics for Statistics* by Bashaw. Educational and Psychological Measurement, 30:1, 189-190. 1970.

ABSTRACTS OF THESES

Alfonso Fausto Carrillo: "Estimation of Variance after Preliminary Tests of Significance." Ph.D. thesis. Iowa State University Library. November 1969.

Let X_{ij}, $j = 1, 2, \ldots, n_i$, be a sample of n_i independent observations drawn from $N(\mu_i, \sigma_i^2)$ with unknown mean μ_i and variance σ_i^2, $i = 1, 2, \ldots, k$. This research considers the problem of estimating a parametric function of the form

$$\xi = \sum_{i=1}^{k} w_i \sigma_i^2 ,$$

where the w_i's are known weights and the k populations are assumed to be independent. Often there is some reason to believe that the k population variances are homogeneous and under such circumstances, it is a common practice to use the estimate

$$\hat{\xi}_1 = s_W^2 \sum_{i=1}^{k} w_i ,$$

where s_W^2 is the pooled estimate of σ_i^2 based on $\sum_{i=1}^{k} f_i$ degrees of freedom, and f_i is the number of degrees of freedom of the unbiased estimate s_i^2 of σ_i^2. If the assumption of homogeneity of population variances is not likely to be true, it is customary to use the estimate

$$\hat{\xi}_2 = \sum_{i=1}^{k} w_i s_i^2 ,$$

which is of course unbiased. Obviously, $\hat{\xi}_2$ will be the most appropriate estimate of ξ when the variances σ_i^2 are all different from one another, while $\hat{\xi}_1$ will be the most appropriate estimate of ξ when the variances are homogeneous.

However, it may happen that the k variances are neither homogeneous nor different from one another. It will generally be the case that the k variances σ_i^2 can be classified into t groups $1 < t < k$, the variances within a group being homogeneous. Since the k variances σ_i^2 are unknown, their classification into t homogeneous groups can be made on the basis of preliminary tests of significance. Under such circumstances neither of the estimates $\hat{\xi}_1$ or $\hat{\xi}_2$ is appropriate.

A solution that suggests itself is to have an estimate of the type $\hat{\xi}_1$ for each of the t groups and then form a suitable weighted estimate. This new estimate based on preliminary tests of significance is proposed in this thesis for the estimation of the parametric function ξ. Expressions are given for the bias and mean square error of the estimate for the cases of two and three populations. Also, the efficiency of the estimate is discussed, with respect to the estimates $\hat{\xi}_1$ and $\hat{\xi}_2$. Finally, the theory is illustrated with some applications to survey sampling.

K. T. deGraft-Johnson: "Some Contributions to the Theory of Two-Phase Sampling." Ph.D. thesis. Iowa State University Library. August 1969.

This study deals primarily with the comparison of the biases and MSE's of several ratio-type and regression estimators in two-phase sampling. The ratio-type estimators considered are the two-phase analogues to the classical ratio, the Hartley-Ross, Pascual's, Beale's, Tin's and Quenouille's estimators. These are considered under two schemes. In the first scheme, the first phase random sample of size n' is independent of the second phase random sample of size $n(<n')$. In the second scheme, the second phase sample is a simple random subsample of the first phase sample.

The comparisons of biases and MSE's are done by three main methods: asymptotic expansions to $O\left(\frac{1}{(n')^2}\right)$, model assumptions (linear and quadratic regression models) and Monte Carlo simulation.

In addition to this main study, the properties of Srivastava's estimator are examined and its relationship to the regression estimator briefly discussed. Also, the use of the first phase sample to determine the optimal stratification boundaries for the selection of the second phase sample is considered. The method used involves adapting techniques previously developed for a continuous random variable to the case of a discrete random variable.

Using two-phase sampling, the problem of estimating domain means is also discussed. The objective is to subsample to minimize the maximum of their variances. Thus, an unequivocal sampling rule for selecting the second phase sample from the first phase sample is obtained. Methods for choosing the optimal first and second phase sample sizes are also given.

John Jeffery Goebel: "Estimators for the Population Size Based on a 'Snow Balling' Procedure." M.S. thesis. Iowa State University Library. August 1969.

The "snow balling" procedure is a procedure for compiling a list of the members of a human population about which little is known. The procedure utilizes a starter list of persons who are believed to be members of the population, or knowledgeable about members of the population (e.g., the subscribers to a magazine related to the product of interest). These individuals are interviewed for two purposes: (1) to determine if they actually belong to the population, and (2) to list other individuals who may belong to the population. New individuals listed during the interviews are similarly questioned; the procedure continues until a cutoff point is reached. Since typically the entire population will not be identified, an estimator is needed for the percentage of the population that is identified.

In this thesis, methods of estimating the unknown population size are discussed. Estimators are developed by comparing the "snow balling" procedure with the capture-recapture methods used in ecology. Comparisons are also made with the "unknown number of balls in an urn" problem. It is assumed in the derivation of the estimators that sampling is random. The validity of this assumption is discussed.

The estimators in this thesis are applied to four surveys conducted in 1968 by the Statistical Laboratory in cooperation with the United States Department of Agriculture: (1) Christmas tree growers in northwest Illinois, (2) apple tree growers in northwest and west central New Mexico, (3) beekeepers in southwest Oklahoma, and (4) beekeepers in eastern Tennessee.

Bonnie Roberts Hanson: "Computational Methods for Model Building in Multiple Linear Regression." M.S. thesis. Iowa State University Library. August 1969.

The purpose of this thesis is to look at four model building procedures, to consider their computer applications, and to describe a statistical computer package which was developed to provide the user with the means for applying any or all of these procedures.

Model Building Procedure I uses small residual sum of squares as a selection criterion and identifies equations which contain a specified number of independent variates and have a small residual sum of squares when compared to all other possible equations having the same number of independent variables. Procedures II and III are based, respectively, on the Forward Selection and Sequential Deletion algorithms discussed by Larson and Bancroft (1963). In Procedure II the experimenter begins with a basic core prediction equation and adds independent variates to it by order of importance. In Procedure III the experimenter begins with the full k-variate prediction equation and deletes the least important variates, in order, from the end of the equation. Procedure IV selects a subset of the independent variates for inclusion in the regression equation in a stepwise manner.

The detailed description of the computer package which is found in this thesis should provide an invaluable source of information to the personnel who perform the upkeep on the system.

Roger Gerald Haugen: "A Statistical Study of Birth, 30-Day and 90-Day Weights in Lambs." M.S. thesis. Iowa State University Library. May 1970.

The purpose of this study was to measure the influence of accountable environmental factors and develop efficient adjustment techniques so that genetic differences and parameters could be more accurately estimated for birth, 30-day and 90-day weights of lambs.

The environmental sources of variation included sex, type of birth-rearing and age of dam. Least squares analyses of these independent variables and their two-way interactions revealed that none of the interactions accounted for more than 0.5% of the variance for each weight. All the main effects were statistically significant within each weight except for sex in 30-day weight. Of the environmental effects studied, type of birth and rearing was the most important source of variation at all three weights.

Least squares analyses were conducted to determine the significance of genetic-environmental interactions. None of the interactions accounted for more than 1.5% of the total variation within each weight and were therefore ignored in subsequent analyses. The data were then corrected for sex, type of birth-rearing and age of dam using additive correction factors.

A nested analysis was conducted to estimate genetic parameters from adjusted data. Phenotypic, genetic and environmental correlations among birth, 30-day and 90-day weights were all found to be positive.

Charles Dean Ingwell: "A Symmetric Binomial Sequential Design." M.S. thesis. Iowa State University Library. August 1969.

This thesis is concerned with "two-source" sequential binomial design (selection) of experiments and specifically addresses itself to two problems: explicit computation of absorption probabilities and expected step counts, and characterization of the Bayes solutions.

A fairly complete exposition of binomial sequential design of experiments (BSDE) is given and the relation between SPRT's and sequential Bayes solutions in the more usual one-source case are discussed. Explicit solutions are found for absorption probabilities and expected step counts by extending the difference equation approach familiar in the one-source case.

Two two-source SPRT plans are discussed. The first has the general character of Bayes solutions for BSDE problems, but does not have a certain Bayes

property. The second has the Bayes property lacking in the first.

These two plans provide the first, as far as is known to the author, indication that expensive reliable sources are preferable when ignorance is greatest.

William Jo Kennedy, Jr.: "Model Building for Prediction in Regression Analysis Based on Repeated Significance Tests." Ph.D. thesis. Iowa State University Library. November 1969.

A research worker who wishes to use regression to analyze his experimental data may be faced with several problems. The problem of a research worker who is sure that the usual assumptions about error in the full model are applicable in their situation, but is not sure which subset of the total set of independent variates might best be used in a final model, is considered in this thesis.

The purpose of this research is to provide more theoretical basis for some of the commonly used model building procedures and to make recommendations which will allow research workers to use model building procedures to better advantage. It is assumed that the primary objective of the research worker is to obtain a suitable prediction equation.

Two different techniques for model building are studied: the sequential deletion procedure and the forward selection procedure. For each procedure the bias and mean square error of predicted y is derived for the case of unknown error variance. Employing the concept of relative efficiency, the results are given of an extensive numeric study to compare the two procedures with one another, and with the procedure in which all independent variates are used. Recommendations are made as to which model building procedure is best in various circumstances and which level of test of significance to use when applying the model building procedure.

Finally, the problem of interval estimation is considered and confidence limits for the expected value of the predicted y are obtained in the case of known error variance.

Abel Geber Mexas: "Some Computational Aspects of Linear Classification Models." Ph.D. thesis. Iowa State University Library. February 1970.

This study presents various computational problems that arise in the use of linear classification models. The algebraic properties of structures are exhibited and exploited in the development of an algorithm for the analysis of variance of balanced complete data. The algorithm presented is aimed at efficient computations and is described entirely in terms of logical Boolean operations.

A formula is derived for the number of admissible sets corresponding to an arbitrary structure. Classification structures containing a fixed number of factors are given a characterization which facilitates the derivation and evaluation of formulas for the number of possible non-isomorphic structures of various classes. In particular, if q is the number of factors in a structure, nested in factors other than the mean, then for q =2, 3, expressions are derived which allow the computer to determine the number of non-isomorphic structures for any arbitrary N. The result for q = 2 is expressible as

$$M_{N,2} = \sum_{i=1}^{N-2} \sum_{j=1}^{N-1} \sum_{j'=\max(0,\ j-1^1-1)}^{\min(i,\ N-i-2)} .$$

This expression manifests that the structures with q=2 are isomorphic to integer 3-tuples suitably restricted. The results of the computer evaluation of the expression led to a conjecture on the form of the second forward difference, which in turn yields the polynomials

$$M_{N,2} = \begin{cases} 1/24\ (2N^3 + 15N^2 - 62N + 48) & \text{if N is even} \\ 1/24\ (2N^3 + 15N^2 - 62N + 45) & \text{if N is odd.} \end{cases}$$

Expressions were obtained for q=3 and the results of computer evaluations are tabled for $N \leq 30$.

Various computational methods are presented for analysis of factorial arrangements with arbitrary incidence. This work manifests the problem of determining the estimable parametric functions within each factor of classification that may be estimated with data arbitrarily arranged. Several counter-examples are presented which illustrate that existing computer algorithms yield incorrectly the analysis of arbitrary arrangements.

Some conditions equivalent to maximal rank in a two-way classification are exhibited. These include connectedness of the arrangement, irreducibility of the incidence matrix and the strong connectedness property of the graph associated with certain incidence matrices. It is demonstrated that these equivalences do not generalize to higher-way arrangements, but that the concepts are useful in defining an algorithm which identifies all the linear parametric functions within any factor of classification that can be estimated from a data set of arbitrary incidence.

Kenneth Stephan Mount: "Minimax Properties of Likelihood Ratio Tests Related to Goodness of Fit." Ph.D. thesis. Iowa State University Library. August 1969.

This thesis is concerned with optimality properties, both small sample and asymptotic, of likelihood ratio procedures for goodness of fit testing problems. In this connection, a testing problem is related to goodness of fit when the distributions of null and alternative hypotheses differ in "shape."

While considering finite samples, a certain power-slope criterion for goodness of fit testing problems is introduced. This amounts to working with the familiar risk which is the sum of the probabilities of the errors of first and second kinds. Certain Neyman-Pearson procedures are shown to be minimax for this risk. The stringency of these tests is also considered, and the use of a risk equal to the probability of the error

of the second kind, this occurring when the probability of the error of the first kind is fixed.

In order to analyze these tests asymptotically, the risks are rewritten as a sum of bivariate probabilities. This leads to a variety of multivariate large deviations which depend in large part on the properties of cumulant generating functions. The first order portions of relevant results of Bahadur and Rao are extended to the bivariate case. In addition, an aspect of large deviations theory for exponential families is presented.

Having these large deviations results, a condition is given that insures asymptotic optimality, and hence asymptotic minimaxness, of a natural likelihood ratio procedure, $\tilde{\delta}$, in a fairly strong sense. As before, the risk used is the sum of the probabilities of errors of the first and second kind. The approach is illustrated by two examples.

Finally, exponential families and hypotheses that are disjoint subsets of the natural parameter space are considered. Using the same risk, it is shown that $\tilde{\delta}$ is asymptotically minimax in a somewhat weaker sense.

Douglas E. Murfield: "Procedures for Soybean Yield Forecasts." M.S. thesis. Iowa State University Library. November 1969.

The objective of this thesis was to develop a model for forecasting soybean yield. Forecasts were constructed for three dates: August 1, September 1 and October 1. The forecast was composed of three components: plants per standard unit, pods with beans per plant, and bean pod weight.

Research data were collected monthly beginning on August 1 and continuing until harvest in 1964, 1965 and 1966 from soybean objective yield surveys of the United States Department of Agriculture in Illinois and Iowa. The sample fields were selected with probabilities proportional to the soybean acreage reported in the June enumerative survey.

For all maturity categories on August 1, the forecasting model for plants per standard unit was linear. Total plant count within the sampling unit was the recommended independent variable. Parameters were estimated by regressing August 1 plant counts on harvest plant counts. For September 1 and October 1 forecasts, plants with fruit at harvest were forecast as a constant percent of the observed count.

The equation for forecasting yield based upon the components is:

$[(PSU)(PBP)(PW)(\alpha)][1\text{-}HL]$ = Bushels per acre

PSU = Forecast plants per standard unit

PBP = Forecast pods with beans per plant

PW = Average weight of beans per pod at harvest

α = Factor to convert grams per standard unit to bushels per acre.

HL = Historic percent harvest loss.

Panagiotis Constantine Papaioannou: "On Statistical Information Theory and Related Measures of Information." Ph.D. thesis. Iowa State University Library. May 1970.

This study presents a small sample statistical information theory based on three measures of information. It provides also an answer to the problem of constructing functional measures of information covering non-regular distributions.

Twelve statements delimit the concept of statistical information. The following properties are required to be possessed by any measure of information: non-negativity, additivity under independent observations, maximal information, and invariance under sufficient transformations.

The Fisherian information measure is examined in the scalar and vector parameter cases. Conditional information is introduced. It is noted that Fisher's measure does not remain invariant under reparametrizations of the model. The significance of reparametrizations yielding constant information is discussed. Fisher's theory of information is applicable if three regularity conditions are met. The crucial one is: $d[\int_A f(x, \theta)d\lambda]/d\theta = \int_A \partial f(x, \theta)/\partial\theta d\lambda$ for all $\theta \in \Theta$ and every measurable set A, $f(e, \theta)$ is a parametric density and Θ the parameter space. These conditions are found to be satisfied for exponential distributions but not for distributions whose range depends on θ and whose graph does not touch the x-axis.

The trace and the determinant of Fisher's information matrix are suggested as measures of information in the multiparameter case. The former is found to have the desired properties and the latter all but the additivity one. Examples are presented.

Metrics or almost metrics on spaces of probability measures P_θ, $\theta \in \Theta$, and 1 - 1 mappings, k, from Θ onto itself are used to construct new information measures. The modified Kullback-Leibler functional measure of information is given by $I_X^{KL}(\theta;k) = \int f(x, \theta)\ln[f(x, \theta)/f(x, k(\theta))]d\lambda$ provided that $P_\theta << P_{k(\theta)}$ for all θ. It is shown to possess the desired properties and to be invariant under parametric transformations. It provides the appropriate measure for some of the non-regular models for which the Fisherian theory is not applicable. It fails for other non-regular models because of the absolute continuity condition.

The generalized Bhattacharyya functional measure is given by $I_X^B(\theta;k) = -\ln p^*_X(P_\theta, P_{k(\theta)})$, where $p^*x(P_\theta, P_{k(\theta)} = \int[f(x, \theta)]^{1/p}[f(x, k(\theta))]^{1q}d\lambda$, $p, q \geq 0$, $1/p + 1/q = 1$. The following propositions on the affinity, p^*_X, are shown to hold: (i) for any statistic T, $p^*_X \leq p^*_T$ with equality if T is sufficient for θ, (ii) $p^*_{X_1X_2} = p^*_{X_1} \cdot p^*_{X_2}$ if X_1, X_2 are independent. It is found that $I_X^B(\theta;k)$ satisfies the basic properties of information theory. It is free of regularity conditions and invariant under reparametrizations. It covers all models covered by the previous measures and, more-

over, models like the uniform (θ-1, θ+1) family or the Pitman trapezoidal distributions.

The additivity property has been criticized by several authors because it seems to give improper results for certain pathological families of distributions in which θ is estimated without error. Examination of typical cases indicates that the difficulties may arise in regard to the ideas of conditional information. Some of the problems appear pathological from certain viewpoints.

Nell Sedransk: "Contributions to Discriminant Analysis." Ph.D. thesis. Iowa State University Library. November 1969.

The discrimination problem treated in this thesis considers only two populations, π_1 and π_2, assumed to be p-dimensional normal with mean vectors, μ_1 and μ_2, and common covariance matrix Σ. To estimate any unknown parameters, previously obtained observations x_{ij} from π_i are used ($j = 1, 2, \ldots, n_i$; $i=1, 2$). It is required that a linear function of a single observation X be used to determine the assignment of X to either π_1 or π_2. When no parameters need to be estimated, there is a single admissible solution (assuming no *a priori* probabilities or origin are available): to assign X to π_1 and π_2 as the function $V = [X - \frac{1}{2}(\mu_1 - \mu_2)]'\Sigma^{-1}[\mu_1 - \mu_2]$ is positive or negative. The usual discrimination function for the case when μ_1, μ_2 and Σ are unknown is obtained by substituting the usual sample estimates for μ_1, μ_2 and Σ, giving $W = [x-\frac{1}{2}(\bar{x}_1-\bar{x}_2)]'S^{-1}[\bar{x}_1 - \bar{x}_2]$. In these processes the "discrimination point" is zero.

Part I of this dissertation deals with the choice of the discrimination point. The results obtained indicate that a discrimination point of $-\frac{1}{2}(n_1^{-1} + n_2^{-1})(p-2)$, where p denotes the dimensionality of π_1 and π_2, reduces the combined probabilities of misclassification from those associated with a discrimination point of zero.

Once a discrimination procedure has been determined, the probability of a particular kind of error may be of interest. Part II deals with the problem of estimating such a (conditional) probability. At least fifteen different estimators appear in the literature; the most promising are investigated here using unconditional mean square error as the criterion for comparison. (Because of the intractability of the precise distribution of W, the results obtained are based on an asymptotic expansion for the distribution.) Estimators requiring fewer initial assumptions (notably, the normality of π_1 and π_2) have larger mean square errors than do those motivated by knowledge of the distributions involved.

The topic treated in Part III is the characterization of populations π_1 and π_2 to simultaneously reduce the numbers of dimensions to be considered and to differentiate as sharply as possible between π_1 and π_2. It is no longer assumed that the populations are normally distributed; and complex-valued variables are considered. A defense of the extension of usual methods to the complex case is presented.

Jan Philip Shoemaker: "Application of General Linear Model to Specific Tests and Estimates." M.S. thesis. Iowa State University Library. August 1969.

The emphasis of this thesis is on the development of a procedure to handle the analysis of variance for the unequal cell size experiment. This procedure is designed to make use of a type of computer program which is available to most research workers. It is also designed to easily facilitate the estimating and testing of specific comparisons of the model effects. Finally, it is designed so that the basic structure of the procedure can be extended to an n-way classification model.

A regression procedure was developed for the one-way classification model. This procedure was then expanded to the two-way classification model.

A system using five models is outlined for the two-way classification model when performing the analysis by this regression procedure. This system minimizes the amount of hand calculations necessary to perform the complete analysis of variance. This system also makes it possible to apply this procedure to models where it is assumed that the sum of the unweighted effects is zero. Finally, this system makes it easier to understand the analysis when applying this regression procedure to an n-way classification model.

Vincent Anthony Sposito: "Aspects of Duality in Linear Programming." Ph.D. thesis. Iowa State University Library. May 1970.

This study extends the concept of duality in linear programming to problems with cone-type domains. Specifically, for problems of the form

$$\begin{aligned} &\text{minimize } b'y \\ &\text{such that } A'y - c \in L_2^* \\ &\qquad\qquad y \in L_1^* \end{aligned}$$

where y is an m-dimensional variable, L_1^* and L_2^* are closed convex cones. Hence, if $L_1^* = L_2^* = Q^+$: positive orthant, then the classical linear programming problems will be a special case of this study. It is shown that for the class of primal-dual problems with non-degenerate cone domains, subject to a certain extension of Slater's regularity condition for convex programs, the fundamental theorems of linear programming hold as well as all the principal theorems of duality.

As first pointed out in the Kuhn-Tucker theory, the development of the classical duality theory is related to the determination of a saddle-point of the Lagrangian function, $\emptyset(y, x) = b'y + c'x - y'Ax$. Hence, if (y^o, x^o) is a saddle-point solution of the Lagrangian, then y^o and x^o, respectively, are optimal solutions of the primal and dual problems. This principle applies as well, in part, to nonlinear problems where the "y-

P. Papaioannou: "On Statistical Information Theory and Related Measures of Information," at Case Western Reserve University, Cleveland, Ohio, November 24.

P. Papaioannou: "On Statistical Information Theory and Related Measures of Information," at the University of Waterloo, Waterloo, Ontario, Canada, November 25.

P. Papaioannou: "On Statistical Information Theory and Related Measures of Information," at the University of Georgia, Athens, December 11.

P. Papaioannou: "Some Information Theory Aspects of Statistical Inference," at the IMS-ASA central regional meeting in Dallas, Texas, April 7-9.

Edward Pollak: "On Survival Probabilities for Some Galton-Watson Processes," at joint meetings of the ASA, Biometric Society (ENAR and WNAR) and IMS in New York City, August 19-22. Abstract 1641 in Biometrics, 25:4, 799. December 1969.

Justus Seely and **George Zyskind:** "Vector Spaces, Estimability and Bestness for the General Mixed Linear Model," at joint meetings of the ASA, Biometric Society (ENAR and WNAR) and IMS in New York City, August 19-22. Abstract 1650 in Biometrics, 25:4, 802-803. December 1969.

Vincent Sposito: "Solutions of Linear Programming Problems with M.P.S.," at a Computer Science seminar at ISU, April 7.

Vincent Sposito: "The Use of READCOMM and other M.P.S. Procedures," at a Computer Science seminar at ISU, April 9.

Richard D. Warren, Gerald E. Klonglan and Medhat M. Sabri: "The Certainty Method of Scoring: Its Application and Usefulness in Developing Empirical Measures in Social Sciences," at the annual meeting of the Rural Sociological Society, San Francisco, August 28-31.

Richard Byerly, Trevor G. Howe and **Richard Warren:** "A Case for Network (Path) Analysis," at the American Educational Research Association meeting in Minneapolis, March 2-6.

PARTICIPATION IN PROFESSIONAL ACTIVITIES

Dr. Barry Arnold attended statistical society meetings in New York City August 19-22, and the annual meeting of the American Psychological Association in Washington, D.C., August 31-September 4.

Dr. T. A. Bancroft assumed the presidency of the more-than-10,000-member American Statistical Association on January 1, after serving a year as president-elect. The Association is organized to foster, in the broadest manner, statistics and its applications, to promote unity and effectiveness of effort among all concerned with statistical problems, and to increase the contribution of statistics to human welfare. Dr. Bancroft continues to serve as a member of the ASA Advisory Committee on Statistical Policy to the Bureau of the Budget.

Dr. Bancroft was chairman of a luncheon meeting of chairmen of statistics departments at the statistical society meetings in New York City August 19-22. He has been named a member of the steering committee for the 1971 conference of the International Statistical Institute, and a member of the IMS committee on Summer Institutes for Teachers. He continues as a member of the Biometry Committee organized by the Veterans' Administration to work on the establishment of a training grant program in biostatistics.

Professor C. P. Cox continues as chairman of the ASA Biometrics Section. He chaired a meeting of the section and served on the program committee for the statistical society meetings in New York City August 19-22. Professor Cox will teach "Statistical Methods for Biological Assays" at the 12th summer session of Statistics in the Health Sciences at the University of Washington, Seattle, June 22-31.

Dr. H. T. David was renamed a member of the organizing committee for the seventh annual visiting lecturer program in statistics sponsored by the ASA, Biometric Society and IMS. Dr. David has been appointed to two IMS committees: the Committee on Nominations and the Committee on Summer Research Institutes.

Dr. Roy D. Hickman has been elected to a three-year term as representative from the College of Sciences and Humanities to the ISU chapter of Phi Kappa Phi, national scholastic honorary. He attended the annual meeting of the American Educational Research Association in Minneapolis, March 3-4.

Dr. Paul Hinz attended the "Natural Resources Biometry Symposium," sponsored by the National Science Foundation, at Fort Collins, Colorado, April 20-24.

Dr. D. K. Hotchkiss attended the Iowa Academy of Science meetings in Waverly, April 24.

Dr. D. V. Huntsberger served during the year as president of the Iowa State Chapter of Sigma Xi. He was honored with an ISU faculty citation at the alumni convocation June 6. He was cited for "outstanding leadership in developing the statistics program for undergraduates." Dr. Huntsberger is serving on the executive committee of the ASA Section on Training of Statisticians.

Dr. James Hutter was in Chicago for the Midwest Political Science Association convention, April 30-May 4.

Louis Jensen chaired a session on "Contributed Papers-Biometrics I: Genetic Statistics" at the statistical society meetings in New York City, August 19-22.

Dr. David Jowett attended an International Symposium on Statistical Ecology at Yale University, New Haven, Connecticut, August 23-28.

Dr. Oscar Kempthorne served as discussant and was chairman of a session at the statistical society meetings in New York City, August 19-22. He also served as a discussant of many papers and was chairman of a session at a symposium on Foundations of Statistical Inference at the University of Waterloo, Waterloo, Ontario, Canada, April 1-11.

Dr. Richard Mensing served during the year as president of the Iowa Chapter, ASA.

Dr. Edward Pollak was chairman of a session on

Mathematical Genetics at the statistical society meetings in New York City, August 19-22.

Dr. Basilio Rojas has been appointed chairman of the statistics section for the 1971 Congress of the International Society of Sugar Cane Technologists which will meet in Louisiana.

Dr. Richard Warren served as chairman of a session on Innovations in Sociological Measurement at the annual meeting of the Rural Sociological Society in San Francisco, August 28-31.

Teaching

The Department of Statistics in the College of Sciences and Humanities continues to offer B.S., M.S. and Ph.D. degrees. All statistics programs are designed to emphasize the close relationship between sound application and modern statistical theory. Each statistics major is built around a common core of courses in both theory and methods, with supplementary courses which fit individual students' backgrounds and interests.

Three new courses were offered for the first time during the present fiscal year: 499X, Business Administration Statistics; 599X, Intermediate Statistical Methods in Education; and 639, Stochastic and Continuous Programming. Sociometric Statistics, previously offered as 599X, now has been numbered 508 and is listed in the catalog.

Business Administration Statistics was designed primarily for business administration majors, as requested by the Department of Industrial Administration. It combines information presented in Principles of Statistics (201), and Elementary Business Statistics (327), with special application to business situations. A total of 224 students enrolled during the three quarters the course was offered.

In a similar situation, the College of Education requested a course in statistical methods in education, which drew an enrollment of 17 when it was first offered spring quarter. Classwork includes a study of generalized linear regression, specialized regression techniques, analysis of variance, trend analysis and repeated measurements, multiple correlation analysis, covariance and discriminant analysis.

Five students were in the first Stochastic and Continuous Programming class. Course work included distributions of game values and program optima, generalized Tchebycheff inequalities and continuous programs.

This year for the first time, a statistics laboratory course was offered off-campus. Statistics 401, which presents the role of statistics in research and includes an introduction to the methods of analyzing data from experiments and surveys, was taught spring quarter in Creston on a once-a-week basis for a ten-week period. The laboratory session, normally a part of the course, was organized during the last 45 minutes of the class. Students wrote basic instructions for resolving statistical problems on the computer. This computation work was brought to Ames for key punching and processing and the corrected results were mailed back to the students for review before their next class.

In another extension offering, Introduction to Theory of Probability and Statistics was taught one day per week during the winter to a special group of employees of Collins Radio Company in Cedar Rapids. It is expected that these off-campus courses will continue as the university strengthens its extension programming to make information available to the largest number of persons.

The OMNITAB programming system which was adapted for use on ISU's IBM 360/65 by Statistical Laboratory personnel, has been adopted for teaching statistical methods in many introductory statistics courses. This procedure introduces students to the computer at an early stage in their training, and prepares them for more advanced statistical use of computing facilities.

COURSE OFFERINGS IN STATISTICS

The courses offered by the Department of Statistics during the academic year 1969-70 were as follows:

	Courses for Undergraduate Students Only		
201,	**Principles of Statistics**	**5 FWS***	Meeker, Tang, Veale
201A,		3 WS,SS_1	Huntsberger, Lin, Madsen, Meeker, S. Sukhatme, Tang
201B		3 FS	Linn, Madsen, Pyne, Tang
327	Elementary Business Statistics	3 F	Jobson
341,	Introduction to Theory	3 F	Huntsberger
342,	of Probability	3 W	Huntsberger
343	and Statistics	3 S	Huntsberger
380	Statistical Applications of Digital Computers	3 FWS	Smith, West
	Courses for Graduate Minors and Undergraduates		
401,	Statistical Methods for Research Workers	4 FW,SS_1	C. P. Cox, Dickinson, Hickman, Hotchkiss, Hutter, Jowett, Meeden, Mensing, Rojas, S. Sukhatme, Veale, Walsh, Warren

*Because the fiscal year began July 1, 1969, and ended June 30, 1970, the courses taught in the second summer session of 1969 through the first summer session of 1970 are reported here. Symbols indicate the quarter each course was taught: SS_2 - Second Summer Session, F - Fall, W - Winter, S - Spring, SS_1 - First Summer Session.

402		4 SS_2WS	C. P. Cox, Dickinson, Hickman, Hotchkiss, Huntsberger, Hutter, Jowett, Meeden, Walsh, Warde, Warren
411	Experimental Design for Research Workers	3 S,SS_1	Dickinson, Jowett, Rojas
421	Survey Design for Research Workers	3 SS_2S	Baker, B. Sukhatme
431	Elementary Statistical Quality Control	3 S	Mensing
436	Genetic Statistics for Research Workers	3 S	Jowett
446,	Statistical Theory for	3 F	Hinz
447,	Research Workers	3 W,SS_1	Hinz
448		3 SS_2S	Hinz, Mensing
481,	Processing of Statistical	2 W	Mosier
482	Data	2 S	Mosier
499	Special Problems	Arr. SS_2WS, SS_1	D. Cox, Hickman, Jowett, Smith, Strand
499H	Special Problems, Honors Program	2 W	Warren
499X	Business Administration Statistics	5 FWS	Jobson

Courses Primarily for Graduate Students, Major and Minor

501	Intermediate Statistical Methods	3 F	Bancroft
506	Factor Analysis	3 F	Walsh
508	Sociometric Statistics	3 F	Warren
511,	Design of Experiments	3 W	Kempthorne
512		3 S	Zyskind
521,	Design of Surveys	3 W	Han
522		3 S	Han
531	Industrial Statistics: Sampling Inspection	3 F	Mensing
535	Biological Statistics	3 S	C. P. Cox
536,	Genetic Statistics	3 F	Pollak
537		3 W	Pollak
538	Econometric Statistics	3 F	Fuller
539	Operations Research	3 W	Mensing
540	Operations Research Methods and Economic Analysis	3 S	Sengupta
541,	Theory of Probability	3 F	Arnold
542,	and Statistics	3 W	Meeden
543		3 S	Arnold
544	Statistical Decision Theory	3 S	David
545	Stochastic Processes	3 SS_2	Isaacson
580	Computational Techniques in Statistics: Methods	3 W	Kennedy
581	Computational Techniques in Statistics: Theory	3 S	Kennedy
599	Special Topics	Arr. SS_2FWS,	
	A. Theory	SS_1	Bancroft, Han, Kempthorne, Meeden, Mensing, S. Sukhatme, Warren
	B. Methods		Bancroft, C. P. Cox, Fuller, Hickman, Hotchkiss, Jowett, Pollak, Sposito, Walsh, Warren
	C. Design of Experiments		Zyskind
	D. Design of Surveys		Hickman
599X	Intermediate Statistical Methods in Education	4 S	Hickman

Courses for Graduate Students, Major and Minor

601	Advanced Statistical Methods	3 F	C. P. Cox
621	Advanced Design of Surveys	3 W	B. Sukhatme
622	Seminar on Design of Surveys	3 S	B. Sukhatme
638	Advanced Econometric Statistics	3 W	Fuller
639	Stochastic and Continuous Programming	3 F	David
641	General Theory of Linear Hypothesis	3 F	Zyskind
642	Probability and Distribution Theory	3 W	Isaacson
643	Theory of Estimation and Testing of Hypotheses	3 S	Meeden
647	Multivariate Analysis	3 F	Han
648	Seminar on the Theory of Statistics and Probability	3 S,SS_1	Arnold, Isaacson, Meeden
649	Theories of Inference	3 SS_1	Kempthorne
699	Research	Arr. SS_2FWS, SS_1	Bancroft, C. P. Cox, David, Fuller, Han, Isaacson, Jowett, Kempthorne, Mensing, Pollak, B. Sukhatme, Warren, Zyskind

SUMMER INSTITUTES

NSF Institute

Thirty-one mathematics teachers from secondary schools throughout the United States participated in a summer institute, "Probability, Statistics and Computer Science," sponsored by the National Science Foundation.

The institute, which met June 30 - August 8, focused on statistics and computer science, with the aim of enabling teachers to acquire knowledge and competency in teaching these areas at the high school level. Ten graduate credit hours were granted for completing the program.

Three courses were included. Statistical Methods was an introduction to statistical concepts used in the

evaluation of experiments and sample surveys, illustrated by examples related to engineering, the biological and social sciences.

Theory of Probability and Statistics focused on concepts of probability and distribution theory, leading to an introduction to the mathematical basis of modern statistical methods.

Components of modern high-speed computers, logic of operation and problem-oriented programming were included in the Computer Organization and Programming course. Emphasis was on solution of statistical problems using the computer.

Dr. T. A. Bancroft served as director of the institute, with Dr. Roy Hickman and Dr. D. K. Hotchkiss assisting with administrative responsibilities. Participants were required to have at least a bachelor's degree with a major or strong minor in mathematics, and be currently employed teaching mathematics in secondary school.

A second NSF-sponsored institute began June 29, following the same format as the highly successful 1969 session. It will be reported in next year's Annual Report.

Survey Demonstration Course

Eight students from four foreign countries participated in a Demonstration Course in Survey Methods conducted by the Statistical Laboratory during the second summer session.

This was the sixth year for such a summer program, which gives the students an opportunity to perform all the steps of a small-scale sample survey while studying survey design and methodology.

Participants were from Indonesia, Korea, Lebanon and Thailand. Dr. Roy Hickman directed the course, assisted by Mrs. Hazel Cook, Harold Baker and other members of the Statistical Laboratory's Survey Section.

GRADUATE STUDENTS

Information released by the American Statistician in its October issue showed that Iowa State University ranked first in the number of Ph.D. degrees in statistics awarded during 1968-69. The Statistical Laboratory has consistently been among the leaders in the awarding of graduate degrees.

Four graduate students represent statistics in the Graduate Student Senate, which became an official organization March 9. The Senate is defined in its constitution as "an elected body through which the graduate students express their concern for the welfare of the graduate students at the university, develop and disseminate ideas for the improvement of graduate education, and contribute to the formation of relevant university policy."

The four statistics senators are G. L. Ghai, Dennis Mar, Jeff Meeker, and Ken Offord.

Ph.D. Candidates

Ray Bailey
George Battese
Leon Burmeister
Alfonso Carrillo
Richard Chamberlain
Kuo Ming Chien
Isidoro David
K. T. deGraft-Johnson
Pamela Doctor
Thomas Fears
G. L. Ghai
John Goebel*
Soner Gonen
Joseph Grimes
Her Tzai Huang*
Cary Isaki
Ronald Jacobson
Louis Jensen
William Kennedy
John Kinney
G. Nick Lauer
John Lin
Roland Loup
Charles MacIsaac
Brian Macpherson
Richard Madsen
Angel Martinez
Donald McElhone
Jeff Meeker
James Mellon
Abel Mexas
Kenneth Mount
Nimmagadda Murthy
Peter O'Brien
Martin J. O'Connell
Ken Offord
Panagiotis Papaioannou
C. M. Patel
Ahmed Salem
J. R. Schmid
Nell Sedransk
Wendell Smith*
V. Solomon
Vincent Sposito
Richard Stein
Malte Sund
Victor Tang
Lonnie Vance
James Veale
George Wang
William Warde
Eric West
Milton Winger
Franklin Wolf
Irene Yung

*received M.S. during the year

M.S. Candidates

Guillermo Adames-Suari
Yahia Ahmed
Forrest Aspengren
Pathra Chatkeo
Eliahou Cohen
Jean-Denis Desrosiers
Peampan Davivongs
Sudha Desai
Richard Dorsch
Amiri Gamshadzahi
A. Ronald Gallant
John Goebel
Omer Gucelioglu
Bonnie Hanson
Nancy Heath
Norman Hoesly
Elizabeth Hsu
Her Tzai Huang
Charles Ingwell
Jairoj Jayavadhanangkur
J. D. Jobson
Thomas Keefe
Henry Kelker
Geung Ho Kim
Hae Ja Chung Kim
Lyle F. Lautenschlager
Muhammad A. Malik
Dennis Mar
Kenneth Merritt
Douglas Murfield
Yupha Onthuam
David Pyne
Marcela Salinas
Ivan Sampaio
Jan Shoemaker
Wendell Smith
Terrance Svejda
Jane Toben
Dennis Tsai
Choosak Udomsri
James Whipple
Chartsee Xumsai
Robert Young

Degrees Granted and Positions Taken

Students who were graduated during the year and their location after graduation are listed here. Abstracts of their theses, written as partial fulfillment of the requirements for advanced degrees, appear in the publications section of this report.

Recipients of the Ph.D. Degree

Alfonso Carrillo (November 1969, under B. V. Sukhatme) returned to his position at the Colegio de Postgraduados, Escuela Nacional de Agricultura, in Chapingo, Mexico.

K. T. deGraft-Johnson (August 1969, under J. H. Sedransk) is with the Central Bureau of Statistics in Accra, Ghana.

W. J. Kennedy (November 1969, under T. A. Bancroft) remains at the Statistical Laboratory as an assistant pro-

fessor in charge of the Statistical Numerical Analysis and Data Processing Section.

Abel Mexas (February 1970, under George Zyskind) is an assistant professor with the Statistical Laboratory, currently on assignment with the Ford Foundation project in Chapingo, Mexico.

Kenneth Mount (August 1969, under H. T. David) went to the University of Manitoba, Winnipeg, Canada, as an assistant professor.

Panagiotis Papaioannou (May 1970, under Oscar Kempthorne) accepted a position as assistant professor with the Statistics Department at the University of Georgia.

Nell Sedransk (November 1969, under Masashi Okamoto) is an instructor of clinical oncology at the University of Wisconsin.

Vincent Sposito (May 1970, under H. T. David) remains at the Statistical Laboratory as an assistant professor.

Recipients of the M.S. Degree

Yahia Zakaria Ahmed (May 1970, non-thesis, under Paul Hinz) expects to be working toward his doctorate.

Sudha Desai (August 1969, non-thesis, under H. T. David) returned to Baroda, India.

Jean-Denis Desrosiers (November 1969, non-thesis, under Wayne Fuller) is a professor at the University of Quebec, Canada.

J. Jeffrey Goebel (August 1969, under Wayne Fuller) remains at Iowa State, working toward his doctorate.

Bonnie Roberts Hanson (August 1969, under T. A. Bancroft) accepted a position as an instructor at Drake University, teaching statistics.

Roger Gerald Haugen (May 1970, joint animal breeding-statistics major, under R. L. Willham, animal breeding, and George Zyskind) remains at Iowa State, working toward his doctorate in animal breeding.

Her Tzai Huang (February 1970, non-thesis, under Wayne Fuller) remains at Iowa State working toward his doctorate.

Charles Dean Ingwell (August 1969, under H. T. David) accepted a position as information systems designer with Western Electric in Chicago.

Dennis Ronald Mar (May 1970, non-thesis, under Paul Hinz) is now serving with the Army as a 2nd lieutenant.

Douglas Murfield (November 1969, under Wayne Fuller) went to Austin, Texas as a mathematical statistician with the United States Department of Agriculture.

Jan Philip Shoemaker (August 1969, under Richard D. Warren) joined American Oil Company in Chicago as a senior systems analyst/technical systems.

Wendell Smith (November 1969, non-thesis, under Richard Mensing) remains at Iowa State, working toward his doctorate.

Jane Marie Toben (August 1969, non-thesis, under D. F. Cox) is a scientist with Westinghouse Bettis Atomic Power Laboratory in West Mifflin, Pennsylvania.

The George W. Snedecor Award in Statistics

A. Ronald Gallant was selected by the graduate faculty in statistics as winner of the 1970 George W. Snedecor Award, named in honor of the Statistical Laboratory's founder and first director. Gallant came to Iowa State in 1967 and has held a National Defense Education Act Fellowship.

This was the 17th annual presentation of the award, which consists of a year's membership in the Institute of Mathematical Statistics, a subscription to the Institute's Annals, and a cash gift.

UNDERGRADUATES

Dr. D. V. Huntsberger and Dr. D. K. Hotchkiss served as undergraduate advisers during the year, advising 28 students. Charlotte Bentley, who was graduated in August, was recognized at the May Honors and Awards dinner for having the highest average of any undergraduate during the fiscal year.

This year for the first time a special freshmen orientation introduced new students to basic statistical topics. Normally students are not exposed to these topics until their sophomore year. The program, conducted by Dr. Hotchkiss during fall quarter, helped students preclassify for the following quarter, explained the history of the Statistical Laboratory, and included sessions on statistical methods, experimental design and sampling. Guest speakers Dr. David Jowett and Dr. Roy D. Hickman spoke on "The Designing of Scientific Investigations" and "Collecting Data Through Surveys," respectively.

Recipients of the B.S. Degree

Nancy Lee Allen, May
Charlotte Bentley, August
Charles Bindewald, November
Arthur M. McMahon, May
Charles Richard Miller (joint economics-statistics), May
Sandra Sue Smith (joint mathematics-statistics), May
Shirley Summy (joint economics-statistics), November
Charles Wells, November

Iowa State University Statistics Club

During the year the Statistics Club expanded its scholarship program by instituting a Science Council-Stat Club scholarship which was offered to an outstanding scholar who had been active in statistics activities. The initial award was presented to Nancy Allen, who served during the year as vice president of the Statistics Club. The scholarship previously instituted by the Statistics Club and IBM was awarded this year to Linda Gorman.

The Club continued its sponsorship of professional programs, field trips and social events to introduce members to new areas of statistics and help them get acquainted with each other.

Robert Bowles, an ISU statistics graduate who is now a development consultant with B. F. Goodrich Chemical Company, discussed job opportunities and applications of statistics used in his company. Dr. David Jowett spoke on "Statistics and Conservation."

In February club members toured the human nutrition research facilities in home economics. They had an opportunity to note techniques used to determine food preference and its nutritional value, and learn about the use of statistical principles in the design and conduct of experiments.

A field trip to the Maytag Company in Newton presented statistics in action in an industrial setting. Students saw and heard about assembly line quality control and product development.

Statistics Club officers for 1970-71 are: Kirk Mattes, president; Doug Kopp, vice president; Linda Gorman, secretary and Richard VanderLeest, treasurer.

Mu Sigma Rho

Charlotte Bentley served during the year as president of the Iowa State Chapter of Mu Sigma Rho, national statistical honor society. Four undergraduates and 14 graduate students were initiated. Dr. D. K. Hotchkiss continues as national president of the society. The national officers and directors are working with four universities who have inquired about establishing chapters of the society on their campuses.

SEMINARS

Statistical Laboratory—Department of Statistics Series

Weekly seminars sponsored by the Statistical Laboratory and the Department of Statistics provide a showcase for reporting current research conducted by staff members in statistics. Available on a non-credit basis to all students and faculty members, they provide an opportunity to explain and discuss statistical theory and methodology. The series includes frequent appearances by guest speakers from other institutions, thus enlarging the range of topics.

Richard Mensing, C. P. Cox and Glen Meeden were in charge of seminars for 1969-70. Topics and speakers presented during the year included:

Fall Quarter 1969

September 10	Outlook in Statistics at Iowa State University. T. A. Bancroft.
September 17	Estimation of a Mean when One Observation May Be Spurious. James R. Veale.
September 24	A Technique for Solving some Diffusion Equations in Population Genetics—Applied to Problems in Selection. Louis Jensen.
October 1	Probability, Statistics and the Knowledge Business. Oscar Kempthorne.
October 8	On Statistical Information Theory and Related Measures of Information. Takis Papaioannou.
October 15	Fitting Segmented Curves whose Join Points Have to Be Estimated. Derek J. Hudson, Bell Telephone Laboratories.
October 22	On Survival Probabilities for Some Galton-Watson Processes. Edward Pollak.
October 29	Ratio and Regression Estimation for Sample Surveys. Wayne Fuller.
November 4	(joint Statistical Laboratory, Department of Mathematics and Department of Computer Science) So What? Who Cares? The Question of Scientific Relevance. Robert Hansen, Department of Computer Science, ISU.
November 13	Chance Constrained Linear Programming. J. K. Sengupta.

Winter Quarter 1970

December 3	Scale Estimation in the Presence of a Nuisance Translation Parameter. Barry Arnold.
December 17	Aspects of Game Theory. H. T. David.
January 7	Comparison of Two Sampling Procedures with an Application to Successive Sampling. B. V. Sukhatme.
January 14	Classical Likelihood Theory, Slightly Updated. Oscar Kempthorne.
January 21	Distribution of Discriminant Function in Circular Models. Chien-Pai Han.
January 28	Experiences with Summer Institutes in Statistics at Iowa State. D. K. Hotchkiss and Roy D. Hickman.
February 4	Heredity and Environment: The Statistical Problems. Oscar Kempthorne.
February 12	(joint Statistical Laboratory and Department of Economics) Bayesian Inference in Econometrics. Arnold Zellner, University of Chicago.
February 18	The Development of Statistics in Mexico. Basilio Rojas.

Spring Quarter 1970

March 11	An Alternative Derivation of a Result Due to Srivastava and Bancroft. Barry Arnold.
March 19	Information and Questionnaires in Statistical Inference. George T. Duncan, University of Minnesota.
March 25	Is Research on the Latin Square Finished? Walter T. Federer, University of Wisconsin.
April 1	Selection Procedures Relating to Multinomial, Gamma, Normal and Certain Incompletely Specified Populations. Peter O'Brien.
April 8	A Generalization of the Mean Square Successive Difference Statistic, with Applications. Roland Loup.
April 15	The Anchor Effect on Voting: How Wallace Helped Nixon Become President. James Hutter.
April 29	Introduction to Structural Inference. Leonard Steinberg, Princeton University.
May 6	Estimating the Mean of a Finite Population. Cary Isaki.
May 13	Aspects of Duality in Linear Programming. Vincent Sposito.
May 18	Observational Studies. William G. Cochran, Harvard University.

Applied Statistics Series

The applied statistics series emphasizes the application of statistics to all fields. It is especially planned for research workers who use inferential statistics as a primary research tool. The special assistance available through the statistical numerical analysis, data processing and consulting activities of the Statistical Laboratory were highlighted during the year. Richard D. Warren, Roy D. Hickman and James A. Walsh were in charge of the series. Programs included:

September 22	Analysis of Multiple Corn Fertility Experiments. Ronald E. Voss, Department of Agronomy, ISU.
October 6	Grafted Polynomials as Approximating Functions. Wayne Fuller.
October 20	Immediate Memory and Dichotic Presentation. Wayne Bartz, Department of Psychology, ISU.
November 3	Application of Kalman Filter Theory in Integrated Navigation Systems. R. G. Brown, Department of Electrical Engineering, ISU.
December 8	Some Examples of Nonlinear Fitting. J. I. Mellon.
January 12	An Orientation to Statistical Numerical Analysis and Data Processing Section. W. J. Kennedy and B. J. Meador.

January 26	An Orientation to Statistical Numerical Analysis and Data Processing Section. W. J. Kennedy, E. N. West and V. A. Sposito.
February 9	An Orientation to Statistical Numerical Analysis and Data Processing Section. V. A. Sposito, R. A. Stein and K. E. Merritt.
March 23	Statistical Contributions to Ecology. David Jowett.
April 13	Reduced Dimensionality in Analysis of Variance of Multivariate Data. Eli Cohen.
April 27	Consulting Areas and Techniques. Harold Baker, Roy Hickman, Paul Hinz, D. K. Hotchkiss, James Hutter, Richard Mensing, James Walsh and Richard Warren.
May 4	(joint seminar with Department of Nutrition) Statistical Appraisal of the Protein Problem. P. V. Sukhatme.
May 14	(joint seminar with Department of Nutrition) Incidence of Undernutrition: Calories. P. V. Sukhatme.
May 21	(joint seminar with Department of Nutrition) Incidence of Malnutrition: Protein-Calories. P. V. Sukhatme.

Quantitative Genetics Series

Oscar Kempthorne and Edward Pollak are in charge of these seminars, which are planned especially for staff members and graduate students in statistics, genetics, animal science, poultry science, agronomy and horticulture. Programs presented during the year included:

December 16	Concepts of Fitness in Mendelian Populations. Oscar Kempthorne.
February 10	A Monte Carlo Check on Index Selection Theory When Population Size is Finite. Ben Bereskin, USDA Regional Swine Laboratory.
March 24	The Role of Major Genes in Egg Production and Livability in Poultry. James Arthur, Hyline Poultry Farms.
May 5	Malthusian Parameters in Genetic Populations. Part I. Haploid and Selfing Models. Edward Pollak.

FORD FOUNDATION MEXICAN PROJECT

Abel Mexas joined the staff at the Statistical Center at the National School of Agriculture in Chapingo, Mexico, at the beginning of the fiscal year. Dr. Barry Arnold and Dr. Richard Lund completed their assignments in the fall. This has been the fifth year of cooperation between the Statistical Center and the Statistical Laboratory, sponsored by the Ford Foundation.

During the year Dr. Mexas taught Experimental Design I, Statistics Theory II and Special Problems. The teaching program in Chapingo is well developed and proceeding smoothly. Although classes are primarily composed of statistics majors, some other students with special interests in statistics are enrolled.

Dr. Mexas is directing the theses of four students. The research problems they've selected include the following areas: estimation with models of classification, non-parametric methods, and algebra of design structures.

The Chapingo statistics staff continues to present seminars on statistics topics. Dr. Mexas discussed "Some Computational Aspects of Analysis of Variance" for the Colegio de Mexico, and presented "The Use of Transformations in the Analysis of Variance" to the Department of Botany.

Consulting assistance is provided to many Chapingo staff and students and governmental research workers in setting up research projects, data collection methods, and analysis and processing of results produced.

The following projects indicate the scope and nature of consulting activities: the sampling design and data analysis methods were determined for a study whose purpose is to relate native pasture grass composition and characteristics to soil conditions; a questionnaire was designed, sampling procedures and punched card formats determined for a study of the feasibility of locating a milk processing plant in a particular area; a program was specified to analyze data from an experiment on cows. A special program was written for the construction of the design matrix reparametrized to full rank, for a study of the effects of saline water application and some physical and chemical properties of soils.

Work at the computing center has been expanded and an IBM 1130 has been installed. The center has become a self-supporting enterprise which serves both the campus and outside agencies.

Dr. Mexas has submitted a proposal to initiate an M.S. program in computing science this fall, and has been assigned by the Dean of the Graduate College in Chapingo to oversee the implementation of this program. The outline of courses has been developed and additional computing staff members now are being sought.

A publication of the Statistical Laboratory
Iowa State University
Susan Alice Brown, editor