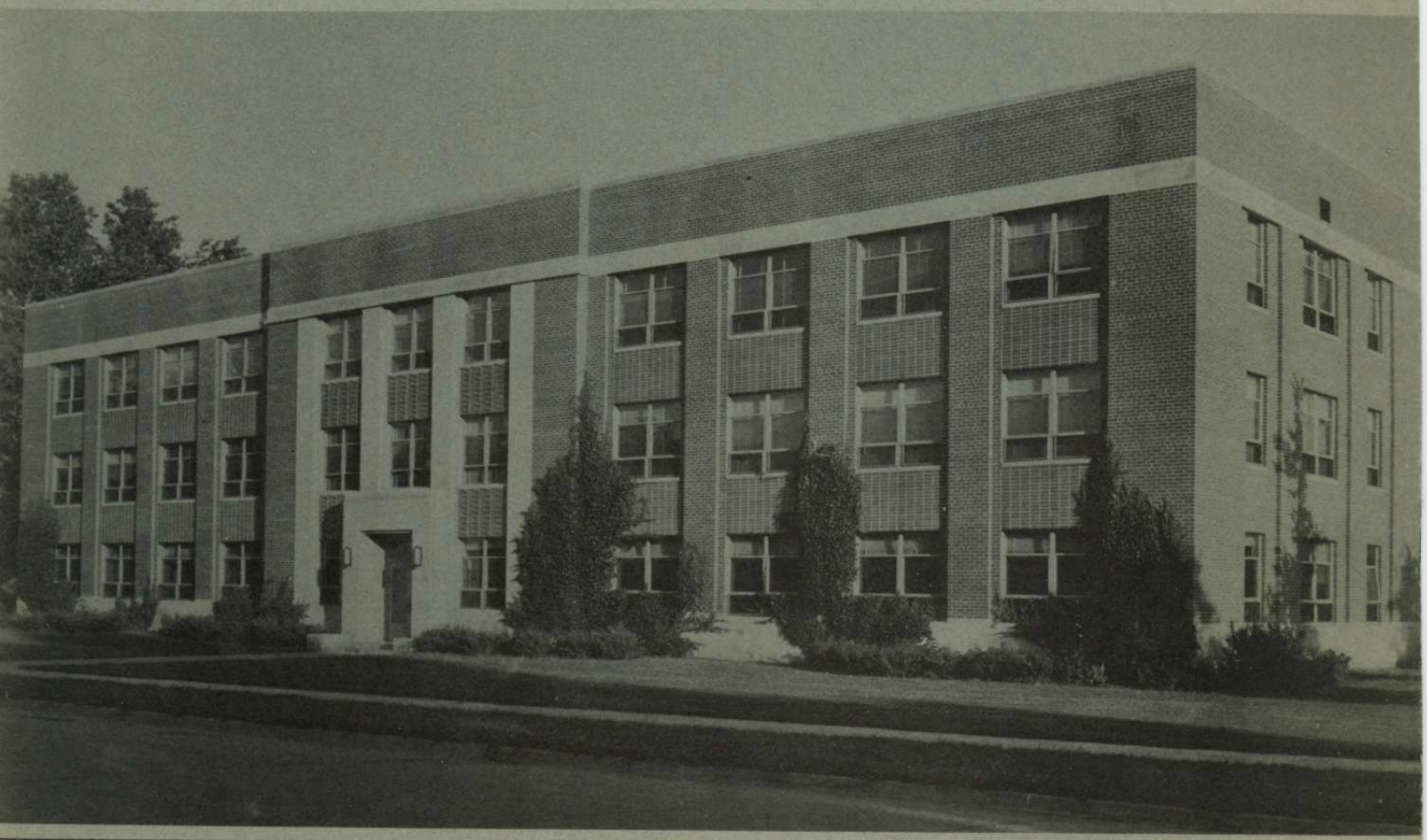
13/24/0/1

1949-50 ANNUAL REPORT

# The STATISTICAL LABORATORY of

## Iowa State College



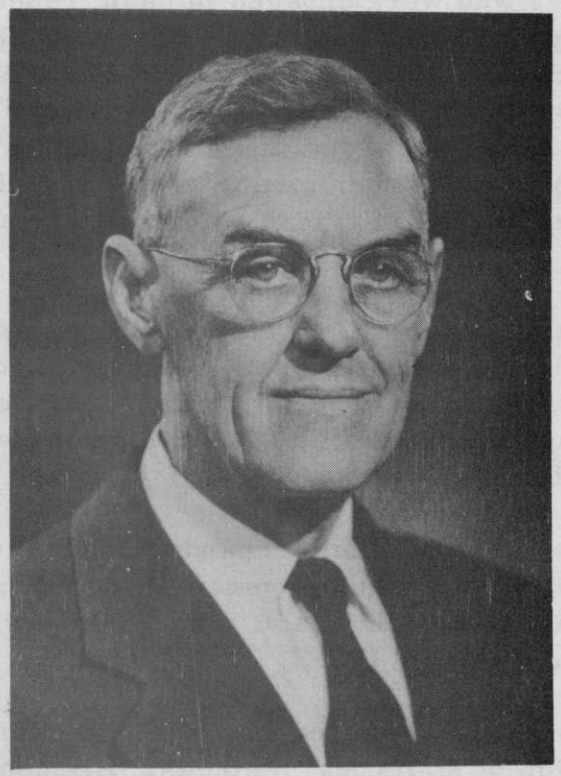
SERVICE BUILDING

★ RESEARCH
★ TEACHING
★ CONSULTATION
★ SERVICES

Cooperating Government Agency

Bureau of Agricultural Economics

CONTENTS



Professor George W. Snedecor: founder and first director of the Statistical Laboratory.

## HISTORY OF THE STATISTICAL LABORATORY

In 1924 some twenty members of the faculty of Iowa State College met on Saturday afternoons to study multiple regression under Dr. Henry Wallace, then editor of Wallaces' Farmer. Especially interested in these early conferences were Dr. R. A. Pearson, president of the college, Dean S. W. Beyer of the Science Division, Dr. E. R. Smith, head of the Department of Mathematics and Mr. George W. Snedecor of the same department.

The bulletin "Correlation and machine calculation" by Wallace and Snedecor was published in 1925, as a result of these conferences, and distributed widely both in this country and abroad. This small book has now become practically a collector's item.

With the further development of this early interest in the significance of experimental results, there arose a growing interest in other statistical techniques. Among the first departments to express their interest in this work were the Department of Animal Husbandry, represented by Dr. Evvard, working in animal nutrition, and the Mathematics Department, represented by such persons as Drs. Brandner and Allen.

This early recognition of the importance of the use of proper statistical methodology led to the establishment of the Statistical Laboratory in 1933, under Professor Snedecor. From the Laboratory, statistical advice was

given to persons working on research projects in the Agricultural Experiment Station and to graduate students working on theses, as well as to other persons and agencies both on and off campus. Also, within the Statistical Laboratory there was developed a staff of persons who were engaged in basic research. Members of the staff in the early days included George W. Snedecor, W. G. Cochran, Gertrude M. Cox, A. E. Brandt and C. P. Winsor. Their earliest contributions were mainly in the direction of the unification, systematization and simplification of the methods of R. A. Fisher of the Galton Laboratory, London, and Frank Yates of the Rothamsted Experiment Station in England. A considerable amount of work was also done in extending the mathematical foundations of statistical theory.

The visits and residencies of the statisticians mentioned above, as well as those of John Wishart of Cambridge University, Jerzy Neyman, Egon Pearson and others were always a source of inspiration to the staff of the Laboratory, as well as to the many other scientists on the campus to whom the statistical processes were an indispensable tool.

Research aspects of the Laboratory were richly expanded under a number of cooperative agreements set up with governmental agencies--first, in 1938, with the Bureau of Agricultural Economics, USDA, and subsequently with the Weather Bureau and the Bureau of the Census, USDC.

Professor Snedecor is recognized throughout the world for his pioneer work in the development and utilization of statistical methods. His book "Statistical Methods", first published in 1937 and now in its fourth edition, has been translated into Spanish and Portuguese. A Japanese translation is slated for publication about the first of 1951. The work of his students alone has given Snedecor a worldwide reputation as one of the foremost teachers of statistics.

The Laboratory has attained a widespread recognition more recently for the work performed under the direction of Professors Arnold J. King and Raymond Jessen in the development of the Master Sample of Agriculture and the compilation of the master sampling materials for most of the United States. These materials form the basis of many of the "area" samples designed by the Laboratory. Various members of the staff have done important work in research in sampling based on these "area" sample designs.

With the assumption of the acting directorship of the Laboratory, Jessen (with his staff) has given considerable attention to the type of training needed in statistics by both those students whose majors are in statistics and those working in other fields which require a facility with statistics as a research tool. Graduate students and research workers have come from India, Australia, China, Argentina, Mexico, Puerto Rico, the Philippines, Canada and all sections of the United States to study.

The Statistical Laboratory now includes (1) a Department of Statistics in the Division of Science, (2) a section of the Agricultural Experiment Station, (3) a unit which participates in the work of the college Industrial Science Research Institute and (4) a research field office of the Bureau of Agricultural Economics, United States Department of Agriculture.

Iowa State College
Statistical Laboratory

## 1949-50 STAFF

Charles E. Friley, LL.D., Sc.D. ..... President
Floyd Andre, Ph.D. ..... Director, Agricultural Experiment Station
Dean, Division of Agriculture
Harold V. Gaskill, Ph.D. ..... Dean, Division of Science
Director, Industrial Science Research Institute
Raymond J. Jessen, Ph.D. ..... Acting Director, Statistical Laboratory
Acting Head, Department of Statistics
Acting Head, Statistics Section, Agricultural Experiment Station

| | |
|---|---|
| Clifford C. Hildreth | ...Visiting Professor of Statistics and Economics |
| Paul C. Homeyer | ..... Professor |
| Leonid Hurwicz | ..... Professor of Economics and Statistics |
| Raymond J. Jessen | ..... Professor |
| George W. Snedecor | ..... Professor |
| Gerhard Tintner | ..... Professor, joint appointment in Statistics, Economics and Mathematics |
| Theodore A. Bancroft | ..... Associate Professor |
| Emil Jebe | ..... Associate Professor |
| Oscar Kempthorne | ..... Associate Professor |
| Gobind Ram Seth (1) | ..... Associate Professor |
| Campbell C. Mosier | ..... Assistant Professor |
| Bernard Ostle | ..... Assistant Professor |
| Norman V. Strand | ..... Assistant Professor |
| Dorothy S. Cooke | ..... Research Associate, Instructor |
| George Darroch | ..... Research Associate, Instructor |
| Arthur Dutton | ..... Research Associate, Instructor |
| John Hofmann | ..... Research Associate, Instructor |
| Daniel Horvitz | ..... Research Associate |
| Garnet McCreary | ..... Research Associate |
| John Monroe | ..... Research Associate |
| Robert Spears | ..... Research Associate |

(1) On leave 1949-50.

Graduate Assistants

Om P. Aggarwal
Virgil Anderson
Robert Bowles
Helen Conners
Eleanor Matsis
Max Ray Mickey
Richard Schrimper
J. E. Keith Smith

Other Graduate Students

William Barclay
Robert D. Branstetter
Edward C. Bryant
Ching-yao Chu
David Gosslee
Franklin Graybill (graduate assistant in Mathematics Department)

Other Graduate Students (cont.)

Donovan Thompson
Leo Tick
John Wiesen
David Huntsberger
Dale Lary
Ralph Madison
Carl Marshall
Bertil Matern
Merlin Miller
Burton Onate
Reuben Hansen
Thomas L. Healy
Norma Howard
George W. Reynolds
William Ritts
Basilio Rojas
Warren Sladky
William Tater
Heng Lih Wang
Richard Wilde

Margaret Kirwin ..... Secretary
June Duffield ..... Stenographer
Margaret Ann Hofmann ..... Stenographer
Mary Lou Armstrong ..... Typist

Jauvanta Walker ..... Research Helper
Margaret Willey ..... Machine Lab. Supervisor

Statistical Clerks

Jessie Bryan
Mary A. Clem, Chief
Bertha Eastman, Asst. Chief
Mary Hayenga
Madge Manning
Douglas Robson
Lucile White
John A. Zoellner

Clerks

Donna E. Browning
Eunice Campbell
Barbara Carl
Joan Randall
Mabel Tilton

Key Punch Operators

Dorothy Alcott
Colleen Blagney
Betty Jean Hall
Janet Lee Shrake

Tabulating Machine Operators

Wanda Anderson
Donald De Roos
Joyce Johnson
Eleanor Perry
Thelma Wignall

Dr. Clifford Hildreth, research associate with the Cowles Commission for Research in Economics, who had resigned from the Iowa State College staff June 30, 1949, commuted to and from Chicago throughout the spring quarter, 1950, to serve as visiting professor in the Statistical Laboratory.

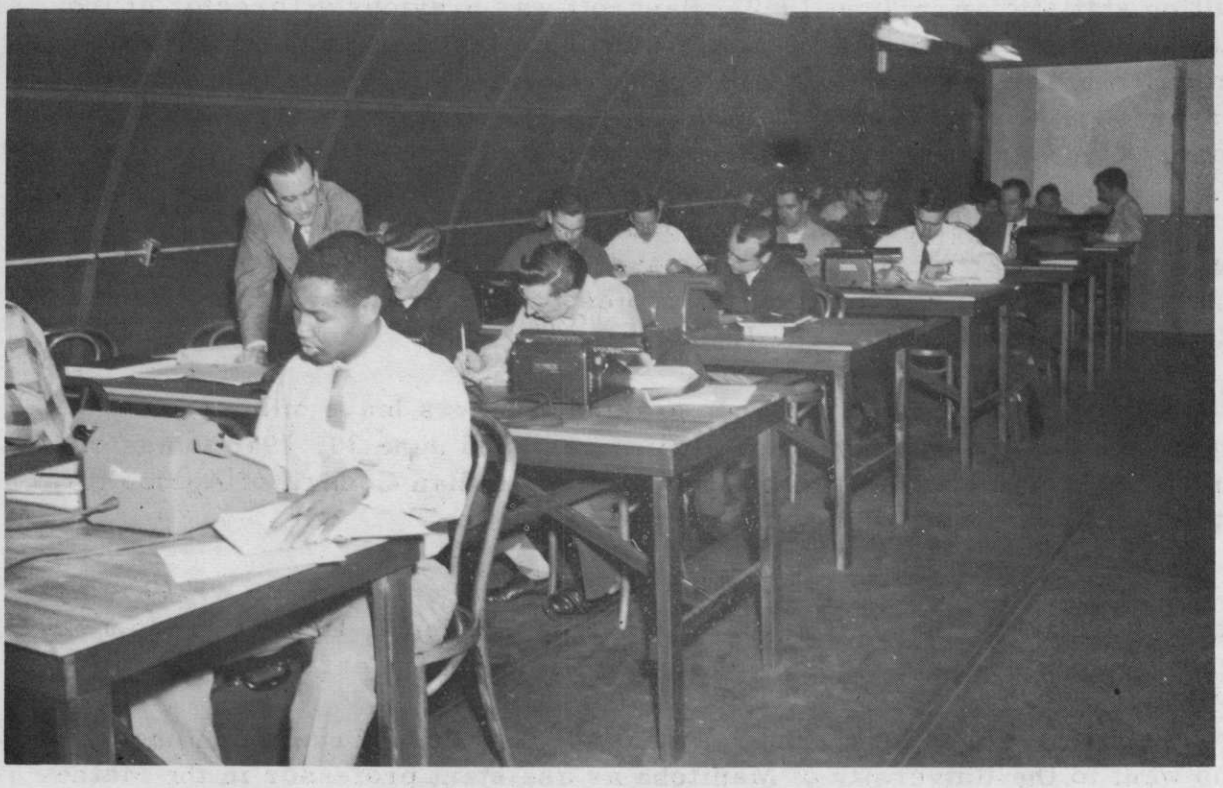
Dr. T. A. Bancroft rejoined the staff as associate professor and consulting statistician in June 1949. Bancroft was previously Director of the Statistical Laboratory of Alabama Polytechnic Institute, Auburn, Alabama. Emil Jebe returned to the Laboratory in November 1949 as associate professor, after completing work toward a Ph.D. in experimental statistics at North Carolina State College. He received his degree at Raleigh in June 1950.

R. J. Jessen resigned as Acting Director of the Laboratory and Acting Head of the Department of Statistics and the experiment station section on June 30, 1950, but is continuing as professor of statistics. T. A. Bancroft will take over as Director and Head.

Gobind Ram Seth, who had been granted a year's leave of absence to work with Dr. P. V. Sukhatme in India, resigned June 30, 1950 to assume the position of professor of statistics with the Indian Council of Agricultural Statistics, New Delhi, India.

Professor Hurwicz resigned from the staff on September 15, 1949 to assume positions on the staff at the University of Illinois, Urbana, and as research consultant with the Cowles Commission for Research in Economics.

Other departures during the year include Garnet McCreary in June, 1950, who went to the University of Manitoba as assistant professor in the Mathematics Department.



Professor Homeyer assisting students in the laboratory section of Statistics 402.

## TEACHING AND TRAINING

In 1915, the first courses were offered in statistics at Iowa State College by the Departments of Economics and of Mathematics. Interest in this work as represented by the numbers of courses offered and students registered has been of steady increase.

By 1940, there were twenty courses being given to three hundred and fifty senior college and graduate students working in more than thirty departments. These courses were being given in the Departments of Economics, Mathematics and Vocational Education.

The exigencies of the war greatly expanded the demand for the services of the trained statistician. The policies of the Procurement Divisions of the various government agencies and departments have assured the use of quality control procedures by most large industrial concerns. The Applied Mathematics panel is credited with the organization of a broad statistical consulting service under the Office of Scientific Research and Development, as well as with the provision that research projects of the federal government should be subject to the review of statisticians. The wide attention attracted by the various public opinion polls stimulated a vast amount of market research as well as extended the use of survey methodology in economic and social research.

As a result of this greatly expanded need for trained statisticians in industry, in government, and in social and economic research agencies, as well as in agricultural experiment stations, there has been a great pressure on the supply of statisticians available. It was in answer to this demand that Iowa State College expanded its teaching program and established a Department of Statistics, July 1, 1947. The new department was fortunate in having the guidance of Professor Snedecor during its formative years. Dr. Raymond J. Jessen has acted as Head of the Department since Snedecor's retirement as Director of the Laboratory.

Since the formation of the Department of Statistics, courses have been offered to students specializing in statistics and to students in other fields in which statistical methods are important research tools, such as Economics, Genetics, Plant and Animal Breeding, Climatology, etc. The department offers the Bachelor's, the Master's and the Doctor's degree in Statistics and in Mathematical Statistics.

The Laboratory had available a number of assistantships, varying from $810 to $1125 for nine months, for the purpose of encouraging graduate study in statistics. During the 1949-50 academic year, there were eleven students serving as assistants. All graduate students are incorporated in various ways with the activities of the Laboratory. They thus receive invaluable experience in the application of statistics as evidenced in the work of the Statistical Laboratory.

The graduate students also participate with the staff in the well-rounded seminar programs of the Laboratory, which draw speakers from the staff of the Laboratory, the staff of other departments or research groups on the campus, as well as visiting statisticians from other institutions and from other countries.

Students graduating with the B.S. degree in statistics are prepared to direct clerical personnel engaged in processing data, to assist professional research workers, and to perform routine statistical investigations. A major in statistics is obtained by taking a series of three courses. The first series (301, 302, 303) introduces the student to the role of statistics in modern society, and to such concepts as variation, averages, probability, the elements of statistical inference, as well as to the procedures of statistical surveys and experiments. A second series of courses (401, 402, 403) deals with the application of statistical methods to biological, physical and social data, tests of significance, linear regression, correlation, analysis of variance and covariance and statistical inference. A third set of courses (447, 448) deals with statistical theory, including such concepts as combinatorial probability, mathematical expectation, population distributions, location and scale parameters, transformations, sampling distributions, estimation of parameters and tests of hypotheses. The latter rounds out the program for these statistics majors in that it gives them an understanding of the validity and limitations of the statistical techniques which they have learned in their other courses.

The curriculum for graduate students is designed for persons desiring to work in research organizations where a background is needed in experimental

design or in sampling techniques, in addition to those who are interested in teaching or doing research in statistical theory.

Receiving Ph.D. degrees during the 1949-50 year were: Bernard Ostle, in mathematical statistics, at the end of the second summer session 1949-- (thesis, "On certain criteria for optimum estimation"); Garnet E. McCreary in statistics, June 1950 -- (thesis, "Cost functions for sample surveys"). M.S. degrees were given to Leo Tick (thesis, "Past and present status of statistical inference") in December 1949 and to David Gosslee (thesis, "Statistical analysis of maximum temperature series for Mason City, Iowa") in June 1950.

## COURSES

The following courses were offered during the year by staff members:

### Primarily for undergraduate students

| | | | |
|---|---|---|---|
| 301,302,303 | Principles of Statistics | three quarters | Snedecor; lab assistants Cooke, McCreary, Mickey, Thompson (303, only 3 lab sections) |
| 401 | Statistical Methods I | fall | Bancroft, Darroch, Hofmann, Homeyer |
| | " " " | winter | Jebe |
| 402 | " " " | winter | Darroch, Homeyer, Ostle |
| | | spring | Homeyer |
| 403 | " " " | spring | Darroch, Jebe |
| 437 | Statistical Quality Control | spring | Hofmann |
| 447,448 | Theory of Statistics I | winter, spring | Bancroft |
| 481,482,483 | Processing of Data (lab) | three quarters | Mosier |

### For advanced undergraduates and graduate students

| | | | |
|---|---|---|---|
| 504 | Analysis of Variance | fall | Homeyer |
| 511,512 | Design of Experiments | winter, spring | Kempthorne |
| 521 | Design of Surveys | winter | Jessen |
| 524 | Theory of Sampling | spring | Horvitz |
| 531,532 | Industrial Statistics | fall, winter | Ostle |
| 536 | Special Statistical Methods in Genetics | fall | Dutton |
| 538 | Economic Statistics | spring | Hildreth |
| 541,542,543 | Theory of Statistics II | three quarters | Ostle |
| 599 | Special Topics (credit as arranged) | each quarter | Homeyer, Horvitz, Jessen, Kempthorne |

For graduate students

| | | | |
|---|---|---|---|
| 611, 612 | Design of Experiments II | fall, winter | Kempthorne |
| 641 | Theory of Statistics III | fall | Kempthorne |
| 699 | Research (credit as arranged) | each quarter | Bancroft, Jessen, Kempthorne, Tintner |

Courses offered during summer sessions 1949-50

| | | | |
|---|---|---|---|
| 401 | Statistical Methods I | 1st ss. 1950 | Ostle |
| 402 | " " " | 2nd ss. 1949 | |
| 403 | " " " | 1st ss. 1950 | Darroch |
| 447 | Theory of Statistics I | 1st ss. 1950 | Bancroft |
| 599 | Special Topics (credit as arranged) | 2nd ss. 1949 | |
| | | 1st ss. 1950 | Bancroft, Darroch, Kempthorne, Ostle |
| 699 | Research (credit as arranged) | 2nd ss. 1949 | |
| | | 1st ss. 1950 | Bancroft, Kempthorne |

## SEMINARS

Throughout the academic year a series of staff seminars were held on Friday afternoons for statistics graduate students, members of the Statistics Department and other interested persons. Seminar topics included the following:

Fall Quarter:

Some formal relations in multivariate analysis: Gerhard Tintner.
Proposed techniques for testing homogeneity of behavior: Dr. Charles C. Neidt (Department of Psychology).
Bias due to omission of independent variables in linear regression analysis: T. A. Bancroft.
Classification techniques: Dr. Carl F. Kossack, Director of the Statistical Laboratory, Purdue University.
Statistical problems in home economics research: Paul Homeyer.
Some aspects in statistical research in Sweden: Fil. Lic. Bertil Matern, Forest Research Institute, Stockholm, Sweden.
Some problems in estimation in survey sampling: D. G. Horvitz.
Cost functions in survey sampling: Garnet McCreary.

Winter:

The function of mathematics in statistics: George W. Snedecor.
Foundations of probability and statistical inference: Gerhard Tintner.

An application of the analysis of variance in factors affecting the feathering growth and body weight in chickens: Arne W. Nordskog (Poultry Husbandry Department).

The selection of primary units for sampling an agricultural population: Emil Jebe.

The changing emphasis in industrial statistics: Bernard Ostle.

Past and present of statistical inference: Leo Tick.

Estimation of genetic parameters: L. N. Hazel (Department of Animal Husbandry)

Some uses of regression in fish growth studies: Kenneth D. Carlander (Department of Zoology & Entomology)

Recent advances in experimental design: Oscar Kempthorne.

Spring:

Sampling the television audience in Iowa: Dorothy S. Cooke.

Problems in genetic estimation: Bruce Griffing (Genetics Department).

Some applications of punched cards in the calculations for statistical analyses: Campbell C. Mosier.

Application of simultaneous equation approach to the study of the demand for meat in the United States: Burton L. French (Department of Economics & Sociology).

Some new properties of the distribution of the ratio of two random variables whose distribution is bi-variate normal: Max Ray Mickey.

Analyses of long-term experiments: Arthur Dutton.

Crop-cutting experiments on a random sampling basis for the estimation of yield of crops in India: Om P. Aggarwal.

Combining cross-section and time series data in economic studies: Clifford Hildreth.

Application of fractional replication in an evaluation of the Adams Consistometer: Robert G. Tischer (Department of Food Technology).

## GUEST SPEAKERS

Several guest-seminars were offered in addition to the regular statistics seminar program. Dr. William Gregory Madow, professor of statistics at the University of Illinois, spoke February 7, 1950, to members of the Graduate College faculty and graduate students on "Systematic sampling" and February 8 to Statistical Laboratory personnel on "Communications problems." Mr. Henry Teicher, graduate student at Columbia University, also gave two talks May 9-10: "On the factorization of distribution" to the Economics, Mathematics and Statistics Departments and "On decision functions" to the college Research Council and the Laboratory.

Professors Oscar Kempthorne (left) and Raymond Jessen conferring over a research problem.

## RESEARCH IN THE STATISTICAL LABORATORY

The research program of the Laboratory is inevitably closely related to the other activities of the staff members and consequently emphasizes both the development and extension of basic theory and its application in the development of new statistical tools and techniques. Many of the research projects arise as a direct result of needs made apparent to the staff in their roles as consulting statisticians, whose experience covers problems evolving from research being done by other agencies and industrial concerns throughout the United States and abroad, as well as problems from campus research jobs. Various government agencies, such as the Bureau of Human Nutrition & Home Economics, Bureau of the Census, Weather Bureau and Office of Naval Research, have sought advice in areas where it became obvious that statistical research needed to be done.

One important area for research is in the design of experiments. The considerations which have led to the necessity for planned observations in the biological sciences are the impossibility of reducing variation in the experimental material so that errors of measurement are the main source of error, and the necessity for the study of many factors at the same time because of interactions. The biologist working with animals can never obtain experimental units which are exactly alike genetically, with the exception of identical twins.

This type of variation in biological material can never be eliminated, in contrast with physical material which can frequently (though by no means always) be purified to any desired extent. The planning of observations is the main problem of the design of experiments. To overcome the impossibility of controlling variation Fisher introduced the principle of randomization and the statistical treatment of randomized block and Latin Square layouts. Fisher and later Yates developed factorial designs for the combined study of several factors. Another method of reducing the effect of variation is the use of incomplete block designs, due primarily to Yates. Continuous development of the field is necessary to deal with the new experimental situations. Research is also necessary on methods of analyzing experimental data.

A second major area for research is in the design of samples for survey use and development of new techniques for handling other phases of survey operations. Such problems as the effects of different types of missing data and differential degrees of response from various groups of respondents, choice of the most efficient sample design and method of estimation for a fixed cost and certain kinds of desired information, effects of different ways of asking questions, etc., are always of concern.

The Master Sample Project was generated from the need on the part of federal and state governments to obtain more accurate and more timely information than was possible under the system of mailed responses and complete census enumeration. There was need to know such things as the estimated quantity of farm production, the movements of the labor force, the amount of distress after floods and the estimated amounts of farm goods going to market. Some of this information was needed quickly for the formation of administrative policy; some of the information was needed with greater accuracy than had formerly been possible. The element of the expense of a complete census as opposed to a partial enumeration was also an important consideration in devising ways of obtaining reliable sampling methods. If relatively less time and money were spent on trying to obtain a complete census, more time and money could be spent on selecting and training good interviewers to obtain information, the reliability of which could be estimated.

For these and other reasons, the Laboratory, Bureau of the Census and Bureau of Agricultural Economics set up the Master Sample Project at Ames in 1943, to process the vast amount of material that was needed before a national area-type sample could be designed. These materials included road maps, aerial photographs and Sanborn maps of all parts of the country. They have been so prepared that samples of many types may be rather quickly drawn, whether the observation units be farms throughout the United States, families in the rural areas or the spending units for families in open country, towns and cities. These materials have already been used extensively for various kinds of samples and for research in sample design. Upon the completion of the project in 1949 and the return of the resident collaborators of the two bureaus to Washington, D.C., a set of the sampling materials was left with the Laboratory for further use and research. A number of recent modifications in techniques for dealing with the materials have increased their usefulness in sampling operations.

## IOWA AGRICULTURAL EXPERIMENT STATION

Two research projects have been established in the statistics section of the experiment station. Project 890, "Design of experiments and analysis of the data," is principally staffed by Kempthorne and Dutton. During the past year particular attention has been given to the following studies:

(1) Incomplete block designs (mainly on the effects of inaccuracies in the estimation of weights on the variances of estimated treatment differences).

(2) A design for testing inbred lines. The testing of a set of n inbred lines in the past has been accomplished by comparing all the possible $n(n-1)/2$ first crosses in an experiment. This procedure rapidly becomes unfeasible with a large number of lines, say 100. A design, which it is believed is due to A. M. Mood, has been worked out and is being tested in conjunction with the farm crops section of the experiment station.

(3) The design and analysis of long-term experiments. The particular case to which most attention has been given is the evaluation for a long-term experiment on a single crop. This has been examined in the light of the existence of correlation in the deviations of yields in successive years from the true curve. A model in which the deviations follow a simple autoregressive scheme has been used. This approach has been extended, to some extent, to the comparison of different continuous cropping systems, evaluated over the same period of years, and to other more complicated cases.

(4) The development of mathematical models for the study of quantitative inheritance. Some work has been done on the design and analysis of experiments investigating quantitative inheritance. This has consisted primarily of two parts: (a) the scope and validity of simple gene models, (b) the possibility of devising models the parameters of which can be estimated from experimental data.

(5) The fractional replication of factorial experiments. The range of applicability of fractional replicated factorial experiments has been clarified and the designs are being tested, in cooperation with the food technology section of the experiment station, for the problem of assessing the factors involved in measurement of characteristics of food.

(6) Randomization test. A preliminary empirical examination of the sensitivity of the randomization test in randomized block experiments has been made, and it is hoped to extend this investigation theoretically.

In addition, a book on the design and analysis of experiments is being prepared by O. Kempthorne.

Project 1005, "Design of surveys and analysis of the data," under Jessen, is concerned with the extension of sampling theory and the testing of the merits and feasibility of theoretically sound sampling designs by practical applications in survey work--as well as with other features of survey design such as questionnaire construction and problems of interviewer-respondent communication.

Devising new survey designs has lead to: (1) improvements on the "fixed-take" sampling scheme (a scheme requiring the number of farms or households to be taken at each sampling point to be fixed, and to be constant from point to point) as regards its use in the field, statistical efficiency and sample data analysis; (2) development of a new scheme for sampling farms in regions where "large" farms cause considerable trouble with standard procedure--this scheme should be appropriate for a number of types of farm surveys west of the 100th meridian and other regions; (3) exploratory use of a scheme for subsampling on aerial photos, rather than on maps or by "map-listing", for a survey in Tennessee; (4) development of sample techniques for getting samples of areas with a number of desired characteristics independently of the Master Sample Project materials.

Many of the samples or surveys mentioned under the Survey Service activities in this report were designed by members of Project 1005. Such practical applications in survey work have resulted in further development of the Master Sample Project materials. Thus, development of techniques for dealing with any predetermined desired size of sampling unit (in terms of the number of households or farms it contains, for example) with the present nature and organization of the project materials has considerably increased the effectiveness of these materials. Similarly development of techniques for avoiding bad sampling unit boundaries (which are sometimes found in the Master Sample) has increased the value of these materials. Some progress has been made in dealing with the problem of defining without bias boundaries between open country, rural place and urban zones, in terms of sampling units.

## COOPERATIVE RESEARCH BETWEEN THE INDUSTRIAL SCIENCE RESEARCH INSTITUTE AND THE ROCKEFELLER FOUNDATION

During the summer of 1948 members of the Laboratory went to Greece to work for the Rockefeller Foundation on the design and administration of a socio-economic survey of the Island of Crete. The purpose of the survey was to obtain information of an economic, social and cultural nature, to be used in planning broad social programs for the betterment of the island. Crete was chosen for the Rockefeller study because it represented a relatively self-contained economy, backward in nature, and was small enough geographically to make feasible a test of such procedures as were needed. The Foundation was also interested in developing techniques for sampling in countries where little up-to-date accurate information of any kind was available as was true in Crete after World War II.

Because of the nature and amount of information wanted, it appeared advisable to obtain it on several bases or units of observation, so a sampling scheme was devised which permitted all these units--the household, individual person, farm, community, local market and local physician--to be dealt with simultaneously. The design for the general survey thus called for a sample of communities and then, within those communities, household samples drawn so that they comprised a 1/150 sample of all households in existence on the island. Laboratory members handled the supervision and training of

interviewer teams and aided in the development and testing of bilingual questionnaires. Data from the socio-economic survey were completely processed and analyzed for the Rockefeller Foundation in the Laboratory. During the 1949-50 year, supplemental analyses were made. A statistical appendix is being written by Jessen and Strand for the report on the survey.

## COOPERATION WITH THE BUREAU OF AGRICULTURAL ECONOMICS

The Laboratory has a cooperative agreement with the BAE to pursue certain theoretical and methodological research in agricultural statistics. Under this agreement a field office of the BAE has been set up within the Laboratory, and a number of staff members have been regularly assigned to that office on a part-time basis. During the past fiscal year, research work was concentrated mainly on four projects:

(1) Cost functions for agricultural and general purpose sample surveys. The completion of a doctoral dissertation by Garnet McCreary in June 1950 brings to a close work in this field. One of the principal problems involved has been to obtain functions which would represent the travel distance between a particular number of randomly located points in areas of different sizes and shapes. McCreary's thesis presents theory for making estimates of the field costs of various possible sample designs for a given investigation. Cost data from previous surveys (BAE 1948 April Survey, Wallaces' Farmer Survey of Farmer Readership, and the Farm Media Survey of Iowa) were examined.

(2) The design and analysis of long-term experiments. This investigation parallels work in Research Project 890 in the statistics section of the experiment station.

(3) The estimation of ratios in sampling surveys. Results were obtained on the bias and variance of the ratio of the sample means, say $z = \bar{y}/\bar{x}$, in terms of the first and second moments of the distribution of $1/\bar{x}$. This was for random samples from joint continuous distributions $f(x, y)$ which satisfy the requirements that

(a) the true mean square regression of y on x is linear, and

(b) the variance of y for fixed x is proportional to $x^{\delta}$ for $\delta = 0$ and $\delta = 1$.

The conditions on the distribution of $\bar{x}$ necessary for z to have a distribution with finite first and second moments were stated. These conditions, of course, applied to joint distributions satisfying the above requirements. In particular it was noted that the mean and variance of $1/\bar{x}$ are finite for the Pearson Type III distribution, so that exact formulae for the bias and variance of z were obtained for $f_1(x)$ distributed in this manner. These formulae were used to make statements on the required sample sizes for negligible bias and negligible underestimation of the variance when the usual approximate formula for the variance of a ratio is used, the above conditions prevailing.

As regards the underestimation of the variance with the usual approximate formula in these circumstances, a good rule is to choose $n \geq 100\, c_x^2$ where $c_x$ is the coefficient of variation of the denominator variable. The underestimation by the approximate formula will be less than 4% if in fact the regression of y on x is linear with constant variance within arrays for which x is fixed and the marginal distribution of x is Pearson Type III. This rule should be a practical guide even if these conditions are not exactly satisfied.

(4) Missing data in agricultural surveys. An extensive bibliography has been prepared which contains references on the extent, effect and treatment of missing data in various kinds of sample surveys and with different types of investigations. Some tests have been made on survey data available in the Laboratory to see if there were actual differences, for items of agricultural interest, between interviewed and noninterviewed groups and among groups interviewed on different calls. It is hoped that this study will result in development of techniques for handling missing data in analysis.

## CONSULTATION AND COOPERATION

The Statistical Laboratory was founded and attained its greatest reputation as a source of information and guidance for persons doing research requiring quantitative analysis.

### STUDENTS

Undergraduates as well as graduate students bring in their problems to staff members for review. Each year a large number of students planning their theses come to the members of the staff for advice about the design of their experiments. At various times, student organizations on campus have asked for advice on local sampling and interviewing problems connected with their activities.

### AGRICULTURAL EXPERIMENT STATION

In numerous cases, projects of the Experiment Station include members of the staff who work along with the research at every stage. In other cases, projects are brought to the Laboratory when plans are being made, to ensure that the maximum amount of information will be obtained with the allocated resources. Even when advice is sought only when final analysis is desired, the members of the staff will still render consultative service.

Service project 1-S, "Statistical service to staff members of the Agricultural Experiment Station," was formally established in 1947. Objectives for 1-S were to offer to station workers such help and advice as they might wish in the design of experiments; to supply, if needed, help in summarizing data, in directing computations and in interpreting the statistical analyses; to assist in specialized analyses of incomplete data; and to review project outlines and manuscripts. Under Paul Homeyer, leader-in-charge, particular attention has been recently given to the following: (1) the design and analysis of feeding experiments, particularly for swine; (2) the relative efficiencies of different sizes and shapes of plots for sweet potato experiments; (3) the comparisons of two or more scales of measurements; (4) the design, analysis and interpretation of frozen-food storage experiments involving subjective evaluations; (5) problems in evaluating diets of school children; (6) the design and analysis of pasture experiments and (7) a study to develop efficient methods of analyzing data with multiple classification and with unequal and disproportionate frequencies in subclasses.

Consulting activities also included collaborative work between the statistical section of the experiment station and the National Corn Borer Laboratory located at Ankeny, Iowa.

## EXTENSION SERVICE FOR AGRICULTURE AND HOME ECONOMICS

A member of the staff is a consulting participant on the Extension Service Committee on Studies, Training and Methods, to make recommendations concerning proposed evaluative studies of projects or programs in which the home economics branch of the extension service participates. This member's activities include helping in outlining patterns for specific state and regional extension studies, advising on appropriate techniques and areas for studies, and serving as contact person between extension staff members and Laboratory personnel whenever certain specialized knowledge is needed.

## OFF-CAMPUS CONSULTATION AND CORRESPONDENCE

Research organizations and industrial concerns, as well as government organizations, are invited to bring their problems to the Laboratory for consultation. It is felt that the continual contact of the staff members with the varied problems of such outside groups and experience with handling them are useful in guiding and stimulating research within the Laboratory. Persons who are unable to visit the Laboratory to bring their problems for consideration are frequently helped by correspondence addressed to the Laboratory.

One interesting problem submitted recently by the forestry section of another experiment station concerns the design of 60-year experiments in the production of even-aged hardwood stands on previously cutover and burned-out forest lands. The following are illustrative of the types of problems or questions received by the Laboratory and handled mainly through correspondence and occasional direct conference:

| Source | Wanted to know: |
|---|---|
| Research worker in Guatemala | How to make tests of hypotheses about individual treatment differences in a randomized blocks experiment. |
| Agricultural experiment station | How to make appropriate comparisons of the total yield of grain and forage on experimental plots. |
| Chemical company | How to design a sample to determine the effect of a set of radio programs on awareness of a brand name of a chemical product sold to farmers. |
| Psychologist | How to test a null hypothesis concerning students' interests in various scientific subjects. |
| Midwestern university school of medicine | How to handle a methodology problem on determination of the effect of anesthesia on three different brain tissues, as measured by oxygen consumption of excised tissue samples. |

| | |
|---|---|
| Agricultural experiment station | Advice about the handling of data from an organoleptic experiment in which none of the treatments were determined unambiguously. |
| Home economics department | How to obtain information on the effectiveness of television in teaching women how to make a dress. |
| Man in Argentina | Appropriate statistical methods to be used in the comparison of varieties of hybrid corn. |
| Canadian experimental station | About the analysis of variance of some data on a poultry experiment. |
| Research institute | About an appropriate method for testing independence in samples from a binomial distribution. |
| Psychologist | A test of independence on correlated data. |
| A cooperative association | How to sample opinions of members in regard to the format and content of the association paper. |
| Agricultural experiment station | How to evaluate treatment relations in a randomized blocks experiment with an incomplete list of combinations of a factorial design. |
| Student in an Eastern college (who was writing a thesis) | How to extract information from an improperly designed experiment on methods for determining porosity of sandstone. |
| Director of a laboratory in a sanitorium | Opinion concerning the evaluation of experimental data on disinfectants used in the hygiene of tuberculosis. |

## QUERIES

A portion of the questions which are asked through correspondence are dealt with in the department, "Queries", which has been edited by George W. Snedecor for the Biometrics Bulletin ever since that publication was begun in 1946. Not only present staff members of the Laboratory but also other competent authorities help to furnish answers to these queries.

# SERVICES OFFERED BY THE LABORATORY

There are other aspects of the work of the Laboratory that may be considered as service activities. These fall under the headings of a survey design service, the processing of survey data, a computing service, the rental of calculating machines, the maintenance of a machine room for the use of students and other persons on the campus, and a library of reprints of articles and of journals deemed important in the field of statistics.

Editing of questionnaires in the Computing Room.

Survey Service: processing a sample for field use.

## SURVEY SERVICE

The function of the Survey Service is to advise, assist or collaborate on such aspects of sample survey work as: the development of the sampling plan, preparation of the questionnaire, training of the interviewers, and supervision of the field work, as well as the calculation of population estimates and sampling errors from survey data. The Computing Service may be called upon, in connection with a particular survey, to edit and code questionnaires and process data on the various IBM machines.

Limited opportunity is provided for graduate students to gain practical experience in survey methodology on current sampling and survey problems. For example, the Survey Service prepared a sample for a state-wide Survey of Food Habits of Women over Age 30 in South Dakota, together with sets of instructions on survey procedures in the field for interviewer use. A three-day interviewer training school was held under the supervision of Laboratory personnel. One feature of the sample design was the extensive use of cruising operations within rural places and cities in the construction of approximately equal-sized small clusters of households as sampling units. One of the graduate students from the Laboratory worked with the interviewers during cruising and handled the probability selection of household clusters in those towns and cities for the sample. He then remained in South Dakota for a month's time to administer the field work. Another student who was interested in the theory of sampling devised several approximate methods of estimation of sampling errors for a farm housing survey.

### Designation of samples

The following samples were drawn and processed by the Survey Service during the 1949-50 year:

Samples of quarter-section areas for a survey of soils samples in six North Central states (for the U.S. Soil Conservation Service).

A 1600-farm sample in the Tennessee Valley area (for the Tennessee Valley Authority).

A sample of farms 50 acres and over in a number of North Central states to get accurate information relative to the purchase and use of petroleum products by the farm operators (for Market Facts, Inc.)

A sample of farms in Johnson county for a survey to determine uses of lumber on the farm and importance of farm woodlots (for the forestry section, Iowa Agricultural Experiment Station).

A sample of farms for a Farm Roofing Survey in Iowa (for the agricultural engineering section, Iowa AES).

A sample for an appraisal of egg marketing methods in Iowa together with various economic aspects of the markets; a 250-farm sample in the Ida-Monona soils area of Iowa for a study of risk and uncertainty in farming; and a sample of open country farms 30 acres and over in size in ten southern Iowa counties for an economic study (all for the economics section, Iowa AES).

Several rural population samples in Oklahoma (for the Industrial Agricultural Development Service, Oklahoma A&M).

A 19-county sample of areas for a Livestock Marketing Survey of farms in Georgia; a sample of rural households for a Tennessee Rural Spending Survey; and a sample of open country households in York county, South Carolina for a study of level of living of white rural families (for the Institute of Statistics, North Carolina State College at Raleigh, which designed these three samples).

## Other survey operations

Other operations performed by the Survey Service included: sample checks from the completed schedules of a Study of the Food Habits of Women over 30 in Iowa, the Nebraska Farm Family Housing Study and the Western Farm Housing Requirements Study; a sample check and tabulations for the South Dakota food habits study; tabulation of data and writeup of the sampling plan for the Teacher Opinion Survey conducted by the Iowa State Education Association. Special tabulation work (for the University of Iowa), relating to items of readership information and household characteristics, was done from data of an Iowa farm survey.

The Laboratory continued cooperation on a study of the media of dissemination of market news and their use by Iowa farmers, Project 1031 of the Iowa Agricultural Experiment Station, in collaboration with the Production & Marketing Administration. This study deals with farmers' exposure to market news and its penetration or effect on their marketing behavior. It also shows the relative use farmers made of the various media disseminating market news. The problem was approached through a study of the ways farmers used information when they marketed their most recent lots of hogs, cattle, corn, soybeans, cream or milk, and eggs if any of these were sold in 1948.

Members of the Laboratory have been closely connected with this project in planning, questionnaire design and construction, survey sampling, training the interviewers, supervision of field work, editing and coding the questionnaires, preparation of preliminary sampling errors and estimates and analysis. The study will be written up in the form of three manuscripts entitled "How do Iowa farmers obtain and use market news?", "What does the Iowa farmer want from radio market news?" and "What does the Iowa farmer want from newspaper market news?"

A writeup of the sample methodology used in an Iowa farm tenure study (for inclusion in an Iowa Agricultural Experiment Station bulletin) and a description of sampling procedures and measures of reliability for the Michigan Medical Needs Survey (Michigan State College) were prepared. Sampling errors were calculated for the North Central Farm Family Housing Study, and preparation of the technical appendix on survey methodology was continued. A bibliography on the extent, effect and treatment of missing data in surveys was compiled.

## COMPUTING SERVICE

At the left are shown two of the late model tabulating machines acquired by the Computing Service.

The Laboratory offers the service of a staff of trained statistical clerks on any data that have been submitted by students, faculty or off-campus persons or organizations and that have been checked over by a qualified staff member. Facilities include IBM key punching, tabulating and auxiliary equipment with a staff of trained operators.

## MACHINE RENTALS

The Laboratory is in charge of the servicing of calculating machines that are placed in the various offices around campus. A fee, used to pay for annual service charges, is assessed for use of the machines. Computing machines which may be taken from the Lab for long- or short-term periods include a number of different current makes and models of calculating and adding machines.

## MACHINE LABORATORY

A computing room equipped with various models of calculating equipment is maintained by the Laboratory for the special use of students, laboratory classes in statistics and other campus personnel who wish to do their own computations. Graduate assistants are assigned to this computing laboratory to give advice on mechanical or statistical procedure. Students in laboratory sections of Stat. 301-302 and Stat. 401-402 are familiarized with new methods of machine computation.

## OTHER ACTIVITIES OF THE STAFF MEMBERS

Staff members have participated in various professional meetings throughout the year. Ray Jessen attended the meetings of the American Farm Economics Association, August 18-20, 1949, at the University of Wyoming, Laramie. The meetings of the Institute of Mathematical Statistics at Boulder, Colorado, August 29-September 1, 1949, were attended by Virgil Anderson. A paper, "Theory of econometrics", by Leonid Hurwicz was presented at the 26th International Statistical Conference, Berne, Switzerland, in September. Part I was published in the Proceedings of the Conference. Bernard Ostle attended the Fourth Midwest Quality Control Conference of the American Society for Quality Control at St. Louis, November 10-11. The 109th annual meeting of the American Statistical Association in New York City, December 27-30, was attended by George W. Snedecor, T. A. Bancroft who read a paper on "Bias due to the omission of independent variables in ordinary multiple regression analysis", and R. J. Jessen. Gerhard Tintner gave a paper, "An econometric investigation of the British labor market", at the winter meeting of the Econometric Society held in New York concurrently with the annual ASA meeting. On February 1, 1950, Jessen spoke on "The development of area sampling" at a joint meeting of the Chicago chapters of the American Statistical Association and the American Marketing Association.

Dorothy Cooke attended the 1949 summer session in Survey Research Techniques given by the Institute for Social Research at the University of Michigan. Snedecor was a guest lecturer at the summer session of the North Carolina Institute of Statistics June-July 1949.

In October Bancroft spoke on "Job opportunities in statistics" during the Freshman Science Orientation Series on campus.

O. Kempthorne was a guest of the Committee on Statistics, University of Chicago, November 10-11. While there he presented two talks, one on "The functions of the design of experiments" for the Statistics Club and one on "Factorial designs". November 21, Kempthorne lectured on "Recent developments in the design of experiments" at the University of Illinois, Urbana.

Paul Homeyer served as statistical consultant at the Swine Nutrition Conference for Collaborators, February 18, 1950, a meeting on campus of the private organizations working jointly with the college on swine nutrition research. In connection with talks given on experimental results of current studies, Homeyer described methods of analysis used and statistical implications for future research. He also aided in preparation of a written report to collaborators, presenting results of research completed during the year and tentative results for research still in progress. On April 20 Homeyer met with some of the same group to discuss and lay out plans for future research to be undertaken.

At the annual meeting of Pi Mu Epsilon, mathematics honorary society, St. Louis University, April 19th, to which the St. Louis chapter of the ASA was invited, Snedecor gave a talk on "Life with statistics".

Tintner was appointed to the Board of Editors for Metroeconomica (International Review of Economics, published every four months) beginning with its first issue in April 1949.

Homeyer spent a month in Washington, D.C., in early summer 1950, with the Division of Housing and Household Equipment, Bureau of Human Nutrition and Home Economics, USDA. This was in part a continuation of his activities the previous summer with the bureau, advising on regional farm housing surveys, washing machine experiments, etc.

During June 1950, Professors Jessen, Strand and Jebe, as consultants with the Bureau of the Census, attended a two-week training school in Washington, D.C., for training field supervisory personnel, observers and consultants for various quality checks on the 1950 census.

Members of the Department of Statistics participated in the seminar programs of other departments of instruction on campus and elsewhere to the following extent:

| | |
|---|---|
| November 16 | Paul Homeyer: The role of statistics in research. . . to the Dairy Industry Seminar on campus. |
| February 20 | Paul Homeyer: The role of statistics in agricultural research. . . . to the Horticulture Department on campus. |
| February 22 | Bernard Ostle: Industrial applications of statistics. . . to the General Engineering Faculty Seminar, Iowa State College. |
| April 13 | Gerhard Tintner: External economics in consumption--a study in welfare economics. . . for a Cowles Commission Seminar. |
| April 13 | Gerhard Tintner: Some formal relations in multivariate analysis. . . . to the Statistics Seminar group, University of Chicago. |

# 1949-50 PUBLICATIONS[1] OF THE STAFF

## WITH SUMMARIES

*BANCROFT, T. A., "Probability values for the common tests of hypotheses." Jour. Amer. Stat. Assn. 45:250, 211-217. June 1950.

Sources and formulas for obtaining probability values for the common tests of statistical hypotheses are listed with an indication of the scope and utility of each. Examples are given which illustrate methods of calculating probability values outside the range of available tables for a large group of important tests. The examples were selected from some given in Professor George W. Snedecor's text, Statistical Methods, 4th edition.

*BANCROFT, T. A., "Some recurrence formulae in the incomplete beta function ratio." Annals Math. Stat. 20:3, 451-455. September 1949.

Several important recurrence formulae are derived. Uses of the formulae in extending Karl Pearson's Tables of the Incomplete Beta Function are demonstrated.

COCHRAN, WILLIAM G. (Johns Hopkins University; Iowa State College) and GERTRUDE M. COX (University of North Carolina; Iowa State College), Experimental Designs. New York: John Wiley & Sons, Inc. 1950.

Work on this book was begun when both Cochran and Cox were members of the staff at Iowa State College as consultants to research workers. In 1944 a mimeographed draft was completed and distributed for reader comment. Completion of the book was delayed until after the war. The book describes in some detail the most useful of the experimental designs that have been developed, with accompanying plans and an account of the experimental situations for which each design is most appropriate. It includes a brief review of the basic theory of analysis of variance and an extensive set of worked examples of the analysis for a number of types of design. Examples are taken from real-life research situations. Frequent references are made to the work of other former and present members of the (Iowa) Statistical Laboratory staff. The book will be useful both as a handbook for research workers who have some knowledge of statistics and as a textbook in experimental design.

HILDRETH, CLIFFORD C., "Combining cross-section data and time series". Cowles Commission Discussion Paper: Statistics No. 347. Dittoed. 13 pp. May 16, 1950

---

1/ * Reprints available.
** Reprints available in limited supply.

The joint use of cross-section data and time series may be expected to yield two kinds of contributions to the problem of estimation of economic relations from empirical data--e.g., (1) "the investigator can expect to work with large numbers of observations, thus reducing his sampling errors and making tests of significance more powerful;" (2) in addition he is in a position to choose from a wider selection of statistical models and thus has a better chance of constructing a model which is both realistic and manageable. In this paper, a few problems of joint use of these types of data are considered and illustrations of the two contributions are given.

HILDRETH, CLIFFORD C., (Abstract) "The use of survey data in econometric studies"--paper presented September 1, 1949, before the Econometric Society, Boulder, Colorado. Econometrica. 18:2, 186-187. April 1950.

A discussion of the relative merits of time-series and survey data for the estimation of structural economic relationships. The problem is discussed mainly from the point of view of identification and aggregation.

HILDRETH, CLIFFORD C., "Problems in the estimation of agricultural production functions." Cowles Commission Discussion Paper: Economics No. 260. Dittoed. 8 pp. 1950. (also Abstract: Econometrica. 17:2, 163-164. April 1949).

A linear model consisting of two equations is constructed for agricultural production functions. The first equation includes crop output, labor, equipment and cost of land. There is also a similar equation for livestock production. Another model is linear in the logarithms. The method of limited information is used in order to estimate the parameters in both models.

HILDRETH, CLIFFORD C., "Preliminary considerations regarding time-series and/or cross-section studies." Cowles Commission Discussion Paper: Statistics No. 333. Dittoed. 13 pp. July 18, 1949.

The general model of production is constructed and stochastic elements are introduced. The author discusses the comparative advantages of a time series on a cross-section study of such a model, mainly from the point of view of the problem of identification.

HURWICZ, LEONID. The following articles were completed by Hurwicz prior to June 30, 1949, although they were not published until 1950:

"Statistical inference in dynamic economic models: (Part One) IV. Generalization of the concept of identification." Cowles Commission Monograph No. 10. New York: John Wiley & Sons. 1950.

"Statistical inference in dynamic economic models: (Part One) VI. Prediction and least squares." Cowles Commission Monograph No. 10. New York: John Wiley & Sons. 1950.

"Statistical inference in dynamic economic models: (Part Two) XI. Variable parameters in stochastic processes: trend and seasonability." Cowles Commission Monograph No. 10. New York: John Wiley & Sons. 1950.

"Statistical inference in dynamic economic models: (Part Two) XV. Least-squares bias in time series." Cowles Commission Monograph No. 10. New York: John Wiley & Sons. 1950.

"Statistical inference in dynamic economic models: (Part Three) XVIII. Systems with nonadditive disturbances." Cowles Commission Monograph No. 10. New York: John Wiley & Sons. 1950.

Since abstracts of these articles have already been printed, in the 1946-47 Annual Report of the Laboratory where they were reported "in press", no further discussions will be given here.

JEBE, EMIL H., "The theory and application of the selection of primary units for sampling an agricultural population." Ph.D. thesis. North Carolina State College Library. June 1950.

The work follows the theory and application described by Hansen and Hurwitz in the Annals of Mathematical Statistics for December 1943. The counties of North Carolina were considered as the primary units. The state was divided into four, ten and twenty strata for examination of various methods of estimation of selected state totals of agricultural items. A number of linear and ratio estimates were used. Both the between-primary sampling unit and the within-sampling unit components of the total survey error were considered in studying the several estimates.

Linear estimates were generally found to be unsatisfactory because of the large between-primary unit component of the error but show some improvement with respect to reduction of the within-county component of the survey error. A ratio estimate using the ratio of the current sample total to a total for the same units of observation at a previous date will yield a satisfactory estimate (coefficient of variation of estimate total of the order of 5%) for a reasonable sample size. This procedure, however, requires repeat enumeration of the same sampling units (farms or Master Sample segments). As an alternative a "semi-ratio" estimate is proposed which uses the ratio of the sample estimate of a primary unit total to a known total for the primary unit at a previous date. This estimate has a lower between-county contribution to the total

error but exhibited a large within-county contribution for North Carolina because of the extreme variability of the Master Sample units in that state. Some attention was also directed toward the problem of the selection of more than one primary sampling unit per stratum.

McCREARY, GARNET E., "Cost functions for sample surveys". Ph.D. thesis. Iowa State College Library. June 1950.

The present study is an attempt to present theory that would enable one to make fairly accurate estimates of the field costs of various alternative sample survey designs being considered for a particular investigation.

It is assumed that travel follows an existing road network which forms a grid pattern and that n random points (representing s.u.'s) have a uniform distribution over the region being studied. A formula was obtained for the expected minimum grid distance for n = 3 in a square. It was found that moderate changes in shape of regions, given constant area, have a minor effect on expected distance. The mean grid distance may be found between two random points in any connected region where the points may have any space distribution over the region. Mean airline distance can be predicted fairly accurately from mean grid distance.

The following formulas, results of empirical investigations on minimum average distance, were derived: (1) an upper bound to the expected minimum grid distance among all n points in a convex region of area A; (2) the expected grid distances for a stratified and an unstratified sample, if the path among the points does not reverse a certain direction; (3) the expected distance of a random point from (a) the center of the arc which forms a circle, semicircle or quadrant, (b) any fixed point, inside or outside a rectangular region; (4) the mean square distance between any pair of random points adjacent in either a clockwise or a counterclockwise direction, for a circle, semicircle or quadrant (although there is reason to believe that the square root of this mean square distance has an upward bias of 6.7 to 9.5 percent over corresponding mean airline distance).

Certain conclusions were drawn regarding the most efficient design with respect to total distance to be traveled. Detailed mileage records of three Iowa farm surveys were compared with theoretical estimates of distance, and a prediction equation was set up for certain travel components (such as for Night's Headquarters, etc.) If the cost is balanced against the losses resulting from errors in estimate for a particular design, the problem of determining sample size can be broached by using Wald's minimax principle.

MOOD, ALEXANDER McFARLAND (RAND Corporation; Iowa State College). Introduction to the Theory of Statistics. McGraw-Hill Book Company. 1950.

This book was primarily written while the author was on the staff of the Statistical Laboratory. Although the book is a text in the theory of statistics, Dr. Mood has not lost sight of the fact that statistics is a branch of applied mathematics and theoretical statistics should concern itself with theory motivated by practical considerations.

Mood's book covers a wide range of topics at the introductory theoretical level. Beginning with concepts and fundamental laws of probability, it proceeds to a detailed discussion at the introductory level of discrete and continuous parent populations and their descriptive properties, derived sampling distributions, point and interval estimation, and tests of hypotheses. A briefer discussion is given of regression and linear hypotheses, experimental designs and the analysis of variance. A later chapter is an exposition of sequential tests of hypotheses. The final chapter contains some new material in distribution-free tests which were developed jointly by George W. Brown and Mood while they were at Iowa State College, on a project sponsored by the Office of Naval Research. The inclusion of this wide range of topics enhances the value of the book for research workers in various fields who have the required modest mathematical background.

OSTLE, BERNARD, "On certain criteria for optimum estimation." Ph.D. thesis. Iowa State College Library. August 1949.

The paper discusses Pitman's concept of closeness in the theory of statistical estimation. If X and Y are two unbiased estimates of a parameter $\theta$, the closer estimator will be the one possessing the smaller variance except when the ratio of the variance is in a certain finite interval. If $S = X/\sigma_1$ and $T = Y/\sigma_2$, where $\sigma_1^2$ and $\sigma_2^2$ are the variances of X and Y respectively, the $P(|s|<|t|) = \frac{1}{2}$ is a necessary and sufficient condition for the equivalence of closeness and smaller variance, provided X and Y possess a continuous probability density function. Certain extensions were possible to (i) biased estimators, (ii) the multi-parameter case, and (iii) asymptotic distributions. The results indicate a further justification for using unbiased minimum variance estimates of an unknown parameter.

*SETH, GOBIND RAM, "On the distribution of the two closest among a set of three observations." Annals Math. Stat. 21:2, 299-301. June 1950.

The joint distribution is obtained of the two closest observations $x'$, $x''$ $(x' < x'')$ of the set $x_1$, $x_2$, $x_3$, $(x_1 \leq x_2 \leq x_3)$, when the distribution of $x_1$, $x_2$, $x_3$ is given or can be obtained. It is assumed that

in general the density function, $f(x_1, x_2, x_3)$, is continuous in the variables involved. The joint distribution of $(x'' - x')$ and $(x'' - x')/(x_3 - x_1)$ is also obtained. Applications are discussed for normal distributions.

*SNEDECOR, GEORGE W., "The statistical part of the scientific method." Annals New York Acad. Sciences. 52: Art. 6, 792-799. March 1950.

"A striking feature of the statistics of the scientific method, the measurement of the uncertainty of the conclusion, has been discussed. The measure rests on the data from the experiment. The uncertainty measured is that due to imprecision, not inaccuracy. The measurement of uncertainty directs the scientist's attention to methods of increasing the sensitivity of succeeding experiments, to designs which enable him to avoid some sources of inaccuracy and to measure others, and finally to estimates of sample size which give a specifiable assurance of an unambiguous conclusion in the succeeding cycle of his research."

**TINTNER, GERHARD, "Some formal relations in multivariate analysis." Jour. Royal Stat. Soc., Series B (Methodological). 12:1, 95-101. 1950.

The formal relations between the following types of multivariate analysis are investigated: canonical correlation, principal components, weighted regression, discriminant analysis. It is shown that they can all be reduced to two fundamental types and that there is a close relationship between those.

**TINTNER, GERHARD, "Die identifikation: ein problem der okonometrie." Statistischen Vierteljahresschrift. 3, 7-12. 1950.

The problem of identification deals with the question, how to derive economically meaningful relationships from the data of economic statistics. This is discussed and exemplified by the derivation of demand and supply functions for American agriculture.

**TINTNER, GERHARD, "La position de l'econométrie dans la hiérarchie des sciences sociales." Revue d'Économie Politique. Vol. . ., p. 634-641. 1949. (Also U. of Cambridge, Dept. of Applied Economics. Reprint Series No. ? in preparation).

Relationships between econometrics and economic theory are discussed. It is pointed out that econometrics is an empirical science and that it can frequently be used for the numeric evaluations of economic laws and the verification of hypotheses. Finally, the relationship between econometrics and statistics is discussed from the point of view of the difficulties of the applica-

tion of methods of modern statistical inference in the field of economics. These propositions are illustrated by examples from earlier investigations of the author.

**TINTNER, GERHARD, "Foundations of probability and statistical inference"— paper read before the Royal Statistical Society, March 16, 1949. Jour. Royal Stat. Soc., Series A (General). 62: Part III, 251-286. 1949. (Also U. of Cambridge, Dept. of Applied Economics. Reprint Series No. 24. 1950).

According to Carnap, there are two concepts of probability: probability as a degree of confirmation and probability as a limit of the empirical frequency. This paper deals essentially with the first concept and its particular formulation by Carnap. Carnap's ideas are explained and illustrated and applied to the following fields of statistical inference: prediction, estimation, testing hypotheses and choice between statistical hypotheses.

## SUMMARIES OF OTHER RECENT STAFF PUBLICATIONS[1]/

*FEDERER, WALTER T. (Cornell University; Iowa State College), "The general theory of prime-power lattice designs: III. The analysis for $p^3$ varieties in blocks of p plots with more than 3 replicates." Biometrics. 5:2, 144-161. June 1949.

Designs for $k^3$ varieties in blocks of k plots were given by Yates and called three-dimensional lattices. In general a multiple of three basic replicates must be used. In the case where k is a prime number or a power of a prime, however, any number of different replicates up to $(p^2 + p + 1)$ can be used. This paper describes and illustrates by an example the analysis of an experiment on $p^3$ varieties in blocks of p plots with 4 replicates. The procedure can be easily extended to the use of more than 4 (different) replicates.

HOMEYER, PAUL G., (Abstract) "Sampling replicated field experiments for yield determinations." Proc. Auburn Conf. on Statistics, Sept. 7-9, 1948. Lithoprinted. Pp. 45-46. Edwards Brothers, Inc. Ann Arbor, Michigan. 1949.

The report gives a brief summary of some results obtained from analyzing the sampling data on yield determinations taken on corn and oats in India during the past several years. For drilled corn, samplings of every $k^{th}$ producing stalk were found to be much more efficient than those of 8-foot segments taken at random. For hill corn similarly, samplings of every $k^{th}$ hill were more efficient than those of 2x2 and 2x5 hills selected at random. Gains in efficiency were obtained for both drilled and hill corn by making a count of the total number of stalks in each plot and using the per stalk method of estimation rather than the per sampling unit method. For oats, a sampling unit 3'x3' was slightly more efficient than a 2'x2' one, in terms of amount of information per hour of sampling, and the smaller sampling unit overestimated plot yields by 8% more than did the larger one.

HURWICZ, LEONID, (Abstract) "Linear programming and general theory of optimal behavior." Econometrica. 17:2, 161-162. April 1949.

The problem is to maximize a weighted sum under linear restrictions. The difficulties of solution are discussed and compared to more conventional economic models. The importance of convexity is emphasized.

---

1/ Summaries of these articles have not previously been printed in the Laboratory Annual Reports. * indicates reprints available; **, reprints available in limited supply.

HURWICZ, LEONID, "Some results on identification in lagged shock-error models." Cowles Commission Discussion Paper: Statistics No. 324. Dittoed. 6 pp. January 28, 1949.

The paper presents results which form a part of a paper by T. W. Anderson and Hurwicz. These results indicate conditions (necessary and/or sufficient) for identification in models with both shock and errors present. It is assumed that the disturbances, as well as the exogenous variables, are a normal universe with mean zero. Also, covariance matrices are independent of time. This reduces the identification problem to one of finding whether the second order moments of the distribution of the observed variables provide a unique determination of the structural parameters. A number of propositions are developed for two models.

SNEDECOR, GEORGE W., "Some principles of experimental design." Lecture, Proc. Auburn Conf. on Statistics, Sept. 7-9, 1948. Litho-printed. Pp. 47-51. Edward Brothers, Inc. Ann Arbor, Michigan. 1949.

The two principles discussed are: (i) To provide unambiguous answers to the questions asked, and (ii) To get the answers with the minimum expenditure. The questions asked should be specific and specified. Unambiguous answers are sought by randomization, adequate replication and appropriate experimental controls including the measurement of correlated variables. Economy may be attained by the selection of homogeneous experimental material, by the use of a suitable design and by the statistical control of extraneous variables. The investigator should seek, with specifiable uncertainty, to know in advance that his experiment will, without excessive cost, yield unambiguous answers to the questions asked.

**SNEDECOR, GEORGE W., "On the design of sampling investigations." Amer. Statistician. 2, 6-9. December 1948.

In order that a sample may lead to correct inferences about the population sampled, every individual in the population must have a chance of being drawn in the sample and the choice of individuals must be random. If the population is heterogeneous, stratification may be employed or other devices may be used for insuring wide scattering of the units. Efficiency in stratified samples may be attained, if the cost per interview is constant, by allocating the sampling to the strata in joint proportion to their sizes and standard deviations. Moderate deviations from optimum allocations do not greatly affect the resulting estimates but errors in stratum sizes can be disastrous. Areal sampling units may be randomly selected if lists of individuals are not available. This may be associated with cluster sampling, usually with decreased costs.

* Reprints are now available of Snedecor's "Increasing the efficiency of sampling investigations." Jour. Ala. Acad. Sci. 20, 53-57. December 1948.

**TINTNER, GERHARD, "Scope and method of econometrics," (paper read before the Statistical and Social Inquiry Society of Ireland March 21, 1949). Jour. Stat. & Social Inquiry Soc. of Ireland. 14:1, 15 pp. 1949. (Also U. of Cambridge, Dept. of Applied Econ. Reprint series No. 22. 1949).

A general methodological discussion of the relationship of econometrics to mathematical economics and statistics. The difficulties involved are exemplified by the statistical derivation of a demand and supply function for American agriculture.

**TINTNER, GERHARD, "Static macro-economic models and their econometric verifications." Metroeconomica. 1:1, 48-52. April 1949.

A comparison of the Keynesian and classical economic models. The question of statistical verification of those models is discussed.