STATISTICS

In 1968 CUPM appointed a Panel on Statistics for the purpose of providing guidance to departments of mathematics at smaller colleges and universities on instruction in statistics. Two concerns of general interest were identified for study by the Panel: a program to prepare students for graduate study in statistics and a basic service course in statistics for students who have not studied calculus. The Panel pointed out that these two topics represent curricular extremes for statistics instruction in most undergraduate programs and that many students' program of study will lie between these extremes.

The Panel's first report, Preparation for Graduate Work in Statistics, was issued in 1971. This document describes the type of training and experiences which undergraduates contemplating graduate study in statistics ought to have. It outlines a basic one-year course in probability and statistics, indicates those mathematics courses which are valuable for pregraduate preparation in statistics, and comments on computer requirements and experience with data.

The Panel's second project involved a study of the introductory, noncalculus statistics courses which are offered by practically every college and taken by students in a wide variety of fields. Prompted by the fact that many of the existing courses are unsatisfactory for a variety of reasons, the Panel developed a set of objectives for such a course and made concrete suggestions for realizing the objectives. A detailed list of topics for a conventional course in introductory statistics, as well as some suggested alternate approaches, appears in the 1972 publication Introductory Statistics Without Calculus.

PREPARATION FOR GRADUATE WORK IN STATISTICS

A Report of

The Panel on Statistics

May 1971

TABLE OF CONTENTS

...Both Computer Science and Statistics have dual sources of identity and intellectual force, only one of which is mathematical; hence they are more accurately described as partly mathematical sciences. ...

Modern statistics could not operate without mathematics, especially without the theory of probability. Equally, it could not exist without the challenge of inference in the face of uncertainty and the stimulus of the quantitative aspects of the scientific method ... statistics is both a mathematical science and something else.

...It is true that undergraduate preparation for majors in mathematics, with its traditional emphasis on core mathematics, provides an excellent foundation of knowledge for potential graduate students in statistics. It does not, however, provide nearly enough students with either motivation to study statistics or an understanding of the extra-mathematical aspects of statistics.*

This report consists of three main sections: (1) Introductory comments on the field of statistics and its study at the graduate level; (2) A recommended undergraduate program for prospective graduate students in statistics; and (3) Implications of the recommendations for departments of mathematics and their students. Our recommendations are addressed to departments of mathematics of four-year colleges and smaller universities which have no specialized departmental programs in statistics. At these institutions, the department is unlikely to have an experienced or trained statistician although it is often called upon to offer statistics courses as a service for students majoring in other fields.

I. STATISTICS AND GRADUATE STUDY

In our modern technological society there is a continually increasing demand and necessity for quantitative information. This requires planning and skill in the collection, analysis, and interpretation of data. Statisticians deal with inherent variation in

* The Mathematical Sciences: A Report, by the Committee on Support of Research in the Mathematical Sciences of the National Research Council for the Committee on Science and Public Policy of the National Academy of Sciences, Washington, D. C. Publication 1681 (1968), pp. 84 and 157.

nature and measurement and are concerned with the planning and design of experiments and surveys, with methods of data reduction, and with inductive decision processes.

Statistics has made significant contributions to many fields, most notably to the experimental sciences, agriculture, medicine, and engineering. It has also had an important role in the development of other fields such as economics, demography, and sociology. Statistics and quantitative methods are assuming major roles in business and in the behavioral sciences, roles destined to receive more and more emphasis. The widespread use of computers in these fields increases the need for statisticians at all educational levels.

Although a demand exists in government and business for persons with only undergraduate training in statistics (especially in conjunction with training in computer technology and a subject-matter field), the attainment of competence in statistics at a professional level necessarily requires graduate study. Our recommendations deal with minimum undergraduate preparation for this study. It is generally agreed that broad knowledge of mathematics is required to proceed with graduate work; currently about two thirds of all graduate students in statistics were undergraduate mathematics majors.* Advanced study in some field (physical, biological, or social science) in which data play an important role is also very helpful.

A significant part of graduate study in statistics is the attainment of a sound understanding of advanced mathematics and the theory of statistics and probability, since these are necessary for research in statistics and also for competent consultation on applications of statistics. The consultant seldom encounters textbook applications and is regularly required to modify and adapt procedures to practical problems.

Graduate study in statistics, at both the M.A. and Ph.D. degree levels, is available in a substantial number of universities. According to a survey published in The American Statistician (October, 1968), there are in the U. S. and Canada approximately 85 departments which offer undergraduate degrees in statistics. There are approximately 160 departments at 110 universities which offer programs leading to graduate degrees in statistics or subject-matter fields having a statistics option. These graduate programs of study lead to frontiers in both theory and application. The M.A. degree provides suitable qualifications for many positions in industry and government, often as a consultant or team member in research and development, as well as for positions in teaching at junior colleges.

---

* Aspects of Graduate Training in the Mathematical Sciences, Vol. II, page 62, 1969. A report of the Survey Committee of the Conference Board of the Mathematical Sciences, 2100 Pennsylvania Avenue, N.W., Suite 834, Washington, D. C. 20037.

It is also useful for persons who will pursue advanced degrees in fields such as psychology and education in which statistical methodology plays a significant role. The Ph.D. degree provides additional preparation for research and teaching careers in universities as well as in government and industrial organizations.

II. THE RECOMMENDED PROGRAM OF STUDY

Our recommendations for undergraduate courses are subdivided into four areas: (A) probability and statistics, (B) mathematics, (C) computing, and (D) other requirements. In this section we describe recommended courses in each of these areas.

A. Probability and Statistics Requirements

We recommend that students take at least a one-year course in probability and statistics and gain experience with real applications of statistical analysis.

1. Probability and Statistics Course (6 semester hours)

A description of this course (Mathematics 7) can be found in Commentary on A General Curriculum in Mathematics for Colleges, page 79.

2. Experience with Data

We believe that students should have experience with realistic examples in the use of the statistical concepts and theory of the key course. They should work with real data, consider the objectives of the scientific investigation that gave rise to these data, study statistical methods for answering relevant questions, and consider the interpretation of the results of statistical analyses. These goals are not easy to achieve, but various approaches are discussed below.

An appropriate course is not likely to be already available in a department of mathematics. The typical one-semester precalculus elementary statistics course, usually serving students from departments other than mathematics, does not fulfill the goals that we recommend.

Courses in research methodology or applied statistics in other subject matter areas may meet, at least in part, some of the objectives of exposing students to realistic statistical problems. For example, courses in biological statistics, research methods in behavioral science, economic statistics or econometrics, survey methods, sociological statistics, etc., can meet our intended objective if they are offered by practicing scientists who are familiar with the way data are generated, the complexities they usually exhibit, and the methods that help in their analysis. Although such a course will provide coverage of some topics in elementary statistical methods, it is important that the course be more than a catalog of methods. The student should see how some one field of science generates experimental data and copes with uncertainty and variability. Insight gained from relatively few examples or ideas can be considerably more valuable for the student than information obtained by covering a large number of separate topics. Modern computers could be useful in such a course.

If a faculty member with training and experience in applied statistics is available to the mathematics department, he can devise a data analysis course. This course could be based on a number of data sets of interest to the student, and it could use books on statistical methods as reference material for the appropriate statistical techniques. Basic texts on statistical methods which contain a variety of examples of applications of statistical methods include:

Dixon, Wilfrid J. and Massey, F. J., Jr. Introduction to Statistical Analysis, 3rd ed. New York, McGraw-Hill Book Company, 1969.

Fisher, R. A. Statistical Methods for Research Workers, 13th ed. New York, Hafner Publishing Company, 1958.

Guttman, Irwin and Wilks, Samuel S. Introductory Engineering Statistics. New York, John Wiley and Sons, Inc., 1965.

Li, C. C. Introduction to Experimental Statistics. New York, McGraw-Hill Book Company, 1964.

Natrella, Mary G. Experimental Statistics, Handbook 91. U. S. Department of Commerce, National Bureau of Standards, 1966.

Snedecor, George W. and Cochran, W. G. Statistical Methods. Ames, Iowa, Iowa State University Press, 1967.

Walker, Helen M. and Lev, Joseph. Statistical Inference. New York, Holt, Rinehart and Winston, Inc., 1953.

Wallis, Wilson A. and Roberts, Harry. Statistics: A New Approach. New York, Free Press, 1956.

Wine, Russell L. Statistics for Scientists and Engineers. Englewood Cliffs, New Jersey, Prentice-Hall, Inc., 1964.

Winer, B. J. Statistical Principles in Experimental Design. New York, McGraw-Hill Book Company, 1962.

Examples of data sources and/or statistical critiques of major scientific investigations are:

Tufte, Edward R. The Quantitative Analysis of Social Problems. Reading, Massachusetts, Addison-Wesley Publishing Company, Inc., 1970.

Kinsey, A. C.; Pomeroy, W. B.; Martin, C. E. Sexual Behavior in the Human Male. Philadelphia, Pennsylvania, W. B. Saunders Company, 1948.

Report on Lung Cancer, Smoking, and Health. Public Health Bulletin 1103. Superintendent of Documents, U. S. Government Printing Office.

Cochran, William G.; Mosteller, Frederick; Tukey, John W. Statistical Problems of the Kinsey Report. Washington, D. C. American Statistical Association, 1954.

"The Cochran-Mosteller-Tukey Report on the Kinsey Study: A Symposium." Journal of the American Statistical Association, 50 (1955), p. 811.

Cutler, S. J. "A Review of the Statistical Evidence on the Association Between Smoking and Lung Cancer." Journal of the American Statistical Association, 50 (1955), pp. 267-83.

Brownlee, K. A. "Statistics of the 1954 Polio Vaccine Trials." Journal of the American Statistical Association, 50 (1955), pp. 1005-1014. (An invited address on the article "Evaluation of 1954 Field Trial of Poliomyelitis Vaccine: Summary Report." Poliomyelitis Vaccine Evaluation Center, University of Michigan, April 12, 1955.)

The following books use experimental and survey data to illustrate statistical concepts and techniques:

Bliss, Chester I. Statistics in Biology. New York, McGraw-Hill Book Company, 1967.

Cox, David R. Planning of Experiments. New York, John Wiley and Sons, Inc., 1958.

Davies, Owen L. Design and Analysis of Industrial Experiments, 2nd rev. ed. New York, Hafner Publishing Company, 1956.

Ferber, Robert and Verdoorn, P. J. Research Methods in Economics and Business. New York, The Macmillan Company, 1962.

Stephan, Frederick F. and McCarthy, Phillip J. Sampling Opinion--An Analysis of Survey Procedures. New York, John Wiley and Sons, Inc., 1958.

Youden, William J. Statistical Methods for Chemists. New York, John Wiley and Sons, Inc., 1951.

When an experienced applied statistician is not available, the mathematics department may be able to develop a seminar with the assistance of faculty members in other disciplines. Typical experimental areas in a discipline may be discussed and illustrated with data from on-going faculty research or from student term projects.

Another opportunity to provide students with exposure to problems in data analysis is through laboratories associated with the key course in statistics. In such a laboratory, students may be asked to analyze data sets, followed by class discussion. Alternatively, term project assignments incorporating study, review, and critique of statistical studies with major data sets could be utilized. Still another opportunity would be through projects developed for independent study following the key course.

The important element in this recommendation is that the student obtain understanding of the role played by statistical concepts in scientific investigations and be motivated to continue the study of statistics.

3. Additional Courses

We recommend that additional courses in probability and statistics be offered to follow the key course whenever possible. Such courses could explore in detail a few topics which were omitted or treated lightly in the key course, e.g., analysis of variance, experimental design, regression, nonparametric methods, sampling, sequential analysis, multivariate methods, or factor analysis.

Other subjects which could serve as useful enrichment are stochastic processes, game theory, linear programming, and operations research. Courses in these subjects would be useful to students with specialized interests and also would help widen the knowledge and capabilities of the prospective graduate student in statistics.

B. Mathematics Requirements

We recommend that students take at least a complete 9-12 semester hour sequence in calculus, a course in linear algebra, and a course in selected and advanced topics in analysis.

1. Beginning Analysis (9-12 semester hours)

This sequence includes differential and integral calculus of one and several variables and some differential equations. It is desirable that prerequisites for calculus, including a study of the elementary functions and analytic geometry, be completed in secondary school.

For detailed course descriptions, we refer the reader to Commentary on A General Curriculum in Mathematics for Colleges, page 44. This beginning analysis sequence is adequately described by GCMC's courses numbered 1, 2, and 4. We think it important to note that elementary probability theory is a rich source of illustrative problem material for students in this analysis sequence.

2. Elementary Linear Algebra (3 semester hours)

This course, which may be taken by students before they complete the beginning analysis sequence, includes the following topics: solution of systems of linear equations (including computational techniques), linear transformations, matrix algebra, vector spaces, quadratic forms, and characteristic roots. An outline for such a course (Mathematics 3) can be found on page 55.

3. Selected Topics in Analysis (3 semester hours)

The GCMC course Mathematics 5 is not particularly appropriate for statistics students, and it is recommended that a course including the special topics listed below be offered for these students in place of Mathematics 5.

Such a course should give the student additional analytic skills more advanced than those acquired in the beginning analysis sequence. Topics to be included are multiple integration in n dimensions, Jacobians and change of variables in multiple integrals, improper integrals, special functions (beta, gamma), Stirling's formula, Lagrange multipliers, generating functions and Laplace transforms, difference equations, additional work on ordinary differential equations, and an introduction to partial differential equations.

It is possible that the suggested topics can be studied in a unified course devoted to optimization problems. Such a course, at a level which presupposes only the beginning analysis and linear algebra courses and which may be taken concurrently with a course in probability theory, would be a valuable addition to the undergraduate curriculum, not only for students preparing for graduate work in statistics but also for students in economics, business administration, operations research, engineering, etc. Experimentation by teachers in the preparation of written materials and textbooks for such a course would be useful and is worthy of encouragement.

4. Additional Courses

The following courses from Commentary on A General Curriculum in Mathematics for Colleges are not required but are desirable as choices for students who wish to have more than minimal preparation. A strong course in real variables is especially recommended for students interested in working for the Ph.D. in statistics.

Mathematics 6M.	Introductory Modern Algebra.
Mathematics 8.	Numerical Analysis.
Mathematics 10.	Applied Mathematics. [See also the CUPM report Applied Mathematics in the Undergraduate Curriculum.]
Mathematics 11-12.	Introductory Real Variable Theory.
Mathematics 13.	Complex Analysis.

In general, the stronger a student's background in undergraduate mathematics, the better prepared he is for graduate work in statistics. If faculty members with special interests and competence are available, additional courses or seminars in the areas mentioned on page 466 would be valuable additions to the curriculum for a student interested in advanced work in statistics.

C. Computing Requirements

We recommend that students be familiar with a modern highspeed computer and what it can do, and that they have some actual experience analyzing data which are sufficiently obscure to require the use of a computer.

It is quite clear that the computer is becoming increasingly important to almost every academic discipline; in addition, it is an integral part of business, government, and even everyday affairs. These are strong enough reasons for every student who receives a baccalaureate degree to be acquainted with the computer and its potential. However, it is even more important that someone who intends to be a professional statistician know and understand the modern computer. Anyone who trains in statistics, who will handle data or work with and advise people who handle data must have a certain minimum competence in the use of a computer. While it is true that increased competence will be developed as the need arises and that some of the more sophisticated applications can be learned while doing graduate work, it is recommended that a student who comes into a graduate program in statistics begin his training in this area at the undergraduate level.

The most desirable way to be sure that sufficient competence is acquired would be through taking a regular course in computing such as Introduction to Computing, the course C1 described in the CUPM report Recommendations for an Undergraduate Program in Computational Mathematics [page 563]. It is, of course, possible to obtain an acceptable level of competence by attending the lectures or informal courses given in many computing centers, supplemented by sufficient additional computing experience. A student should understand and be able to use one of the major programming languages. He should learn enough of the nomenclature and characteristics of a computer to be able to stay abreast of the developments that will surely occur during his working lifetime. He should be made aware of and have some experience with library programs that are available. And, perhaps most important for our special purposes, students should have experience with actual data, with the numerical analysis and statistical problems they generate, and with the use of the computer for simulation. Some of this experience could be obtained by the techniques recommended in Section A2 above. The course CM1 of the above-mentioned report on computational mathematics [page 551] is also based on simulation techniques and would therefore be appropriate for students of statistics.

Finally, we suggest that in the next decade the availability of computers will change many subject-matter areas. If the statistician is to be an effective consultant in these areas, he must be aware of the way in which the computer is shaping the disciplines with which he will be associated.

D. Other Requirements

Statistics deals with the drawing of inferences from data and, in its applications, involves the statistician in working jointly with subject-matter specialists in the framing of relevant questions, developing appropriate methodology for drawing inferences, and assisting in the analysis of final results. Whether a person will principally do research in statistical methodology, teach statistical applications, or consult on statistical applications, knowledge of one or more areas of application and an understanding of the nature of statistical problems in them is highly desirable.

Undergraduate preparation for work in statistics should therefore include study of a variety of areas of application, with one studied in some depth. This will insure that the student, upon graduation, will have an acquaintance with fundamental concepts in a variety of areas and technical competence to a moderate extent in at least one of the physical, life, or social sciences. Courses selected for study of a field in depth may include a statistical or research methodology course offered by that field, in which the student will develop an understanding of data collection and data handling problems. The student's adviser may be particularly helpful in identifying such a course. Students who wish to undertake graduate study in specialized areas of statistics, such as econometrics or biostatistics, will find it desirable to take at least some advanced work in these areas as undergraduates.

III. IMPLICATIONS OF THE RECOMMENDATIONS

We are aware that at the present time most mathematics departments have few advanced courses in statistics available and few, if any, trained statisticians on their staffs. Because of the late date when a student may discover the field of statistics, he may not have time to elect many of our recommended courses even if these are available. Finally, some of the courses he takes will be designed not only for prospective graduate students but for students with other majors and interests as well.

Were these and other limitations not present, we would expect our recommended program and the mathematics department to serve the needs of undergraduate students by not only imparting to them knowledge of the field of statistics but also by enabling them to discover their abilities and interests and, if appropriate, by arousing their interest in graduate study in statistics or a related field. But limitations do exist, and it is not to be expected that all of our recommendations can be implemented quickly or that all the needs of students will be met by the programs in mathematics departments. It is difficult to meet all student needs under the best of circumstances. Realistically it is to be expected that severe limitations in time and facilities will prevent the student from obtaining a well-rounded understanding of the subject at the undergraduate level, but at least he can be exposed to some of the basic ideas in statistics. Perhaps most important, a program designed to achieve limited goals with relatively few courses, even if it falls short of the full program we recommend, can arouse the interest of students in statistics and related fields.

In the light of the recommendations in this and other CUPM reports (especially GCMC) dealing with courses in probability, statistics, and related areas, <u>it is highly desirable, and we recommend that each department of mathematics review its course offerings so as to establish appropriate courses in probability and statistics and arrange that these courses be staffed by a person competent in these fields</u>. Ph.D.s in statistics are increasingly available to four-year colleges and smaller universities. With the establishment of courses in probability and statistics taught by a person competent in these fields, the mathematics department can serve the needs of prospective graduate students in statistics by (1) arousing interest in and demonstrating the nature of the field of statistics; (2) giving students an acquaintance with statistics, its theory, its applications, its traditions, even some of its open problems, and its relation to other fields such as probability, pure and applied mathematics, computer science, and operations research; (3) counseling students as to courses, curricular choices, and graduate and career opportunities in probability, statistics, pure and applied mathematics, computer science, and operations research.

In creating this report, the Panel on Statistics confined its attention to recommending undergraduate programs for students who intend to do graduate work in statistics. In the development of appropriate recommendations, however, it became apparent that a broad program of study in the mathematical sciences was emerging, a program suitable, we believe, not only for graduate study of statistics but also for graduate study in the quantitative aspects of the social sciences and business and in newer areas such as operations research and computer science. The recommendations developed, therefore, can form the basis for an innovative degree program in the mathematical sciences different from the traditional programs in pure or applied mathematics.

The recommended program accomplishes important secondary objectives. These include: (1) A decision by students on the nature of their future graduate study can be made at a later point in the undergraduate program. (2) Mathematically gifted students are exposed to a wider range of potential careers than is presently the case. (3) The possibility is created for a substantial emphasis in the mathematical sciences to be used as part of an undergraduate major, not only in mathematics and statistics but also in other departments and in interdisciplinary programs.

The needs and opportunities for an innovative undergraduate program in the mathematical sciences are great. It can provide options in computer science, applied mathematics, econometrics, operations research, statistics, probability, and pure mathematics--all built around a solid core of training in mathematics. Early exposure to the concepts and possibilities of a variety of these options can lead to better choices of areas of concentration later on. A curriculum within this framework, combining mathematics, statistics, computing, and at least one field of application, has great potential for continued study in various graduate programs as well as value as a terminal degree program. The student may proceed from such an undergraduate program to advanced study in statistics, operations research, econometrics, psychometrics, demography, or computer science. He also will have excellent qualifications for advanced work in sociology, political science, business, urban planning, or education. If the undergraduate program in mathematical science is a terminal one, the student will have employment opportunities in computing, business, industry, and government, with qualifications to meet many social needs.

The minimum preparatory program outlined in Section II can be supplemented in a variety of ways with additional work leading to undergraduate majors in (1) mathematics, (2) statistics, (3) computational mathematics [see Recommendations for an Undergraduate Program in Computational Mathematics, page 528], (4) other fields (e.g., psychology, political science, economics, sociology, engineering, linguistics, business administration--especially management sciences, biological and physical sciences).

INTRODUCTORY STATISTICS WITHOUT CALCULUS

A Report of

The Panel on Statistics

June 1972

TABLE OF CONTENTS

PREFACE

Colleges and universities offer a great variety of courses in introductory statistics with no calculus prerequisite. Courses of this kind are frequently offered by several departments within a single institution, and therefore this report should be of interest to instructors in many departments, not merely to those in mathematics and statistics.

Section I contains a discussion of the background of the report and some review of past and current approaches to the general introductory course. Section II contains recommended objectives for the course and a discussion of their implications. A conventional course is developed in detail in Section III, followed by recommendations for alternative approaches in Section IV. Sections V through VII contain suggestions for improving the effectiveness of the course, useful in all of the various approaches to it. Selected bibliographies are included in each section, and a list of additional resource materials appears at the end of the report.

The Panel is indebted to many colleagues for their participation in a fact-finding conference and for their thoughtful comments on a draft version of this report.

I. INTRODUCTION

This report is concerned with a general introductory statistics course without a calculus prerequisite, which is typically a one-semester or one-quarter course offered at the sophomore or junior level in college. For many students this is a terminal course, although some students may elect additional courses in statistics or in research methods. In four-year colleges and smaller universities it is often taught by the mathematics department as a service to other departments.

An introductory statistics course without a calculus prerequisite is often required of students majoring in many different fields, such as business administration, psychology, sociology, forestry, and industrial engineering. In addition, this course serves as an elective subject for other students. An understanding of statistical concepts is important for students in any subject where data play an important role. Knowledge of basic concepts also permits students to use data more effectively in making everyday decisions as citizens and consumers, and it might stimulate them to learn more about statistics in order to obtain the competence needed for research and analysis in their major fields of interest. Some students' interest might even be aroused sufficiently by this introductory course to encourage them to prepare for a program in statistics at the graduate level.

Concern for a noncalculus-based introductory statistics course has been frequently expressed. Many studies have been conducted to consider how to make this course a worthwhile, challenging intellectual enterprise that provides students with some understanding of basic statistical concepts. (See, for instance, "Interim Report of the Royal Statistical Society Committee on the Teaching of Statistics in Schools." Journal of the Royal Statistical Society, Series A, 131 (1968), pp. 478-497.)

One point of interest has been the proliferation of introductory statistics courses that has taken place on most large campuses. These courses may differ not only in content but also in emphasis and method of approach. Some of the factors responsible for this are the desire by different disciplines to have courses tailored to their needs, the varying backgrounds of persons teaching statistics, and the different objectives for the basic statistics course.

Proliferation in itself is not necessarily undesirable and indeed may be desirable if introductory courses directed to specific disciplines serve to motivate the basic concepts of statistics more effectively or if substantially different objectives are being served. However, no widespread satisfaction has been expressed by professional statisticians concerning any of the introductory statistics courses that have been developed. These courses have been criticized as follows:

1) Their emphasis is mathematical or probabilistic without providing sufficient insights into statistical concepts.

2) They provide little insight into the variety and usefulness of the applications of statistics.

3) They are too technique-oriented, overemphasizing computation and underemphasizing the fundamental ideas underlying statistical reasoning.

4) They give insufficient attention to drawing statistical inferences from real data.

5) They fail to provide intellectual stimulation.

6) They fail to allow for the "symbol shock" suffered by some students.

7) They provide inadequate motivation for an interest in stochastic (random) models and fail to differentiate them from the deterministic models with which students are more familiar.

The discussion and recommendations that follow are designed to counter some of these criticisms through suggestions on course objectives, alternative approaches, curricular content, increased motivation, computer use, data sets, and student-generated data.

II. RECOMMENDED OBJECTIVES AND THEIR IMPLICATIONS

1. Recommended Objectives of Introductory Statistics Courses

When one considers the variety and extent of the demands for an "ideal" first course in statistics, one recognizes the impossibility of having any one course come even close to the ideal. No one course can (1) serve as an appreciation course so that students can understand the underlying ideas of statistical methodology and statistical inference, (2) discipline students to think quantitatively and to appraise data critically, (3) give students the facility to analyze data for everyday problems, (4) train students to understand probabilistic models and their uses in a variety of situations, and (5) enable students to master the basic techniques of statistical methodology and to use these techniques flexibly in their own applications.

Hence, it is not possible for the Panel to prescribe an ideal introductory statistics course. The special needs of the departments which require this course, the abilities and interests of the instructor, and the different characteristics of the students will influence the nature of the course. However, the following guidelines are recommended by the Panel:

a. <u>The introductory statistics course must have limited objectives</u>. Otherwise, it is likely that none of the objectives will be met adequately.

b. <u>The primary objective of the introductory statistics course should be to introduce students to variability and uncertainty and to some common concepts of statistics</u>; that is, to methods such as point and interval estimation and hypothesis testing for drawing inferences and making decisions from observed data.

c. <u>A secondary objective of the introductory statistics course should be to teach the student some common statistical formulas and terms and some of the widely used statistical techniques</u>; e.g., the t-test.

The primary objectives may be met in many ways. Much of the report considers the conventional introductory course but with a view to highlighting desired emphases and utilizing a variety of approaches to enhance interest and motivation. Utilization of computers, data sets, experiments, and demonstrations for greater clarity and motivation are discussed. The report also contains suggestions for implementing the primary objective through less conventional approaches, e.g., nonparametric, decision-theoretic, and problem-oriented approaches. The intent of much of the report is to suggest possible ideas for syllabus and presentation; the instructor must seek that combination suited to his needs and those of his students.

The Panel recommends the primary objective of the introductory statistics course, stated above, because an understanding of the basic statistical concepts is essential to an intelligent and flexible use of statistical techniques. The use of techniques without understanding of concepts can be dangerous. If the introductory statistics course is largely devoted to concepts, students can obtain additional knowledge of statistical techniques in two basic ways. They can take a second-level statistics course which focuses on a variety of statistical techniques or a course in a subject in which specialized statistical techniques are introduced in a context in which they are used.

2. Implications of the Recommended Objectives

A number of implications follow from the declared objectives.

a. Since the main objective of the course is understanding of basic statistical concepts, it follows that proofs and extensive manipulations of formulas should be employed sparingly. While statistics utilizes these, its major focus is on inferences from data.

b. The course should not dwell on computational techniques. Rather, the amount of computation and whether it is done on a high-speed computer, a desk calculator, or by hand should be determined by the extent to which it helps the students to understand the principles involved.

c. While probabilistic concepts are essential for an understanding of statistical inference, probability theory should not constitute a dominant portion of the course.

d. In order to illustrate the application of statistical methodology in making inferences, the course must be data-oriented and must incorporate analysis of real-world data.

Three critical dimensions of the introductory statistics course are inferential philosophy, mathematics, and data analysis. Among these components we believe that the emphasis should be strongly on inferential concepts and data analysis, and less on mathematical elements.

3. Recommended Topics to be Covered

As we have stated earlier, no single introductory statistics course will be suitable for all situations. Within the framework of the objectives for the introductory statistics course just recommended, a wide variety of courses can be designed. In Sections III and IV we consider several different courses which could meet the stated objectives: (1) a more or less standard statistics course, (2) a decision-theory oriented statistics course, (3) a statistics course embodying the nonparametric approach, and (4) a statistics course utilizing a problem-oriented approach.

Despite the variety of possible approaches, most courses will include the elements of probability, the binomial and normal distributions, the distinction between sample and population, descriptive statistics (such as mean, median, variance, standard deviation, and frequency distribution), and statistical inference. The study of inference will treat hypothesis testing, point estimation, and confidence interval estimation and will include the use of the t-statistic.

We present two lists: one includes important topics from which a limited selection should be made since no single course can reasonably be expected to cover a large fraction of these topics; second is a list of topics which should be avoided unless they are used as fundamental pedagogical tools or are logically essential (such as Bayes' theorem for a Bayesian approach).

a. Important topics from which a selection should be made

1. Probability: sample space, mutually exclusive events, independent events, conditional probability, random variable (expected value, variance, standard deviation)

2. Samples: frequency distribution, histogram, ogive, percentiles, mean, median, variance, standard deviation, mode, range

3. Distributions: normal, binomial, Poisson, exponential, rectangular, geometric

4. Sampling theory: Law of Large Numbers, Central Limit Theorem

5. Estimation: point estimates, confidence intervals

6. Hypothesis testing: alternative and null hypotheses, power function; errors of types I and II, significance levels, one-tail and two-tail tests

7. One-sample tests: for the mean of a normal distribution (t-test), for the proportion of a binomial distribution

8. Two-sample tests: t-test, Mann-Whitney test, sign test

9. Chi-square and contingency tables

10. Regression and correlation

11. Analysis of variance

12. Decision theory: minimax and Bayes' strategies, admissibility

b. Topics to be avoided (except when they serve as useful teaching devices)

1. Combinatorics

2. Bayes' theorem

3. Partial regression

4. Sequential analysis

5. Maximum likelihood and likelihood ratio

6. Compendia of variations of a given procedure like the t-test

7. Inference on variances using chi-square distributions (because they involve nonrobust procedures)

III. A CONVENTIONAL COURSE IN INTRODUCTORY STATISTICS

An approach to the teaching of elementary statistics whose popularity is reflected in the most widely used textbooks may be called the conventional approach. It is characterized in part by a logical development in which basic tools are developed slowly and in some detail before serious statistical problems are attacked.

Because of the popularity of the conventional approach and the existence of a wide variety of texts oriented toward this approach, it is the one most likely to be used in two-year and four-year colleges and in smaller universities in the near future. We therefore believe it is wise to focus primarily on this kind of course, particularly since significant modifications and enrichments can make it a worthwhile intellectual experience.

This traditional approach has potential defects, as noted earlier. Typically, the major ideas of statistical inference are introduced too late and in haste, and they are often illustrated by examples which are not compelling. Frequently, courses of this kind suffer from the inclusion of irrelevant concepts and excessive mathematical derivations. Finally, the attempt to break up the subject into small digestible bits may cause the student to miss the unifying concepts of statistical inference. The course outline and comments which follow attempt to provide guidance on topics and suggestions for avoiding these defects.

1. Course Outline

A recommended conventional course outline, which permits early treatment of the ideas of statistical inference and which stresses concepts rather than a proliferation of statistical techniques is given in this section. The suggested pace has been indicated by assigning a number of hours to each group of topics. A standard semester contains 42 to 48 class meetings, and we arbitrarily allowed 36 hours of class time for the presentation of new material; the slack time that we have left provides for tests, review, etc. More detailed suggestions of what to mention only briefly and what to emphasize are included in the Comments, Section 2.

CONVENTIONAL COURSE OUTLINE

	Topics	Lectures
0.	Introduction	1
1.	Statistical Description Frequency distributions, cumulative frequency distributions; measures of location and variation	3
2.	Probability Concept; sample space; addition theorem; marginal probability; conditional probability; multiplication theorem; independence	3
3.	Random Variables and Probability Distributions Concepts; simple discrete univariate probability distributions; expectation and variance of discrete random variables; functions of discrete random variables; mean and variance of functions of discrete random variables	2
4.	Special Probability Distributions Binomial probability distribution; continuous probability distributions; normal probability distribution	2

5. Sampling Distributions — 3

Random sampling; mean and variance of sum of independent random variables; sampling distribution of mean; Central Limit Theorem; sampling distribution of proportion

6. Estimation of Population Proportion — 4

Point estimation of population proportion; confidence interval for population proportion based on large samples

7. Tests Concerning Population Proportion — 5

Formulating hypotheses; statistical decision rules; types of errors; power of a test; construction of one-sided and two-sided tests (small and large samples)

8. Inferences Concerning Population Mean — 4

Point estimation of population mean; properties of estimators: unbiasedness, consistency, efficiency; confidence intervals and one- and two-sided tests of hypotheses for the mean of a population whose variance is also unknown, based on small and large samples

9. Additional Topics — 4-9

Selection from the following:

(a) Inferences concerning differences of two population means and proportions
(b) Inferences concerning population variance
(c) Chi-square and contingency tables
(d) Regression and correlation
(e) Analysis of variance
(f) Nonparametric methods
(g) Survey sampling
(h) Quality control
(i) Bayesian methods
(j) Decision theory

The number of lectures assigned to topics 1-8 is a minimal number. Use of computers, case discussions, and student-generated data will require additional time. Since these activities will vary from course to course, we have indicated a number of lectures for each topic without considering such activities. Consequently, the number of lectures available for additional topics usually will be less than the maximum of nine. We do recommend, however, that at least four lectures be devoted to one or more of the additional topics.

2. Comments on the Course Outline

General. Since the above outline contains only a list of topics, it does not recognize four important ingredients that we believe will enhance the conventional approach.

a. Continuing motivation of the student through examples and problems that he finds interesting and important. Data sets and student-generated data are discussed in Sections V and VII.

b. Use of computers to help develop basic concepts such as the Central Limit Theorem as well as to remove the drudgery of statistical calculations. Use of computers is discussed in Section VI.

c. Emphasis on the basic ideas underlying statistical inference and a demonstration of how these recur from one application to another. This requires skillful teaching and demands that the teacher see the broad picture of statistical inference.

d. Student use of programmed learning materials for developing mastery of important but routine topics. We believe it is wasteful to use precious class time having students practice constructing frequency distributions, calculating measures of location and variation, using binomial and normal tables, etc. Class time is better used for explanation and discussion, not for routine calculations and drill. Examples of self-help programmed materials are:

[1] Elzey, Freeman F. A Programmed Introduction to Statistics. Belmont, California, Brooks/Cole Publishing Company, 1966.

[2] Gotkin, Lassar G. and Goldstein, Leo S. Descriptive Statistics: A Programmed Textbook, vols. 1 and 2. New York, John Wiley and Sons, Inc., 1965.

[3] McCollough, Celeste and Van Atta, Loche. Statistical Concepts: A Program for Self-Instruction. New York, McGraw-Hill Book Company, 1963.

[4] Whitmore, G. A., et al. Self-Correcting Problems in Statistics. Boston, Massachusetts, Allyn and Bacon, Inc., 1970.

The suggested time allocation in the above outline makes it clear that, in our view, the early material should be covered quickly so that statistical inference can be reached in the first half of the course.

A nonexhaustive, illustrative list of texts fitting the pattern of the conventional course follows.

Illustrative List of Texts for an Introductory Course:

[5] Alder, Henry L. and Roessler, Edward B. Introduction to

Probability and Statistics, 4th ed. San Francisco, California, W. H. Freeman and Company, 1968.

Reviewed in Journal of the American Statistical Association, 64 (1969), p. 675. The first edition is reviewed in The American Mathematical Monthly, 68 (1961), p. 1018.

[6] Freund, John E. Modern Elementary Statistics. Englewood Cliffs, New Jersey, Prentice-Hall, Inc., 1967.

Reviewed in Journal of the American Statistical Association, 62 (1967), p. 1504.

[7] Guenther, William C. Concepts of Statistical Inference. New York, McGraw-Hill Book Company, 1965.

Reviewed in Journal of the American Statistical Association, 61 (1966), p. 529.

[8] Hoel, Paul G. Elementary Statistics, 3rd ed. New York, John Wiley and Sons, Inc., 1971.

Third edition is reviewed in Journal of the American Statistical Association, 67 (1972), p. 497.

[9] Huntsberger, David V. Elements of Statistical Inference, 2nd ed. Boston, Massachusetts, Allyn and Bacon, Inc., 1967.

[10] Mendenhall, William. Introduction to Statistics, 2nd ed. Belmont, California, Wadsworth Publishing Company, Inc., 1967.

Second edition is reviewed in School Science and Mathematics, 67 (1967), p. 754.

[11] Walker, H. M. and Lev, J. Elementary Statistical Methods, 3rd ed. New York, Holt, Rinehart and Winston, Inc., 1969.

First edition is reviewed in Journal of the American Statistical Association, 54 (1959), p. 699.

Following is a list of texts which may be helpful to the instructor as resource materials.

Illustrative List of References for an Introductory Course:

[12] Blackwell, David. Basic Statistics. New York, McGraw-Hill Book Company, 1969.

Reviewed in Journal of the American Statistical Association, 65 (1970), p. 1398, and in The American Mathematical Monthly, 77 (1970), p. 662.

[13] Dixon, Wilfrid J. and Massey, F. J. _Introduction to Statistical Analysis_, 3rd ed. New York, McGraw-Hill Book Company, 1969.

Reviewed in _Journal of the American Statistical Association_, 65 (1970), p. 456. The second edition is reviewed in _The American Mathematical Monthly_, 64 (1957), pp. 685-686.

[14] Hodges, J. L., Jr. and Lehmann, E. L. _Basic Concepts of Probability and Statistics_, 2nd ed. San Francisco, California, Holden-Day, Inc., 1970.

Reviewed in _Journal of the American Statistical Association_, 65 (1970), p. 1680. The first edition is reviewed in _The American Mathematical Monthly_, 72 (1965), p. 1050.

[15] Natrella, Mary G. _Experimental Statistics_, Handbook 91. U. S. Department of Commerce, National Bureau of Standards, 1966.

[16] Neyman, J. _First Course in Probability and Statistics_. New York, Henry Holt and Company, 1950.

Reviewed in _Journal of the American Statistical Association_, 46 (1951), p. 386.

[17] Savage, I. Richard. _Statistics: Uncertainty and Behavior_. Boston, Massachusetts, Houghton Mifflin Company, 1968.

Reviewed in _Journal of the American Statistical Association_, 64 (1969), p. 1677.

[18] Snedecor, George W. and Cochran, W. G. _Statistical Methods_, 6th ed. Ames, Iowa, Iowa State University Press, 1967.

Reviewed in _Applied Statistics_, 17 (1968), p. 294.

[19] Wallis, W. Allen and Roberts, Harry V. _Statistics: A New Approach_. New York, The Macmillan Company, 1956.

Reviewed in _Journal of the American Statistical Association_, 51 (1956), p. 664.

We turn now to topic-by-topic comments on the outline. These comments contain occasional references to books in the above lists.

Topic 0. Introduction

This lecture should be devoted to a discussion of the nature and importance of statistics, including the difference between the inferential nature of statistics and the deductive nature of mathematics. The first lecture should illustrate the existence of statistical problems in everyday life in order to emphasize that statistics is problem-oriented.

Source material which may be helpful includes:

[20] Careers in Statistics. The Committee of Presidents of the American Statistical Association, the Institute of Mathematical Statistics, and the Biometric Society. The American Statistical Association, 806 15th Street, N. W., Washington, D. C. 20005.

[21] Kruskal, William, ed. Mathematical Sciences and Social Sciences. Englewood Cliffs, New Jersey, Prentice-Hall, Inc., 1970.

[22] The Mathematical Sciences: A Collection of Essays. Edited by The National Research Council's Committee on the Support of Research in the Mathematical Sciences (COSRIMS) with the collaboration of George A. W. Boehm. Boston, Massachusetts, MIT Press, 1969.

[23] Sills, David L., ed. International Encyclopedia of the Social Sciences. New York, The Macmillan Company and the Free Press, 1968. Listings under:
Statistics
Statistical analysis
Survey analysis

[24] Tanur, Judith, et al., eds. Statistics: A Guide to the Unknown. San Francisco, California, Holden-Day, Inc., 1972.

[25] Wallis, W. Allen and Roberts, Harry V. Statistics: A New Approach. New York, The Macmillan Company, 1956.

Topic 1. Statistical Description

General comments. The discussion should begin with a problem that will motivate the need for statistical data and their analysis through frequency distributions and descriptive measures. A good problem may be used to give an overview of the course raising questions relating to later topics. Instructors will be able to find problems for which data of local interest from registrars' records are appropriate. Examples: Are student entrance test scores higher today than they were five years ago? Is student family economic status higher today than it was five years ago? Are college-age students today taller than they were 30 years ago? Do college aptitude test scores for men differ from those for women?

Our suggestion departs from the usual approach in which data are provided without the context of a problem, and a frequency distribution or a variety of descriptive measures are calculated for their own sake. The suggested approach emphasizes problems of inference from the start; the instructor should not hesitate to raise questions about inferential problems that will be considered later in the course.

A comparison of a frequency distribution with an expected, theoretical distribution, such as the birth pattern distribution in R. A. Fisher's Statistical Methods for Research Workers* (p. 67), is useful for introducing ideas of statistical inference. So is a bimodal distribution, the problem being whether it is made up of two different groups, each with a different pattern of variation.

Specific suggestions.

a. Class-generated data, from which one or several frequency distributions are constructed, often interest students.

b. In discussing frequency distributions, one should emphasize the variation inherent in the phenomenon under study and the pattern of this variation. Examples of different phenomena (e.g., life of light bulbs, number of calculators out of order per day) should be utilized to demonstrate widespread presence of variation.

c. Data for constructing frequency distributions should be simple, so that little time is spent on developing class limits.

d. The discussion of frequency polygon and histogram can be used to distinguish between continuous and discrete variables.

e. In taking up the descriptive measures of location (mean, median) and of variation (standard deviation, range), alternate computational forms of the mean and standard deviation should be called to the student's attention, but no derivations should be given. A useful means of pointing out the applicability and limitations of these measures is through a comparison of two or more frequency distributions (e.g., income distribution for 1970 and 1971).

f. Class time is better devoted to discussion of the meaning and limitations of various measures of location and variation than to drilling students on their calculation.

g. Mean and median should be explained as alternate descriptive measures of location, neither of which is perfect for all situations.

Topic 2. Probability

General comments. The way in which introductory probability is presented can vary greatly. It is a broad topic, intensive examination of which leads to deep philosophical problems. The approach of

* [26] Fisher, R. A. Statistical Methods for Research Workers, 13th ed. New York, Hafner Publishing Company, 1958.

Blackwell [12] in the first three chapters is elementary. The concept of area is used when extended ideas are required. The basic laws of probability are developed, making effective use of tables and diagrams. A major advantage of the presentation is that it stays consistently at a modest mathematical level without disruptive digressions.

A more complete exposition of finite probability theory is presented by Hodges and Lehmann in [14]. Chapters 1 and 4 contain material appropriate for an introduction to probability in the elementary statistics course. We suggest this reference as additional reading to augment the material in Blackwell.

An interesting aspect of the modern emphasis in probability is that of assigning personal or subjective probabilities to events. For an elementary introduction, Savage [17] seems appropriate. The discussion includes a great deal of gambling terminology which may not appeal to all students. At the end of the chapter there are many stimulating notes and problems that are used to extend the theory and to show examples of probabilistic reasoning occurring in everyday life. Hodges and Lehmann [14] have a short section on subjective probability beginning on page 127.

Specific suggestions.

a. Since only three lectures are devoted to probability, discussion must be restricted to the major concepts, and proofs will need to be kept to a minimum. Only finite sample spaces should be used for explaining the probability concepts.

b. The relative frequency interpretation is a useful way to motivate the concept of probability. The equally-likely notion should not be given major emphasis, though it may be helpful if extended to real-life situations such as taste experiments or genetic problems.

c. Whether set notation is to be used depends on the mathematical training of the students. At any rate, the allotted time in the suggested outline does not permit the teaching of set theory or even of basic set concepts. The teacher may, however, wish to point out that probability theory is based on set theory.

d. In teaching statistical independence, the conditional definition:

$$\text{If } P(A|B) = P(A), \text{ then } A \text{ and } B \text{ are independent,}$$

is easier for the student to understand than the product definition:

$$\text{If } P(AB) = P(A)P(B), \text{ then } A \text{ and } B \text{ are independent.}$$

e. The small amount of time allotted to probability makes it imperative that teachers not get bogged down in combinatorial probability problems. Simple, though realistic, examples should be used to illustrate the probability concepts covered.

Topic 3. Random Variables and Probability Distributions

General comments. Random variables, probability distributions, and expectations are treated in fairly standard ways in most texts. A simple presentation is given by Blackwell [12], but his presentation is very brief.

Specific suggestions.

a. Do not give a formal definition of random variables; a statement that "the value of a random variable is a number determined by the outcome of an experiment that varies from trial to trial" should be adequate.

b. Introduce probability functions and cumulative distribution functions in terms of simple examples with a small number of points and associated probabilities. Functions of random variables also should be considered only in terms of such simple examples.

c. Stay on the intuitive level. Many definitions and results are intuitively reasonable for most students; for instance:

$$E(X) = \Sigma x_i P(X = x_i)$$

$$E(kX) = kE(X)$$

$$\sigma(kX) = k\sigma(X), \quad k > 0$$

d. Use simple everyday examples, with three or four outcomes only; e.g., the number of calculators needing repair in a group of four calculators; the number of males in families of five persons.

e. The mean and variance of a linear function of a random variable should be mentioned and their usefulness illustrated by an example, but formulas should not be derived.

Topic 4. Special Probability Distributions

General comments. Student participation can be developed both through access to a computer on an individual basis and through elementary class exercises involving coin tossing, tables of random normal deviates, and exercises of this kind. The results of student exercises can be combined to demonstrate sampling variation over the various samples obtained individually by students, thus serving as an introduction to Topic 5, Sampling Distributions.

Specific suggestions.

a. State the conditions for a probability distribution to be binomial but do not derive its formula.

b. Formulas for the mean and variance of a binomial distribution should be given and illustrated by a simple example but should not be derived.

c. Tables of binomial probabilities should be used to demonstrate characteristics of binomial distributions such as skewness. Students should not be asked to calculate binomial probabilities repeatedly.

d. Use a variety of realistic examples. Do not confine examples to coin tossing; e.g., use the number of defectives in a sample from a shipment of parts or the number of voters favoring an issue in an opinion poll.

e. The transition to a continuous random variable, for which area represents probability, can be facilitated by considering a histogram and shrinking the width of the classes.

f. The normal distribution can be introduced as:

(i) an approximate description of many real-life phenomena, such as weights of the contents of cans of fruit or the Scholastic Aptitude Test scores of students;

(ii) an approximation to certain discrete distributions, such as the binomial distribution or the distribution of the sum of the digits on three dice.

The text used will influence the approach taken, but both notions should be introduced.

g. In discussing the normal distribution, one should emphasize the variety of shapes encountered for different values of μ and σ and the method for transforming any normal random variable to a standard normal one.

h. Realistic examples should be used for the normal distribution. Students will probably be interested in SAT scores, for which the mean is 500 and the standard deviation is 100.

i. Unless the text makes the continuity correction an integral part of the normal approximation to the binomial distribution, this topic should not be taken up or should be mentioned only briefly.

Topic 5. Sampling Distributions

General comments. Computers are effective in demonstrating sampling variability and sampling distributions. Exercises can be developed for many distributions where the computer takes repeated samples of fixed size and, for example, provides information on the distribution of the sample mean through plots and histograms. The student learns of the nature of sampling variation and can proceed through variation in sample sizes to obtain insight into the inherently increased stability of sampling distributions as sample sizes are increased. The student can be given insight thereby into the behavior of the variance of a mean and into the concepts of the Central Limit Theorem.

Specific suggestions.

a. The concept of the sampling distribution of the mean can be explained initially by considering a small finite population and sampling with or without replacement with a small sample size ($n = 2$ or 3). This permits the exact sampling distribution of the mean to be developed by enumeration. The mean and variance of this sampling distribution can then be studied.

b. Since the Central Limit Theorem should only be stated and not proved, evidence of its operation will need to be given to the student. Some texts contain exact sampling distributions for different sample sizes or the results of sampling experiments. Printouts of computer runs simulating sampling distributions for different sample sizes can also be distributed to students and discussed. If a computer is not available on campus, printouts could be obtained from a computer located elsewhere.

These approaches are helpful but, in our opinion, are not as effective as having students participate in sampling experiments. A simple experiment is to sample a rectangular distribution, either from a table of random numbers, by drawing chips from a bowl, or by computer. If a computer is used, it will also be easy to sample other kinds of populations. Sampling a moderately skew population may help convince students of the Central Limit Theorem in the absence of symmetry. Indeed, the use of several populations (e.g., rectangular, exponential) can demonstrate to the student that the rapidity with which the sampling distribution of $(\bar{x} - \mu) \div (\sigma/\sqrt{n})$ approaches a normal distribution as n increases depends on the population from which the samples are selected.

c. It should be pointed out to students that the sample proportion is a mean, and hence the Central Limit Theorem applies directly to sample proportions.

Topic 6. Estimation of Population Proportion

General comments. If the teacher prefers, he can take up estimation of the means of continuous populations and tests concerning them before he discusses inferences on population proportions. Inferences on proportions are, in our opinion, somewhat simpler to present; e.g., binomial tables can be used to derive the power of tests. Hence, we place inferences on proportions before inferences on means of continuous populations.

Also, testing hypotheses can be taken up before estimation. We recommend taking up statistical estimation first because we believe it is easier for the student to understand than hypothesis testing.

The most common interpretation of a confidence interval has been that of a random interval which contains a fixed parameter with given probability over repeated applications. This point of view has been presented traditionally in most texts. Through use of subjective probability and Bayes' theorem, one can interpret a confidence interval as a fixed interval (for given observed data) which contains a random parameter with given posterior probability. Savage [17] puts forward both concepts (see page 209 and page 260) and points out basic differences in philosophy.

Specific suggestions.

a. The large sample confidence interval for the population proportion p can be developed readily by relying on an extension of the Central Limit Theorem which states that $(\hat{p} - p) \div \sqrt{\hat{p}(1 - \hat{p})/n}$ is approximately normal for large sample size n, $\hat{p}$ being the observed proportion.

b. An experiment with repeated sampling of a known population, setting up a confidence interval each time, can be helpful in conveying the idea of the confidence coefficient and in illustrating that the location and width of the confidence interval varies from sample to sample. This can be done either by computer or by class-generated samples. Such an experiment is particularly desirable since it will provide evidence of the working of the extension of the Central Limit Theorem.

c. Realistic examples, such as estimation of the proportion of voters favoring a candidate or the proportion of persons in the labor force who are unemployed, will be helpful in illustrating the pervasiveness of estimation problems. The subject of sample surveys in general may be discussed in connection with statistical estimation.

d. The discussion of confidence intervals should consider the usefulness of the particular confidence interval obtained. For instance, in a close election race involving two candidates, a confidence interval for the proportion of voters favoring one of the candidates which ranges from .45 to .53 may not be useful. Discussion of several such examples can help the student to recognize that different problems may call for different levels of precision.

e. The determination of the confidence coefficient should be discussed in general terms. Several examples may be used to illustrate that important problems, such as estimating the unemployment rate for purposes of determining national economic policy, call for higher confidence coefficients than do less important problems.

f. The determination of the sample size required to yield a confidence interval of sufficiently small width for a given confidence coefficient provides a useful means of introducing the student to the notion that statistical investigations should be planned in advance.

g. In discussing interval estimation of the population proportion for small sample sizes, reference should be made to tables or to the Clopper-Pearson charts which appear in many texts and books of tables. See, for example,

[27] Owen, Donald B. Handbook of Statistical Tables. Reading, Massachusetts, Addison-Wesley Publishing Company, Inc., 1962.

Topic 7. Tests Concerning Population Proportion

General comments. This subject can be discussed either by considering small sample sizes first and then large sample sizes, or vice versa. An advantage of beginning with small sample sizes is that power calculations can readily be made by using widely available binomial tables. A disadvantage is that the power obtained for small sample sizes will be relatively poor so that the student may feel that the statistical test is almost useless. If small sample sizes are taken up first, it is incumbent upon the instructor to note early that many practical problems require much larger sample sizes.

An advantage of beginning with large sample sizes is that many problems can be treated much more realistically. A disadvantage is that power calculations will be more tedious.

Neyman [16] gives a complete, systematic analysis of several stages involved in testing hypotheses (pp. 268-271).

Specific suggestions.

a. A basic difficulty for most students is the proper formulation of the alternatives H_0 and H_1 for any given problem and the consequent determination of the proper critical region (upper tail, lower tail, two-sided).

b. Practice with problems containing a verbal statement of the situation and requiring the students to develop H_0 and H_1 can be most helpful. Indeed, students should be competent in formulating the alternatives H_0 and H_1 and in designating the appropriate type of critical region before they begin the detailed calculations for determining the exact critical region.

c. Realistic problems should be used when the student is asked to formulate the appropriate alternatives H_0 and H_1.

d. Students can often be helped by introducing the critical region in terms of the statistic $\hat{p}$, the sample proportion. For instance, if:

$$H_0 : p \leqq p_o$$

$$H_1 : p > p_o,$$

the student can readily see that the appropriate decision rule is of the form:

$$\text{if } \hat{p} \leqq A, \text{ accept } H_0,$$

$$\text{if } \hat{p} > A, \text{ accept } H_1,$$

where A is the cut-off point to be determined. It can then be explained how to determine A if, for example, the probability of concluding H_1 when $p = p_o$ is to be α.

The standardized statistic $(\hat{p} - p_o) \div \sqrt{p_o(1 - p_o)/n}$ for determining the critical region is a confusing way to introduce the student to testing hypotheses. After the student gains some familiarity with testing hypotheses, the standardized statistic then is more easily used.

e. In developing the properties of a statistical decision rule for testing hypotheses, attention should not only be given to the power curve of the decision rule but also to the error curve which shows the probabilities of error directly. Many students understand the implications of a decision rule for testing more easily through the error curve than through the power curve [e.g., if $H_0 : p \leqq p_o$, $H_1 : p > p_o$, then $P(\text{Error}) = P(\text{Accept } H_1 \mid p)$ for $p \leqq p_o$ and $P(\text{Error}) = P(\text{Accept } H_0 \mid p)$ for $p > p_o$)].

f. When discussing the two-sided test, one should explain the correspondence between the confidence interval for the population proportion and the testing approach.

g. When small sample sizes are considered first, the power of a test and the error probability for a given population proportion should be obtained using binomial tables rather than by making actual calculations. Binomial tables can then be used to set up a decision rule with specified control on type I and/or type II errors. At this point a transition can be made to the case of large samples and the use of the normal approximation.

h. Students often do not understand how the p_o level is determined in an actual problem. This situation can be clarified by illustrating a variety of situations, for instance:

(i) In deciding whether a company should accept or reject an incoming shipment, p_o may be the break-even proportion of defective items for which it is equally costly to accept or reject a shipment.

(ii) In an experiment to determine whether a person has extrasensory perception by having him indicate whether coins flipped in a different city are heads or tails, p_o may be .5, the level expected if the person were guessing and no ESP were present.

These examples illustrate two commonly encountered situations, namely those for which p_o is the break-even proportion and those for which p_o is determined by theoretical considerations as the level where no effects, or no effects of practical significance, are present.

i. Discussion of the determination of the probability of a type I error (size of test) α may usefully be combined with the problem of how to determine sample size, as follows:

(i) For given n and α, find the probability of type II error β at an appropriate value of p. If β is satisfactory, use the given n and α. If β is too high, increase the sample size. If no increase in sample size is possible, raise α until a suitable balance between α and β is found.

(ii) If, for given n and α, β turns out to be smaller than necessary, the sample size may be reduced or α lowered.

j. Sampling experiments to demonstrate the behavior of a given decision rule for different levels of p may be a helpful supplement for many students.

Topic 8. Inferences Concerning Population Mean

Specific suggestions.

a. Some major properties of estimators (unbiasedness, consistency, and efficiency) could be discussed here, but only briefly. If they are discussed, we recommend that it be done here rather than in Topic 6. We feel that students should first become acquainted with the general concept of statistical estimation before they become concerned with properties of point estimators.

b. In discussing the small-sample confidence interval based on the t-distribution, emphasis should be placed on the robustness (insensitivity of properties of the procedure to departures from the assumption of normality) of the t-statistic and the consequent wider applicability of this confidence interval to populations which are not exactly normal.

c. In developing inferences concerning population means, students should not be asked to perform extensive calculations to find the sample mean and standard deviation.

d. Sampling experiments for estimation and testing with small sample sizes can be helpful in several ways:

(i) When sampling a normal population, the experiment will give the student further confidence in the t-distribution and will illustrate once again the meaning of the confidence coefficient or the level of significance and power.

(ii) When sampling a population moderately different from a normal population (e.g., a rectangular population), the experiment will illustrate the robustness of the t-statistic.

e. We recommend bypassing the case of known σ unless inferences concerning the population mean are considered before inferences concerning the population proportion. The reason for this recommendation is that the case of known σ does not arise often in practice and only serves to introduce an unnecessary repetitive element for the student. Reliance should be placed on $(x - \mu) \div (s/\sqrt{n})$ being approximately distributed as t for small n for populations not departing excessively from a normal population and being approximately distributed as a standard normal variable for larger n for almost all populations.

If there is time available to discuss planning the sample size or to determine the power of the test, we recommend the use of charts such as those in Guenther [7]. For the case of unknown σ these charts could be entered by using two values of σ within which the standard deviation is expected to fall, thereby obtaining bounds on the power or sample size.

Topic 9. Additional Topics

Students in the social sciences might benefit most from topics a, c, d, f (see page 481); in the physical and biological sciences, from topics a, c, d, e; in management science, from topics g and h; in economics, from topics d and h; in education, from topics a, d, f.

IV. SOME ALTERNATE APPROACHES

In this section we discuss three alternate approaches to the introductory statistics course. Each approach provides great potential for innovation, and some instructors may wish to try one or several of these in lieu of the conventional approach. For two of the alternate approaches (decision theory, nonparametric), a number of possible textbooks are available. Since the textbook used will have a major effect in the determination of the precise contents, the order of presentation, and the amount of time to be devoted to the various topics, no effort is made here to present outlines. Instead, we discuss the principles underlying the use of the approaches. This will serve to indicate potential advantages and disadvantages and to provide a foundation in terms of which the instructor may interpret various texts that are now and will later become available.

No text designed specifically for the third approach (problem-oriented) is available, to the best of our knowledge. The instructor wishing to implement this approach must be prepared for substantial developmental efforts.

1. A Decision Theory Course in Introductory Statistics

Basic elements of the course. Decision theory is a formulation of statistical problems in which the statistician has available a choice of actions, the consequences of which depend on an unknown state of nature. To help decide on an appropriate action, one must perform an experiment which will yield relevant data to help determine the state of nature. Since the data generally depend not only on the state of nature but also on chance, uncertainty appears in two places--the effect of chance, i.e., random variation, and the initial ignorance of the state of nature.

A decision theory course in introductory statistics should include the following concepts.

> Concepts: Philosophical principles of decision-making under uncertainty; averages and measures of variability; probability and expectation; utility; Bayes' strategies and posterior probabilities; the parameters of distributions relevant for optimal actions; testing hypotheses, significance levels; estimation, confidence intervals.

The development of these ideas requires the presentation and use of (1) elementary properties of probability and mathematical expectation, (2) sets and functions on a relatively simple level, and (3) properties of convex sets such as the separating hyperplane theorem.

Traditional statistical techniques of handling data such as histograms, cumulative frequency polygons, means and standard deviations are not vital to the decision theory approach. Their introduction has considerable potential pedagogic value, however, in preparing the student for the analogous probabilistic concepts of density, cumulative distribution function, and expectation. The standard statistical methods in estimation, hypothesis testing, and confidence intervals are not easily blended with this approach.

Results of decision theory have implications in business and other real-life situations where decisions must be made in the face of uncertainty, and a wise choice must consider consequences of the available actions.

Advantages and disadvantages of the decision theory approach. One major advantage of the decision theory approach over traditional approaches lies in the problem-solving orientation. Each example considered involves a problem in decision-making which requires a good and sensible answer.

The decision theory approach should appeal to students who are interested in the philosophical foundations of scientific inference and who are curious about the rationale for coping with random variation and uncertainty. Such students would find the decision theory approach natural and easy to follow. It would be esthetically pleasing for the student who likes mathematics and enjoys the application of basic ideas such as sets, functions, and convexity to clarify nontrivial results in inference. (It is possible to present these results without, or with a minimum of, formal derivations.) For the student who anticipates further serious study of statistics, this course can serve as a valuable complement to his other work in statistics.

This course will not serve for a student who is expected to learn some of the major tools of statistics in the first course. Although the course can present an exciting view of statistics and scientific inference, the decision theory point of view gives a limited view of statistics and does little to acquaint the student with actual statistical practice; it does not give him experience with data analysis.

Another limitation of the decision theory approach is that it tends to involve an overformulation of statistical problems in the sense that the statistician is presumed to know precisely the set of available actions, the set of possible states of nature, the consequences of the actions, and all the relevant probability distributions. Little allowance is made for the possibility of error in specifying the model. No provision is made for learning from data.

While a minimum of mathematical background is required, the student who is very weak in mathematics will fail to enjoy some of the esthetic values of this approach and may find that coping with the elementary arithmetic and algebra distracts him from the main ideas.

Finally, the opportunities to apply the ideas and methods of such a course in other academic work are relatively few, although such opportunities seem to be increasing steadily as decision-making becomes used more and more in diverse fields such as business, engineering, and medicine.

Illustrative List of Texts and References:

Aitchison, John. Choice Against Chance: An Introduction to Statistical Decision Theory. Reading, Massachusetts, Addison-Wesley Publishing Company, Inc., 1970.

Reviewed in The Australian Journal of Statistics, 13 (1971), p. 123.

Chernoff, Herman and Moses, Lincoln E. Elementary Decision Theory. New York, John Wiley and Sons, Inc., 1959.

Reviewed in Journal of the American Statistical Association, 55 (1960), p. 291, and in The American Mathematical Monthly, 67 (1960), p. 487.

Hadley, G. Introduction to Probability and Statistical Decision Theory. San Francisco, California, Holden-Day, Inc., 1967.

Reviewed in Journal of the Royal Statistical Society, Series A, 131 (1968), p. 437.

Lindgren, B. W. Elements of Decision Theory. New York, The Macmillan Company, 1971.

Lindley, D. V. Making Decisions. New York, John Wiley and Sons, Inc., 1971.

Raiffa, Howard. Decision Analysis: Introductory Lectures on Choices Under Uncertainty. Reading, Massachusetts, Addison-Wesley Publishing Company, Inc., 1968.

Reviewed in Journal of the American Statistical Association, 64 (1969), p. 1668.

Schlaifer, R. Probability and Statistics for Business Decisions: An Introduction to Managerial Economics Under Uncertainty. New York, McGraw-Hill Book Company, 1959.

Reviewed in Journal of the American Statistical Association, 54 (1959), p. 813.

2. A Nonparametric Course in Introductory Statistics

Basic elements of the course. Using nonparametric methods it is possible to introduce substantial problems in inference very early in the course, with impressive solutions which can be easily comprehended. The preparation required before beginning standard nonparametric methods consists mainly of some work in rather elementary discrete probability. No time need be devoted to the usual development of histograms and the computation of means and standard deviations. With the early introduction of meaningful problems, the discussion of sampling distributions is interwoven in the development of statistical inference in a way which leads to ready understanding by the student. It is possible to avoid a digression into combinatorics by presenting simple illustrations and relying on appropriate tables.

A nonparametric course in introductory statistics should include the following concepts.

> Concepts: Probability, probability distributions, hypothesis testing, estimation, two-sample tests, chi-square tests and contingency tables, correlation, robustness.

While it is most desirable that the introductory course using a nonparametric approach concentrate on nonparametric methods, it is possible toward the end of the course to introduce the student to some parametric statistics. If the instructor wishes to cover some parametric topics, only a few should be taken up. To be in line with our recommended objectives, the course should not become technique-oriented; one should avoid making the course a compendium of nonparametric and parametric methods.

Advantages and disadvantages of the nonparametric approach. The use of nonparametric methods provides several potential advantages. With the availability of appropriate tables, the detailed working of numerical examples, regarded by many as essential to a thorough grasp of principles as well as techniques, is simple and brief. The methods are easily interpreted and have natural and sensible justifications. The methods have the advantage of robustness. Thus the student is quickly introduced to useful and simple techniques which have wide applicability. This exposure to statistical concepts throughout the course provides the student with a foundation for a good understanding of standard parametric procedures.

For the student who is required to take statistics to satisfy the demands of a major field in which statistics is used, this course is likely to provide an introduction to some useful nonparametric methods at an elementary level.

A major disadvantage of this approach is that students will have little or no exposure to those parametric methods needed for later work in other subjects. Another disadvantage is that current books in nonparametric statistics do not focus on either point or interval estimation. Also, there are few texts at the elementary level.

Illustrative List of Texts for an Introductory Course:

Conover, W. J. Practical Nonparametric Statistics. New York, John Wiley and Sons, Inc., 1971.

Reviewed in Journal of the American Statistical Association, 67 (1972), p. 246.

Kraft, Charles H. and van Eeden, Constance. A Nonparametric Introduction to Statistics. New York, The Macmillan Company, 1968.

Reviewed in Journal of the American Statistical Association, 66 (1971), p. 223, and in The American Mathematical Monthly, 77 (1970), p. 207.

Noether, Gottfried E. Introduction to Statistics: A Fresh Approach. Boston, Massachusetts, Houghton Mifflin Company, 1971.

Reviewed in Journal of the American Statistical Association, 67 (1972), pp. 496-497, and in The Mathematics Teacher, 64 (1971), p. 630.

Other Selected References in Nonparametric Statistics:

Bradley, James V. Distribution-Free Statistical Tests. Englewood Cliffs, New Jersey, Prentice-Hall, Inc., 1968.

Reviewed in Journal of the American Statistical Association, 64 (1969), p. 1090.

Gibbons, J. Nonparametric Statistical Inference. New York, McGraw-Hill Book Company, 1971.

Hájek, Jaroslav. Nonparametric Statistics. San Francisco, California, Holden-Day, Inc., 1969.

Reviewed in the Australian Journal of Statistics, 13 (1971), p. 123.

Noether, Gottfried E. Elements of Nonparametric Statistics. New York, John Wiley and Sons, Inc., 1967.

Reviewed in Journal of the American Statistical Association, 63 (1968), p. 728.

Siegel, Sydney. Nonparametric Statistics for the Behavioral Sciences. New York, McGraw-Hill Book Company, 1956.

Reviewed in Journal of the American Statistical Association, 52 (1957), p. 384, and in The American Mathematical Monthly, 64 (1957), pp. 690-691.

3. A Problem-oriented Course in Introductory Statistics

Basic elements of the course. Envisioned here is a relatively radical departure, in the direction of a "case study" approach, from any of the courses described previously. Such a problem-oriented approach, which some may advocate for a second rather than a first course, is not entirely new. R. A. Fisher's classic Statistical Methods for Research Workers [26], which had a profound influence on a generation of scientists, is in large part problem-oriented. Although that book attempts to build up sophistication and complexity gradually, there are many places where a previously uninformed reader would find imposing gaps. Nevertheless, the audience, consisting largely of people with meaningful problems upon which to apply the methods and with the professional backgrounds to evaluate how sensible the conclusions were, found this book to be an invaluable guide. The level of maturity demanded was considerably higher than can be expected of students taking a first course in statistics. On the other hand, these students will have help from a teacher and other students.

We propose a modification of Fisher's approach. Let meaningful and nontrivial problems be presented (initially with a suggested solution, later without one). The detailed examination of the problems and solutions should evoke questions which lead to a discussion of fundamental concepts and methods. The order of the concepts to be discussed will depend on the problem and on the questions concerning a proposed solution. Generally, the order will not produce a simple and systematic course development of concepts. The student will need to be intelligent enough to have a quick grasp of new concepts, to recognize a reasonable approximation to a sensible solution, and to be satisfied with a nonsystematic development of the course.

The practice of scientists in presenting their work is to remove many of the untidy traces of false starts and preliminary studies which led to final clear-cut results. But a major part of the excitement of research lies in these lost traces, and much of this work involves statistics in its formal and informal aspects. It is to be hoped that a case study approach would spark this kind of excitement and interest on the part of the students and also lead to an understanding of statistical practices.

In a course of this kind, the computer and computer simulations can play an important role. When one inquires about the properties of a procedure, it is possible to use a computer simulation to see how the procedure works. The output of a simulation can also serve as a source of valuable data for further problems. Results of simulations can be reported by printouts (especially where there is no access to a computer). When adequate computer facilities are available, however, students can be encouraged to make their own independent investigations.

To illustrate the problem-oriented approach and to show how different cases are useful for different configurations of concepts, we present two examples. As these examples will make clear, the envisioned problem-oriented course differs sharply from present courses. In most current textbooks, problems are used simply to provide examples with which to illustrate the method under study. Often the problems are artificial and noncompelling, and the intellectual challenge to the student is limited to the study of how to carry out the details of the method correctly rather than of what method to apply and how well it works in the problem.

The problem-oriented course will not be easy to teach. As the need for more material on statistical concepts arises, assignments will have to be made in supplementary texts, programmed materials, or handouts developed by the instructor.

Example 1: The Peach Crop. A cling peach crop has been insured against frost damage by a growers' association. A frost occurs and a court must adjudicate disagreement between the growers and the insurance carrier on the amount of damages.

One possible way for deciding the compensation to be paid to growers is to examine each of the 26,793 trees in the orchards belonging to members of the association. An impartial expert could assess the damage to each tree. If he assesses the peaches on a tree as 15% damaged, the damage would be 37 × .15 = 5.50 dollars, where 37 dollars is the estimated average value of the peaches on undamaged trees. Theoretically, all 26,793 trees in the orchards could be examined in order to obtain a total damage figure.

Clearly, it is unreasonable to examine all the trees. Indeed, the cost of doing so would far exceed the value of the crop. A statistical consultant proposed examining a sample of 100 trees randomly selected from among all 26,793 trees in the orchards and using the average damage to the crop per tree in this sample as an estimate of the average damage per tree in the total population. His advice was taken, and the sample resulted in damages of 5.50, 7.10, 9.90, ..., 1.57, with a mean of 6.31. When this mean is multiplied by the number of trees in the orchards, an estimate of 6.31 × 26,793 = $169,063.83 is obtained for the value of the damages suffered by the growers.

The growers are concerned about this procedure. They wish to recover the full damages caused by the frost, but they are anxious to avoid being criticized by the court for using the wrong method of estimating the damages to the peaches. They ask, "Is the above estimate the correct value for the damages to the peach crop?" No, it is not. It is an estimated damage; it is impractical to obtain the actual data needed to compute the "correct" damage. "The law is not clear on how to treat estimates. Would a mathematician say that this is the correct estimate?" This is not a mathematical problem in the sense that there exists a "correct" way to select and analyze the data. It is a statistical problem. The question should be whether the procedure used to arrive at this estimate is a reasonable procedure for an unbiased statistician. It seems reasonable, but to answer this question in a meaningful way, we must know how an estimate obtained by this method is related to the unknown "damage" and how estimates from alternative methods would compare with this.

Up to this point the concepts of random sample and average have been used. These terms can and probably should be explained in detail after some of the more pressing problems are discussed. It is necessary now to point out that the estimate is random and to illustrate how the estimate would have varied if the experiment had been repeated 10 times. It would be valuable to consider the consequence of selecting different sample sizes and the cost of sampling. The necessity of dealing with variability is now clear, and the Central Limit Theorem can be hinted at or discussed briefly.

An analysis of the data from simulation would use histograms, means, and standard deviations. These could be presented in a matter-of-fact way. Instructions on how to carry out the presentations and analyses need not be given immediately if these concepts are easily enough appreciated in this specific context.

At this point one can extend the problem in several directions. One can make it more realistic by adding the fact that the orchards are located in two different river valleys and that the location and elevation of the orchards can be used as a basis for stratified sampling. This raises the point that the more one knows, the better one can do; it raises the question of how to profit from vague information.

Alternatively, one can discuss the advantages of randomization to avoid hidden bias.

Suppose that, due to a clerical error, it was reported that there were 700 trees in a certain orchard, whereas the actual figure was 70. The statistician might not wish to decrease his original estimate accordingly, because the 630 nonexistent trees had had a chance of being sampled and could have yielded a 0. An adjustment would make his estimate biased, and he would not understand the sampling properties of such _ex post facto_ procedures. On the other hand, if he were to testify in court, his opinion might be criticized if the estimate were not adjusted.

Another question that may arise is how much the insurance carrier should pay for damage to the crop if they are confident that it is between $150,000 and $190,000. Should the carrier pay $2,000 for increased sampling which is likely to reduce the length of the confidence interval to $10,000? Should one regard a payment of $170,000 in the first case as a gamble in which the insurer risks as much as $20,000? Would the $2,000 qualify as a prudent expense to reduce the amount at risk from $20,000 to $5,000? What is a reasonable trade-off?

It is apparent that this simplified example has enormous potential for illustrating statistical concepts in a context which is meaningful and interesting. Indeed, there is some danger of dwelling too long on one example, as students may become bored with it.

Example 2: Death Takes a Holiday. Do famous people, people whose birthdays are likely to be celebrated publicly, put off dying until after their birthdays? To answer this question, David Phillips studied the birth and death days of over 1,200 famous people (Tanur, Judith, *et al*., eds. Statistics: A Guide to the Unknown. [23]). To make the classification less tedious, he examined only the birth and death months.

There are two conflicting hypotheses. One states that there is no relation between birth and death months. According to this hypothesis, someone born in April is as likely to die in November as someone born in December. The alternate hypothesis states that indeed there is a relation between birth and death months, that someone born in December is less likely to die in November than someone born in April.

Phillips does not present a formal test of these hypotheses in his paper. He shows that for four different sets of data, the month before birth is less often a death month and the four months after birth are more often death months. The consistency of this result in all four sets of data is interpreted as increasing the plausibility of the alternative hypothesis, specialized somewhat to include the four-month death rise after the birth month.

What concepts are raised by this problem? The student is introduced to two-way contingency tables (month of birth and month of death) and the use of observed proportions as estimates of probabilities. Confidence intervals can be introduced to assess whether an observed proportion is consistent with a theoretical expectation. The idea of testing hypotheses and the use of the chi-square test enter naturally, as does the notion that the chi-square test is an all-purpose test whose power can be improved upon in the presence of a sufficiently specific alternative hypothesis. This point gives one an opportunity to discuss problems about designing hypotheses after studying the data. Finally, there is the problem of describing independence carefully and indicating how this hypothesis differs from a uniform distribution in the length of time between birth and death months.

This problem emphasizes the need to supplement informal heuristics by a sound theoretical framework which serves to guarantee that plausible conclusions are sound. Further, this problem is not closed. Several additional questions are open to study. The death rate in the birth month is also high. Should this enter the analysis? What would a detailed study of death dates within a month of the birth date show, and can this be studied effectively in view of the potentially small samples involved? Does the ability of modern medicine to prolong life for a few days if necessary have an effect on the death rate very near the birth date? There seem to be low death rates four and six months before the birth month. Does this mean anything?

Advantages and disadvantages of the problem-oriented approach. Some major advantages of the problem-oriented approach are:

a. Discussion of meaningful and nontrivial problems evokes interest and raises fundamental statistical questions.

b. Sensible, even if only partial, answers to realistic questions provide reinforcement for understanding of key concepts and motivation for further study.

c. The opportunity to use computer simulations is valuable and may help to motivate students.

d. A feeling for the excitement of research is derived when the traces of false starts and preliminary studies are not erased.

Some major disadvantages of the problem-oriented approach are:

a. This approach leads to the use of methods and ideas that have not been carefully digested in advance. There will be some need to operate at levels of less than complete understanding. The uninvolved and unmotivated student may find it difficult to salvage anything from such a course.

b. Live problems are often very complicated. It will be difficult but necessary to simplify without throwing away too much of the essence.

c. While this approach should be exciting to those who are involved, it may be frustrating to students who expect to be told what to do and how, and who find open-ended questions or nonresolved philosophical issues difficult to tolerate.

d. In this approach, concepts and methods arise in context and not in a systematic framework for review by the student without undue repetition.

e. The heuristic approach suggested here sometimes leads to incorrect results or paradoxes. When these are corrected, the student may feel insecure. How is he to know when his reasoning is sound without being told by the teacher? Some rigorous follow-up may be necessary.

f. Problems that seem easy may require a good deal more maturity from the students than we think. It may be difficult to transform unmotivated students, required by their major departments to take a statistics course, into participants in statistical inquiry.

Selected Sources for Cases:

Brownlee, K. A. "Statistics of the 1954 polio vaccine trials." Journal of the American Statistical Association, 50 (1955), pp. 1005-1014. (An invited address on the article "Evaluation of 1954 field trial of poliomyelitis vaccine: Summary Report." Poliomyelitis Vaccine Evaluation Center, University of Michigan, April 12, 1955.)

Cochran, William G.; Mosteller, Frederick; Tukey, John W. Statistical Problems of the Kinsey Report. American Statistical Association, 806 15th St., N.W., Washington, D. C. 20005, 1954.

Coleman, James S., et al. Equality of Educational Opportunity. Washington, D. C., U. S. Government Printing Office, 1966. (A set of Correlation Tables, separately bound, is also available for the use of research workers.)

Cutler, S. J. "A review of the statistical evidence on the association between smoking and lung cancer." Journal of the American Statistical Association, 50 (1955), pp. 267-83.

Heermann, Emil F. and Braskamp, Larry A. Readings in Statistics for the Behavioral Sciences. Englewood Cliffs, New Jersey, Prentice-Hall, Inc., 1970.

Kinsey, A. C.; Pomeroy, W. B.; Martin, C. E. Sexual Behavior in the Human Male. Philadelphia, Pennsylvania, W. B. Saunders Company, 1948.

Mosteller, Frederick, et al., eds. Statistics by Example, Part III, Detecting Patterns. Reading, Massachusetts, Addison-Wesley Publishing Company, Inc., 1973.

Mosteller, Frederick, et al., eds. Statistics by Example, Part IV, Finding Models. Reading, Massachusetts, Addison-Wesley Publishing Company, Inc., 1973.

Moynihan, Patrick. The Negro Family: The Case for National Action. Washington, D. C., U. S. Government Printing Office, 1965. Reprinted in Rainwater, Lee and Yancy, William L. The Moynihan Report and the Politics of Controversy. Cambridge, Massachusetts, MIT Press, 1967.

Peters, William. Readings in Applied Statistics. Englewood Cliffs, New Jersey, Prentice-Hall, Inc., 1969.

Report on Lung Cancer, Smoking, and Health. Public Health Bulletin 1103, Superintendent of Documents, U. S. Government Printing Office.

Steger, Joseph A., ed. Readings in Statistics for the Behavioral Sciences. New York, Holt, Rinehart and Winston, Inc., 1971.

Tanur, Judith, et al., eds. Statistics: A Guide to the Unknown. San Francisco, California, Holden-Day, Inc., 1972.

"The Cochran-Mosteller-Tukey Report on the Kinsey Study: A Symposium." Journal of the American Statistical Association, 50 (1955), p. 811.

Tufte, Edward R. The Quantitative Analysis of Social Problems. Reading, Massachusetts, Addison-Wesley Publishing Company, Inc., 1970.

V. USE OF DATA SETS

The use of problem material involving real data in teaching the introductory course in statistics overcomes some of the recognized defects of the traditional systematic course. By focusing attention on problems involving data which have intrinsic interest for the student, it is possible to introduce important statistical ideas and concepts in a context that will serve to facilitate student understanding. Use of examples and cases from real life can also be of great help in motivating students. Further, they can serve as vehicles for illustrating the application of statistical concepts and methods in a variety of settings.

Data sets useful for the introductory statistics course can be small or large. Small data sets can be extracted from investigations in a variety of fields (e.g., sample surveys of voters or consumers, federal statistical reports, annual business reports, and the like), or they can be generated by various class experiments or investigations. They are characterized by limited amounts of data suitable for illustrating a particular statistical concept or technique. An example is the annual earnings of a company during the past ten years, which might be used for comparing average earnings in the first five years with those in the second five years.

Large data sets are characterized by large volumes of interrelated data. For example, a data set on voting behavior may contain information about a variety of personal characteristics of the voter (e.g., age, sex, income, education, marital status, attitudinal data) as well as a variety of voting behavior (e.g., voting behavior in elections during the past five years at national, state, and local levels). Another example of a large data set is medical information about a segment of the population (e.g., age, sex, height, weight, blood pressure, cholesterol level, etc.). Large data sets are often available on tapes and can be stored in local computers for ease of access by students.

For the introductory statistics course that features a systematic development of statistical concepts, large and small data sets are used in essentially similar fashion. For example, to illustrate the construction of a confidence interval, one might use data on consumer food expenditures or on blood pressure of patients. These data could be given to the student in the form of a small data set, or he could be asked to perform a limited task of analysis with these data. (This is in distinction to the use of large data sets as a means of developing understanding of statistical concepts unsystematically in a problem-oriented fashion.)

Use of problem material based on real data sets requires that the instructor devote class time to discussion. The setting of the problem should be discussed initially so that students understand it clearly. After the statistical analysis has been completed, the meaning and interpretation of the results should be discussed, as well as any further analyses that might be required for investigation of the problem.

The use of data sets is not restricted to illustrating the application of particular statistical methods. They can also be used as a meaningful vehicle in sampling experiments. The data set for this purpose would be viewed as a population, and repeated samples would be selected from it by a table of random numbers or by computer. An advantage of this use of data sets over sampling from artificial populations is that real data sets permit the students to interpret the results more meaningfully. For example, the demonstration that the sampling distribution of the sample mean becomes more concentrated with increasing sample size is more meaningful for many students when they can interpret the numbers involved in real terms, e.g., as incomes or blood pressures.

The use of data sets requires computational effort. Although we strongly believe that the use of real data sets adds a valuable dimension to the introductory statistics course, we also strongly believe that students should not engage in computational drudgery. Consequently, adequate computational facilities need to be provided if large data sets are to be employed. If these facilities are not available, small data sets with simplified numbers should be employed.

Experience suggests that students find data sets most interesting when they come from an area of interest to them. Thus, if the introductory statistics course is taught in a number of sections, it would be desirable to set up one or more sections for social science students, one or more sections for physical science students, and one or more sections for biological science students. Cases, examples, and data sets from these subjects can then be used so that students will be motivated by illustrations from situations with which they are familiar. If the introductory statistics course is taught in only one section, examples and cases should be selected that are simple and easy to understand by most of the students in the class.

Data sets can be obtained from diverse sources. One major source, containing demographic, sociological, economic, and many other kinds of data, is:

U. S. Bureau of the Census. Statistical Abstract of the United States, 1970. Washington, D. C., U. S. Government Printing Office, 1970.

Data on consumer behavior and voting preferences, and other types of sample survey data are available from various survey research centers. A list of sources for such data sets is:

Social Science Data Archives in the United States, 1967. Council on Social Science Data Analysis, 605 West 115th Street, New York 10025.

Three books sponsored jointly by the National Council of Teachers of Mathematics and the American Statistical Association contain a wealth of statistical examples and applications:

Mosteller, Frederick, et al., eds. Statistics by Example, Part III, Detecting Patterns. Reading, Massachusetts, Addison-Wesley Publish-Company, Inc., 1973.

Mosteller, Frederick, et al., eds. Statistics by Example, Part IV, Finding Models. Reading, Massachusetts, Addison-Wesley Publishing Company, Inc., 1973.

Tanur, Judith, et al., eds. Statistics: A Guide to the Unknown. San Francisco, California, Holden-Day, Inc., 1972.

In addition to using data sets furnished by the instructor, students can obtain data sets of their own from their particular field of interest, or the entire class can develop data by class-organized demonstrations and experiments (see Section VII) or by surveys conducted by members of the class.

VI. USE OF COMPUTERS

A computer can serve three broad roles in the implementation of the goals of an introductory statistics course. It can: (1) clarify certain key ideas of the course, (2) perform routine numerical calculations, and (3) facilitate more active student participation in the development of statistical concepts. Each of these roles is illustrated below. The question of whether a student should learn computer programming inevitably arises in discussions of the role of computers in education. It is this Panel's view that so much is demanded in the ordinary introductory statistics course, that the taking of additional time to teach programming is not justified. This Panel does not feel that the learning of programming will enable the student to obtain sufficiently greater analytical skills to warrant taking this time away from other topics. There may be reasons in the context of a student's entire curriculum for his being taught programming. If so, a decision must be made in the context of the curriculum as to which course should contain the teaching of programming.

We turn now to illustrations of the various ways in which computers can be used to clarify important statistical concepts. The long-run relative frequency interpretation of probability can be illustrated by having a computer repeatedly simulate an experiment which results in "success" or "failure" on each trial and print the relative frequencies of success after 10, 20, 30, ... trials. The experiment should not be such a trivial one that students can find the probability by the examination of a few equally likely outcomes.

One of the primary objectives of this introductory course is to introduce students to the notions of variability and uncertainty; these concepts can be illustrated in a number of ways on the computer. The notion of a confidence interval, for example, can be brought to life with a computer. A large number of 80% confidence intervals obtained from samples of size 20 from a binomial distribution with $p = 0.6$ are easily obtained on a computer and the proportion of intervals containing 0.6 determined. The confidence level should be low enough so that students will see cases where $p = 0.6$ does not lie within the interval. This demonstration can be linked to the testing of statistical hypotheses and can provide a sound basis on which to discuss that topic.

Properties of estimators can also be studied on the computer by obtaining sampling distributions of estimators, e.g., distributions of $\frac{\Sigma(x-\bar{x})^2}{n}$, $\frac{\Sigma(x-\bar{x})^2}{n-1}$, $\frac{\Sigma(x-\bar{x})^2}{n-2}$ for fairly small n, to illustrate what is meant by an unbiased estimator. Variances of different estimators also can be studied by examination of sampling distributions generated by the computer.

An additional benefit can result from the use of random number generators in an introductory course. The necessity of knowing that pseudo-random digits generated by a computer possess characteristics expected of random digits presents an opportunity to introduce the study of statistical tests such as the chi-square test and tests of runs.

A second role of the computer in an introductory statistics course is to perform routine numerical calculations that are required in the analysis of the problems. This Panel believes that the computational aspects of the subject should be held to a minimum, consistent with the clarification of statistical ideas and principles and with the motivation of students by the introduction of interesting, contemporary problems. If an analysis of a situation that interests an instructor and his class requires computational assistance, then by all means some computational aid should be used, whether it be from an electronic computer or a desk calculator. Students expect a course to contain a touch of realism, and they are frequently disappointed if they do not attain some experience with handling large amounts of data. Students should learn the procedures for calling up a statistical package because that knowledge is often used by those who analyze data. Three widely used statistical packages are:

Dixon, W. J., ed. Biomedical Computer Programs, University of California Publications in Automatic Computation #2, Second Edition. Los Angeles, California, University of California Press, 1970.

Hogben, David, et al. OMNITAB II User's Reference Manual. NBS Technical Note 552, U. S. Department of Commerce. Washington, D.C., U. S. Government Printing Office, 1971.

Nie, Norman H., et al. Statistical Package for the Social Sciences. New York, McGraw-Hill Book Company, 1970.

Even when major reliance for computation is placed on a large electronic computer, there are occasions when the understanding of a concept or a computational technique requires that students perform computations on a desk calculator or by hand. One or two simple problems illustrating the Student t-test, for example, should require students to calculate sample variances. Experience indicates that it is unwise to assume that students always understand what computations are required with raw data and know that statistical analysis is not always based on large amounts of data requiring the use of a computer.

A computer can also serve instructors who wish to break away from presenting all the ideas of the course through formal lectures. We refer to this third function of computers as their "interactive" role. For example, the instructor need not lecture students as to what constitutes a "large" sample in order to be able to apply the Central Limit Theorem; instead, an empirical investigation can be undertaken. First, the instructor can have the students guess minimum values of n for which sample means tend to be normally distributed when sampling from a number of populations with different characteristics. The computer can then be used to generate sampling distributions of sample means for different sample sizes, thereby providing a basis for refined estimates of minimum n and ultimately enhancing the students' understanding of the Central Limit Theorem and the degree to which the shape of the population density and sample size affect the shape of the sampling distribution. Working with data of these kinds, students usually raise new questions. In the case cited, students naturally want to know how to decide whether their sampling distributions are normal, and the instructor has the motivation for a new investigation or at least a mention of inferential problems. Another way of encouraging student participation through interaction with the computer is the selection of n and α to obtain a reasonable power curve for testing hypotheses about binomial p. Students also can be asked to select those independent variables in a set of variables which are important in a regression analysis; economists in most colleges and universities can provide the instructor with suitable data for this purpose.

The interactive role of the computer is enhanced by the existence of terminals and a sufficient number of screens in the classroom, but this equipment is not essential. Students can be required to make investigations of the sort illustrated above as homework exercises using the available computer facility.

In the use of a computer in any of its three functions, it is desirable to give the student a feeling of control over the operations of the computer. This can be done by requiring the student to specify parameters, desired output, and the like. For instance, in a sampling experiment the student might specify the population to be sampled, the sample size, and the number of trials, as well as the nature of the output (e.g., frequency distribution, histogram). Some students also benefit by writing simple programs using subroutines of basic statistical operations.

The printout of a computer run will be of interest to the student to the extent that he is interested in the original problem. Hence, our earlier comments on the importance of interesting and meaningful data sets are equally relevant whether the calculations are performed on a computer or by the student.

Even when a computer is not available, students can still realize some of its benefits. For instance, an instructor can obtain a complete set of computer printouts for his students. Students can use this set of printouts in the same fashion as a laboratory manual. At appropriate times the instructor can refer to one of the printouts and explain how a particular analysis is carried out with the computer, explain the types of information given on the computer printout, and indicate how this information would be used for analysis of the data. In this way the students can obtain some of the benefits of learning how a computer can assist in statistical analysis and also obtain computer illustrations of some of the basic statistical concepts such as sampling distributions without actually having access to a computer. Dixon discusses the use of a "side inch" of computer printout in Review of the International Statistical Institute, 39 (1971), pp. 315-339. The "side inch" can be obtained for $3.00 by writing to Professor W. J. Dixon, Department of Biomathematics, School of Medicine, University of California, Los Angeles, California 90024.

The use of computers in teaching introductory statistics can be expensive. In addition to the costs of the equipment, the use of computers requires much planning by the instructor, as well as substantial monitoring during the course. A modern programmable minicomputer provides a relatively inexpensive way of performing a wide variety of statistical operations.

The impression has been gained by many instructors using computers in teaching introductory statistics that their use is of interest to students and motivates them. Despite their wide use, however, only limited formal evaluation of the effectiveness of computers in the introductory statistics course has been carried out to date. Undoubtedly, the coming years will provide important information on the best way in which computers can be used in introductory statistics courses as well as on their cost-effectiveness as teaching and learning devices.

None of the above mentioned uses of computers are to be confused with computer-assisted instruction. This method of instruction is expensive for use in the introductory course and, when compared with a well written programmed learning text, it may not be economically justified at this time. That is not to say, however, that technology will not advance sufficiently far in the next few years to make computer-assisted instruction an important alternative to existing modes of instruction.

Selected References:

Buford, Roger L. Statistics: A Computer Approach. Columbus, Ohio, Charles E. Merrill Publishing Company, 1968.

Carmer, S. G. and Cady, F. B. "Computerized data generation for teaching statistics." The American Statistician, 23 (1969), pp. 33-35.

Computers in Undergraduate Education: Mathematics, Physics, Statistics, and Chemistry. Proceedings of a Conference sponsored by the National Science Foundation and conducted at the Science Teaching Center of the University of Maryland, College Park, 1967.

Development of Materials and Techniques for the Instructional Use of Computers in Statistics Courses. Department of Statistics, University of North Carolina, Chapel Hill, North Carolina, 1971.

Foster, F. G. and Smith, T. M. F. "The computer as an aid in teaching statistics." Applied Statistics, 18 (1969), pp. 264-270.

Freiberger, W. and Grenander, U. A Short Course in Computational Probability and Statistics, Applied Mathematical Sciences, Volume 6. New York, Springer-Verlag New York Inc., 1971.

Lohnes, P. R. and Cooley, W. W. Introductory Statistical Procedures: With Computer Exercises. New York, John Wiley and Sons, Inc., 1968.

Milton, Roy and Nelder, John, eds. Statistical Computation. New York, Academic Press, Inc., 1969. The following articles:

1. "Computer-assisted instruction in statistics," William W. Cooley, pp. 337-347.

2. "Computers in the teaching of statistics: Where are the main effects?" David L. Wallace, pp. 349-361.

3. "Time sharing and interactive statistics," P. M. Britt, et al., pp. 243-265.

Proceedings of a Conference on Computers in the Undergraduate Curricula. The University of Iowa, Iowa City, Iowa, 1970. The following articles:

1. "Use of computers in undergraduate statistics instruction," Mark I. Appelbaum and Donald Guthrie, pp. 2.1-2.3.

2. "The computer as an instructional tool for the statistics course," Young O. Koh, pp. 2.4-2.17.

3. "Using the computer in basic statistics courses," Richard L. Wikoff, pp. 2.18-2.24.

4. "Use of computers in teaching statistics to engineering students," Armit L. Goel, pp. 2.25-2.31.

5. "SAMDS: A program to generate empirical sampling distributions of the mean," Henry G. Garrett, pp. 2.32-2.38.

6. "Computer-assisted teaching of experimental design: The development of a 'Master Experimenters' Program," Albert R. Gilgen and Michael A. Hall, pp. 2.39-2.46.

Proceedings of the Second Annual Conference on Computers in the Undergraduate Curricula. Dartmouth College, Hanover, New Hampshire, 1971. The following articles:

1. "Use of computers in statistical instruction," Richard N. Schmidt, pp. 437-443.

2. "Inclusion of explanatory material in computer programs for statistical analysis," Jolayne Service, pp. 444-455.

3. "STATLAB, a simple programming system for the statistics laboratory," G. F. Atkinson, pp. 456-462.

Review of the International Statistical Institute (2 Oostduinlaan, The Hague, Netherlands.) Vol. 39, Number 3 (1971). The following articles:

1. "Notes on available materials to support computer-based statistical teaching," W. J. Dixon, pp. 257-286.

2. "Side Inch," W. J. Dixon, pp. 315-339.

3. "On using a conversational mode computer in an intermediate statistical analysis course," D. Quade, pp. 343-345.

4. "Development of materials and techniques for the instructional use of computers in statistics courses," D. Quade, pp. 361-362.

5. "Survey sampling in a computerized environment," T. E. Dalenius, pp. 373-397.

Schatzoff, Martin. "Application of time-shared computers in a statistics curriculum." Journal of the American Statistical Association, 63 (1968), pp. 192-208.

Sterling, T. and Pollack, S. "Use of the computer to teach introductory statistics." Communication of the Association of Computing Machinery, 9 (1966), pp. 274-276.

VII. EXPERIMENTS, SIMULATIONS, DEMONSTRATIONS, AND TEACHING AIDS

Since planning of experiments and analysis of actual data are major concerns of most practicing statisticians, it would seem that one might use relatively simple experiments to develop basic statistical concepts, to obtain data for future analysis, and to give some experience in designing an experiment. There are arguments against this approach. It can use up a considerable amount of time for the individual student, for the class, and especially for the teacher. It can sometimes be frustrating. The arguments for such an approach center about the personal involvement of the student, his increase in interest, and his discovery of important principles or relationships. There is considerable evidence that such an approach can be successful. Practically all of the basic training courses in statistical quality control used experiments and demonstrations to introduce statistical concepts, and these uses were judged successful. Two helpful references on these uses are:

Olds, E. G. and Knowler, L. A. "Teaching statistical quality control for town and gown." Journal of the American Statistical Association, 44 (1949), pp. 213-230.

Scherwin, R. L. "Teaching aids for statistics and quality control." Industrial Quality Control, 23 (1967), pp. 654-660.

1. Demonstrations and Verification Experiments

For out purposes we would like to distinguish among several types of experiments. The first type we shall designate as a "demonstration." The student is given an assignment for which the instructor has the theoretical solution. If the student also knows or guesses the expected outcome, then this approximates the typical high school science experiment which simply illustrates a law stated or derived in the text book. Requiring students to toss a coin 100 times and to record the total number of heads would be a typical demonstration. The students know that the probability is close to ½ and most of them also know that one does not really expect 50 heads and 50 tails.

An assignment that the student count the number of rolls of a die until all six digits occur or count the number of digits in a random digit table until all ten digits are found is another type of experience. The instructor, with his knowledge of the geometric distribution and the theorems on sums of independent random variables, may know the mean and variance of the two distributions. He also may know that the distributions are skewed. The student does not have this knowledge and in this first course is unlikely to acquire it. To the student this assignment is a venture into the unknown. On the other hand, the instructor, while he may be surprised by an individual result, is quite sure that he will not be surprised by the entire set of results. We will call an activity such as this a "verification experiment."

There are many demonstrations and verification experiments from which the students may discover or better understand basic statistical concepts. The student might be asked to toss a coin 64 times, recording his results (0 or 1) in a 4 by 16 table, and to find the average for each column, for each row, and for the entire table. These results could be used to justify statements about the distribution of the sample mean as n increases. The student might be asked to draw samples, using a sampling paddle, from a box of beads or balls of at least two different colors. Bowls of numbered chips might be used, made up to represent different distributions, either of different types, such as rectangular, triangular, exponential, and normal, or of one type with different parameters. The student might draw small samples from a population, construct a confidence interval for each sample, and determine the proportion of confidence intervals which include the population mean or proportion. In general, the distribution of any sample statistic might be motivated by a demonstration or experimental verification. To a great extent, the value of such activities depends upon the student's interest and enthusiasm. In any case, the activities should be used with moderation.

Selected References:

Berkeley, Edmund C. Probability and Statistics--An Introduction Through Experiments. New York, Science Materials Center, 1961. This book describes some 27 experiments. It is written to accompany a kit.

Dixon, Wilfrid J. and Massey, Frank J., Jr. Introduction to Statistical Analysis. New York, McGraw-Hill Book Company, 1969. Most chapters have a set of class exercises based on random number tables, bead drawings, etc.

Harrison, R. D. "An activity approach to the teaching of statistics and probability." (in three parts) Mathematics Teaching, 34 (1966), pp. 31-38; 35 (1966), pp. 52-61; 36 (1966), pp. 57-65.

Malpos, A. J. Experiments in Statistics. Edinburgh, Scotland, Oliver and Boyd, Ltd., 1969.

2. Open Experiments

An "open experiment" is quite different from a demonstration or an experimental verification. Neither the instructor nor the student is able to predict the outcome. As an extreme example, let us consider a different version of the coin-tossing demonstration. The class, in some manner unknown to the instructor, will split into two groups of equal size. Each member of the first group will actually toss a coin 100 times and record the outcomes (0 or 1) in order. Each member of the second group will write out a sequence of 100 numbers (0 or 1) without tossing a coin, in such a fashion that the sequence, in his opinion, simulates the results of actual coin tossing. Using statistical tests (which may be developed over the term), the instructor will attempt to identify in which group each sequence belongs. The teacher might classify a sequence as group two if the results come too close to 50-50, or if there are too many runs, or if the longest run is suspiciously small, or if the cumulative fraction of heads stays too close to $n/2$. But now the experimental situation is completely reversed; the students know the answers and the teacher is the discoverer.

Actual classroom experiments of the open type can cover a wide range of difficulty. A relatively simple one is to ask the class the probability of getting a head on a coin if the coin is spun instead of tossed. First, one must specify the essential conditions. What kind of coin? How long a spin? If the coin is spun by holding it with the finger of one hand and flicking it with a finger of the other hand, does the initial position of the coin make a difference?

Some apparently simple experiments can suddenly develop complications. Consider the experiment of empirically determining the approximate probability that a thumbtack will fall point up. Obviously one must first specify at least the particular make of thumbtack. But does one need to specify the type of surface it falls on? And can one speed up the experiment by taking ten thumbtacks in a container and shaking them up? In one class, two students shook the thumbtacks in a china cup with a gentle lateral motion and then poured them on a table. Result--almost 100% points up! At this point another student decided to investigate the effect of different grades of sandpaper for the receiving surface and discovered an interesting regression problem.

An open experiment that can display individual differences is the "coin shove" experiment (see the Jowett reference, below). On some smooth poster board 30 inches long, draw a target line (designated 0) approximately 6 inches from one end. Also draw lines an inch apart (designated +1, +2, ...; -1, -2, ...), parallel to the target line. Then draw a starting line approximately 6 inches from the other end of the poster board. A trial of this experiment consists of shoving a coin from behind the starting line so that it stops as close to the target line as possible. The purpose of the experiment is to study the frequency distribution of scores that result from n repeated trials. Many questions can be investigated:

Is there a significant improvement in the scores if 5 practice crials are allowed? Is there a significant difference between the abilities of individuals? Is it reasonable to assume independent trials? Can the scores of individuals be improved by allowing the use of a slingshot-type coin shooter? At what value of n is student interest replaced by boredom?

Selected References:

"Interim Report of the R. S. S. Committee on the Teaching of Statistics in Schools." Journal of the Royal Statistical Society, Series A, 131 (1968), pp. 478-497.

Jowett, G. H. and Davies, H. M. "Practical experimentation as a teaching method in statistics." Journal of the Royal Statistical Society, Series A, 123 (1960), pp. 10-35.

Mosteller, Frederick, et al. Probability with Statistical Applications, 2nd ed. Reading, Massachusetts, Addison-Wesley Publishing Company, Inc., 1970. Manual describes a variety of experiments, demonstrations, and teaching aids. Pages 449 to 454 of the text suggest a set of projects for highspeed computers.

Rade, Lennart, ed. The Teaching of Probability and Statistics. New York, John Wiley and Sons, Inc., 1970.

3. Simulations

Another type of laboratory experience might be classified a "simulation." Here there may exist a theoretical solution to the problem, but it may not be at the competence level of the class. One example might be a simplified epidemic where each day 20 persons line up in random order. Suppose that of the $n = 20$ persons, $x = 2$ persons are contagious, $y = 3$ persons are immune, and $z = 15$ are susceptible. A susceptible person who stands next to a contagious person catches the disease. The duration of the disease is one day, so that a contagious person is immune the next day. To simulate the epidemic, shake a container having two red beads labeled C for contagious, three yellow beads labeled I for immune, and 15 white beads labeled S for susceptible, and pour the beads in a trough or a long-stemmed funnel. (A computer may be used instead of beads.) Suppose the result for the first trial is:

S S C S S I C S S I S S S S S I S S S S

Then on the next trial we would have three infectious persons (places 2, 4, 8), five immune persons (places 3, 6, 7, 10, 16), and 12 susceptible persons. Variables which might be investigated are the number of trials until the epidemic is ended, the number of susceptible left at the end of the epidemic, and the largest number of contagious persons at any time. The simulation study can be enlarged by investigating the behavior of the epidemic for various values of x, y, and n.

Another simulation is the server problem. Customers arrive at integral values of time t from 1 to n (for example, let $n = 50$). The probability of a customer's arriving at any time t is a constant p (for example, $p = 1/5$). The service time for a customer is a constant c less than $1/p$ (for example, $c = 4$). At the end of n time intervals the gates are closed but the customers in line must still be served, and of course the persons serving them must be paid overtime. For equipment one can use a die, a pack of playing cards, a table of random digits, or a computer. The first decision that needs to be made concerns what data to record. After one class tried the experiment, they decided that they should record: number of arrivals, total service time, idle time, overtime, sum of the delay times, and sum of squares of delay times. This last number they named the "riot index."

Films

A number of films pertaining to probability, statistics, and quality control are available. Of those known to the Panel, the film "Random Events" of the PSSC Physics Series seems most suitable for the kind of course we are considering.

Selected Films:

1.	Probability and Statistics 25 minutes B & W 1957	ASSN Films 347 Madison Avenue New York, New York 10036
2.	Mathematics and the River 19 minutes Color 1959 Horizons of Science Series	Educational Testing Service 20 Nassau Street Princeton, New Jersey 08540
3.	Probability and Uncertainty 56 minutes 1965	Educational Service, Inc. 40 Galen Street Watertown, Massachusetts 02172
4.	Mean-Median-Mode 13 minutes Color, B & W 1966	McGraw-Hill Text Films 330 West 42nd Street New York, New York 10018
5.	Probability 12 minutes Color 1966	McGraw-Hill Text Films 330 West 42nd Street New York, New York 10018
6.	Matter of Acceptable Risk 30 minutes B & W 1967	Indiana University Audio-Visual Center Bloomington, Indiana 47401
7.	Random Events 31 minutes B & W 1962 PSSC Physics Series	Modern Learning Aids 315 Springfield Avenue Summit, New Jersey 07901

8. How's Chances
30 minutes B & W 1957
Westinghouse

Association Films
347 Madison Avenue
New York, New York 10017

9. It's All Arranged
30 minutes B & W 1957
Westinghouse

Association Films
347 Madison Avenue
New York, New York 10017

10. Photons
19 minutes B & W 1960
PSSC Physics Series

Modern Learning Aids
315 Springfield Avenue
Summit, New Jersey 07901

11. Predicting Through Sampling
10 minutes Color 1969

Bailey Films Associates
11559 Santa Monica Blvd.
Los Angeles, California 90025

12. The Probabilities of Zero and One
11 minutes Color 1969

Bailey Films Associates
11559 Santa Monica Blvd.
Los Angeles, California 90025

13. Probability, An Introduction
9 minutes Color 1969

Bailey Films Associates
11559 Santa Monica Blvd.
Los Angeles, California 90025

Visual Aids and Other Materials

There are available a variety of visual aids and other helpful materials. Various manufacturers of games offer large dice, roulette wheels, chuck-a-luck cages, etc. Biased and mismarked dice are available, although it is sometimes difficult to locate a source.

Selected Sources:

a. Lightning Calculator Company, Box 6192, St. Petersburg, Florida 33736.

1. Quincunx or Galton Board. A device for generating a binomial or normal distribution. This model features an adjustable outlet which enables one to shift the population mean, a sliding control which enables one to drop from one to five beads at a time, and a control for examining small samples before they are accumulated into a large sample; completely self-enclosed so that the beads cannot drop out.

2. Sampling Demonstrators. Consists of six different colors of beads, plastic container, and sampling paddles. Available in two different models, one with 2000 beads and three sampling paddles, a second with 1000 beads and two sampling paddles.

3. Control Chart Simulator. Device consisting of various distribution patterns, dowel rods for indicating limits of variability, etc., and frame. Reverse side has set of horizontal wires with beads to demonstrate observations over time, either individual or grouped.

b. Ray R. Lilly, 30 Lilly Road, Wanaque, New Jersey 07465.

1. Sampling demonstrators, plastic balls in various colors and lot sizes, sampling paddles. Available in different models or to specification.

2. Demonstration Board. One side consists of various models for distributions, with templates (attachable) to show distribution of average for $n = 5$ and $n = 20$. Also a model to show change in variability. Dowel rods show specification limits, control limits, modified control limits, etc. Reverse side consists of one set of horizontal wires for distribution over time of either individuals or subgroups, plus overlay set of wires to demonstrate distribution of medians.

3. Three-dimensional models. One consists of a peg-board with plastic rods and beads to demonstrate multivariable control charts, bivariate distributions, etc. A second has a set of sliding channels with pegs, which can be used to demonstrate correlation, regression, etc.

c. E. S. Lowe Company, Inc., 200 Fifth Avenue, New York, New York 10010. Manufacturer of games and accessories. Items such as large dice (up to 4 inches), roulette wheels, bingo games, chuck-a-luck, miniature slot machines, etc.

d. Hunter Spring Company, Lansdale, Pennsylvania 19446. Normal frequency distribution template. For use in illustration work. Small template with $\sigma = 10$ mm.

e. Tell Manufacturing Company, 200 South Jefferson Street, Orange, New Jersey 07050. Manufacturer of plastic beads and balls in various size ranges from 6 mm to 23 mm diameter. Some 13 different colors. Smaller sizes in units of 100.

f. Walco Products Company, 37 West 37th Street, New York, New York 10018. Manufacturer of wooden beads in various shapes and sizes. Spherical beads come in sizes from 3 mm to 20 mm. Cylindrical and barrel shapes up to 1 inch in height. Small beads must be ordered in units of 1000, larger in units of 100.

g. Quality Service Foundation, Weena 700, Rotterdam, Netherlands.

1. Polyhedra numbered 0 to 9 for generating random digits or selecting random samples.

2. Drafting triangle with two normal templates for distribution of individuals and distribution of averages. (n = 5)

h. Japanese Standards Association, Ginza Higashi 6-1, Chuo-ku, Tokyo, Japan. Icosahedron numbered 0 to 9 for generating random digits or for selecting random samples.

Course Organization

Inclusion in the introductory statistics course of the varied motivational and instructional activities, such as the use of data sets, computers, sampling experiments, and films, is time-consuming. The instructor may therefore wish to add another class meeting each week to permit more extensive use of this pedagogy.

One method of including these activities is by scheduling them in class when they naturally arise. An advantage of this approach is that the student can understand the activity in its context. A disadvantage is that many of the activities, such as sampling experiments, are time-consuming and therefore may disrupt the continuity of the overall development.

Another approach is to schedule formal laboratory periods from time to time, during which these activities take place. Then the development of the course material is not disrupted, but the timing of events can not always be made to coincide with the class developments.

Learning Resources Centers

Demonstrations, verifications, experiments, and simulations require physical facilities and organization. The Panel has been encouraged to learn of various institutions that have developed learning resources centers where students can listen to tapes, view films, find programmed material for remediation, carry out experiments, and plan their own group investigations. Such a center provides an opportunity for small group discussions both at the planning stage of an investigation and at the analysis stage. For some of the experiments discussed, such as the server problem or the epidemic problem, it is advisable to have students work in teams of two to four, with each team repeating the experiment a small number of times. The results can then be pooled. For other experiments there will need to be serious consideration given to the question of pooling. For the coin shove experiment, for instance, the question of pooling results will raise serious problems and can lead to a discussion of individual and group differences.

Organizing students into small groups to carry out investigations can be an effective means of stimulating student interactions and can thereby aid the learning process. Student teams should be reasonably small so that each student can participate effectively.

The teacher may meet with each team from time to time or be available for assistance when a team requires it.

Some learning resources centers have a time-shared console or are adjacent to a computation facility having such consoles. The Panel suggests that mathematics departments offering statistics courses consider the establishment of such a facility.

ADDITIONAL RESOURCE MATERIALS

Folks, LeRoy. "Some prior probabilities on the future of statistics." The American Statistician, 24 (1970), pp. 10-13. References cited in the Folks article:

1. "Some aspects of the teaching of statistics." John Wishart, 1939. JRSS, 102, pp. 532-564.

2. "The teaching of statistics." Harold Hotelling, 1940. Annals of Math. Stat., pp. 457-471.

3. "On the future of statistics." M. G. Kendall, 1942. JRSS, 105, pp. 69-91.

4. "The teaching of statistics." Institute of Mathematical Statistics Committee on the Teaching of Statistics, 1948. Annals of Math. Stat., 19, pp. 99-115.

5. "The teaching of statistics." John Wishart, 1948. JRSS, A, 111, pp. 212-229.

6. "Why statistics?" P. C. Mahalanobis, 1948. Sankhyā, 10, pp. 195-228.

7. "The future of data analysis." J. W. Tukey, 1962. Annals of Math. Stat., 33, pp. 1-67.

8. "On the future of statistics--a second look." M. G. Kendall, 1968. JRSS, A, 131, pp. 182-194.

Kirk, Roger E., ed. Statistical Issues: A Reader for the Behavioral Sciences. Belmont, California, Brooks/Cole Publishing Company, 1972.

Klein, Morris. Mathematics in Western Culture. New York, Oxford University Press, Inc., 1953. The following articles:

1. "The mathematical theory of ignorance: The statistical approach to the study of man," pp. 340-358.

2. "Prediction and probability," pp. 359-375.

3. "Our disorderly universe: The statistical view of nature," pp. 376-394.

Mathematical Thinking in Behavioral Sciences (Readings from *The Scientific American*). San Francisco, California, W. H. Freeman and Company, 1968. The following articles:

1. "Chance," A. J. Ayer, October, 1965.

2. "What is probability?" Rudolph Carnap, September, 1953.

3. "Subjective probability," John Cohen, November, 1957.

4. "Probability," Mark Kac, September, 1964.

Newman, James R., ed. *World of Mathematics*, 4 vols. New York, Simon and Schuster, Inc., 1956. The following articles:

Vol. 2, Part VII: The Laws of Chance

1. "Concerning probability," Pierre Simon de Laplace, p. 1325.

2. "The red and the black," Charles Sanders Peirce, p. 1334.

3. "The probability of induction," Charles Sanders Peirce, p. 1341.

4. "The application of probability to conduct," John Maynard Keynes, p. 1360.

5. "Chance," Henri Poincaré, p. 1380.

Vol. 3, Part VIII: Statistics and the Design of Experiments

1. "Foundations of vital statistics," John Graunt, p. 1421.

2. "First life insurance tables," Edmund Halley, p. 1437.

3. "The law of large numbers," Jacob Bernoulli, p. 1452.

4. "Sampling and standard error," L. C. Tripett, p. 1459.

5. "On the average and scatter," J. C. Moroney, p. 1487.

6. "Mathematics of a lady tasting tea," Sir Ronald A. Fisher, p. 1512.

7. "The vice of gambling and the virtue of insurance," George Bernard Shaw, p. 1524.

Preparation for Graduate Work in Statistics, 1971. CUPM, P. O. Box 1024, Berkeley, California 94701.

Review of the International Statistical Institute. 2 Oostduinlaan, The Hague, Netherlands. Vol. 39, Number 3 (1971). The following articles:

1. "New techniques of statistical teaching: Opening remarks," pp. 253-256.

2. "The first course in statistical methods and the use of teaching aids," K. Austwick, J. Hine, G. B. Wetherill, pp. 287-306.

3. "Post college continuing education activities in statistics," J. Stuart Hunter, pp. 307-311.

4. "Comments about a general audience TV course on statistics," J. Hemelrijk, pp. 312-314.

5. "The joint American Statistical Association-National Council of Teachers of Mathematics Committee on the Curriculum in Statistics and Probability," F. Mosteller, pp. 340-342.

6. "Operation of the Centre for Applied Statistics in Medicine and Medical Biology," D. Schwarz, pp. 346-347.

7. "The teaching of statistics," T. Yoshizawa, pp. 348-350.

8. "New techniques of statistical teaching," H. C. Hamaker, pp. 351-360.

9. "The jackknife," F. Mosteller, pp. 363-368.

10. "Round table meeting: Recommendations," pp. 369-372.

The American Statistician. 805 15th Street, N.W., Washington, D. C. 20005. "The Teacher's Corner" frequently contains articles on topics suitable for an introductory statistics course.

The Statistician. 55 Park Lane, London, W 1, England. See Vol. 18, Number 3, 1968:

The Teaching of Statistics

1. "Innovation in teaching," D. T. Page, p. 207.

2. "Teaching management statistics by programmed instruction," D. Kitchen, pp. 209-226.

3. "Practical classes in statistics," H. M. Davies, pp. 227-236.

Varberg, Dale E. "The development of modern statistics," The Mathematics Teacher. Part 1, 56 (1963), pp. 252-257; Part 2, 56 (1963), pp. 344-348.