INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeib Road
Ann Arbor, Michigan 48108
SITUMEANG, Robinson Panahatan, 1931-
MOTIVATIONAL SUPPLEMENTS TO LECTURE DISCUSSION
IN INTRODUCTORY COLLEGE STATISTICS COURSES.

The Ohio State University, Ph.D., 1976
Statistics

Xerox University Microfilms, Ann Arbor, Michigan 48106
MOTIVATIONAL SUPPLEMENTS TO LECTURE DISCUSSION
IN INTRODUCTORY COLLEGE STATISTICS COURSES

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Robinson Panahatan Situmeang, B.Sc., M.Sc.

* * * * *

The Ohio State University
1976

Reading Committee:  
Jagdish S. Rustagi
F.Joe Crosswhite
Douglas A. Wolfe
Harold C. Trimble

Approved By

Harold C. Trimble
Advisor
Interdisciplinary Program
ACKNOWLEDGEMENTS

The author is indebted to Dr. Harold C. Trimble for his guidance and encouragement during the writing of this dissertation. He wishes to thank his professors in the Department of Statistics, Drs. Douglas A. Wolfe and Jagdish S. Rustagi in particular, for the help they have given him over the past four years, and for making the study at The Ohio State University a valuable and worthwhile experience. Thanks are also due to Dr. F. Joe Crosswhite for his valuable suggestions. Finally, the author wishes to thank his wife and son for their patience, understanding, and encouragement during the completion of his study.
VITA

October 19, 1931  Born - Pangururan, Indonesia

1957  B.Sc., BI/BII College, Bandung, Indonesia

1957 - 1961  High School Teacher in Bandung and Medan, Indonesia

1962  M.Sc., State University of New York, Albany, New York

1963 - 1970  Instructor, Udayana University, Bali, Indonesia
Assistant Dean for Academic Affairs, Medical School, Udayana University, 1964-1966 and 1966-1968

1970 - 1972  Graduate Teaching Associate, Department of Mathematics, Miami University, Oxford, Ohio

1972  M.Sc., Miami University, Oxford, Ohio

1973 - 1976  Graduate Teaching Associate, Department of Statistics, Ohio State University, Columbus, Ohio

FIELDS OF STUDY

Major Fields: Statistics - Mathematics Education

Studies in Sampling Theory. Professor Ramesh C. Srivastava

Studies in Nonparametric Statistics. Professor Douglas A. Wolfe

Studies in Experimental Design. Professor Jagbir Singh
LIST OF TABLES

Table 5.1 Number of Students in Statistics 125 Fall Quarter 1975 64
Table 5.2 Result of Preliminary Question Set 66
Table 5.3 Comparison of the Means of the Test Scores of the Two Groups for the Six Experiments 69
Table 5.4 Comparison of the Means of the Effect of Class Experiments and Lecture Discussions in Explaining the Concepts 1, 2, and 4 70
Table 5.5 Comparison of the Means of the Effect of Class Experiments and Lecture Discussions in Explaining the Concepts 3, 5, and 6 71
Table 6.1 Result of the Initial Tests Winter Quarter 1976 73
Table 6.2 Results of Preliminary Questions Set Winter Quarter 1976 74
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS.</td>
<td>ii</td>
</tr>
<tr>
<td>VITA.</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES.</td>
<td>iv</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Statement of the problem</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Development of statistics</td>
<td>1</td>
</tr>
<tr>
<td>1.3 Growth of application</td>
<td>3</td>
</tr>
<tr>
<td>1.4 Development of statistics education</td>
<td>4</td>
</tr>
<tr>
<td>2 REVIEW OF THE LITERATURE</td>
<td>11</td>
</tr>
<tr>
<td>2.1 Computer assisted instruction as a supplement to live teaching.</td>
<td>11</td>
</tr>
<tr>
<td>2.2 The use of teaching aids and class experiments.</td>
<td>16</td>
</tr>
<tr>
<td>2.3 General curriculum and pedagogical suggestions.</td>
<td>19</td>
</tr>
<tr>
<td>2.4 Improving student background through pre-college instruction.</td>
<td>25</td>
</tr>
<tr>
<td>2.5 Conclusion</td>
<td>32</td>
</tr>
<tr>
<td>3 REPORT OF THE SURVEY AT THE OHIO STATE UNIVERSITY</td>
<td>35</td>
</tr>
<tr>
<td>4 DESIGN OF THE EXPERIMENT</td>
<td>52</td>
</tr>
<tr>
<td>4.1 Preliminary questions</td>
<td>54</td>
</tr>
<tr>
<td>4.2 Design and analysis of class experiments</td>
<td>56</td>
</tr>
<tr>
<td>4.3 Computer programming</td>
<td>60</td>
</tr>
<tr>
<td>4.4 Student survey projects</td>
<td>61</td>
</tr>
<tr>
<td>5 REPORT ON FALL 1975 TEACHING OF STATISTICS</td>
<td>125</td>
</tr>
<tr>
<td>5.1 Results of preliminary questions technique</td>
<td>65</td>
</tr>
<tr>
<td>5.2 Results of the use of class experiments</td>
<td>67</td>
</tr>
<tr>
<td>5.3 Results of the use of computer programming</td>
<td>73</td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS (contd.)

6 REPORT ON WINTER AND SPRING 1976 TEACHING OF STATISTICS 125  
6.1 Results of the preliminary questions  
6.2 Results of the use of computer programming  
6.3 Results of the students surveys  
6.4 Results of Spring quarter 1976  

7 CONCLUSIONS AND RECOMMENDATIONS  
7.1 Conclusions from the investigation  
7.2 Suggestions for the teaching of introductory statistics courses  
7.3 A suggestion for the organization of introductory statistics courses  
7.4 Recommendations  

APPENDICES  
A List and outlines of courses in statistics and courses related to statistics offered at The Ohio State University in 1975-76  
B Complete course outline of Statistics 125  
C Preliminary questions  
D Class experiments and results  
E Initial tests and results  
F Computer programs for students in Statistics 125  
G Samples of student surveys  
H Comparison of final scores  

BIBLIOGRAPHY

vi
1.1 Statement of the problem

From its early beginning in the fourth century until the end of the nineteenth century statistics developed very slowly. Since then it has developed rapidly and has become so central to investigations in many fields of scholarship that a statistics requirement has become a standard portion of many college curricula. This has aroused great interest from many authors and wide concern of educators and statisticians for the teaching of statistics. One of the challenges in the teaching of statistics is to find methods to make the course more meaningful, intellectually worthwhile and stimulating to the students, and to induce them to see its relevance to their various fields of study. In this study several approaches to a partial solution of this general problem were attempted. More details will be given later in this introductory chapter.

1.2 Development of statistics

Walker (1958) identified the four waves of ideas in the development of statistical theories and methods.
The first period was marked by the publication of Galton's "Natural Inheritance" in 1889 and the publications of Karl Pearson starting in 1891. At this time the interest in statistical matters was restricted to the collection of public statistics such as population data, vital statistics and economic data, the adjustment of observations in astronomy and meteorology, and so on. Outside these areas the movement was hindered by a lack of interest and inadequate theory. There were a few theoretical developments during this period, such as the theories related to correlation coefficient, Chi-square, etc. Work was also done on the characterization of frequency distributions and the construction of tables needed by statisticians and biometricians. The second period began in 1915 when R.A. Fisher published his first paper on the exact distribution of correlation coefficient. This period was also marked by the use of small samples, sampling distributions, the foundation of the theories in testing hypotheses and the design of experiments, the invention of techniques of analysis of variance, the choice of best estimators, the extension of multivariate analysis, and the introduction of the concept of fiducial probability. The third period began about 1928 with the publication of joint papers by Neyman and Egon Pearson. Many methods and theories were invented during this period including the applications in quality control for industries. The concepts of nonparametric methods were started during this period.
The fourth period was characterized by the invention of sequential analysis and the foundation of the basic ideas of decision theory which had profound effects on statistical inference.

1.3 Growth of application

The rapid growth of application of statistics in many different fields has been made possible by the development of the theories, by the availability of persons having knowledge of the theories, and by increasing public awareness in this matter. Probability and statistics have become an integral part of a scientific society. Scientists in almost every field design experiments and analyze and interpret the results by using statistical methods. Modern statistics is used in the design of agricultural experiments to help to find ways of obtaining better results at less cost. Medical research uses statistics to find out whether a certain treatment, for example polio shots, reduces or prevents the incidence of certain diseases. Business and industry use statistical methods to check the quality of their products and to decide on the most economical ways to organize their operations and processes. Insurance companies make use of the laws of probability and statistics in the study of life expectancy. Engineers and production managers must know aspects of probability and statistics basic to operations research and, in particular, to quality
control. Psychologists must have a background in probability and statistics in order to design appropriate models for their experiments and to be able to interpret the results. Likewise, geneticists need to be acquainted with such probability and statistical concepts as regression analysis, correlation analysis, analysis of variance, and so on. Decision theory finds its application in many areas such as in military affairs in designing strategies, and in business. Statistical mechanics has been a standard course in the preparation of engineers and physicists. Astronomers apply the theory of probability to the statistical study of the distribution of star galaxies.

1.4 Development of statistics education

The increasing public awareness of the importance of statistics in almost every walk of life, and the growing need for persons with statistical knowhow create the contemporary demand for a sound statistics education. Yet students in many colleges take a statistics course only because it is a required course, without having any concept as to its relevance to their various fields of study and its importance in real life problems. Realizing those two opposing facts, statistics educators have expressed their concern for making statistics more than just a requirement to be endured by students, and for getting students to see the need for and the relevance of statistics in various fields.
To this end the Joint Committee on the Curriculum in Statistics and Probability of the American Statistical Association and the National Council of Teachers of Mathematics suggested that authors of textbooks in statistics and probability move strongly in the direction of real life problems and provide more explanatory text. In response to this, a series of statistics textbooks was published such as "Statistics By Examples" by Mosteller et.al. in 1972.

Mathematics educators voiced their opinions about the teaching of mathematics, including statistics. Professor Polya (1963) stated that the foremost aim of teaching is to teach the learner to think. The learners should be active, yet the learners will not act if they have no motive to act. They must be induced to act by stimuli, by the hope of some reward. The learners should be interested in the material to be learned and find pleasure in the activity of learning. Professor Halmos (1975) stated that the best way to teach mathematics is to arrange the teaching in such a way as to induce students to ask and to do. The best way to learn to solve problems is to solve the problems. Give the students problems they are interested in, let them think and ask how to solve, and let them do it.

As to the content of an introductory statistics course, many articles have been written pertaining to it. It would be worth mentioning that, normally, a typical
university would offer two kinds of introductory statistics courses; one designed for students without a calculus prerequisite, and the other for those with a calculus prerequisite. For students in the first category a course in introductory statistics might be a terminal course, although some of them might have to take continuing courses in statistics or in research methods. For those in the second category, this course would be a foundation course to be continued with other applied or theoretical statistics courses offered by the department of statistics or other departments. A typical introductory statistics course would be a one or two-quarter course offered to students at the sophomore or junior level. Students majoring in areas such as psychology, education, biology, forestry, sociology, social work, nursing, etc., would be advised to take the introductory statistics course without a calculus prerequisite. Those in business administration, economics, engineering, mathematics, statistics, etc., would take a course with a calculus prerequisite. Regardless of the fields of interest and of which introductory statistics course the students might have to take, an understanding and sound knowledge of the basic statistical concepts is necessary to permit the students to analyze data more effectively and to make everyday inferences and decisions more meaningful. At various times concern has been expressed to modify the content of such courses and many
studies have been conducted to find methods to make them more meaningful and challenging to the students. The major concern was for the introductory statistics course without calculus background since there has been considerable criticism about it. Some of the criticism said that the course was generally too technically oriented, contained too much probability and not sufficient material for statistics. Many introductory statistics courses have been criticized for providing little insight into the variety of the subject matter area in which statistics finds its applications, and for not illustrating the diversity of problems in which statistics plays a useful role. Some have been criticized for overemphasizing computations and not stressing sufficiently the concepts and basic ideas underlying statistics. On the other hand, many have been criticized for placing too much emphasis upon the basic concepts underlying statistics rather than teaching students how statistical techniques could be used to draw inferences from real data (see Committee on the Undergraduate Program in Mathematics 1972).

The investigator shares the views and opinions of the educators mentioned above in that the introductory statistics course should be designed to stimulate the students to learn more about statistics and to gain competence needed for the research in their major fields of interest. He also believes that success in an introductory
statistics course might even arouse some students' interest to pursue a program to become a statistician. The quality of instruction the students obtain in class and the methods by which the materials are presented have potential bearing on the interest of the students in the subject as well as on their perseverance which was defined by Carrol (1963) as the time the learner is willing to spend in active learning.

1.5 Plan of dissertation

In his attempt to explore ways to make an introductory statistics course more meaningful to the students and relevant to their various fields of study, the investigator assumed several objectives for the teaching of statistics: (1) to introduce to the students the variability and uncertainty of many kinds of measurements. (2) to introduce to the students some common statistical formulas, terminologies, and some widely used statistical methods. (3) to get the students to recognize the availability of many statistical techniques that they can use to help them to solve many problems in their future careers. Upon the achievement of the objectives the students could be expected to think statistically and to have developed the ability to make inferences and decisions from the very limited information obtained in a sample. They would also have developed an appreciation toward the role of statistics in
their own field as well as in other areas.

In his investigation the investigator was mainly concerned with the teaching of introductory statistics at the university level, although the findings might be found useful for advanced courses as well. The investigations which were conducted at The Ohio State University during the Spring and Fall quarters of 1975 and the Winter and Spring quarters of 1976 consisted of three stages:

(1) a careful study of literature for the purpose of reviewing the ideas of statistics educators and the work that has been done by previous investigators in the field of methodology. Report on this is presented in Chapter 2.

(2) a survey which was conducted at The Ohio State University in the form of interviews with persons teaching statistics in many departments. The purpose was to find out the objectives of statistics courses with respect to the departments offering the courses, and to learn about the methods employed. Report on this is presented in Chapter 3.

(3) experimentation for the purpose of evaluating four techniques in the teaching of introductory statistics such as the inclusion of a "preliminary questions set", class experiments, computer programming, and student surveys as supplements to the systematic lecture method. The subjects in this experimentation were students from a variety of schools and colleges at the Ohio State University who were
taking Statistics 125 which was an introductory course in mathematical statistics offered by the Department of Statistics. The experimentation was conducted during the Fall quarter of 1975 and the Winter and Spring quarters of 1976. The complete report of this experimentation is presented in Chapters 5 and 6. The interpretation of the findings followed by recommendations for further investigations are presented in Chapter 7.
Chapter 2

REVIEW OF THE LITERATURE

The search for ways to make an introductory college course in statistics seem relevant to the students led to explorations in several directions. For the purpose of reviewing the literature the investigator grouped these explorations into four categories: computer assisted instruction as a supplement to live teaching; use of teaching aids and class experiments; general curricular and pedagogical suggestions; and improving students' background through pre-college instruction. The elaboration of each of the categories is presented in the following sections of this chapter.

2.1 Computer assisted instruction as a supplement to live teaching

Genetics 650 or Analysis and Interpretation of Biological Data I is a one-quarter introductory statistics course offered by the Department of Genetics at The Ohio State University. This course provides credits for both graduates and undergraduates. It was originally introduced with one required two-hour laboratory period and four
required one-hour lecture periods per week in a quarter of ten or eleven weeks. Professor Skavaril (1974) modified the course by moving to computer based instruction. In this modified form the students can follow the course by the Computer Assisted Instruction (CAI) using one of the computer terminals available in many locations at the university. They can follow the course at their own pace and at a time which is convenient to them. They need only attend six half-hour meetings per quarter for the administration of laboratory quizzes and four one-hour lecture periods to obtain introduction and instructions to enable them to follow the course. The remaining lecture periods have been converted to discussion session at which attendance is optional; those who wish to have their questions answered by the instructor come to these sessions. This computer based instruction is a combination of the technology of CAI, computer generation of exercises and answers, and the use of computer by students for the analysis of problems. The Computer Assisted Instruction has 29 modules written in the Coursewriter III, version 3, and author language. The author also wrote nine exercise generating programs in PL/I and FORTRAN IV which produce computer generated exercises used by the students in connection with the laboratory portion of the course. The programs also produce correct answers for the various exercises. In addition to the above mentioned programs,
Professor Skavaril wrote 21 analysis programs in Conversational Programming System or CPS language. These programs are self-instructive in that each will produce a description of the program, its limitations, and the nature of the input data if it is desired. He believes that this system has advantages for students and the instructor: it provides convenience of choice of time for the students and the pace of the course which is suitable for a student's background, and it has the flexibility for students to obtain supplementary materials if they feel they need them in order to proceed in the course. Professor Skavaril reported that the time required by the students to complete the prescribed amount of materials is much less than the time needed using the conventional lecture method. He pointed out that the CPS analysis program is especially advantageous because it enables the students to solve problems using realistic data rather than trivial or hypothetical data and small sample sizes. He further commented that even graduate students use this program to analyze their own research data. He also noted that there have been many attempts to use computers as aids to instruction during the last ten years. Most of them used either CAI, or exercise generating programs for laboratory work, or analysis of data, but not the combination of the three.

Johnson (1965) introduced and discussed the Michigan Algorithm Decoder or MAD program which can
generate normally distributed numbers.

Tarter (1967) analyzed the use of computer programmed instruction in statistics. The advantages cited were: the faceless programmed learning that permits students to study unobserved, the reduction of a professor's load, and its use to supplement and train teaching assistants.

Schatzoff (1968) discussed the use of time-shared computers in a statistics curriculum. He discussed the use of four computing systems such as Compatible Time-Sharing System, Culler on Line Computer System, Console Orient Model Building as a mean to "comb" a set of data and as a teaching aid, and Console Oriented Statistical Matrix Operator System as a general purpose for statistical computations. He also suggested that computer training is mandatory for every statistics student.

Kitchen (1968) discussed the teaching of management statistics by programmed instruction in a course offered in the Training and Education Section of the British European Airways. A great deal of interest and effort were generated among the students by this type of presentation.

Foster and Smith (1969) stated that in the London School of Economics an experiment has been started to use a computer system as an aid in teaching introductory statistics. They reported that it showed advantages both
pedagogically and computationally.

Carmer and Cady (1969) developed the explicit use of a group of exercise-generating program in the teaching of statistics. They pointed out that these programs conserved the instructor's time spent on searching for and preparing problems as well as provided him with the correct final calculations. Eight programs were developed and used in statistical methods courses for non-statistics majors in classes with large enrollment of both undergraduate and graduate students.

Dixon (1971) stated that computer utilization is expensive but it helps the students to understand the materials better and faster. Based on the Dixon and Massey text, Dr. Allan Forsythe prepared a question and answer type of instruction for elementary statistics. Dixon further suggested that an interactive teaching program should be prepared by professional programmer and the instructor.

Quade (1971) discussed the use of the Conversational Mode in the teaching of several statistics courses through teletypes using Basic, FORTRAN, and PL/I languages. Several courses had been offered using this program and the result was very successful.

Wagner and Motazed (1972) used "Proctorial System of Instruction" (P.S.I.) utilizing computer terminals. There are four main features of this system. i) go-at-your-own pace
which permits the students to move through the course at a speed commensurate with their ability. ii) the unit-perfection requirement for advancement, which lets the students proceed to a new material only after they demonstrate their mastery of previous material. iii) the use of lectures and demonstrations only as a vehicle to motivation, rather than the source of information, and iv) the use of proctors, which permits repeated testing and immediate grading.

2.2 The use of teaching aids and class experiments

There have been many attempts to increase students' interest in studying introductory statistics. Investigators have tried numerous means to make the approach to and the discussion in the course as close as possible to the students' environment and interest. In trying to achieve this objective the investigators have explored the use of teaching aids and class experiments.

Rubin (1963) successfully conducted simple experiments in his introductory statistics class. In so doing he selected experiments that were familiar to the students and were as simple as possible in order that each student could perform them in a short period of time. He insisted that experiments must be performed without any additional cost. Thus he made familiarity, simplicity, and zero-cost the three requirements for his experiments. He concluded that students' participation in providing the
basic materials, collecting data, performing computations, analyzing the results and reporting them as research information could enliven the interest of the students in the study of elementary statistics.

Jowett and Davis (1960) applied and discussed the advantages of teaching an introductory statistics course by using experiments. They concluded that experiments, besides arousing the interest of the students, also gave the bright students an opportunity to display the superiority of their skills while allowing the less bright students to follow at a good pace.

Spitnagel (1968) conducted an experimental approach as a supplement in the teaching of probability. Each student performed two experiments both involving large numbers of trials. He found this easy to do and pedagogically effective.

Davis (1970) reported the role of practical experimentation in the teaching of probability and statistics. The purpose was to obtain the data for analysis and to work with concepts in realistic ways. The students also had the opportunity to design experiments for themselves.

Holmes (1971) heightened the interest of the students by classroom activities. He conducted an analysis of two-sample problems using Monte Carlo methods with the data collected in the classroom. One of the problems was to compare boys' and girls' successes on a test. The students
understood the method easily.

Greenberg (1964) developed the use of field training for biostatistics students. He asserted that field training has several advantages: it provides an opportunity to apply the theory, it helps students to understand the purposes of health agencies, it encourages improved skills in consulting, it opens up opportunities for becoming familiar with vital events, and it gives students familiarity with the collection and analysis of data.

Butterbaugh (1959) suggested that the American Statistical Association begin to look into the field of audio-visual education and training through films. He said that not many films have been produced that can be used to teach statistics.

Austwick and Wetherill (1971) stated that the objective of statistics as a service course is not to teach the technique of computations but to improve the understanding of the underlying phenomena that give rise to the problem such as sampling variability. For this purpose, teaching aids using programmed learning in statistics, closed circuit television material, filmstrips, film loops, slides and video-tapes were found to be very helpful.
2.3 General curriculum and pedagogical suggestions

The panel on Statistics of the Committee on the Undergraduate Program in Mathematics (1972) reported the considerations and suggestions of the committee on the general introductory course in statistics. The emphasis was on the basic concepts of statistics and not on the mathematics of statistics. The primary objective of the course was to introduce the students to variability and uncertainty and to give them illustrations of inferential reasoning from observed data. Much of the report considered the traditional systematic introductory course with a view to emphasizing and utilizing a variety of approaches to enhance interest and motivation. The alternative approaches suggested were the decision theory approach, the nonparametric approach, and problem solving approach.

The systematic approach is the most widely used, perhaps because of its simplicity and the existence of a wide variety of textbooks oriented toward this approach. It is therefore the most likely to be used in small colleges and universities even though it has potential defects. Typically the major ideas of statistical inference are introduced too late and in haste, and are often illustrated by examples which are not compelling to a majority of the students. Frequently, courses of this type suffer from the inclusion of irrelevant concepts and excessive mathematical derivations. Of course, these defects
can be reduced by careful preparation and selection of materials to be covered that are relevant to the interest of the majority of the group.

In a decision theory approach one formulates problems in which a choice of actions whose consequences depend on the state of nature has been made available to him. To help decide on an appropriate action, one must perform an experiment which will yield relevant data to help determine the state of nature. One major advantage of this over the traditional approach lies in the problem solving orientation. This approach is usually appealing to the students who are interested in the philosophical foundation of scientific inference and who are curious about the rationale for coping with random variation and uncertainty. These students find this approach natural and easy to follow. It is pleasing to the students who like mathematics and enjoy applying the basic ideas of statistical inference. The limitation of this approach is that the students get minimal acquaintance with actual statistical practice in data analysis.

A nonparametric approach is gaining adherence of more persons interested in the teaching of introductory statistics. Perhaps the main reason is that it permits one to introduce substantial problems in inference very early in the course, and to provide impressive solutions that are easy for teachers to realize and for students to comprehend.
The methods of verification are easily explained and the students can be led to accept them as natural and sensible justifications. The approach also has the advantage of robustness in testing hypotheses. The students can be introduced quickly to useful and simple techniques which have wide applicability. Exposure to statistical concepts throughout the course provides the students with a foundation for a good understanding of standard parametric procedures. The major disadvantage of this approach is felt to be that the students have little exposure to those familiar parametric methods which are widely used for later work. Another disadvantage is that there are only a few introductory textbooks which present nonparametric procedures in a simple way.

A problem-oriented approach in an introductory statistics course is a "case study" approach. It is said that this approach demands a more mature level of the students in order to solve the problems. The advantages of this approach are given as: (1) it provides the discussion of meaningful and nontrivial problems that evokes student interest and raises fundamental statistical questions. (ii) it provides the answers to realistic questions resulting in the reinforcement of understanding of key concepts. (iii) it provides the opportunity to use computer simulation which is valuable and which may help the students in continuing their investigations. The disadvantages given are: (i) this approach involves the use of methods and ideas
which have not been carefully digested in advance; there will be some need to operate at levels of less than complete understanding. (ii) this approach presents life problems which are often very complicated; it will be difficult but necessary to simplify data without throwing away too much of the essence of the information. (iii) in this approach, concepts and methods arise in context and not in a systematic framework for assimilation by the students without undue repetitions. (iv) this approach deals with problems in life that seem easy in retrospect but may require a good deal more maturity from the students than teachers are inclined to realize.

There have been attempts to investigate the best approach to improve the teaching of introductory statistics. Hoffman (1961) commented on the problem of teaching statistics to medical students. He suggested that the method should involve reading assignments, and the theoretical part be presented in lectures. The homework should be arithmetically very simple, but should lead to solving real problems and preparing a report.

Hoffman (1962) discussed several commonly used methods of teaching elementary statistics. He suggested that one good method to teach statistics is to discontinue the formal lectures. The students should determine their own pace, and choose the part of a subject in which they are interested. This may be just the theoretical part, or
just the applied sections, or both. He considered it most
important for students to devise, study and report upon the
solutions for simple problems. The instructor is needed only
to answer questions; attendance is optional.

Blom and Råde (1969) explored the need of probabi-
licity and statistics for students in engineering at the
iversity level. They felt that such students should be
familiar with how to deal with uncertainty and variability,
how to make inferences, how to choose a course of action,
and how to control production. They suggested the presenta-
tion of problems from engineering as illustrations of the
ories and methods.

Rao (1969) advocated teaching statistics not as
a separate subject but through an interdisciplinary approach
combined with economics, biology, mathematics, etc. and
supplemented by laboratory field work.

Holm (1970) listed a collection of interesting
problems in probability and statistics to amke a course
more appealing to students. He listed 18 problems and
examples in probability such as games, lottery, who baked
the cake and so on, and 13 problems in statistics such as
problems in ecology, traffic, quality control, and so on.

Noether (1969) suggested the nonparametric
approach to teaching elementary statistics. He found the
subject easier to teach and the students found the course
more meaningful; for example, students did not need to
handle complicated formulas like those for a two sample t-test. The students understood the logic of statistical procedures more naturally and used the table easily.

Mosteller (1971) urged that the purpose of teaching statistics is not learning to do statistics but rather to see how statistics or probability contributes to society or science, how quality control charts make ball bearing better, how sampling procedures could help us to find a more accurate estimates, and so on.

Urquhart (1971) discussed the use of a "Nonverbal Instructional Approach" for introductory statistics. Computers, models, pictures, puzzles, games, laboratory exercises, field trips, and colorful examples are excellent ways to stimulate mathematical concerns. He emphasized the use of an indirect approach ("nonverbal"). The main consideration should always be "how the problems should be approached" rather than "what is the numerical answer".

Hamaker (1971) considered the teaching of statistics as a difficult and complex problem in which one has to be concerned with the audience, the field of application, the selection of materials, the teaching techniques, and the terminologies to be used. He noted that the need for some knowledge of statistics is universally recognized. Therefore he suggested that elementary statistics be taken by everyone at the college level. The main problem, as he mentioned, is the amount of material to be presented and
the method of teaching to be employed.

Hogg (1972) stated that a common agreement in the method of teaching statistics is that the students become involved with the collection of real data. He suggested that the teachers use a variety of sources of problems and projects that can be done in the classroom where students' participation plays an important role.

Hemelrijk (1971) recommended that a television course in statistics be provided for students, users of statistics, and professional statisticians. He suggested that there should be more interesting examples, better prepared experiments, and that the discussion on probability be minimal.

2.4 Improving student background through pre-college instruction

In agreement with the several committees and individual investigators cited later in this section, this investigator believes that an early introduction of probability and statistics to high school students is very important in terms of its relevance to the numerous activities in a modern society. The standpoint people take on almost every argument based on numbers and uncertain information will have to be statistically valid. This writer also believes that probability and statistics instruction in secondary schools can have a highly stimulating effect on
the study of mathematics since it presents to the students ample opportunities to discover interesting applications of mathematics in real life situations. He also believes that the teaching of probability and statistics in high schools is relatively simple as compared to some other branches of mathematics. Many statistical concepts are derived from cases that the students are familiar with; hence they can be comprehended easily, logically, and often intuitively.

David A. Page (1959) wrote about the need and the possibility of teaching probability to students in pre-college level. He said that many of the notions that are important in probability theory can be grasped in a clear and intuitive way by the children. These can be taught to them with at least three purposes: (i) to familiarize the students with some of the important and fundamental ideas of probability and to expose them to the real and imagined experiments which should precede a thorough and rigorous study of the subject. (ii) to provide an interesting context for standard mathematical ideas and theories in order to deepen the students' understanding of and interest in these ideas. (iii) to give the students and teachers a new and refreshing method for, in essence, drilling on mathematical ideas which are already taught as a part of the standard curriculum in mathematics, for example, solving verbal problems.
Pieters (1959) also discussed the need of statistics for students at pre-college level. In this discussion he stated that the students can become familiar with methods of solving problems which are involved in making decisions in the face of uncertainties due to incomplete information. This includes determining what data are pertinent to the problems, what type of sampling is the most suitable, how to collect data as needed, how to organize, calculate, interpret and finally report the result. He suggested that the idea of histogram should be given not later than the 7th or the 8th grade of junior high school. He said many statistical concepts can be introduced to students in high schools without any difficulties; this, in turn, will help them reduce their difficulties in mathematics. Students who have difficulties or have little ability to utilize and interpret symbols and indices correctly will find statistics very helpful; thus statistics is one of the subjects which richly utilizes symbolism, especially capital letters, subscripts, primes, Greek letters, summation notation, and many others. He pointed out that approximation is a very important concept in our life, since all measurements are approximate. He also suggested that mensurational formulas involving addition, subtraction, multiplication, division, and square roots be emphasized. Discreteness and continuity which are essential in statistics can be introduced with relative ease to children when they are working with graphs,
plots, dots, lines, curves, and graphical estimations. Permutations and combinations, which are very interesting concepts in mathematics as well as in probability and statistics, can be introduced in the 11th and 12th grades. At those levels the students will have ample time to work on these concepts as applied to many interesting problems in their own lives. He strongly believes that statistics provides a variety of methods which guide the students in making better estimates, better judgements, and more accurate conclusions.

Morris (1968) pointed out that statistics could be used as a remedy for students who had a pronounced dislike for mathematics and who were often classified as failures. His method was to give those students a modified mathematics instruction in which the students provided the class with data related to themselves, such as how many children in the family. The findings from the students' investigations served as excellent topics for class discussion, and the students showed a marked interest in what they were doing. Unknowingly, he said, the students learned a number of mathematical concepts in the process.

Rade (1965) reported on the opening of a new gymnasium in Sweden in 1966 in which the curriculum branched out into four lines of study; humanities and social science, economics, science, and technology. Many of the courses were the same for all the different lines and probability and
statistics was one of them. Descriptive statistics was given in grade ten, and the main part of a course in probability and statistics, including probability and its applications, estimation and decision theory, were given in twelfth grade. He also pointed out that the two basic reasons for teaching probability and statistics in the new gymnasium were: (i) The findings obtained from a survey, conducted on the students graduated from the old gymnasium, showed an overwhelming desire of students to have a basic knowledge in probability and statistics upon graduation, and (ii) The demands from universities and other institutions required students to be familiar with reasoning using probability and statistics.

Sparrow (1964) reported his success in teaching statistics in high schools by having the students do a survey to find some correlation between the amount of liquid intake per day and its effect on personal characteristics, such as catching colds and influenza, among teenagers. He stated that in doing the survey the students learned to collect data from their friends, summarize the data, interpret them, draw some useful conclusions, and to report the conclusions in an interesting manner.

O'Toole (1966) who visited schools in thirteen countries in Western Europe found that many of those countries have started to teach probability and statistics to their high school students. The main difficulty, he
mentioned, was the lack of teachers and textbooks in that area.

Hume (1970) stated that high school students in Western Australia have learned elementary statistics well. Their knowledge of simple concepts such as hypothesis testing, Chi-square test, and goodness of fit test provides them with good background for their study at the university level as well as in their own life situation.


Kruskal (1970), at the meeting of the Joint Committee of the American Statistical Association and the NCTM, discussed numerous examples in the area of probability and statistics that could be introduced to students at a pre-college level.

Douglas (1970) reported that he has successfully taught descriptive statistics and probability in many pre-college schools in Australia.

Beckenbach (1970) wrote the historical background of the teaching of statistics and probability at pre-college level. He strongly advocated that statistics and probability should be introduced in the K-12 program with emphasis on laying a sound foundation for the subject including the development of intuitive feelings and the technical skills.

Bell (1970) taught probability and statistics to sixth grade and ninth grade students with great success. He commented that nonparametric procedures were relatively simple to teach as compared to parametric procedures.

Since 1959, on many occasions, educators, mathematicians, statisticians have met for international and national conferences to discuss the teaching of mathematics and came up with strong recommendations that probability and statistics be included in high school mathematics curriculum. They believed that probability and statistics provide ways to study, understand, and to control uncertainties. Many of the newer applications of mathematics, they agreed, use the theory of probability and statistical reasoning, and the increase in the use of probabilistic descriptions of phenomena in modern science is phenomenal.

The Commision on Mathematics (1959) expressed its strong desire for the introduction and the inclusion of statistics in the high school curriculum. It pointed out the more and more important role statistics plays in a
society of educated men and women where it serves as a supplement to deductive thinking.

The Report of the Royamount Seminar (1959) suggested that elementary probability and statistics be recognized as an integral part of mathematics for secondary schools. Statistical inference was viewed as applied mathematics that contributes, in an essential way, to the decision processes in the spirit of the "scientific method" basic to many fields in both physical and behavioral sciences, and as important knowledge in the field of public affairs. It further suggested that suitable preparatory courses for teachers of these subjects be introduced in Normal Schools and other Teacher Training Institutions.

In the report of the Cambridge Conference on School Mathematics (1963) it was suggested that probability and statistics be taught in four stages, in the elementary school, in the junior high school, in the senior high school after the first work on limits and series, and in the senior high school after integral calculus has been discussed. Suggestions were made as to the outlines of the course for each stage.

2.5 Conclusion

The use of computer programs as a substitute for a good teacher in the teaching of statistics is increasing and developing very rapidly. This investigator believes
that this method is good in that it attempts to pace the course according to the students' personal background, and it enables the students to focus their time and effort on the theories and methods rather than on mathematical computations. This method makes it possible for the students to solve many problems easily and accurately. However, this method is known to be expensive and its use is feasible only in universities where computer facilities are available. Another limitation, in the view of this investigator, is that it emphasizes explanations through reading, and thus omits the audio-explanations that many good teachers are able to give in very pleasant and appealing ways.

The investigator agrees that the use of teaching aids and class experiments in the teaching of statistics is indeed very helpful. The employment of this method requires additional preparations from the instructor for the presentation; this is especially true because, ideally, activities should be varied from quarter to quarter according to the mode of interest in the class.

It seems that the three alternative approaches, the decision theory, the nonparametric, and the problem solving approach have advantages in common in that they emphasize students' involvement, and they expose students to realistic problems. Since the decision theory approach and the problem solving approach require a good deal more maturity from the students, it is less likely that those approaches
will be employed in the teaching of introductory statistics course. The investigator feels that a systematic lecture method with a nonparametric approach is suitable in the teaching of introductory statistics.

In regard to the improvement of students' backgrounds through pre-college instruction, this investigator believes that this is probably the ideal long range approach. This ideal situation, where students have had pre-college statistics instruction, is not existent at the university during this investigation and may not be for many years to come. Therefore during the investigation he will deal with the students' backgrounds as they are currently found at the university.

It appears that all categories have one basic problem in common which roots from the varied interests of the students taking the course. In an attempt to find some resolution to the problem, the investigator conducted interviews with professors from several departments at The Ohio State University who teach statistics. The report of the discussions with those professors is presented in the following chapter.
In searching for methods to make statistics more meaningful and relevant to the students, the investigator had the opportunity to have discussions with many professors teaching statistics in various departments at The Ohio State University. The following is a summary of the discussions pertaining to the curriculum and the methodology employed in the respective departments.

From the variety of names given to the statistics courses offered by these departments, from the methods of teaching employed, and from the assortment of textbooks recommended, it was obvious that these departments have tried to make statistics more meaningful to the students and more relevant to their fields of study. It was felt that the data collected from these departments might provide guidelines for organizing and teaching an introductory statistics course, such as Statistics 125, for students from two or three departments. To this end, each person interviewed was asked about the objectives and methodology for the statistics course or courses in his field. The outlines of the courses are presented in Appendix A, pages
3.1 Statistics in the Department of Genetics

The main objective stated for the teaching of statistics in the Department of Genetics was to familiarize the students with the basic elements of statistics and their applications in many areas such as Agriculture, Home Economics, and Biological Sciences.

Genetics 650, Analysis and Interpretation of Biological Data I, as was explained in Chapter 2, is a computer based course. It is designed to discuss statistical concepts such as descriptive statistics, normal distribution, Central Limit Theorem, testing hypotheses, t-test, Chi-square test, estimation theory, F-test, one-way and two-way analysis of variance, linear regression, correlation, multiple comparisons, and so on. This basic statistical knowledge is needed in Genetics 651 which is the application of statistics in experimental design; it includes latin-square design, split plot, factorial design, analysis of covariance, and multiple regression analysis.

3.2 Statistics in the Department of Animal Science, Dairy Science, and Poultry Science

These departments do not offer any special courses in statistics, but the students are asked to take introductory work in statistics such as that given in Genetics 650
and 651. These courses provide them with a good foundation of statistical knowledge for the discussion of its application to their own fields, for example to gene frequency, population statistics, mating systems, breeding plans, and so on. The students are expected to familiarize themselves with the variance-covariance models, simple correlation, regression analysis, multiple regression including curvilinear regression analysis, and some experimental design.

3.3 Statistics in the Department of Agronomy

The objective of the teaching of statistics in this department is to familiarize the students with statistical procedures which will enable them to design appropriate experiments, analyze the data, interpret the results in relation to agronomical problems such as field-crop analysis and soil management. For this purpose the students usually take Genetics 650 and 651. It is further recommended that they take a continuation course, Agronomy 887, Techniques of Experimental Agronomy. This is a lecture course followed by discussions with the students concerning the application of experimental designs, such as split plot design, incomplete block design, factorial design, confounding, and so on, to problems in agronomy. The emphasis is on qualitative discussions of the advantages and disadvantages of several designs, their utilization and relative efficiency, and the analysis and statistical interpretation of data.
3.4 Statistics in the Department of Agricultural Economics

It is recommended that students, mostly graduates, take a course in elementary Economics Statistics, and Applied Regression and Correlation Analysis. These courses are expected to provide the students with a good foundation for participation in the discussions in a Seminar in Problems in Agricultural Economics, Ag.Econ.801.

3.5 Statistics in the Department of Preventive Medicine

In this department, the objective of the teaching of statistics is to provide the students with sufficient background in statistical procedures to enable them to read and understand scientific articles, design experiments, and interpret the data obtained from their own investigations.

There are two statistics courses provided by the department: Design of Biomedical Investigations, and Biostatistics and Computers in Medical Research. Statistics and computer programming are presented concurrently in an integrated manner; one serves as an example of the other. The basic statistical concepts include descriptive statistics, probability distribution, sampling methods, estimation, testing hypotheses, decision theory, ANOVA, some experimental design, and simple and multiple regression analysis. By giving and discussing the necessary proofs of the theorems, it is hoped that the students will understand the concepts better. Multiple regression analysis plays an important role
in Preventive Medicine since the problems with which the students deal involve many variables.

3.6 Statistics in the Department of Veterinary Physiology and Pharmacology

The main objective is to study the theory and application of statistical concepts to veterinary medical research. The department offers a two-quarter statistics course, Design and Analysis of Comparative Biomedical Research I and II, which includes the basic elements of statistics, regression analysis, analysis of variance. Since the course is usually offered to graduate students, it offers them the opportunity to bring their own problems and data and to discuss them using statistical considerations.

3.7 Statistics in the Political Science Department

The objective of the teaching of statistics in this department is to provide the students with a conceptual understanding of statistical methods. It is mainly to introduce the students to some methods in the interpretation and the application of quantitative analysis of political data, and descriptive and inferential statistics with emphasis on multivariate analysis relevant to the data obtained in political science.

The department offers two statistics courses; Elementary, and Intermediate Methods of Quantitative
Analysis. Both are prerequisites for the higher level courses, Causal Analysis and Dimensional Analysis. In these courses the students select projects they are interested in, collect the data with one dependent and many independent variables, analyze and interpret them, and report the results. To this end, occasionally the instructor provides special sessions for discussion of problems in the analysis of the data.

3.8 Statistics in the Department of Sociology

The main objective of teaching statistics in this department is to introduce the students to the variety of statistical techniques that can be utilized in sociological research. This research pertains to the description of social statistical information related to birth, health and death, marriage, divorce and the family, education, labor force, social security and welfare, delinquency and crime, housing and construction, outdoor recreation, elections, public opinion polls, and so on.

The department offers seven courses in introductory statistics and the applications of statistics to sociological problems and research methods. The students are expected to have a minimum ability to use statistical techniques to interpret data obtained in a sociological survey. The mathematical derivations of the theories and formulas are seldom discussed in class; rather the students are assigned many problems. In this way it is felt that the students will
have sufficient practice in the use of formulas and theories, and in gaining better understanding of the concepts. The main topics discussed are sampling procedures in sociological research, construction of questionnaires, interview techniques, analysis of research designs, the assessment of error variance in sociological research, the exposition of census data and vital statistics, demographic rates, and life tables; seminars related to research methodology are offered. Also included in the courses are correlation, regression, analysis of variance, factor analysis, and nonparametric procedures.

3.9 Statistics in the Department of Psychology

The main objective is to familiarize the students with existing statistical techniques and to provide them with an understanding of basic statistics to enable them to read professional articles intelligently. It is also intended to develop in the students the ability to use some of the common techniques in designing their own experiments, collecting the data that they need, and interpreting the results. It is strongly recommended that the students in the Psychology Department, especially those who intend to do graduate work, take an introductory course in statistics. This introductory course can be taken in the Department of Psychology, Psy 220 and 221, or in another department. It is expected that the undergraduate students be familiar with
such basics as elementary probability, descriptive statistics, theory of estimation, and testing hypothesis by either parametric or nonparametric procedures. Generally speaking, only the simplest derivations of the formulas are discussed in class. The more complicated concepts are usually explained intuitively, supplemented by many examples and assigned homework problems. The graduate students are usually advised to take some graduate courses in applied statistics like Psychology 825 and 826, Statistics in Psychology I and II. They are also encouraged to participate in discussions in seminars related to the application of statistics to psychology, especially the topics concerning the analysis of variance, experimental design, factor analysis, and non-parametric procedures. It is expected that the students, especially those who are interested in quantitative or experimental fields of study, take additional courses offered by the Department of Psychology to become conversant with such advanced concepts as simple and multiple regression analysis, test selection, item analysis, factor analysis, and canonical correlation. Graduate students are usually assigned to a few large projects that utilize statistical techniques and computer programs. The projects can be done by individuals or by group of students.
3.10 Statistics in the Department of Education

It is suggested that the students in the Department of Education familiarize themselves with basic statistics. The students may take courses in the Department of Psychology (Psy 220 and 221), or the Department of Statistics (Stat 125 or 123), or the Faculty of Educational Development in the College of Education (Ed Dev 785 and Ed Dev 786). The first part, Ed Dev 785, is titled Introduction to Inquiry, a survey course designed to acquaint master's level students with the basic principles relating to the planning, execution, and reporting of results of experimental research in education. Special emphasis is directed toward (i) the identification and delimitation of the research problem, (ii) the logical design of experimental studies, (iii) the principles of measurement, and (iv) the design of descriptive research. The quantitative part of this course, Ed Dev 786, is designed to familiarize the students with treatment of quantitative data, methods of calculation and use of formulas relevant to educational settings. Generally speaking, graduate students, especially those who are involved in science or mathematics, are encouraged to take two quarters of Experimental Design (Ed Dev 808 and Ed Dev 809). The objectives are to (i) enable the students to design and analyze factorial experiments, (ii) acquaint the students with the more advanced topics in experimental design such as hierarchial design and analysis of covariance, (iii) enhance students' ability to read
research literature. In addition to those courses, the students are also urged to participate in discussion of the applications of statistics to educational problems through seminars.

3.11 Statistics in the Department of Agricultural Education

The students in Agricultural Education, mostly graduates, are advised to take a basic statistics course such as the Introduction to Inquiry offered by the Faculty of Educational Development or other comparable introductory statistics courses. This is thought to be a good foundation for them to continue their study in research methodology courses provided by the Department of Agricultural Education (Ag Ed 885, 886, and 887); these courses are Research Design and Analysis and Interpretation of Data. The students are expected to be familiar with a variety of experimental designs, correlation analysis, sampling theory, the application and interpretation of descriptive and inferential statistics, estimation theory, testing hypothesis, simple and multiple correlation analysis, and the adaptation of a variety of computer programs to their problems. The students are also expected to be familiar with both parametric and nonparametric procedures.
3.12 Statistics in the Department of Business Administration

The students in the Department of Business Administration are usually required to take an introductory course in statistics from the Department of Statistics (Stat 123). This course includes the basic probability concepts, probability distributions for discrete and continuous variables, bivariate distributions, estimation theory, testing hypotheses, decision theory, and some basic knowledge in simple regression and correlation analysis. The theorems and formulas are usually explained intuitively. The students are advised to take continuation courses Decision Theory I and II (Bus Adm 390 and 490) which lead them to use statistical techniques in the managerial decision making process and to study sampling procedures, design and analysis of experiments. In this way it is believed that the students will have a better judgement in planning, organizing, and controlling the output of an industry. The graduate students are advised to take more advanced Business Statistics courses which are mainly related to the application of statistics to problems in business such as price and production indices, time series, and correlation analysis. The students are also encouraged to use computer programs to help them calculate the necessary statistics.
3.13 Statistics in the Department of Economics

The undergraduate students are required to take the course Elementary Economic Statistics, Econ 442. The objective is to familiarize them with the role of statistics in economic problems. This course contains basic statistics such as descriptive statistics, probability and application, probability distributions, estimation theory, testing hypotheses, bivariate distributions, and decision analysis; it is an applied statistics course. The derivation of formulas are limited to simple cases, but the application of more advanced techniques to economic problems is strongly emphasized. The students are then advised to take several additional courses in probability and statistical decision theory, applied regression, correlation analysis, and econometrics.

3.14 Statistics for students in Geodetic Science

Usually the students are advised to take Mathematical Statistics courses, Stat 520 and 521, and Regression Analysis, Stat 645. These courses will provide them with a sufficient foundation for the discussion of statistical problems in geodetic science such as in Adjustment Computation I and II. The advanced adjustment computations
require correlational analysis, multivariate statistical analysis, generalized matrices, and many other statistical procedures. Numerous graduate students in geodetic science take several more courses offered by the Department of Statistics to obtain additional knowledge related to their field of study.

3.15 Statistics in the Department of Civil Engineering

There is an elementary statistics course offered in this department which is called Observational Analysis. This is a basic course that introduces the students to a variety of probability distributions, some goodness of fit tests, estimation theory, and hypothesis testing. The objective of the course is to train the students to use statistical techniques in engineering problems. In addition, the students are advised to take mathematical statistics, either the Stat 425, 426 series or the Stat 520, 521 series.

3.16 Statistics in the Department of Industrial and Systems Engineering

Most students are required to take the basic probability and statistics courses, Stat 425 and 426; the students who are working toward a Ph.D. are advised to take the more rigorous courses, Stat 520 and 521. These courses are planned to provide students with sufficient basic
understanding of probability and statistics for use in their future study. There are many courses offered by the Department of Industrial and Systems Engineering which contain 10%, 20%, 50% or more statistics and probability. The students are required or expected to know such probability and statistics concepts as estimation theory, experimental design, forecasting and estimation techniques, multivariate analysis, decision theory, queing theory, stochastic processes, quasi experiments, and the use of statistics in Operation Research.

3.17 Statistics in the Department of Speech Communication

The graduate students are expected to know some elementary statistical methods as a basis for the discussion of research methodologies in speech communication. They are expected to know a variety of statistical concepts and the logic and design of experiments.

3.18 Statistics in the Department of Physical Education

The graduate students are expected to have some familiarity with the basic elements of statistics. This department provides a course in statistics, Statistics for Physical Education and Health Education. This is an applied course in which the techniques for interpreting research publications in the field of physical education are discussed. It is also intended to equip the students with
sufficient knowledge of those techniques for their own research.

3.19 Statistics in the Department of Social Work

The undergraduate students are usually advised to take one course in introductory statistics, Stat 125, and one applied course in the Department of Social Work (Soc Work 380), the Introduction to Research Methods in Social Work. The graduate students are usually expected to take Research Methods in Social Work (Soc Work 680), a course that utilizes some statistical techniques.

3.20 Statistics in the Division of Biostatistics

The students in biostatistics are required to take many mathematical statistics and probability courses as applied to biological sciences such as Stochastic Processes in the Biological Sciences, Population Dynamics, Bioassay, and Laboratory and Computer applications in the Biomedical Sciences. The basic philosophy of the training in Biostatistics is to provide trained personnel not only for the academic profession, but also for industries and the government.

3.21 Statistics in the Department of Statistics

The general objectives and philosophy of the Department of Statistics are: (i) to provide courses for
students in many areas, ranging from freshmen to graduate students from almost every field of study, (ii) to provide courses primarily for students working toward a degree in statistics, (iii) to do research in statistics which can range from applications in business or biology to the more abstract formulation of problems of inference and decision theory. In addition, the Department of Statistics also provides a statistics laboratory for consultation and data analysis service to faculty and student researchers from a variety of areas.

3.22 Conclusions

The most common objective of the teaching of statistics stated by many departments cited in this chapter is to familiarize the students with statistical procedures as needed. The most common method of teaching statistics employed is the systematic lecture method. In some cases, such as in advanced classes, lectures are supplemented by the use of statistical computer packages to help the students in mathematical computations. In view of the objectives, the presentation of the materials is often kept mathematically simple, i.e. the derivations of theorems and formulas are kept to a minimum. The theorems which involve complex derivations are usually explained intuitively or illustrated by examples. In regard to the selection of materials and examples for the course, it is obvious
that each department tries to relate it as closely as possible to the interests of the students.

The result of the survey and the course outlines (see Appendix A, pages 101-24) reveal that there are many elements of statistics which are of common interest to many departments. This raises a question as to whether it would be more advantageous to the students and the University if the introductory statistics courses were offered in cluster forms as suggested in Chapter 7 of this dissertation.

The survey also made the investigator even more aware of the need to search for a method to teach introductory statistics to students from various fields of study in such a way as to make the course more meaningful and relevant to the students. The design of the experimental part of this study is presented in the chapter that follows. It describes how several techniques could be used to supplement the systematic lecture method in the teaching of Statistics 125, Elementary Mathematical Statistics, at The Ohio State University.
Chapter 4

DESIGN OF THE EXPERIMENT

From the study of the literature with regard to the methods or techniques of teaching statistics as presented in Chapter 2, and from the survey conducted at The Ohio State University as discussed in Chapter 3, the investigator recognized two problems pertaining to the teaching of statistics: (1) how to make the course more meaningful and interesting to the students and more relevant to the various fields of study, (2) what methods or techniques of teaching might be employed to achieve this. These problems led the investigator to try several techniques or combinations of techniques in his investigation. The techniques employed for the study were (i) the use of preliminary questions, (ii) the use of class experimentation, (iii) the use of computer program, and (iv) the use of student survey projects.

When this investigation was conducted, Statistics 125 (Stat 125) was an introductory course in mathematical statistics offered by the Department of Statistics at The Ohio State University. The course was provided for students
without calculus background. However, the mastery of the basic arithmetic operations comparable to first year high school Algebra was essential. The course was attended by students from a variety of schools and colleges such as University College, School of Nursing, School of Allied Medical Professions, College of Arts, Division of Continuing Education, College of Agriculture, School of Social Work, School of Architecture, and so on. The level of the students varied from Freshmen to Seniors, but the majority were Sophomores and Juniors. The course outline is presented in Appendix B, pages 125-127.

Given the responsibility to teach this class, the investigator found it an appropriate context in which the search for a meaningful method of teaching statistics with relevance to varied fields of study was possible. The experimental methods were designed with the restriction that the departmental course outline be followed closely and that there be no adverse effect on the course coverage. The investigator strongly believed that a careful integration of additional activities such as the use of preliminary questions, class experiments, computer programming, and student surveys with the formal systematic lecture method would help make the course meaningful to the students and relevant to their fields of study. Hopefully this would be true even for those students taking this as a terminal course.
4.1 Preliminary questions

The objective of preliminary questions was to direct students' attention to the relevance of statistics to many problems in various areas of scholarship. The questions set consisted of ten questions corresponding to the chapters covered during the ten weeks of the quarter, and was designed to meet the following requirements:

(i) the problem should be very practical, if possible with real data, and relevant to the interest of the students in the group.

(ii) the questions should be appealing to such an extent that the students would answer the questions with enthusiasm either by intuition or by guessing.

(iii) each question should be representative of the characteristic problem of each concept or chapter to be discussed in the course; i.e. each question should be a model question. If a student could answer this model question correctly, he should be able to answer many questions of similar type.

(iv) the questions should be well selected and their answers should not be very obvious.

The following is an example of the questions which relates to the concept of a one sample test of hypothesis.
Eskimos' Body Temperature

A doctor is interested in checking to see if the average body temperature of Eskimos is significantly lower than man's normal temperature, which is 98.6°F Fahrenheit. After selecting ten Eskimos at random, the following temperatures were recorded: 99.0 97.8 99.8 98.6 98.7 98.3 97.9 98.0 98.9 and 98.0. The average is 98.0. Do you think these results give evidence to reject the claim that the Eskimos, on the average, have normal body temperature?

The complete set of preliminary questions is presented in Appendix C, pp 128 - 136.

The questions were given to the students at the beginning of the quarter as an in-class exercise, and the students were asked to do the best they could in answering each question. The instructor graded, recorded, and kept the answers until the end of the quarter. At the beginning of the discussion of a new chapter, the students were referred to one or two questions from the set. The instructor led the students to find solutions to the problems. While the students' interest in the search for correct answers was aroused, the instructor presented and explained the concepts contained in the chapter. After this presentation, the students were expected to be able to solve the problems correctly.

At the end of the quarter, the preliminary questions were reassigned as homework for which the students were asked to give complete answers, including the necessary mathematical computations. Their work was graded, recorded,
and returned to them together with their answers at the beginning of the quarter. A follow-up discussion was then conducted in class. This gave the instructor an opportunity to explain the problems and to comment upon the answers of the students. As for the students, they had the opportunity to compare their own two answers, before and after the course.

4.2 Design and analysis of class experiments

The purpose of each experiment was to help explain a certain concept by doing. It was hoped that the students would see the relevance of the problems in reality and the theory in statistics which would result in better understanding of the concept. The effect of each experiment in terms of the understanding of the corresponding concept was measured by means of students' performances in solving problems related to the concept. Six experiments were conducted during the quarter, and each took ten to fifteen minutes. Description of the class experiments is as follows.

Experiment number 1: Empirical probability

The objective was to show how probability could be determined empirically. In this experiment each student was asked to flip a coin ten times and record how many heads appeared. A frequency table for the whole class was
constructed which was used in a discussion to determine empirical probabilities. This experiment was also used for the discussion of approximation to Binomial Distribution.

Experiment number 2: Random variable and probability

The objective was to explain the meaning of random variable and to find the corresponding probabilities both empirically and mathematically. Each student was asked to roll a die twice and record the results. By defining a random variable $X$ as the sum of the dots, each student was asked the value of $X$ he or she had by rolling the die twice. By tabulating the results of the whole class the students were able to determine the probability for $X$ empirically. They were also asked to list all the possible outcomes of rolling the die twice with the corresponding values of $X$, and to try to formulate the mathematical probability distribution. Following the experiment, a discussion was conducted in order to compare the empirical and the mathematical probabilities.

Experiment number 3: Central Limit Theorem

The objective was to explain the meaning of the averages and the sums of samples, and to show the resemblance of their histograms to the normal curve. Each student was asked to roll a die ten times, record the results, and
calculate the sum and the average. The results for the class were tabulated and the histograms of the sums and the averages were graphed.

Experiment number 4: Normal approximation to Binomial

This experiment was preceded by a discussion of the need for normal approximation to the Binomial for large $n$. Each student was asked to flip a coin forty times and record how many heads appeared. The results of the whole class were recorded, tabulated, and the histogram was graphed. The shape of the histogram was examined and compared to a normal curve.

Experiment number 5: Two sample t-test

A simple question "Who are taller, men or women?" was used in a discussion preceding the experiment. Each student was asked to write down his or her weight, height, and the ratio of the weight and the height. The data from the whole class was collected and tabulated, and the students were asked to test the equality of the means of the heights, the means of the weights, and the ratios for males and females using the two sample t-test following the method described in the book.
Experiment number 6: Regression line

In this experiment the heights and the weights of the students chosen at random were tabulated on the board, plotted, and the eye-ball fit regression line was graphed. This was followed by a discussion on the interpretation of the regression line obtained.

Each experiment was followed by a homework assignment pertaining to it. Two groups of students were randomly formed at the beginning of the course. For each of the six experiments one group was chosen at random to be the treatment group and the other served as the control. On the days that a treatment group performed an experiment, the control group learned the concept by the usual lecture method using the data from the textbook or other sources. Each of the two study groups served as the treatment group for three experiments. In order to evaluate the effectiveness of the experiments, achievement tests were given following each experiment and results were compared and analyzed using a two sample test - see Scheffe (1959) page 59, or Hollander & Wolfe (1973) page 675 for the procedures. In order to determine initial comparability of the two study groups, the first two weeks were used as an observation period, after which an achievement test was given. The results were analyzed using two sample test.
4.3 Computer programming

The investigator believed that the use of computer programs in statistics instruction would serve several purposes. The first was to make the course more appealing to the students by giving them additional activities. The second was to give the students some experience in working with the computer and to familiarize them with some computer procedures; this might be the first and the last such college experience for some of them. The third was to help them realize that a computer can be used as a tool for numerical computation, and that statistics is involved with the interpretation of the results. Knowing that the numerical computations can be done by the computer, the students may be less apprehensive about finding the mean, variance, or other necessary statistics from real data which may involve many observations, large numbers with decimal points, negative numbers, and so on. The interpretation of real data, the investigator believed, would be more interesting and challenging to the students than interpreting simple, hypothetical data.

Included in the experimental treatment were only two computer-related problem sets using Omnitab (see Appendix F, pp. 172-8). It was felt that the two sets in a quarter were sufficient for the above stated purposes. The first problem set was designed to obtain a statistical
analysis of a set of data, particularly the mean, the median, the range, the variance, the standard deviation, and a 95% confidence interval for the population mean if the data were obtained from a normal population. The second set was designed to perform a linear regression analysis for a set of data pertaining to the heights (X) and the weights (Y) of the students in class. Students' attention was especially directed to the plot of the weight versus height, the plot of the corresponding ratio weight over height, the equation of regression line, and the list of numerical values for height, weight, ratio and estimated weight according to the regression obtained. At the end of the course the students were asked to give their opinion about the effectiveness of this technique in terms of the achievement of the objective.

4.4 Student survey projects

The purpose was to have the students gain an experience working with realistic problems, be actively engaged in collecting data needed, and utilize some statistical procedures in solving these problems. This was felt to be especially needed in a terminal course in order for the students to see the application of statistics in many areas of interest.
The investigator believed that there were at least seven statistical procedures in Statistics 125 that could be utilized in conjunction with the student survey projects; (1) to estimate a population mean $\mu$, (2) to estimate a population proportion $p$, (3) to test a hypothesis concerning a population mean $\mu$ or population proportion $p$, (4) to test the equality of means or proportions of two populations, (5) to use a goodness-of-fit test, and (6) to test the independency of two variates using the Chi-square distribution.

In this technique, the instructor explained the procedures intuitively and gave two or three illustrative examples for each procedure. The students were then asked to decide upon the survey projects they were interested in which were related to their own fields of study. For example, a student in Nursing may conduct a survey to compare the means of the pulse rates of male and female students. They submitted their plans to the instructor for approval prior to collecting data. The final, complete reports were submitted during the eighth week. This gave the instructor ample time to read, record, and give comments on the reports. During the ninth week some students were asked to present their papers in class to show the results of their surveys and to share their experience with the other students.
The chapter that follows contains a report of the investigation conducted during the Fall quarter 1975 using preliminary questions, class experiments, and computer programs.
Chapter 5

REPORT ON FALL 1975 TEACHING
OF STATISTICS 125

During the Fall quarter of 1975 there were seventy nine students enrolled in Statistics 125 class. They were all undergraduates from many departments and from all levels. The distribution of the students in terms of their fields of study was as follows:

Table 5.1 Number of Students in Statistics 125
Fall Quarter 1975

<table>
<thead>
<tr>
<th>School / College</th>
<th>Number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>University College</td>
<td>21</td>
</tr>
<tr>
<td>School of Nursing</td>
<td>8</td>
</tr>
<tr>
<td>School of Natural Resources</td>
<td>26</td>
</tr>
<tr>
<td>College of Arts and Sciences</td>
<td>7</td>
</tr>
<tr>
<td>Continuing education</td>
<td>2</td>
</tr>
<tr>
<td>College of Agriculture</td>
<td>6</td>
</tr>
<tr>
<td>School of Social Work</td>
<td>6</td>
</tr>
<tr>
<td>College of Engineering (Ceramic)</td>
<td>2</td>
</tr>
<tr>
<td>School of Architecture</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>79</strong></td>
</tr>
</tbody>
</table>

Of the 79 students in this class, 8 were freshmen, 33
sophomores, 17 juniors, 19 seniors, and 2 special classifications.

Upon students' agreement it was decided to divide the class into two groups. On the day when an experiment was to be performed each group was to attend class at different time. The group that performed the experiments was considered the treatment group, while the other the control group. The latter received the explanation of the concept through lecture method. Further explanation on class experiments is presented in Section 5.2

5.1 Result of preliminary questions technique

A comparison of the percentages of correct answers to the preliminary questions made by the students at the beginning and at the end of the quarter is shown in Table 5.2. The result showed that the students are generally very good at making guesses, especially in cases where chance is involved; problems 1 and 3 are examples of such cases. Almost all of the students gave correct answers to those problems, although they did not know how to verify the answers mathematically. The results also suggest that special attention be given to theories such as Central Limit theorem, tests of hypotheses concerning one or two sample problems, test of independency, goodness-of-fit test, and regression analysis. In response to the
investigator's questions concerning their opinion about the technique employed by the students stated "very helpful", 12% stated "just another method", and 4% stated "the problems are too difficult".

Table 5.2 Result of Preliminary Questions Set

<table>
<thead>
<tr>
<th>Question number</th>
<th>Problems</th>
<th>B</th>
<th>A *)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>True and False Examination</td>
<td>75%</td>
<td>95%</td>
</tr>
<tr>
<td>2</td>
<td>An E S P Problem</td>
<td>44%</td>
<td>78%</td>
</tr>
<tr>
<td>3</td>
<td>A Magazine Salesman's dream</td>
<td>60%</td>
<td>74%</td>
</tr>
<tr>
<td>4</td>
<td>An Airplane Overweight Problem</td>
<td>19%</td>
<td>88%</td>
</tr>
<tr>
<td>5</td>
<td>Car Repair Expenses</td>
<td>8%</td>
<td>83%</td>
</tr>
<tr>
<td>6</td>
<td>Eskimo's Body Temperature</td>
<td>50%</td>
<td>90%</td>
</tr>
<tr>
<td>7</td>
<td>Which Paint is Better?</td>
<td>31%</td>
<td>90%</td>
</tr>
<tr>
<td>8</td>
<td>Birth by month</td>
<td>8%</td>
<td>97%</td>
</tr>
<tr>
<td>9</td>
<td>MPH-MPG Problem</td>
<td>21%</td>
<td>88%</td>
</tr>
<tr>
<td>10</td>
<td>More Children?</td>
<td>38%</td>
<td>92%</td>
</tr>
</tbody>
</table>

*) B Percentage of students choosing correct answers at the beginning of the quarter
   A Percentage of the students choosing correct answers at the end of the quarter

It can be discerned from the table that there is a marked increase in the percentages of students choosing correct answers. This also means that they are more able to analyze and interpret the information contained in the problems.
5.2 Results of the Use of Class Experiments

For this purpose the class was divided into two groups by randomization at the beginning of the quarter. During the first two weeks no experiment was conducted and the two groups met at a single class. This period was an observation period to see if there was a significant difference in the initial achievements of the two groups such as in the ability to manipulate numbers, to convert data into graphs, to read and interpret graphs, to calculate the means, variance, modes, and so on. A two sample t-test was used in comparing the two groups on the achievement test given to them at the end of the observation period. It was found that there was no significant difference \( \alpha > 5\% \) between their mean achievements. This, in addition to randomization, supports the assumption that the two groups had comparable initial performances (for detailed computations, see Appendix E pp 164 - 71.

It was originally planned that the first group would be the treatment group and the second the control group. However, the students expressed their strong desire to be in "the better group", whichever group that was, be it treatment or control. As a result the experiments were assigned at random to the groups. This randomization resulted in the assignment of experiments 1, 2, and 4 to group I, and experiments 3, 5, and 6 to group II. It was arranged
that on an experiment day group I was to come to the first half of the period and group II to the second half. However, in some meetings, especially for experiments 5 and 6, some students did not adhere to their group assignment. Some from group I attended group II meetings and vice versa; a few of them stayed on for the entire period. Although they had a legitimate reason for not staying in their own group - they wanted to gain as much as possible from the course - this situation caused contamination of the study.

Experiments 1, 2, 3, and 4 as presented in Table 5.3 showed significant differences favoring the treatment group, while experiments 5 and 6 did not. This lack of significance might be due to (i) contamination of the study as mentioned above. (ii) the fact that experiments 5 and 6 dealt primarily with mathematical computations rather than conceptual understanding. (iii) the fact that group I had performed experiments 1, 2, and 4, which might have helped in understanding the concepts conveyed by experiments 5 and 6. (iv) the inappropriateness of experiments 5 and 6 to explain the concepts.
Table 5.3 Comparison of the Means of the Test Scores of the Two Groups for the Six Experiments

<table>
<thead>
<tr>
<th>Experiments number</th>
<th>Sample averages</th>
<th>t-values</th>
<th>$\alpha$</th>
<th>The difference of the means are</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment group</td>
<td>Control group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>6.21</td>
<td>5.22</td>
<td>2.4</td>
<td>.8% significant</td>
</tr>
<tr>
<td>2</td>
<td>5.33</td>
<td>4.16</td>
<td>2.24</td>
<td>1.26% significant</td>
</tr>
<tr>
<td>3</td>
<td>6.27</td>
<td>5.33</td>
<td>2.02</td>
<td>2.2% significant</td>
</tr>
<tr>
<td>4</td>
<td>6.07</td>
<td>5.41</td>
<td>1.85</td>
<td>3.2% significant</td>
</tr>
<tr>
<td>5</td>
<td>5.89</td>
<td>5.54</td>
<td>.08</td>
<td>47% not signif.</td>
</tr>
<tr>
<td>6</td>
<td>6.43</td>
<td>6.05</td>
<td>1.02</td>
<td>15% not signif.</td>
</tr>
</tbody>
</table>

Note: $\alpha$ is the lowest significance level at which one would reject the null hypothesis that the means are equal.

For further calculations see pages
- Experiment 1: Empirical probability
- Experiment 2: Random variable and probability
- Experiment 3: Central limit theorem
- Experiment 4: Normal approximation to Binomial
- Experiment 5: Two sample
- Experiment 6: Regression line

In order to see the effect of the method using class experiments as compared to that of the lecture method, two analyses were performed. The first was to compare the sum of the test scores of each student in group I after conducting experiments 1, 2, and 4 with the sum of the test scores of each student in group II after receiving
lectures related to concepts 1, 2, and 4. The result, see Table 5.4, showed a significant difference in their achievements favoring group I. Calculations for this analysis are presented in Appendix D page 162.

Table 5.4 Comparison of the Means of the Effect of Class Experiments and Lecture Discussions in explaining the Concepts 1, 2, and 4

<table>
<thead>
<tr>
<th></th>
<th>Group I</th>
<th>Group II</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>39</td>
<td>37</td>
</tr>
<tr>
<td>Mean</td>
<td>18.05</td>
<td>14.78</td>
</tr>
<tr>
<td>t = 3.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha = .0006$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance</td>
<td>18.79</td>
<td>19.95</td>
</tr>
<tr>
<td>F = 1.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha = .05$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The second analysis was performed for the purpose of comparing the sum of the test scores of each student in group II after conducting experiments 3, 5, and 6 with the sum of the test scores of each student in group I after receiving lectures related to concepts 3, 5, and 6. The result showed a slightly significant difference in their achievements favoring group II, $\alpha = .0375$ (see Table 5.5). Calculations are presented in Appendix D, page 163.
Table 5.5 Comparison of the Means of the Effect of Class Experiments and Lecture Discussions in Explaining the Concepts 3, 5, and 6

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Variance</th>
<th>t</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>39</td>
<td>16.92</td>
<td>18.32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>37</td>
<td>18.67</td>
<td>17.98</td>
<td>1.78</td>
<td>1.2</td>
<td>.0375</td>
</tr>
</tbody>
</table>

5.3 Results of the use of computer programming

In general, the students found their experience working with the programs worthwhile and very enjoyable. Their responses to the questionnaire concerning the usefulness of those computer programs in their study showed that 80% found it very useful and interesting, 15% responded "it was O.K.", and the remaining 5% expressed frustration caused by the technical difficulties such as typing up the job control cards where spaces and blanks are of prime importance.

In the following chapter, Chapter 6, a report on the investigation conducted during the Winter quarter of 1976, in which preliminary questions, computer programming, and student surveys were employed, will be presented. It also contains a report on the investigation conducted during the Spring quarter 1976 in which no supplementary techniques were employed.
During the Winter quarter 1976, the investigator was given the opportunity to teach two classes in introductory statistics. One class was held at 12:00 and the other at 3:00 PM; there were about 30 students in each class. The enrollment in both classes increased during the first and second week. As was done in the Fall quarter of 1975, the first two weeks of the quarter were used as observation period for the purpose of finding out the difference in the initial achievement of the two classes. At the end of the second week a test was given to measure the understanding and the achievement of the students in working with numbers, finding descriptive measures, graphing and interpreting the graph, and so on. It was found that at 5% significance level, using a two sample t-test, there was no significant difference between the two classes.

For the sample test calculations, see Appendix E, pp.168-9. The results of the initial test are shown in Table 6.1
### Table 6.1 Results of the Initial Test Winter 1976

<table>
<thead>
<tr>
<th>Class</th>
<th>12:00</th>
<th>3:00</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of students</td>
<td>40</td>
<td>36</td>
</tr>
<tr>
<td>Average score</td>
<td>36.02</td>
<td>35.81</td>
</tr>
<tr>
<td>Sample variance</td>
<td>8.7429</td>
<td>20.6183</td>
</tr>
</tbody>
</table>

Note: since the variances of the two populations are not the same, Smith-Satterwaite procedure was used to test the equality of their means. It was found that at \( t' = .244 \) with 60 d.f., \( \chi^2 \) had a value between 40% and 41%; thus the difference is not significant.

The instruction in the 12:00 class was supplemented by the use of preliminary questions set, computer programming, and the student surveys, while in the 3:00 class only computer programming and student surveys were used. Although class experiments per se were not extensively used in either class, on many occasions data were collected from the class. For example, for the purpose of the calculation and discussion of means, variance, and so on, the students were asked to tabulate the number of credit hours they were carrying that quarter. For Binomial distribution the students were asked to flip a coin ten times, and for estimation theory the heights of male students were tabulated. For goodness-of-fit test the number of birthdays in each
season (Fall, Winter, Spring, and Summer) were recorded.
For contingency table the independency of color preferences,
in this case either green or yellow, versus the male and
female classification was tested. The investigator believes
that collecting data from the class and providing the
students with such an experience will help make the subject
more interesting to them.

6.1 Results of the preliminary questions

Table 6.2 shows a comparison of the percentages
of students choosing correct answers to the preliminary
questions at the beginning and at the end of the quarter.

Table 6.2 Results of Preliminary Questions Set
Winter 1976

<table>
<thead>
<tr>
<th>Question number</th>
<th>Problems</th>
<th>B</th>
<th>A</th>
<th>*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>True and False Examination</td>
<td>73%</td>
<td>97%</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>An ESP problem</td>
<td>68%</td>
<td>95%</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>A Magazine Salesman's dream</td>
<td>65%</td>
<td>90%</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>An airplane Overweight Problem</td>
<td>39%</td>
<td>90%</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Car Repair Expenses</td>
<td>39%</td>
<td>80%</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Eskimo's Body Temperature</td>
<td>34%</td>
<td>97%</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Which Paint is Better?</td>
<td>7%</td>
<td>92%</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Birth by month</td>
<td>5%</td>
<td>95%</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>MPH-MPG problem</td>
<td>25%</td>
<td>97%</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>More Children?</td>
<td>13%</td>
<td>97%</td>
<td></td>
</tr>
</tbody>
</table>

*) B percentage of students choosing correct answers
at the beginning of the quarter
A percentage of students choosing correct answers
at the end of the quarter
The table shows a marked increase in the percentages of students choosing correct answers at the end of the quarter. This could be interpreted that at the end of the quarter the students are better equipped to analyze the information given in the problems. The results also showed the need of such theories as Central Limit theorem, testing hypotheses, two sample test, regression analysis, test of independence, and so on.

Students' responses to the questionnaire were as follows:

**Question 1**

Has the preliminary questions set directed your attention to the relevance of statistics to various fields of study?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>29</td>
<td>2</td>
<td>31</td>
</tr>
</tbody>
</table>

**Question 2**

Were the problems challenging enough to encourage you to think about the solutions as you studied various concepts?

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>29</td>
<td>2</td>
<td>31</td>
</tr>
</tbody>
</table>

**Question 3**

How did these questions affect your attitude toward statistics in terms of its usefulness in real life problems?

<table>
<thead>
<tr>
<th></th>
<th>Positively</th>
<th>Negatively</th>
<th>Just the same as before I took the course</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>28</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>31</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Question 4**

What do you think of these problems

<table>
<thead>
<tr>
<th></th>
<th>Difficult</th>
<th>Fair</th>
<th>Easy</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>19</td>
<td>12</td>
<td>0</td>
<td>31</td>
</tr>
</tbody>
</table>


Question 5

Did you notice the difference in your ability to solve these problems between the beginning and the end of the quarter? Yes No Total
29 2 31

Question 6

Do you have any general comment about the preliminary questions set?

* It is very worthwhile when you see how various formulas are applied to real life situations which is what P.Q. were example of.

* They took an awful lot of time, but it was good, it served as a review for the course in general.

* I realize how ignorant I was in the subject matter; learning how to solve the problems which increase in my self-confidence.

* They were helpful. It gave us a mixture of problems at one time. I like this course very much.

* I like the before and after comparison - it made me feel that I have really learned something. Plus I feel the type of problems definitely show the usefulness of statistics in life - something which I feel is extremely important to do.

* I thought they were too deep for this level of statistics.

* It was a good help

6.2 Results of the use of computer programming

Since computer programming was a novelty to many of the students in these classes, the program was discussed and the problems pertaining to it were assigned early in the quarter. The students were expected to finish their work at the end of the sixth week. In this way they were allowed enough time to punch the cards at their convenience.
There were two sessions provided for discussions of the results; one problem was discussed at the end of the sixth week and the other at the end of the ninth week. These problems were related to the chapters being discussed.

The responses of the students to the questionnaire were as follows:

**Question 1**

How would you rate the experience working with a computer program in this course?

<table>
<thead>
<tr>
<th></th>
<th>Class</th>
<th>Good</th>
<th>Fair</th>
<th>Nothing</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:00</td>
<td>16</td>
<td>14</td>
<td>1</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>3:00</td>
<td>24</td>
<td>9</td>
<td>1</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>40</strong></td>
<td><strong>23</strong></td>
<td><strong>2</strong></td>
<td><strong>65</strong></td>
<td></td>
</tr>
</tbody>
</table>

**Question 2**

Was the experience meaningful to you?

<table>
<thead>
<tr>
<th></th>
<th>Class</th>
<th>Yes</th>
<th>O.K.</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:00</td>
<td>18</td>
<td>9</td>
<td>4</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>3:00</td>
<td>20</td>
<td>9</td>
<td>5</td>
<td></td>
<td>34</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>38</strong></td>
<td><strong>18</strong></td>
<td><strong>9</strong></td>
<td><strong>65</strong></td>
<td><strong>65</strong></td>
</tr>
</tbody>
</table>

**Question 3**

How much time did you spend on this computer program?

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>12:00</td>
<td>average: 2.3 hours</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3:00</td>
<td>average: 2.6 hours</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total average:</strong></td>
<td><strong>2.4 hours</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results showed that most of the students enjoyed working with the computer programs. They found the experience very meaningful without spending too much time on it. The students who had worked with computers found this as an opportunity to use Omnitab set up.
6.3 Results of the student surveys

As was stated in Section 4.4, the purpose of the student survey projects was to provide the students with the experience of working with realistic problems, being actively engaged in collecting the data needed, and utilizing some statistical procedures in solving those problems. The plan for conducting the survey was discussed early, at the beginning of the quarter. The purpose of the survey was explained, the method of collecting the data was discussed, and the types of statistical procedures were explained intuitively. The investigator believes that an early introduction of the student survey plan is important in order to allow ample time for the students to think through the projects, and to plan on the survey and the collection of data, although they have to postpone the statistical analysis and the computations until the related theory has been discussed in class. Several samples of student surveys are presented in Appendix G, pp.179-196.

The responses of the students on the student survey questionnaire were as follows:

Question 1

<table>
<thead>
<tr>
<th>Class</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:00</td>
<td>27</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td>3:00</td>
<td>24</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>51</strong></td>
<td><strong>2</strong></td>
<td><strong>53</strong></td>
</tr>
</tbody>
</table>
### Question 2

Was it relevant to your study?

<table>
<thead>
<tr>
<th>Class</th>
<th>Yes</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:00</td>
<td>25</td>
<td>3</td>
<td>28</td>
</tr>
<tr>
<td>3:00</td>
<td>23</td>
<td>2</td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>48</td>
<td>5</td>
<td>53</td>
</tr>
</tbody>
</table>

### Question 3

What was the effect of your project on your general knowledge of statistics?

<table>
<thead>
<tr>
<th>Class</th>
<th>Good</th>
<th>Fair</th>
<th>No effect</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:00</td>
<td>21</td>
<td>5</td>
<td>2</td>
<td>28</td>
</tr>
<tr>
<td>3:00</td>
<td>21</td>
<td>3</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>42</td>
<td>8</td>
<td>3</td>
<td>53</td>
</tr>
</tbody>
</table>

### Question 4

Do you think your project will result in longer retention of the concept?

<table>
<thead>
<tr>
<th>Class</th>
<th>Yes</th>
<th>May be</th>
<th>No</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:00</td>
<td>22</td>
<td>6</td>
<td>0</td>
<td>28</td>
</tr>
<tr>
<td>3:00</td>
<td>20</td>
<td>3</td>
<td>2</td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>42</td>
<td>9</td>
<td>2</td>
<td>53</td>
</tr>
</tbody>
</table>

### Question 5

How much time did you spend on the survey including the preparation and making the report?

- 12:00 average: 4.5 hours
- 3:00 average: 3.3 hours
- Grand average: 3.8 hours

### Question 6

Do you have any comment or suggestions about the survey?

12:00 class

- * Was interesting
- * It was fun, but I was constantly worried if I had chosen the right test to perform
- * Good idea
- * I thought it was a good thing relating material in class to an area of interest
* I was not quite sure what you were asking for, for example in length, detail, etc. When I handed in the paper, I wasn't quite sure it was satisfactory.

* It should remain a permanent part of the course.

* I think it was a good idea to have a practical test of the "stuff" we had in class. It was good we were given a choice of the subject & tests, too.

* I thought it was very helpful in seeing how statistics is actually applied.

* It was a good idea.

3:00 class

* Once I started on doing it I really enjoyed it and it seemed to help me understand the two sample t-test better.

* The survey project was a good idea because it helped me see how you can use statistics in many situations and why you should or should not always rely on them as fact for an entire population.

* Overall they were fun to do. Should read the more unique ones in class

* It is a good idea.

* I think it might be possible to work for points.

* I think the survey was a relevant one

* It is good for extra credit project.

* I thought it is good to get some extra credit points and pull up the grade in case one test was not as good as others.

* It is a good idea - shows the practical use of statistics.

* I think it was well worth the effort. It gave me a way that I can use what I have learned practically.

* Very interesting and enjoyable

* It was an excellent integration of knowledge I learned in class and practical application.
* It increased my knowledge of practical uses for statistics which help me be more interested in them. I think it was a good idea. I understand the concept better now.

* I thought the survey was very worthwhile in learning statistics and then applying it to everyday. I was able to see how to make statistics work.

* May be to do two surveys, one for a single sample, one for two sample problem.

* Helped to relate formulas with real situations, as story problems but more.

In general, the students were enthusiastic about doing the survey, especially since they had their own choice of topics. The topics they chose were varied in accordance with their fields of study; many of them were very interesting.

6.4 Results of Spring quarter 1976

During the Spring quarter 1976, the investigator was given another opportunity to teach Statistics 125. The class was held at 3:00 PM, M - F, and there were 23 students enrolled. Only the systematic lecture method, without any of the supplementary techniques - preliminary questions, class experiments, computer programs, and student surveys - was employed in the instruction. Hence this group could be considered as an additional control group for the entire experimentation. The textbook used was the same, and the tests and the final examination were similar to the ones given in previous quarters. The results of the
initial test given this group as compared to those of the group of Winter quarter 1976 showed the comparability of their initial performances. For calculations, see Appendix E, pp. 170 - 171.

Results of the test for the entire quarter showed that this control group achieved lower total scores than those obtained by the group of Winter quarter. The latter, the reader might recall, received instruction supplemented by the four teaching techniques. Complete analysis of the results of the tests is presented in Appendix H, pp.198-200. The results obtained from the experimentation conducted during this Spring quarter 1976 indicated the significance of the use of the supplementary techniques in the instruction of Statistics 125.

This completes the study, and in the following pages the investigator will try to summarize and interpret the results of his investigation. The following chapter, Chapter 7, will also contain recommendations for further investigations.
Chapter 7

CONCLUSIONS AND RECOMMENDATIONS

This investigation was conducted to find a partial solution to a general problem in the teaching of an introductory statistics course - to find ways to make it more meaningful, intellectually worthwhile, and challenging to the students, and to help them see its relevance to their various fields of study. This final chapter includes a brief discussion of the study of the literature and the survey at The Ohio State University, a discussion of the techniques employed, and the limitations of the study. Then recommendations are made pertaining to methodology and organization for the teaching of introductory statistics. Finally, directions for further investigation are suggested, particularly ones intended to perfect each of the techniques employed in this study.

7.1 Conclusions from the investigation

A careful review of the literature revealed numerous attempts by previous investigators to explore ways to teach introductory statistics with special
attention to students without calculus background.

Many computerized methods have been introduced including Computer Based Instruction using a combination of Computer Assisted Instruction, exercise generating programs and programs for the analysis of data; the use of computers numerically or graphically in teaching statistics; Proctorial System of Instruction; Personalized System of Instruction; and many others. The study also indicated that classroom experiments, teaching aids, audio-visual materials such as television, filmstrips, slides, or video-tapes have been used extensively.

Several pedagogical suggestions have been made to improve the quality of the systematic lecture method by using a decision theory approach, a nonparametric approach, or a problem solving approach. Each approach appeared to have advantages and disadvantages; each was directed toward the enhancement of students' interest in statistics with the hope that they would profit as much as possible from the course.

Many authors felt that the teaching of statistics could be improved through pre-college level instruction. The teaching of statistics at the pre-college level could be considered as having three purposes. The first purpose was to have the high school students have some knowledge
of statistics prior to entering a university. The second was to give the students some practical knowledge of statistics in preparation for becoming members of a society of educated men and women where statistics is widely used. The third was to help the students to study mathematics, especially those students who have had a certain degree of difficulty in learning mathematics.

From the survey discussions with many professors teaching statistics in various departments at The Ohio State University, the investigator learned that many methods have been employed in the teaching of basic statistics as well as applied statistics. The lecture method was the most common methodology employed, using illustrations and problems that were directly related to students' own fields of study. Another technique was the use of a problem solving approach in which the students presented problems whose solutions called for statistical theories and discussions. Some departments did not offer statistics as such, but integrated it into their own major course. In such an integrated course statistical theories were discussed only when students encountered problems that required statistics for solutions. Most departments have used statistical computer packages which helped the students obtain the statistics necessary for the analysis of their data.
From the methods employed by the departments it became obvious that they were trying to make statistics more relevant by emphasizing its direct utilization in their own areas. Both the study of the literature and the results of the survey conducted at The Ohio State University confirmed the investigator's belief in the importance of exploring and formulating methods to make statistics seem relevant to students in whatever field of study.

It was previously stated that the primary objective of the preliminary questions set in the instruction was to direct the students' attention to the relevance of statistics to many problems in various fields of study. Another objective was to stimulate the students' interest to learn more statistics and to recognize the need of it in their study.

The results of the investigation which were presented in Chapters 5 and 6 in terms of its achievement of the objectives and also in terms of students' response toward the use of this technique indicated some strongly positive points. (1) It gave the students a feeling of accomplishment as they compared their own two sets of answers, one set done at the beginning of the quarter and the other at the end of the course. Normally, it has not been feasible to make an exact comparison such as this because the students were not likely to remember their own
knowledge of the subject at the outset of the course. (2) It affected students' attitude positively in that they recognized the usefulness of statistics in their own fields as well as, more generally, in real-life problems. It helped them view statistics as an important part in their study and not just a requirement to be endured or as a substitute course for mathematics. (3) As a challenge to the students, it encouraged them to think and aroused interest and desire to learn more about the subject. (4) Their scholastic achievement in statistics as was shown by the results of the investigation conducted during the Winter quarter 1976, was slightly improved. The class where the preliminary questions set, computer programming, and student surveys were used showed a little higher achievement than the one with only computer programming and student surveys, although it was not statistically significant; \( \alpha = 5.8\% \). For calculations see Appendix H, page 199.

In his analysis of the results obtained by using the preliminary questions set technique, the investigator assumed three basic reasons underlying its success: (i) the introduction of the subject to the students in the form of problems related to many areas of study helped direct their attention to the role of statistics as an integral part in many fields. In solving the problems the students were actively engaged in thinking; thus the problems made
the class situation more conducive to learning. (2) the frequent reference to the problems set, as a whole at the beginning and at the end of the course and to each individual problem as the related concept was being discussed, served as a constant reminder of the role of statistics. (3) the active participation of the students in relating the theory to the solution of the problems in the set and to some other problems provided them with the opportunity to practice applying statistical techniques and thus promoted the mastery of the subject. The investigator strongly believes that the use of the preliminary questions technique in teaching an introductory statistics course helped him make it more relevant to the students.

Class experiments were used in the instruction to help explain a certain concept through related activities; it was hoped that the students would see the relevance of the concept in statistics to problems real to them.

As presented in Chapter 5 the results of the use of class experiments indicated a positive effect of this technique upon students' comprehension of the concept. The students felt that the experience was worthwhile and very meaningful. Despite the usual argument against class experiments, as expressed in many writings, that they are too time consuming, the investigator found it to be otherwise.
He believed it was worth the effort and, as judged by the results, very rewarding. He agreed that it required more of his time for the selection and the preparation of the class experiments, but he was persuaded by the results that use of experiments enhanced students' understanding of the concepts.

As was stated in Chapter 4, the main purpose of the inclusion of computer programming in this type of introductory statistics course was to provide the students with some experience in working with the computer: setting up the program, punching the cards, submitting the deck, and reading the computer output. Another purpose was to get them to realize the existence of canned programs at The Ohio State University. Programs such as Omnitab, SPSS, P-Stat, SAS, and the like were made available to students to help them perform computations involving large numbers and decimals to calculate the statistics necessary for the analysis of their data.

The students' response to the inclusion of computer programming in the course as presented in Chapter 5 and 6 was strongly positive. Many remarked upon the meaningfulness of the experience in their study. Some students, who at the beginning of the course came to class with fear of arithmetic computations and deep concern for the drudgery of arithmetic manipulations, expressed their
relief and interest in the course.

Based on the results obtained in Chapters 5 and 6, the investigator strongly believes that the use of computer programs in statistics instruction not only served the original purpose, but also helped arouse students' interest and develop better attitudes toward the course.

The student surveys were intended to provide the students with experiences working with realistic problems, becoming actively engaged in collecting the data needed and utilizing statistical procedures in solving these problems.

The results of the student surveys as presented in Chapter 6 indicated very positive outcomes. (1) They helped arouse students' interest in the subject; students who initially did a survey just for the sake of doing it, finally found it very enjoyable, interesting and challenging. The collection of data and its analysis helped them see the relevance of statistics to their fields. They also made the students see some of the difficulties involved in actually collecting the data. Some expressed an unexpected effect of helping them decide on their field of specialization as the result of the interesting interviews they conducted during the collection of their data. Some suggested that the student surveys should become a permanent part of the course. (2) They helped the students to understand the
concepts better and to believe that they would retain them longer.

Those positive outcomes convinced the investigator of the positive effects of using this technique in his teaching in that it has helped him enliven the course and make statistics seem relevant to the students in many areas of study.

The investigator was aware of the limitations of the study. He realized that in order for this type of investigation to have a desirable external validity, several instructors and several groups under the same conditions would be required. However, for this investigation this requirement could not be met. The other limitations encountered in this study were: (1) the difficulty in avoiding contamination during experimentation. This was due to the fact that some students did not adhere to the assigned group; instead, they attended both groups, especially while experiments 5 and 6 were conducted. This was discussed in Chapter 5.2. However, in the analysis of the results these students were considered as belonging to their originally assigned group. (2) the absences of some students from class which was beyond the investigator's control. In the analysis this was not taken into account. (3) the withdrawal of a few students during the quarter. The test scores of these students were not included in
the analysis of the results of the initial and the final tests. (4) the initial test which was intended to measure the comparability in the initial achievements of the groups, seemed to work well in the Fall 1975 and Spring 1976 quarters. However, in the Winter quarter 1976 many students obtained high scores on the test. This could be interpreted as an inappropriateness of the test in measuring their initial achievements. Even so, the fact that each group was represented by students from all undergraduate levels and from various departments could be considered as providing a natural randomization. (5) the unavailability of a standardized test covering the material discussed during the first two weeks which could be used in place of the initial test. (6) the unavailability of a standardized test suited to the scope and content of the course which could be used for the final test.

7.2 Suggestions for the teaching of introductory statistics courses

At the commencement of this exploration to find ways to make an introductory statistics course more meaningful to the students and relevant to their various fields of study, the investigator assumed several objectives for the teaching of introductory statistics. The first was to introduce the students to variability and uncertainty of many kinds of measurements. The second was to introduce the
students to some common statistical formulas, terminologies, and some widely used statistical methods. The third was to lead the students to recognize the availability of many statistical techniques that can be helpful to them in solving many problems in their future careers. The results of the investigation have led him to believe that a systematic lecture method supplemented by preliminary questions, class experiments, computer programming, and student surveys would serve as a partial solution toward the achievement of those objectives.

The plenteous employment of the systematic lecture method in teaching statistics could very well indicate the advantages inherent in the method. One advantage is that it is the simplest way to present the material by which the instructor could prepare his lectures ahead of time, limit the scope of the material for each lecture, and keep up with the schedule. The second advantage is the availability of many textbooks which are normally oriented toward this method. These textbooks are generally written in a well organized manner, containing theories, proofs and derivations, as well as examples and exercises. The difficulties that an instructor might encounter as he attempts to make the course more relevant to many fields of study would be in matters such as the selection of a suitable textbook, examples, illustrations, and homework
assignments. It was toward this goal that this exploration was directed. The supplementary techniques employed in this study were directed toward the enhancement of the relevance of statistics to the students in their fields of study. The reader might recall some points presented earlier as to the advantages of the inclusion of the supplementary techniques, preliminary questions, class experiments, computer programming, and student surveys in the systematic lecture method. Preliminary questions confronted the students from the very beginning of the course with the relevance of statistics to many fields of scholarships; hence it helped to arouse students' interest in the subject. This technique provides ample opportunity for practice toward the mastery of some statistical procedures. Class experiments, which led the students to participate in class by collecting data related to themselves or to their field, provided them with the opportunity to observe the variability and uncertainty of the outcomes which were due to chance. The use of computer programming in the teaching of statistics helped enrich students' experience and develop an appreciation for the relationship between computer and statistics. The student surveys provided the students with the opportunity to collect data related to their fields of interest and analyze them statistically. Thus they also helped the students master the procedures.
7.3 A suggestion for the organization of Introductory Statistics Courses

Many authors and educators have expressed their concern for the organization of the teaching of statistics courses at the university level. Their concern was mainly directed toward the question "Who should teach the course?" Should it be taught by an Institute of Statistics? Or by a Department of Statistics? Or under an Interdepartmental Committee? Or scattered among the departments of application? It was the opinion of this investigator based upon his study that the teaching of basic statistics courses should be centralized in a department of statistics or mathematics. Especially in a new university where there are no established traditions, such an arrangement would appear to offer advantages of efficiency and opportunity to build a strong, balanced staff. He believes, however, that specialized applied courses should be given in the respective departments. His reasons for favoring this centralization of the basic courses are the instructional universality, instructional efficiency, instructional expertise, and economy. The instruction could be kept universal since the introductory statistics course is an appreciation course in which students are introduced to the philosophical background of the invention of statistical concepts, and to the broad scope of its application in other areas. Students
obtain instruction using the same notations, the same definitions, and the same formulas. The centralization could also be more efficient because the department of statistics could provide more intensive supplementary laboratory work where the students would obtain individual assistance from the staff member of the department. It would also help the individual departments to focus their attention upon the wider application of statistics in their own fields, rather than on the teaching of some basic statistics. It is more likely for the statistics department to provide faculty members with instructional expertise, who have a thorough knowledge of the subject, with good mathematical background, and who are familiar with recent advances and current developments in statistical theories. They also have made contributions to the advancement of statistics, and are acquainted with the history of statistics and the philosophical background of statistical concepts. They could also be knowledgeable in the theory of learning and methods of teaching, including the provision of attractive instructional aids, and have genuine interest and understanding of the many applications of statistics. It would probably also be financially more economical if the teaching of basic statistics was centralized on account of the fusion of small numbers of students from various departments into a larger group which could be
taught by one faculty member.

In the light of the centralization of basic statistics courses and their relevance to the students, it would be worth mentioning the idea of clustering the students from several departments having similar statistical objectives. The following is an example of a possible clustering of students in terms of their fields of interest:

(1) Education, Psychology, Physical Education, and Agricultural Education
(2) Political Science, Sociology, Social Work
(3) Business and Economics
(4) Biological Sciences, Dairy Science, Agriculture, Agronomy, Poultry Science, Animal Science, Zoology, and Genetics
(5) Nursing
(6) Preventive Medicine, Pharmacy, Medical Science, Veterinary Science, Pre-medicine
(7) Engineering
(8) Physics, Mathematics, and Statistics
(9) Introductory Statistics for Graduate students without calculus prerequisite
7.4 Recommendations

In this section the investigator wishes to discuss some follow-up problems for future investigations relative to each of the techniques he employed in his study. During his explorations, the investigator has compiled several questions to be used in the Preliminary Questions set. These questions have served the purpose and have been shown to work well. However, to compile a larger and better variety of questions would be desirable for future investigations. He wishes to point out some features of a typical preliminary questions problem which he found very useful: it should reflect a certain area of study, relate or convey a certain concept, be practical, easy to understand, appealing to the students in such a way that they would be inclined to give the answer immediately either intuitively or by guessing, while the correct solution would need further mathematical verifications.

Similar suggestions go for the collection and selection of class experiments. These experiments are intended to help the students understand the concepts by doing. A typical class experiment should be simple, it needs a little time to perform, and if possible without any cost. It should also be related to a certain concept; in fact it should be designed to explain the concept.
At present, many computer programs are available which can be used by students even without prior instruction in computer science. However, these programs are mainly designed to be used as tools in statistical computations, rather than as learning aids to study introductory statistics. The investigator feels it would be very valuable for future investigations to design pedagogically-oriented programs as a supplement to the systematic lecture, by which beginners learn statistics while working with the programs.

Finally, the investigator wishes to present his suggestions on the student surveys. Again, the suggestions are directed toward the collection and selection of the topics for the surveys. A typical student survey should be simple to conduct, yet it should involve the utilization of several statistical procedures. It should be educative, arouse interest, and it should motivate the students to study further the applications of statistics to their fields.

The type of investigation conducted for this dissertation could be continued by any person interested in finding ways to make an introductory statistics course meaningful and relevant to the students, especially those who do not have calculus background. The investigator realizes the limited scope of this study; the problems in
the Preliminary Questions set and the topics for the student surveys have been selected from several areas, but many more fields of study could be reflected in those problems.

The investigator wishes to suggest that other teachers use the techniques employed in this investigation with some modifications in order to suit their teaching strengths and meet the needs of the class. It would also seem relevant at this time to point out the implications of this study for the exploration into the teaching of high school statistics; it would seem sensible that more time could be provided for statistics instruction with emphasis on the activities rather than the lectures. It is hoped that textbook authors would consider modifying the usual, standardized approach to the presentation of the materials in view of the ideas. Finally, the investigator hopes that the ideas contained in this dissertation will be welcomed by statistics educators and will interest persons concerned with the teaching of statistics and with its adaptation to the dynamic quality of the development of the subject.
Appendix A

List and outlines of courses in statistics and courses related to statistics offered at The Ohio State University in 1975-76
Most of the course outlines listed were obtained through interviews and discussions with the professors teaching those courses, and some were quoted from The Ohio State University Bulletin, Book 3, 1974-75, 1975-76.
GENETICS

Gen 650 (5) ANALYSIS AND INTERPRETATION OF BIOLOGICAL DATA I

Methods analyzing biological data including: descriptive statistics, normal distribution, Central Limit Theorem, testing hypothesis about the mean and the variance, the use of t-test, F-test, Chi-square test, one way and two way analysis of variance, linear regression, correlation, probability of type II error, relation between Z, t, X², and F distributions.

Gen 651 (5) ANALYSIS AND INTERPRETATION OF BIOLOGICAL DATA II

Methods used in analyzing data classified in two or more ways including: Latin Square, split-plot and factorial designs, analysis of covariance, data transformations, multiple regression and least squares.

Gen 832 (5) MATHEMATICAL GENETICS

Application of statistics to genetics. The construction of mathematical models, use of path coefficients, least squares and maximum likelihood methods for estimating genetics parameters and breeding values in quantitative genetics.

Note: Gen 650 uses 29 CAI Modules using Coursewriter III, version 3; 9 exercise generating programs using PL/I and FORTRAN IV; 21 analysis programs written in CPS language. This course is offered twice a year with an average enrollment of 200 per year.
The following courses need basic statistics courses as prerequisites because these courses utilize such statistical concepts as: variance-covariance in general, correlation, regression, multiple regression, selection procedure, design and analysis of experiments, estimation, confidence intervals, etc.

An Sc, Poul Sc and D.S. 720 GENETICS OF ANIMAL POPULATIONS

An Sc, Poul Sc and D.S. 820.01 SELECTION INDEX THEORY AND PRACTICE

An Sc, Poul Sc and D.S. 820.02 NON ADDITIVE GENETICS EFFECTS

Note: in addition to the computer programs available at The Ohio State University, those courses also use L S M L G P or Least-Squares and Maximum Likelihood General Purpose program designed by Dr. Harvey in 1968.
PREVENTIVE MEDICINE

Prev Med 764 (5) DESIGN OF BIOMEDICAL INVESTIGATION

Design of studies in biomedical area, formulation of hypothesis, sampling, planning observations and measurements, selection of statistical techniques, testing hypothesis. Also discussed in the course are concepts such as: frequency distribution, measure of location, measure of dispersion, sets, events, sample space, probabilities, normal curves, sampling distribution, estimation, testing hypothesis, power function, decision theory, t-distribution, Binomial and Poisson distributions.

Prev Med 785 (5) BIOSTATISTICS AND COMPUTERS IN MEDICAL RESEARCH

Review of the fundamental concepts in biostatistics including more complex analysis of variance designs, integrated with the electronic computer. The concepts discussed include: Poisson process, estimation of lambda, multinomial distribution and application, ANOVA for fixed and random effect model - one and two ways, multiple regression, correlation, multivariate analysis, and introduction to time series.

Note: Sampling techniques are introduced in Prev Med 861, INDUSTRIAL HYGIENE.
The idea is to provide a conceptual understanding of statistical methods for the analysis of political data.

**Polit Sc 685 (5) METHOD OF QUANTITATIVE ANALYSIS: Elementary**

Explication, interpretation, and application of techniques for quantitative analysis of political data; descriptive and inferential statistics, with emphasis on bivariate analysis. Concepts discussed include: central tendencies, dispersion, distributions, z-score, bivariate table, rank order correlations, Kendall's Tau, Gamma, Somer's d, linear models, correlation, and preliminary multivariate analysis.

**Polit Sc 686 (5) METHOD OF QUATITATIVE ANALYSIS: Intermediate**

Descriptive and inferential statistics with emphasis on multivariate analysis; possible scaling, index construction, sampling and measurement reliability; additional topics offered as desired. Multivariate analysis includes: bivariate, multiple regression, partial correlation, and causal interpretations, standardizing, multiple least squares, Beta coefficient and multiple correlation $R^2$. Also discussed: sampling distribution, hypothesis testing, point and interval estimation, Chi-square test, significance of r.

**Polit Sc 786 CAUSAL ANALYSIS and Polit Sc 787 DIMENSIONAL ANALYSIS**

These courses use numerous statistical techniques including multidimensional scaling and factor analysis.
SOCIOLOGY

Sociol 650 (5) INTRODUCTION TO QUANTITATIVE RESEARCH TECHNIQUES IN SOCIOLOGY

An introduction to the analysis of sociological data; measurement theory and techniques of integration; sampling procedures in sociological research and implication for inference and generalization. Concepts discussed include: central tendency, variability, distribution shape, probability and sampling, statistical estimation, hypothesis testing, one way ANOVA, Spearman correlation, Chi-square test, Phi-coefficient, Kendall's Tau, Goodman and Kruskall Gamma, Somer's d, Guttman's lambda, linear regression and correlation.

Sociol 651 (3) APPROACHES TO SOCIOLOGICAL INQUIRY

Theory and practice in the essentials of the research process, comparison of alternative and design models, questionnaire construction, interview techniques and related problems.

Sociol 706 (3) EXPERIMENTAL RESEARCH METHODS

Survey and analysis of research designs and statistical techniques permitting control and/or assessment of error variance in sociological research by experimental methods.

Sociol 707 (5) PROBLEMS IN QUANTITATIVE ANALYSIS

A survey of advanced problems in the multivariate analysis of sociological data, causal inference in nonexperimental research and path analysis.

Sociol 754 (5) DEMOGRAPHIC ANALYSIS

Census data, vital statistics, demographic rates, life tables, Cohort analysis and similar elementary techniques and data sources in demography.

Sociol 850 SEMINAR IN SOCIOLOGICAL RESEARCH METHODS

Special topics seminars in research methodology, correlation regression, ANOVA, ANCOVA, factor analysis, nonparametric techniques, etc.
EDUCATION

Ed Dev 785 INTRODUCTION TO INQUIRY, PRINCIPLES, STRATEGIES, AND TECHNIQUES

The emphasis is on the conceptualization of educational problems and related statistical techniques. This includes: experimental design, randomization, internal and external validities, quasi experimental design, statistical tests, reliability, item analysis, survey, observation and data collection, and some computer programming.

Ed Dev 786 INTRODUCTION TO INQUIRY: QUANTITATIVE METHODS

An introduction to quantitative techniques with emphasis on the application in educational settings. This has more technical statistics than Ed Dev 785.

Ed Dev 808 EXPERIMENTAL DESIGN IN EDUCATION I

An examination of logical and quantitative principles, especially the analysis of variance for two and three factor completely randomized design, with equal and unequal observations per cell, multiple comparison, post hoc test, etc. underlying basic experimental design employed in educational research. Some computer programs are also used.

Ed Dev 809 EXPERIMENTAL DESIGN IN EDUCATION II

An examination of intermediate quantitative principles underlying experimental design in education, such as repeated measures design, hierarchial design for fixed, random, and mixed models, and the analysis of covariance.

Ed Dev 925.50 MULTIVARIATE ANALYSIS I

Matrix algebra.
Multiple regression: formulations, examination of residuals, moderator and suppressor variables, computer programming.
Two group discriminant function analysis: formulations, interpretation, classification. Linear regression and discriminant function in matrix term.
PSYCHOLOGY

Psych 220 (3) QUANTITATIVE AND STATISTICAL METHODS IN PSYCHOLOGY I

Elementary presentation of probability, descriptive and inferential statistics, and methods of measurements relevant to contemporary psychology

Psych 221 (3) QUANTITATIVE AND STATISTICAL METHODS IN PSYCHOLOGY II

A concentrated examination of the application of statistical tools in inference, and theory of construction in contemporary psychology

Psych 825 (4) STATISTICS IN PSYCHOLOGY I

Inferential statistics, basic concepts of sets, probability, distributions, and foundation of inference and estimation, special application to psychology

Psych 826 (4) STATISTICS IN PSYCHOLOGY II

Theoretical justification and uses of inferential techniques: t, Chi-square, and F distributions, correlation and regression, nonparametric techniques

Psych 827 (4) ANALYSIS OF VARIANCE

A coverage of statistical inference in ANOVA designs including randomized blocks, repeated measures, mixed models, and related contrast tests

Psych 828 (4) CORRELATIONAL ANALYSIS

Technique and rationale of using quantitative and qualitative data for prediction, test and battery analysis and validation. This includes simple regression analysis, Pearson correlation, test for correlation, composite and partial correlations, multiple regression, stepwise regression, Wherry-Gaylord method, item analysis and selection, factor analysis, canonical correlation
PSYCHOLOGY (cont.)

Psych 829 (4) QUANTITATIVE FOUNDATION OF PSYCHOLOGICAL STATISTICS

Principles and techniques for deriving statistical equations, their modification to handle special cases, clarifying assumptions and their application

Psych 831 (2) SEMINAR IN PSYCHOLOGICAL STATISTICS

a. Analysis of variance (Su)
b. Experimental Design (W)
c. Factor analysis (W)
d. Mathematical models and theory (Su)
e. Nonparametric statistics (Sp)
f. Advanced experimental design
g. Advanced multivariate analysis
h. Computer simulation research
i. Current practices and trends (Su)

Psych 833 (3) STATISTICAL PROBLEMS IN DEVELOPMENTAL PSYCHOLOGY

A bridge from formal statistics to current research in developmental and educational psychology with repeated measure, time series data, indices of change, etc.

Psych 815 DECISION PROCESS

Introductory course in the application of statistical decision theory

Note: The program used in some of those courses is "Ohio State University Program for Psychology" by Wherry and Wherry
CIVIL ENGINEERING

CE 405 (4) OBSERVATIONAL ANALYSIS

Theory and application of analysis;
Experiment, outcome, sample space, and random variables
Probabilities: joint, marginal, and conditional independence
Distributions: Bernoulli Geometric, Binomial, Poisson, exponential, normal, lognormal, uniform
Chi-square and Kolmogorov Smirnov goodness of fit tests, sampling, confidence intervals, and theory of estimations

Note: Statistics 525 is needed in CE 874 (4), URBAN TRANSPORTATION NETWORK ANALYSIS.
The following courses are not statistics courses but they require statistics courses as prerequisite

<table>
<thead>
<tr>
<th>Course</th>
<th>Prerequisites</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CIS 652 MODELING OF INFORMATION SYSTEM</strong></td>
<td>Stat 425, 426, or Stat 521</td>
</tr>
<tr>
<td><strong>CIS 730 BASIC CONCEPT IN ARTIFICIAL INTELLIGENCE</strong></td>
<td>Stat 520, 521</td>
</tr>
<tr>
<td><strong>CIS 735 STATISTICS METHODS IN PATTERN RECOGNITION</strong></td>
<td>Stat 520</td>
</tr>
<tr>
<td><strong>CIS 760 SELECTED TOPICS IN MATHEMATICS OF INFORMATION HANDLING</strong></td>
<td>Stat 520, 521</td>
</tr>
<tr>
<td><strong>CIS 765 THEORY OF MANAGEMENT INFORMATION SYSTEM</strong></td>
<td>Stat 520</td>
</tr>
<tr>
<td><strong>CIS 812 COMPUTER AND INFORMATION SCIENCE RESEARCH METHODS</strong></td>
<td>Stat 520, 521</td>
</tr>
<tr>
<td><strong>CIS 835 SPECIAL TOPICS IN PATTERN RECOGNITION</strong></td>
<td>Stat 520, 521</td>
</tr>
</tbody>
</table>
AGRICULTURAL EDUCATION

Ag Ed 886 RESEARCH DESIGN

Development of effective design for research problems in vocational, technical, and extension education, including theory, models, and sampling. This includes: internal and external validity, control of extraneous variables, choosing and developing appropriate designs, appraisal of designs in educational research, correlational research, ex post facto research, survey research using sampling methods.

Ag Ed 887 ANALYSIS AND INTERPRETATION OF DATA

Application and interpretation of descriptive and inferential statistics for research in vocational, technical, and extension education including the use of computer. This includes: scale of measurement, measures of relationship, computer program for descriptive statistics. Also included are: point and interval estimation, testing hypothesis, parametric and nonparametric, simple and multiple correlation, and computer programming.

Note: Computer programs used in those courses are: BMD, NPAR, SPSS, for the need of phi-coefficient, contingency, Cramer's statistics, Spearman rank correlation coefficient, Kendall's Tau, Pearson correlation coefficient, t-test, Mann-Whitney U test, Wilcoxon Signed test, Chi-square test, and Mac Nemar test.
AGRICULTURAL ECONOMICS

Ag Ec 801 SEMINAR IN PROBLEMS IN AGRICULTURAL ECONOMICS

STATISTICS

Prerequisite: Ec 442 ELEMENTARY ECONOMICS STATISTICS
Ec 641 APPLIED REGRESSION AND CORRELATION ANALYSIS

Discussion in the seminars pertain to the application of statistics to problems in Agricultural Economics

AGRONOMY

Agron 887 TECHNIQUES AND EXPERIMENTAL DESIGN

A study of experimental design and their application to agricultural research. This includes: the basic designs as completely randomized design, randomized block design, Latin squares, split-plot design, balanced and partially balanced incomplete block design, factorial experiments, confounding and analysis of data obtained in actual research problems.
VETEROINARY PHYSIOLOGY AND PHARMACOLOGY

Vet Phys 640 DESIGN AND ANALYSIS OF COMPARATIVE BIOMEDICAL RESEARCH I

Theory and application of basic statistical concepts as they affect design, analysis and interpretation of veterinary medical research

Note: This is very similar to Prev Med 764 except for the emphases on several topics

Vet Phys 641 DESIGN AND ANALYSIS OF COMPARATIVE BIOMEDICAL RESEARCH II

This is the continuation of Vet Phys 640, but more emphasis is given on the application of statistics to problems related to students' research
ECONOMICS

Econ 442 (5) ELEMENTARY ECONOMICS STATISTICS
Descriptive statistics, discrete probability and
application including Bayes law, probability dis-
tribution, estimation of parameters, testing hypo-
thesis, stochastic relationships, two variable
model and decision analysis

Econ 640 (4) PROBABILITY AND STATISTICAL THEORY (DECISION)
The theory of probability and stochastic processes,
statistical inference, test of significance and
analysis of variance, statistical decision theory

Econ 641.01 REGRESSION ANALYSIS I
Methods of estimation and properties of estimators,
two variable least squares and bivariate normal
correlation, heteroscedasticity, serial correlation,
errors in variables and problems of measurement.
Multivariate regression, multiple and partial corre-
lation, multicollinearity, analysis of direct and
indirect effects, models with dummy variables

Econ 641.02 REGRESSION ANALYSIS II
Introduction to economic theory, matrix methods,
statistical tools, least squares and standard linear
model, partial and multiple correlation and analysis
of residuals, generalized least squares and linear
constraints, the combination of several linear rela-
tions and asymptotic distribution theory

Econ 740 (4) INERENCE AND DECISION ANALYSIS UNDER
UNCERTAINTY
Distribution theory, point and interval estimation,
statistical hypothesis testing, decision analysis
under uncertainty

Econ 741 (4) GENERAL LINEAR REGRESSION ANALYSIS
Multiple regression analysis, the general linear
models, nonlinear and distribution lag models

Econ 742 (4) ECONOMETRICS
Review of the general linear models, identification,
estimating criteria, single and simultaneous equa-
tion estimation, econometric application

Econ 842 ADVANCED ECONOMETRICS
Theory and application of advanced quantitative
research method, computerized application to
econometrics method developed in Econ 742
BUSINESS ADMINISTRATION

Bus Adm 390 DECISION THEORY I

Examine the use of statistical tools in the managerial decision making process, pay off table and decision trees, sampling and hypothesis testing to decide on appropriate courses of action. This includes a review of discrete and continuous probability distributions, sampling distributions, and estimation.

Bus Adm 490 (2) DECISION THEORY II

The application of statistical methods to design and analysis of experiments with a view on planning, organizing, and controlling the output of industry.

Bus Adm 601 (5) BUSINESS STATISTICS

Price and production indices, analysis of time series, linear correlation applied to economic and business problems.
GEODETIC SCIENCE

Geod Sc 650 ADJUSTMENT COMPUTATION I

Classification of errors, measures of dispersion, variance covariance, propagation of errors, weights, observation, condition and normal equations and examples

Geod Sc 651 ADJUSTMENT COMPUTATION II

A posteriori variance, representation of residuals, combination of observation and condition equations, generalized minimum variance solution for hybrid measuring systems, statistical tests, empirical fitting of polynomials

Geod Sc 762 ADVANCED ADJUSTMENT COMPUTATIONS

Analysis of mathematical models, systematic errors, correlations, inner adjustment, multivariate statistical analysis, generalized matrices in adjustment

Note: Usually the students are required to take statistics courses, Stat 520, 521, 645 (alternatively Stat 525) before taking Geod Sc 650, 651, and 762. Additional statistics courses that the students usually take are Stat 632, Stat 641, Stat 742, and Stat 746
Normally the students are required to take such statistics courses as Stat 425, 426, 623, or mathematical statistics Stat 520, 521. Also Stat 645 or 641, 661, and applied multivariate analysis. The following are applied statistics courses offered by the Department of Industrial and Systems Engineering

**ISE 406 INDUSTRIAL QUALITY CONTROL**

The application of probability theory, statistics and control theory to problems in product inspection and process control, economic evaluation of quality control technique

**ISE 653 ENGINEERING DATA ANALYSIS**

Techniques for estimating parameters, calculating confidence intervals, and estimating probabilities from various non-normal probability distributions. Distributions include Weibull, Gamma, Lognormal, Beta, Pearson Types I-XII, Poisson, Binomial and negative binomial. Graphical techniques for estimation, and goodness of fit tests

**ISE 750.03 TECHNIQUES OF INDUSTRIAL ENGINEERING EXPERIMENTAL DESIGN**

Mathematical analysis and experimental designs involving industrial empirical investigations. Block designs, factorial experiments, orthogonal contrasts and multiple comparison methods, response surface designs in the study of production performance

**ISE 750.09 FORECASTING AND ESTIMATION**

A variety of forecasting and estimation technique with primary emphasis on forecasting. Regression analysis, auto correlation, exponential smoothing, and applications to engineering problems are included

**ISE 835 PRODUCT DEVELOPMENT EXPERIMENTATION I**

The study of useful models relating to the design of surveys and experiments in several phases of product development
ISE 836 PRODUCT DEVELOPMENT EXPERIMENTATION II

The study of application of multivariate models to the analysis of non-machine systems; topics include the multivariate normal distribution of quadratic forms, subject repeated designs-univariate and multivariate analysis, multivariate estimation confidence regions, profile analysis—application to human factors experimentation, multivariate analysis of variance—product experimentation.

The following are courses in ISE that utilize numerous probability theories and statistical techniques:

ISE 650 ANALYSIS FOR INDUSTRIAL ENGINEERING
ISE 652 ANALYSIS OF INVENTORY SYSTEM
ISE 655 SYSTEM RELIABILITY AND AVAILABILITY
ISE 750.05 DECISION THEORY
ISE 750.08 SIMULATION OF COMPLEX SYSTEMS WITH CONTINUOUS STATE MODELS
ISE 813 ADVANCED QUEING THEORY
ISE 814 STOCHASTIC PROCESSES USED IN SYSTEMS ENGINEERING
ISE 815 ESTIMATION OF SYSTEM PARAMETERS FROM THE TIME SERIES DATA
ISE 842, 843, and 844 OPERATION RESEARCH I, II, AND III
ISE 862 BAYESIAN DECISION THEORY (application)
ISE 861 RESEARCH IN DECISION PROCESSES
ISE 694N QUASI EXPERIMENT
SOCIAL WORK

Soc Work 680 RESEARCH METHOD IN SOCIAL WORK

Prerequisite to this course is Soc Work 380, INTRODUCTION TO RESEARCH METHODS IN SOCIAL WORK, and Stat 125 or equivalent. This include: formulating researchable questions and hypothetical answers for testing, measurement, research design, data collection and processing, statistical analysis, and reporting results. Computer programs are also used in this course.

SPEECH COMMUNICATION

Communic 701 RESEARCH METHODOLOGIES IN SPEECH COMMUNICATION

Predominant research methodologies in the field of speech communication

Communic 701.01 STATISTICAL CONCEPTS AND LOGIC

Communic 701.03 EXPERIMENTAL DESIGN

PHYSICAL EDUCATION

Phys Ed 845 STATISTICS FOR PHYSICAL EDUCATION AND HEALTH EDUCATION

A study of techniques for interpreting research publications in the field, and of statistical procedures useful in analyzing data. It includes analysis of variance, regression analysis and multiple regression analysis.
BIOSTATISTICS

Biostat 601 STOCHASTIC PROCESSES IN THE BIOLOGICAL SCIENCES

Introduction to discrete stochastic processes, random walk, Markov Chains, birth and death processes, epidemic process, processes for competing among species, diffusion processes, and applications. Prerequisite: Stat 520

Biostat 605 POPULATION DYNAMICS

Study of birth, death, and growth process, use of rates and rations, force of mortality, competing risks, and selected epidemiological problems

Biostat 607 BIOSTATISTICS LABORATORY

Experience in statistical design and analysis of biomedical studies through individual association with active research workers in medicine

Biostat 610 STATISTICAL BIOASSAY I

Direct assays, dose-response relationships, parallel line and slope ratio assay, special statistical designs in assay, Bayesian bioassay; examples

Biostat 611 STATISTICAL BIOASSAY II
Continuation of Biostat 610

Biostat 800 ADVANCED TOPICS IN BIOSTATISTICS I

Biostat 801 ADVANCED TOPICS IN BIOSTATISTICS II

Biostat 802 ADVANCED TOPICS IN BIOSTATISTICS III

Biostat 999 RESEARCH IN BIOSTATISTICS
STATISTICS

Stat 123 STATISTICS FOR THE BUSINESS, SOCIAL, AND BIOLOGICAL SCIENCES

Stat 125 ELEMENTARY MATHEMATICAL STATISTICS

Stat 180 CHOICE AND CHANCE

Stat 194 GROUP STUDIES

Stat 223 ELEMENTARY DECISION THEORY

Stat 421 INTRODUCTION TO STATISTICS

Stat 425 and 426 PROBABILITY AND STATISTICS I AND II

Stat 494 GROUP STUDIES

Stat 505 and 506 INTRODUCTION TO ANALYSIS, PROBABILITY AND STATISTICS I AND II

Stat 518 and 519 STATISTICAL THEORY IN MEDICAL RESEARCH I AND II

Stat 520 and 521 MATHEMATICAL STATISTICS I AND II

Stat 525 STATISTICAL METHODS

Stat 528 and 529 DATA ANALYSIS I AND II

Stat 532 DISCRETE PROBABILITY

Stat 593 INDIVIDUAL STUDIES

Stat 594 GROUP STUDIES

Stat 600 STATISTICS LABORATORY

Stat 620 and 621 STATISTICAL THEORY I AND II

Stat 623 MATHEMATICAL STATISTICS SUPPLEMENT

Stat 632 APPLIED STOCHASTIC PROCESS

Stat 635 STATISTICAL ANALYSIS OF TIME SERIES

Stat 641 LINEAR MODELS

Stat 645 APPLIED REGRESSION AND DESIGN
STATISTICS
(cont.)
Stat 651 SURVEY AND SAMPLING THEORY
Stat 656 APPLIED MULTIVARIATE ANALYSIS
Stat 661 APPLIED NONPARAMETRIC STATISTICS
Stat 671 SIMULATION AND MONTE CARLO TECHNIQUE
Stat 693 and 694 INDIVIDUAL AND GROUP STUDIES
Stat 725 SEQUENTIAL STATISTICAL METHODS
Stat 742 ANALYSIS OF VARIANCE
Stat 746 DESIGN AND ANALYSIS OF EXPERIMENT
Stat 755 and 756 MULTIVARIATE ANALYSIS I AND II
Stat 761 and 763 NONPARAMETRIC STATISTICS I AND II
Stat 777 OPTIMIZING METHODS IN STATISTICS
Stat 794 GROUP STUDIES
Stat 821 and 822 STATISTICAL INERENCE I AND II
Stat 824 STATISTICAL DECISION THEORY I
Stat 825 STATISTICAL DECISION THEORY II
Stat 828 RANKING, SELECTION, AND MULTIPLE DECISION
Stat 832 APPLIED PROBABILITY MODEL
Stat 847 ADVANCED DESIGN OF EXPERIMENT
Stat 834 STATISTICAL INERENCE FOR STOCHASTIC MODELS
Stat 881 and 882 ADVANCED TOPICS IN MATHEMATICAL STATISTICS I AND II
Stat 888 LARGE SAMPLE THEORY
Stat 999 RESEARCH
Appendix B

Complete course outline of Statistics 125
Stat 125

ELEMENTARY MATHEMATICAL STATISTICS

Course outline:

1. Introduction to statistics; objective, meaning and history of statistics

2. Empirical Frequency Distributions
   Frequency distributions, cumulative frequency distribution, graphic presentation, data and populations

3. Descriptive Measures
   Measures of location: mean, weighted mean, mode, median.
   Measure of variation: range, variance, and standard deviation

4. Elementary Probability
   The meaning of probability, addition principle, multiplication principle, computing probabilities, repeated independent trials, permutation and combination

5. Populations, samples, and distributions.
   Random variables, distribution of random variables, expected value and variance of random variables, normal distributions, Binomial distributions

6. Sampling distributions
   Sampling from normal population, the Central Limit Theorem

7. Estimation
   Point estimation, interval estimation
   Confidence intervals for normal means - known variance; confidence intervals for normal means - unknown variance; confidence interval for Binomial proportion p

8. Test of hypotheses
   Type I and type II error; hypothesis on normal mean (using a t-test); hypothesis on a Binomial p; connection between testing hypothesis and interval estimation
Course outline Stat 125 (continued)

9. Two sample technique: to test equality of two populations; confidence interval for \( \mu_1 - \mu_2 \) and relation with testing the equality of two means, means or population proportions

10. Approximate tests
   Test of goodness of fit
   Contingency tables

11. Simple linear regression and sample correlation coefficient
Appendix C

PRELIMINARY QUESTIONS
Please answer these questions as best as you can. The answers should be done on separate papers. The instructor will read your answers and return them to you at the end of the quarter for comparison. In this way you will be able to see how introductory statistics has helped you to improve your ability to analyze information. Keep the questions with you.

Note: For the "chance" questions, please give the estimate for example 10%, 20%, 70% chance, etc.
1. **True and False exam**

A student, unprepared for an examination, selects the answers to 100 True and False questions by flipping a coin. He puts True if the coin showed a Head and False if otherwise. In order to pass the test, a student should have at least 65 correct answers. What is the chance for this student to pass?

2. **An ESP problem**

An experimenter selects a card from a standard deck and notes the color red or black. A subject, who does not see the card, attempts to read the "experimenter's mind" and call the correct color. The experiment is repeated 40 times, the deck being shuffled before each trial. The subject makes correct identifications 25 times. Can you agree that the subject has an Extra Sensory Perception?

3. **A Magazine Salesman's Dream**

It is known from experience that 50% of all housewives who let a door-to-door magazine salesman into their homes will end up ordering a magazine. What do you think the chance that out of ten housewives who let such salesman into their homes, at least eight will end up ordering a magazine?
4. **An Airplane Overweight Problem**

If the average of all men traveling by air between Columbus and New York City is 150 lbs., and the standard deviation of 25 lbs., what is the chance that the total weight of 50 men traveling on a plane between the two cities will exceed the passenger safety limit of 8000 lbs.?

5. **Car Repair Expenses**

Insurance adjusters were concerned about the high repair estimates that they were getting from garage A. To verify their suspicions, each of ten damaged cars was taken to garage A to obtain a repair estimate and then to another, more reliable garage B. The results are given below in dollars:

<table>
<thead>
<tr>
<th>Car no.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Garage A:</td>
<td>210</td>
<td>450</td>
<td>630</td>
<td>300</td>
<td>120</td>
<td>540</td>
<td>730</td>
<td>950</td>
<td>990</td>
<td>760</td>
</tr>
<tr>
<td>Garage B:</td>
<td>250</td>
<td>380</td>
<td>590</td>
<td>280</td>
<td>130</td>
<td>500</td>
<td>750</td>
<td>860</td>
<td>900</td>
<td>780</td>
</tr>
</tbody>
</table>

Mean estimate of A is 568; Mean estimate of B is 542.
Do you agree that in general the estimates from garage A are higher than those from garage B?

6. **Eskimo's Body Temperature**

A doctor is interested in checking to see if the average body temperature of Eskimos is significantly lower than man's normal average, which is 98.6°F Fahrenheit.

After selecting ten Eskimos at random, the following temperatures (in degrees Fahrenheit) were recorded:
99.0 97.8 98.8 98.6 98.7 98.3 97.9 98.0
98.9 and 98.4
Do you think these results give evidence to reject the claim that the Eskimos, on the average, have normal body temperatures?

7. Which Paint is Better?

To compare the durabilities of two paints for highway use, samples of ten four inch lines of each paint were laid down across a heavily travelled road. The order was decided at random. After a period of time reflectometer readings were obtained for each line. The higher the reading, the better the reflectivity.
The results are as follows:
Paint A: 12.7 11.5 9.9 9.6 10.3
9.6 9.4 11.3 8.9 11.3 Average 10.45
Paint B: 9.4 11.6 9.7 10.4 6.9
7.3 8.4 7.2 7.0 12.1 Average 9.0
Do you think, in terms of reflectivity, paint A is in general better than paint B?

8. Birth by month

Various studies have been done on the season of birth by countries, states, climate regions, etc. Here is a sample of such data on 240 births in families in State A, classified by Birth-Month (see table on the next page). Based on this sample, can you
conclude that in the whole population of State A, the number of births are equally distributed among the 12 months?

Table for problem 8:

<table>
<thead>
<tr>
<th>Month</th>
<th>Number of births</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>23</td>
</tr>
<tr>
<td>February</td>
<td>24</td>
</tr>
<tr>
<td>March</td>
<td>21</td>
</tr>
<tr>
<td>April</td>
<td>19</td>
</tr>
<tr>
<td>May</td>
<td>17</td>
</tr>
<tr>
<td>June</td>
<td>16</td>
</tr>
<tr>
<td>July</td>
<td>14</td>
</tr>
<tr>
<td>August</td>
<td>17</td>
</tr>
<tr>
<td>September</td>
<td>19</td>
</tr>
<tr>
<td>October</td>
<td>23</td>
</tr>
<tr>
<td>November</td>
<td>24</td>
</tr>
<tr>
<td>December</td>
<td>23</td>
</tr>
</tbody>
</table>

Total 240

9 MPH-MPG problem

The following values were obtained to observe the relationship between the speed and fuel consumption on level roadway, for a certain car model.

<table>
<thead>
<tr>
<th>Mean Speed</th>
<th>Fuel Consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed (X)</td>
<td>MPG (Y)</td>
</tr>
<tr>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>25</td>
<td>19</td>
</tr>
<tr>
<td>30</td>
<td>20</td>
</tr>
<tr>
<td>35</td>
<td>20</td>
</tr>
<tr>
<td>40</td>
<td>21</td>
</tr>
<tr>
<td>45</td>
<td>20</td>
</tr>
<tr>
<td>50</td>
<td>19</td>
</tr>
<tr>
<td>55</td>
<td>21</td>
</tr>
<tr>
<td>60</td>
<td>18</td>
</tr>
</tbody>
</table>

Scatter diagram

Based on this sample, can you give the best prediction for the gas consumption if someone drives that model at 70 MPH?
10. More children?

The following data is obtained in a sociological survey on the number of children per family according to the educational level of the father.

<table>
<thead>
<tr>
<th>Educational level of the father</th>
<th>Number of children per family</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Grade school</td>
<td>16</td>
</tr>
<tr>
<td>High school</td>
<td>24</td>
</tr>
<tr>
<td>University</td>
<td>11</td>
</tr>
<tr>
<td>Grand total</td>
<td>300</td>
</tr>
</tbody>
</table>

*) read: there are 58 families out of 100 from "Grade School classification" who have more than 2 children

Based on this data can you conclude that the number of children per family depends on the educational level of the father?

Answers through mathematical calculations: (not given to students)

1) 0.0013  2) No ESP  3) 0.0547
4) 0.0023  5) No  6) No, they have just the normal temperature
7) No, they are equal
8) Yes, equal  9) about 18 MPG
10) Yes they are dependent
Questionnaire

Computer programming

1. How would you rate the experience working with computer program in this course?  good  fair  nothing new

2. Was the experience meaningful to you?
   yes  OK  no

3. How much time did you spend on it?  ......hours

Student survey project (if you performed one)

1. Did your project help increase your interest in learning basic statistics?  yes  no

2. Was it relevant to your study?  yes  no

3. What was the effect of your project on your general knowledge of statistics?  good  fair  no effect

4. Do you think your project will result in longer retention of the concept?  yes  no  may be

5. How much time did you spend on the survey, including the preparation and making report?  ......hours

6. Do you have any suggestion or comment about the survey?

Preliminary questions (only for 12:00 class)

1. Has the preliminary questions set directed your attention to the relevance of statistics to various fields of study?  yes  no

2. Were the problems challenging enough to encourage you to think about the solutions as you studied various concepts?  yes  no

3. How did these questions affect your attitude toward statistics in terms of its usefulness in real life problems?  positively  negatively  just the same as before I took the course

4. What do you think of these problems?  difficult  fair  easy
5. Did you notice a difference in your ability to solve these problems at the beginning of the quarter versus the end of the quarter?  

  yes    no

6. Do you have any general comment about the preliminary questions set?
Appendix D

CLASS EXPERIMENTS AND RESULTS
Experiment number 1: **Empirical Probability**

**Treatment group:**

Objective: 
a. to explain the concept of frequency table and relative frequency 
b. to obtain empirical probability

Two trials:
1. flipping a coin twice. Each student was asked to flip a coin twice and record the results as the number of Heads appeared. The results of the whole class were tabulated on the board. This was followed by a class discussion on the meaning of relative frequency and the construction of empirical probability.
2. flipping a coin ten times. Procedure was the same as trial 1.

Home work assignment: complete the table and answer the questions on the following page.

**Control group:**

A frequency table similar to the one obtained by the treatment group was given to the control group along with the discussion on how to find the empirical probability.

**Checking the results:**

After two days a test was given to both groups. Their test was graded and the results were compared.
Form for experiment number 1

Let X be the number of Heads appeared in "flipping the coin twice" repeatedly. Tabulate the results in the following frequency table:

<table>
<thead>
<tr>
<th>Number of Heads</th>
<th>Freq.</th>
<th>Relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. complete the column of relative frequency

b. find the following empirical probabilities

\[
P(X=0) = \quad P(X=2) = \quad P(X \leq 5) =
\]

\[
P(X \geq 0) = \quad P(X \geq 2) = \quad P(X=12) =
\]

Let Y be the number of Heads appeared "in flipping a coin ten times" repeatedly. Tabulate the results in the following frequency table:

<table>
<thead>
<tr>
<th>Number of Heads</th>
<th>Freq.</th>
<th>Relative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. complete the relative frequency column

b. Find the following empirical probabilities

\[
P(Y=0) = \quad P(Y=1) = \quad P(Y \leq 2) =
\]

\[
P(Y \geq 0) = \quad P(Y \geq 2) =
\]
Test following experiment number 1

Name: ........................

1. Let $X$ be the number of Heads appeared in flipping a coin twice by 35 students

<table>
<thead>
<tr>
<th>$X$</th>
<th>Freq.</th>
<th>Rel. freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>8</td>
<td>a. complete the table</td>
</tr>
</tbody>
</table>

b. Find the following empirical probabilities from the table

$$P(X=3) = \quad P(X \geq 1) = \quad P(X < 1) = \quad P(X > 0) =$$

2. Let $W$ be the number of Heads appeared in flipping a coin ten times by 30 students

<table>
<thead>
<tr>
<th>$W$</th>
<th>Freq.</th>
<th>Rel. freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>a. complete the table</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

b. Find the following probabilities from the table

$$P(W=15) = \quad P(W \geq 7) = \quad P(W < 12) =$$

$$P(W \leq 4 \text{ or } W > 8) =$$
Result of the test following experiment number 1

Treatment group:
8 4 5 6 8 5 6 6 9 4
7 8 6 7 4 7 3 2 6 4
6 6 7 4 6 8 8 5 4 8
9 9 8 5 6 8 7 6 7

Control group:
4 8 10 4 8 4 5 3 6 4
6 4 6 6 .3 6 6 8 4 4
2 6 4 4 4 5 4 5 7 5
2 8 3 4 8 6 7

<table>
<thead>
<tr>
<th></th>
<th>Treatment group</th>
<th>control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>39</td>
<td>37</td>
</tr>
<tr>
<td>mean</td>
<td>6.21</td>
<td>5.22</td>
</tr>
<tr>
<td>variance</td>
<td>3.06</td>
<td>3.51</td>
</tr>
</tbody>
</table>

combined $s^2 = 3.28$

$t = 2.4$ which is significant at $\alpha > .8\%$ for testing the equality of the means

$F = 1.21$ which is significant at about 4% for testing the equality of the variances.
Experiment number 2: Random Variable and Probability

Treatment group:

Objective: a. to introduce the meaning of random variable X that can be defined according to a certain need. In this experiment, X is defined to be the sum of the numbers in each outcome if a die is rolled twice.
b. to introduce the meaning of "possible values of X"
c. to find the probability empirically
d. to find the probability mathematically

Trials: Each student was asked to roll a die twice and record the numbers for X. In order to obtain more observations, the "roll a die twice" was repeated three or more times. The results of the whole class were tabulated on the board, and a discussion followed.

Homework assignment:

a. to complete the table
b. to construct the empirical probability distribution
c. to find the mathematical probability distribution
d. to compare the two probabilities.

Control group:

Similar results were listed on the board followed by a discussion and a lecture on the empirical and mathematical probabilities.

Checking the results:

After two days, a test was given to both groups.
Form for experiment number 2:

If a die is rolled twice, then the sample space is:

\[
\begin{array}{cccccccc}
1,1 & 1,2 & 1,3 & 1,4 & 1,5 & 1,6 \\
2,1 & 2,2 & 2,3 & 2,4 & 2,5 & 2,6 \\
3,1 & 3,2 & 3,3 & 3,4 & 3,5 & 3,6 \\
4,1 & 4,2 & 4,3 & 4,4 & 4,5 & 4,6 \\
5,1 & 5,2 & 5,3 & 5,4 & 5,5 & 5,6 \\
6,1 & 6,2 & 6,3 & 6,4 & 6,5 & 6,6 \\
\end{array}
\]

Probability of each outcome is ........

Suppose we define a random variable \( X \) to be the sum of the numbers in each outcome of rolling a die twice, what would be the possible values of \( X \)?

Answer: ..............

Mathematically, the probability of each outcome can be listed as follows:

If a die is rolled twice and repeat the experiment many times with the following frequency distribution:

Find the empirical probability

\[
\begin{align*}
P(X=1) &= \\
P(X=2) &= \\
P(X=3) &= \\
P(X=4) &= \\
P(X=5) &= \\
P(X=6) &= \\
P(X=7) &= \\
P(X=8) &= \\
P(X=9) &= \\
P(X=10) &= \\
P(X=11) &= \\
P(X=12) &= \\
\end{align*}
\]

Can you find the following probabilities?

\[
\begin{align*}
\text{Mathematically} & \quad \text{Empirically} \\
P(X \geq 11) &= \quad \\
P(X=11) &= \\
P(X < 3) &= 143
\end{align*}
\]
Test following experiment number 2

Suppose a die is rolled twice and the number of dots facing up is recorded.

a) Construct the sample space (work it out below)

b) Suppose $X$ is the number of dots facing up in the first throw minus the number of dots facing up at the second throw, what are the values of $X$?

c) List all the possible values of $X$

d) List the frequency of each $X$

e) If the trial is repeated 500 times, how many times do you expect each $X$ to occur?

Answers:

a) and b)
Possible outcomes: 
Corresponding values of $X$ :


c) possible values of $X$ :

d) frequencies of the $X$'s:

e) expected freq. of the $X$'s :
Result of test following experiment number 2

Treatment group:

8 6 6 4 4 4 10 8 10 8
8 6 8 6 4 4 6 8 4 2
4 2 2 6 8 4 4 8 6 8
4 6 6 10 8 2 4 4 5

Control group

2 4 10 6 4 2 4 2 4 6
8 4 4 8 0 2 6 4 4 2
2 4 2 2 4 6 4 4 6 6
2 6 2 0 6 4 8

Analysis:

<table>
<thead>
<tr>
<th></th>
<th>Treatment group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>39</td>
<td>37</td>
</tr>
<tr>
<td>mean</td>
<td>5.33</td>
<td>4.16</td>
</tr>
<tr>
<td>variance</td>
<td>5.23</td>
<td>5.19</td>
</tr>
</tbody>
</table>

Combined $s^2 = 5.22$

$t = 2.24$ significant at $\alpha \leq 1.2\%$

$F = 1.007$ not significant

Conclusion: the means are significantly different
Experiment number 3: Central Limit Theorem

Treatment group:

Objective:

To introduce the concept of Central Limit Theorem as applied to the sample averages and the sample sums.

Discussion:

First discuss a population which was simulated by the result of rolling a die. Let $X$ be the number of dots appearing face up. Then it is known, or it can be calculated that $E(X) = \mu = 3.5$ and standard deviation $\sigma = 1.7$

Thus the mean and standard deviation are known.

Experiment:

Take a sample of size ten from this population by rolling a die ten times. Let each student do this experiment. Calculate the sample sum and the sample average. Construct the frequency distribution of the result of the whole class for both the averages and the sums. Draw the histogram for each and smoothen to form a curve. The students would see the similarity of the curve with the normal distribution curve. Ask the students to formulate some sort of rule pertaining to the distribution of the sample sums and the sample averages.

Control group:

Explain the Central limit theorem by stating the theorem and illustrate following the example from the textbook.

Checking:

At the end of the week, after some examples of computational procedures have been discussed, a test was given to see the performances of the students applying the Central Limit Theorem.
Test following experiment number 3:

An elevator has a sign saying its maximum safe load is 1700 lbs. If men's weights are assumed to follow a distribution with mean 160 and standard deviation of 20 lbs,

a. What is the distribution of the sum of the weight of ten men in the elevator if they are a random sample from this population.

b. What is the probability that ten men like in a. will overload the elevator?
Result of the test following experiment number 3:

Treatment group:
Scores 9 8 9 8 7 3 5 5 3 6
7 3 7 7 2 6 8 8 6 4
6 8 4 6 4 6 8 6 8 8
5 9 4 7 8 7 7

Control group:
Scores 7 6 6 4 4 5 8 7 8 1 8 8 7 4 5 6 6 8 4 5
4 5 4 3 6 1 3 8 3 5 3 2 7 8 8 1 7 7 6

Analysis:

<table>
<thead>
<tr>
<th></th>
<th>Treatment group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>37</td>
<td>39</td>
</tr>
<tr>
<td>sum</td>
<td>232</td>
<td>208</td>
</tr>
<tr>
<td>average</td>
<td>6.27</td>
<td>5.33</td>
</tr>
<tr>
<td>variance</td>
<td>3.59</td>
<td>4.65</td>
</tr>
</tbody>
</table>

Combined:

\[ s^2 = 4.14 \quad \quad s = 2.03 \]

\[ t = 2.02 \quad \text{which is significant (} \alpha = 2.2\% \text{) for testing the difference between the means} \]

\[ F = 1.22 \quad \text{which is not significant for testing the difference between the variances} \]

Conclusion: The differences are significantly different.
Experiment number 4: Normal approximation to Binomial

Treatment group

Objective: to explain the use of normal distribution to approximate binomial distribution for large n

Discussion: First discuss how to solve a rather difficult problem concerning Binomial experiment with large n. For example, let X be the number of successes in a Binomial experiment with 200 trials. Let the probability of success be .40. What is the probability of 70 or more successes?

Let the students try to solve this using Binomial formula to find out the difficulties in calculations.

Experiment: Each student in treatment group was asked to flip a coin 40 times (this took about 5 minutes) and present the result, the number of heads appeared, to be listed on the board. The result of the whole class were tabulated, the histogram was drawn and smoothened to see the shape of the curve.

Further discussion was held as to how to compute the mean, the variance, and how to apply the theorem.

Control group:

For the control group this concept was discussed according to the presentation in the text-book.

Note:
To save time during the experimentation, a form (see next page) was already provided.
Form for experiment 4

In a Binomial experiment \( B(X,n,p) \), where \( X \) is the number of successes, \( n \) is the number of trials and \( p \) the probability of success which is the same from trial to trial, the mean \( \mu = np \) and the variance \( \sigma^2 = np(1-p) \).

If \( n \) is very large (greater than 30), especially if \( p \) is close to 0.5 , then the number of successes \( X \) has a distribution which is approaching a normal distribution as \( n \) get larger.

Experiment: Toss a coin 40 times and record the result on the following scheme. Put number 1 for Heads and 0 for Tails.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
</table>

Frequency table for the result of the whole class

<table>
<thead>
<tr>
<th>#H</th>
<th>f</th>
<th>Rel. Freq</th>
<th>Rel. Freq</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-45</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-10</td>
<td>.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10-15</td>
<td>.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15-20</td>
<td>.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-25</td>
<td>.20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 10 20 40
A doctor knows, from previous experience, that 20% of all patients who are given a certain medication will have side effects of one kind or another. Find the probability that among 100 patients who are given this kind of medication, at least 10 but at most 30 will have side effects.

Answer: 98.76% by using a normal approximation but without continuity correction.
Result of the test for experiment number 4

Treatment group:

7 5 5 7 5 5 7 5 7 8 8 8 7 5 7 5 7 4 8
7 5 5 4 5 7 7 8 5 7 5 5 4 7 8 1 5 1 0 7

Control group:

7 5 5 7 4 5 4 4 7 4 5 7 7 5 4 5 7 5 4 4 4 7 5 5
7 7 5 4 5 7 5 1 0 4 4 5 5 7

<table>
<thead>
<tr>
<th></th>
<th>Treatment group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>sum</td>
<td>237</td>
<td>201</td>
</tr>
<tr>
<td>n</td>
<td>39</td>
<td>37</td>
</tr>
<tr>
<td>SS</td>
<td>104.77</td>
<td>74.91</td>
</tr>
<tr>
<td>mean</td>
<td>6.07</td>
<td>5.41</td>
</tr>
<tr>
<td>variance</td>
<td>2.76</td>
<td>2.08</td>
</tr>
</tbody>
</table>

Combined: \( s^2 = 2.43 \quad s = 1.56 \)

\( t = 1.85 \) significant at 3% for testing the means \( (\alpha = .032) \)

\( F = 1.25 \) significant at 5% for testing the variances
Experiment number 5: Two sample t-test

Objective: to test the equality of two population means with the assumption of equal variances and the populations are normal.

Treatment group:

Discuss two problems with this group
Problem 1: Which is taller on the average, boys or girls?
Problem 2: Define the ration as the quotient of weight in pounds over height in inches of each student. Which is larger on the average, the ratio for male students or female?

Collect the data from the students. List the result on the board and discuss

Assignment: read the section about the two sample t-test and answer the above problems.

Control group:

Discuss two problems with the data provided by the instructor.
Problem 1: GPA of boys and GPA of girls
Test the equality of the means with the proper assumptions.

Problem 2: Similar problem for the number of hours of study for boys as compared with that for girls.

Note: the students should have the intuitive answers to these problems. They only need to see how statistics may affirm these answers.
Data collected from treatment group

<table>
<thead>
<tr>
<th>Male</th>
<th>Height in inches</th>
<th>weight in pounds</th>
<th>ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>63</td>
<td>120</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>73</td>
<td>190</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>73</td>
<td>155</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>160</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>170</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>154</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>230</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>175</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>180</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>150</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>155</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>125</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>210</td>
<td>3.1</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>150</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>135</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>74</td>
<td>175</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>150</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>190</td>
<td>2.7</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Female</th>
<th>height in inches</th>
<th>weight in pounds</th>
<th>ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>62</td>
<td>105</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>119</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>128</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>135</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>100</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>128</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>135</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>120</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>96</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>128</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>120</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>130</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>137</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>200</td>
<td>3.1</td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>153</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>153</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>150</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>130</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>140</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>125</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>122</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>115</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>145</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>140</td>
<td>1.9</td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>130</td>
<td>1.9</td>
<td></td>
</tr>
</tbody>
</table>
Data provided for control group:

<table>
<thead>
<tr>
<th>GPA</th>
<th>Number of hours of study per day</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2</td>
<td>6</td>
</tr>
<tr>
<td>3.8</td>
<td>2</td>
</tr>
<tr>
<td>2.1</td>
<td>5</td>
</tr>
<tr>
<td>2.7</td>
<td>3</td>
</tr>
<tr>
<td>2.1</td>
<td>3</td>
</tr>
<tr>
<td>3.7</td>
<td>2</td>
</tr>
<tr>
<td>3.6</td>
<td>4</td>
</tr>
<tr>
<td>2.5</td>
<td>3</td>
</tr>
<tr>
<td>2.9</td>
<td>5</td>
</tr>
<tr>
<td>2.5</td>
<td>5</td>
</tr>
<tr>
<td>3.0</td>
<td>2</td>
</tr>
<tr>
<td>2.9</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GPA</th>
<th>Number of hours of study per day</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>5</td>
</tr>
<tr>
<td>2.8</td>
<td>8</td>
</tr>
<tr>
<td>2.9</td>
<td>6</td>
</tr>
<tr>
<td>3.0</td>
<td>8</td>
</tr>
<tr>
<td>3.0</td>
<td>8</td>
</tr>
<tr>
<td>3.0</td>
<td>2</td>
</tr>
<tr>
<td>3.2</td>
<td>3</td>
</tr>
<tr>
<td>3.7</td>
<td>3</td>
</tr>
<tr>
<td>3.1</td>
<td>3</td>
</tr>
<tr>
<td>2.5</td>
<td>5</td>
</tr>
</tbody>
</table>
Test following experiment number 5

In an experiment designed to compare the mean service life of two types of tires it was found that the difference between the means was 2000 miles. The pooled estimate variance is 160,000 miles. There were 20 tires of each type. Assume the mean service life can be considered from populations having equal variances, test the equality of the population means. Use 1% level of significance.

Answer: n=20

   calculated t = 50.7
   tabulated  t = 2.704

Conclusion: the difference is significant
Result of the test for experiment number 5

Treatment group:

\[
\begin{align*}
6 & 6 & 4 & 6 & 6 & 6 & 8 & 10 & 8 & 6 & 10 & 8 & 6
\end{align*}
\]

Control group:

\[
\begin{align*}
4 & 6 & 2 & 8 & 6 & 4 & 6 & 8 & 6 & 4 & 6 & 4 & 6 & 6
\end{align*}
\]

<table>
<thead>
<tr>
<th></th>
<th>treatment group</th>
<th>control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>sum</td>
<td>218</td>
<td>216</td>
</tr>
<tr>
<td>n</td>
<td>37</td>
<td>39</td>
</tr>
<tr>
<td>SS</td>
<td>119.57</td>
<td>147.68</td>
</tr>
<tr>
<td>mean</td>
<td>5.89</td>
<td>5.54</td>
</tr>
<tr>
<td>variance</td>
<td>3.32</td>
<td>3.88</td>
</tr>
</tbody>
</table>

Combined: \( s^2 = 3.61 \quad s = 1.90 \)

\[ t = .08 \] not significant for testing inequality of the means (\( \alpha = .47 \))

\[ F = 1.107 \] not significant i.e. the variances are not significantly different.
Experiment number 6: **Regression line**

**Objective:**

a. to see the relation of two variables by means of regression line

b. to predict the value of $Y$ if $X$ were known

**Treatment group:**

Each student in this group reported his or her height in inches and weight in pounds on a piece of paper - no name. The instructor collected the papers and picked some values at random. The values were plotted on the board, and at the same time, the students also made the graph of the values.

**Discussion:**

1. how to draw an eye-ball fit to those points and how to use the line for the purpose of prediction

2. How to find the best regression line. In this case, the least square method should be used.

**Control group:**

Discuss the regression line procedure using the example from the book

**Assignment:**

Find the regression line of the height vs weight of the students, using the data from the experiment. Predict the weight if the height is given.
Results of observation of treatment group:

<table>
<thead>
<tr>
<th>Height in inches</th>
<th>Weight in pounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>67</td>
<td>122</td>
</tr>
<tr>
<td>64</td>
<td>116</td>
</tr>
<tr>
<td>67</td>
<td>128</td>
</tr>
<tr>
<td>70</td>
<td>148</td>
</tr>
<tr>
<td>69</td>
<td>175</td>
</tr>
<tr>
<td>65</td>
<td>127</td>
</tr>
<tr>
<td>65</td>
<td>127</td>
</tr>
<tr>
<td>62</td>
<td>130</td>
</tr>
<tr>
<td>66</td>
<td>128</td>
</tr>
<tr>
<td>69</td>
<td>140</td>
</tr>
<tr>
<td>63</td>
<td>136</td>
</tr>
<tr>
<td>69</td>
<td>150</td>
</tr>
<tr>
<td>68</td>
<td>195</td>
</tr>
<tr>
<td>65</td>
<td>210</td>
</tr>
<tr>
<td>73</td>
<td>155</td>
</tr>
<tr>
<td>69</td>
<td>130</td>
</tr>
<tr>
<td>62</td>
<td>107</td>
</tr>
<tr>
<td>64</td>
<td>114</td>
</tr>
<tr>
<td>64</td>
<td>120</td>
</tr>
<tr>
<td>66</td>
<td>137</td>
</tr>
<tr>
<td>72</td>
<td>160</td>
</tr>
<tr>
<td>75</td>
<td>170</td>
</tr>
<tr>
<td>69</td>
<td>170</td>
</tr>
<tr>
<td>67</td>
<td>150</td>
</tr>
<tr>
<td>72</td>
<td>155</td>
</tr>
<tr>
<td>63</td>
<td>118</td>
</tr>
<tr>
<td>63</td>
<td>123</td>
</tr>
<tr>
<td>72</td>
<td>175</td>
</tr>
<tr>
<td>70</td>
<td>185</td>
</tr>
<tr>
<td>67</td>
<td>150</td>
</tr>
<tr>
<td>64</td>
<td>120</td>
</tr>
<tr>
<td>65</td>
<td>125</td>
</tr>
</tbody>
</table>

Total 32 students
Test following experiment 6

The following data is the result of an interview to five students pertaining the number of hours spent by each student in studying for the final examination.

X = the number of hours spent

Y = the score of the final examination out of 10
(to simplify the calculation)

<table>
<thead>
<tr>
<th>X (hour)</th>
<th>Y (Score)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
</tr>
</tbody>
</table>

a. Plot the points

b. Find the equation of the least square prediction line

2

c. Graph the line correctly

d. Using the graph, predict the score if some-one had studied for 2.5 hrs.

e. Using the line on b. predict the score of a student if he had studied for 4 hours. State your comment if you compare the result with the answer on the table.

e. State your comment about a student who had studied for 6 hours.

Score

(Y)

5

1 2 3 4 5 X (hour)
Result of the test following experiment number 6

Treatment group: 4 6 8 6 8 4 4 8 4 8 8 6 6 6 4 6 6 8 8 4 4 6 6 8 8 6 8 8 6 10 8 8

Control group: 8 8 6 6 6 6 8 6 6 4 6 8 6 4 6 6 8 10 6 4 4 2 6 6 8 4 6 6 8 8 4 6 6 8 6 4 4 6 6

<table>
<thead>
<tr>
<th>Treatment group</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>37</td>
</tr>
<tr>
<td>sum</td>
<td>238</td>
</tr>
<tr>
<td>average</td>
<td>6.43</td>
</tr>
<tr>
<td>variance</td>
<td>2.697</td>
</tr>
</tbody>
</table>

Combined: $s^2 = 2.66$    \[ s = 1.63 \]

\[ t = 1.02 \] which is not significant for testing the difference between the means

\[ F = 1.08 \] which is not significant, i.e. the variances are not significantly different.

Conclusion: The means are not significantly different.
Comparison of the Mean Effect of Class Experiments and Lecture Discussion in Explaining the Concepts 1, 2, and 4

Concept 1: Empirical probability
Concept 2: Random Variable and Probability
Concept 3: Central Limit Theorem
Concept 4: Normal Approximation to Binomial
Concept 5: Two sample t-test
Concept 6: Regression line

Group: $E_1 + E_2 + E_4$ (E: experiment)
Scores: 23 15 16 17 17 14 23 19 26 17 23 22
22 20 15 18 15 22 11 12 17 11 13 16
20 15 17 24 19 20 13 19 19 26 24 8
15 22 19

Group: $L_1 + L_2 + L_4$ (L: lecture)
Scores: 13 17 25 17 16 11 13 9 17 14 19 15
17 19 7 13 19 17 12 10 8 17 11 11
15 18 13 13 18 18 9 24 9 8 19 14
22

Analysis:

<table>
<thead>
<tr>
<th></th>
<th>Group I</th>
<th>Group II</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>39</td>
<td>37</td>
</tr>
<tr>
<td>sum</td>
<td>704</td>
<td>547</td>
</tr>
<tr>
<td>means</td>
<td>18.05</td>
<td>14.78</td>
</tr>
<tr>
<td>variance</td>
<td>18.79</td>
<td>19.95</td>
</tr>
</tbody>
</table>

Pooled estimate $s^2 = 4.399$
Variance ratio $F = 1.06 < F_{0.05}$ Thus the variances are equal

To test the equality of the means:
$t = 3.25 < t_{0.006}$

Conclusion: the means are significantly different.
Comparison of the Mean Effect of Class Experiments and Lecture Discussions in Explaining the Concepts 3, 5, and 6 (see page 162)

Group I : \( L_3 + L_5 + L_6 \) (L : Lecture discussions)

Scores:

\[
\begin{array}{cccccccccccc}
23 & 20 & 16 & 16 & 14 & 15 & 24 & 19 & 22 & 9 \\
16 & 24 & 17 & 14 & 19 & 18 & 24 & 24 & 18 & 15 \\
12 & 9 & 16 & 11 & 18 & 11 & 11 & 22 & 17 & 17 \\
13 & 16 & 19 & 20 & 20 & 9 & 17 & 17 & 18 & \\
\end{array}
\]

Group II : \( E_3 + E_5 + E_6 \) (E : Experiments)

Scores:

\[
\begin{array}{cccccccccccc}
19 & 18 & 23 & 20 & 23 & 11 & 15 & 21 & 11 & 18 \\
23 & 15 & 19 & 17 & 9 & 18 & 20 & 20 & 16 & 12 \\
20 & 20 & 14 & 18 & 14 & 16 & 18 & 18 & 22 & 22 \\
23 & 21 & \\
\end{array}
\]

Analysis:

\[
\begin{array}{ccc}
\text{Group I} & \text{Group II} \\
\hline
n & 39 & 37 \\
\text{sum} & 660 & 691 \\
\text{average} & 16.92 & 18.67 \\
\text{variance} & 18.32 & 17.98 \\
\hline
\end{array}
\]

Pooled estimate of \( s^2 = 18.397 \)

\( t = 1.78 \) with 74 d.f. \( \alpha = 0.0375 \)

Variance ratio \( F = 1.02 \) \( \alpha > 0.05 \) The variances are equal

Conclusion: The means are significantly different if \( \alpha > 3.75\% \)
Appendix E

INITIAL TESTS AND RESULTS
Test 1 Name: ........................
(initial test)

4 problems, 10 points each

1 The following is the distribution of the final examination grades which 150 students obtained in a course in statistics:

<table>
<thead>
<tr>
<th>Grades</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 19</td>
<td>18</td>
</tr>
<tr>
<td>20 - 39</td>
<td>42</td>
</tr>
<tr>
<td>40 - 59</td>
<td>57</td>
</tr>
<tr>
<td>60 - 79</td>
<td>24</td>
</tr>
<tr>
<td>80 - 99</td>
<td>9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a) Complete the table
b) Draw the histogram for this distribution
c) How many students had at least 60?
d) What percentage of students have grades at most 39?
e) Give your comment about the histogram

2 The following data show counts of yeast cells in 10 squares of hemacytometer (a device used to count blood cells and other microscopic objects)

1 10 8 1 3 4 6 5 6 6

Find a) the median b) the mode c) the mean
d) the range e) the variance

3 A survey was conducted by a forestry service crew yielded measurement of diameters of a sample of 100 oak trees of the same age in a certain area, in inches measures 3 feet above the ground, and that it is desired to construct a table for publication. The result:

<table>
<thead>
<tr>
<th>Class</th>
<th>Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 - 6</td>
<td>10</td>
</tr>
<tr>
<td>7 - 9</td>
<td>30</td>
</tr>
<tr>
<td>10 - 12</td>
<td>40</td>
</tr>
<tr>
<td>13 - 15</td>
<td>20</td>
</tr>
</tbody>
</table>

a. draw the histogram
b. calculate the mean
c. find the variance of the diameters of the trees.
4 It is given that $X = 4, 5, -2, 0$ and $Y = 5, -4, -2, 1$

Find the value of the following:

$$\sum_{i=1}^{4} X_i Y_i$$

$$\sum_{i=1}^{4} (X_i - 2)^2$$

$$\sum_{i=1}^{4} (X_i - Y_i)$$

$$\sum_{i=1}^{4} (X_j - j)$$

$$\sum_{i=1}^{4} (X_k + Y_k)^2$$
Initial test for Fall quarter 1975

Scores for group I
36 26 36 30 36 26 34 29 38 24 38 31 36 26 34
26 30 31 36 16 25 22 32 36 22 37 36 34 17
23 31 37 32 30 30 29 30 30

Scores for group II
37 27 27 34 33 34 30 34 24 34 20 31 33 25
26 38 31 32 24 20 38 36 26 33 34 36 27 32
34 19 36 24 20 33 24 29 27

Analysis:

<table>
<thead>
<tr>
<th></th>
<th>Group I</th>
<th>Group II</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>39</td>
<td>37</td>
</tr>
<tr>
<td>average</td>
<td>30.46</td>
<td>29.76</td>
</tr>
<tr>
<td>variance</td>
<td>32.467</td>
<td>30.25</td>
</tr>
</tbody>
</table>

Pooled estimate $s^2 = 29.78$

$t = 0.56 \quad \alpha > 5\%$

Variance ratio $F = 1.07 \quad \alpha > 5\%$

Conclusion: The initial achievement of the two groups are considered to be equal (based on this initial test)
Initial test for Winter quarter 1976

Scores for group I (12:00 class)
35 38 35 39 34 35 37 31 34 40 33 32 34 34
38 40 38 39 33 34 38 40 40 38 34 32 32 37
40 34 34 39 33 40 36 40 31 36 40 34 34

Scores for group II (3:00 PM class)
39 36 35 32 23 32 37 35 32 40 34 40 37 36
40 40 40 34 34 35 32 40 40 39 40 40 40 40
33 22 40 40 32 33 35 32

Analysis:

<table>
<thead>
<tr>
<th></th>
<th>Group I</th>
<th>Group II</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>40</td>
<td>36</td>
</tr>
<tr>
<td>average</td>
<td>36.02</td>
<td>35.81</td>
</tr>
<tr>
<td>variance</td>
<td>8.7429</td>
<td>20.6182</td>
</tr>
</tbody>
</table>

Pooled estimate $s^2 = 14.36$

t = 0.24 $\alpha < 5\%$

Variance ratio $F = 2.358$ $\alpha = .01$

The variances cannot be considered equal. Therefore the Smith-satterwaite approximation should be used:

$$t' = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = .244$$
with the degree of freedom determined by:

\[
\frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(s_1^2\right)^2}{n_1} + \frac{\left(s_2^2\right)^2}{n_2}} = 59.07 \quad \text{Choose d.f.} = 60
\]

The t-table shows that for \( t' = 0.244 \) with the degree of freedom 60, \( \alpha > 10\% \). Using the Z-table, \( \alpha = 0.4052 \)

Therefore we may conclude that the two samples come from populations with equal means.
Comparison of Initial Tests for Winter and Spring quarters 1976

Group I: Winter quarter 1976; 12:00 class
Scores: 35 38 35 39 34 35 37 31 34 40 33 32 34 34 38 40 38 39 33 34 36 40 31 36 40 34

Group II: Winter quarter 1976; 3:00 class
Scores: 39 36 35 32 23 32 37 35 32 40 34 40 37 36 40 40 40 34 34 35 32 40 40 39 40 40 40 40 33 22 40 40 32 33 35 32

Group III: Spring quarter 1976; 3:00 PM
Scores: 38 32 38 34 39 30 37 38 36 38 36 38 38 40 36 36 38 36 36 40 34 32 23

Analysis

<table>
<thead>
<tr>
<th>Group Number</th>
<th>Mean</th>
<th>Variance</th>
<th>SS of Dev</th>
<th>Sum</th>
<th>Sum of Sq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>40</td>
<td>36.02</td>
<td>340.9750</td>
<td>1441</td>
<td>52253</td>
</tr>
<tr>
<td>2</td>
<td>36</td>
<td>35.81</td>
<td>721.6389</td>
<td>1289</td>
<td>46875</td>
</tr>
<tr>
<td>3</td>
<td>23</td>
<td>35.78</td>
<td>313.9130</td>
<td>823</td>
<td>29763</td>
</tr>
<tr>
<td>Total</td>
<td>99</td>
<td>35.89</td>
<td>1377.7778</td>
<td>3553</td>
<td>128891</td>
</tr>
</tbody>
</table>

Analysis of variance

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>1.251</td>
<td>2</td>
<td>0.625</td>
</tr>
<tr>
<td>Error</td>
<td>1376.527</td>
<td>96</td>
<td>14.339</td>
</tr>
<tr>
<td>Total</td>
<td>1377.778</td>
<td>98</td>
<td></td>
</tr>
</tbody>
</table>

F=0.04      S=3.786  Χ^2>0.05
Since \( F \) is between 3.07 and 3.15, (.05,2,96)
and \( F = .04 \), then the means of those three calculated
groups can be considered equal.

To test the equality of the variances, consider the following:

\[
F_{21} = \frac{s_2^2}{s_1^2} = 2.36 \quad \alpha = .01 \text{ The variances are not equal}
\]

Therefore Smith-Satterwaite approximation was used on page 168.

\[
F = 1.63 \quad \alpha > 5\% \text{ The variances can be considered equal}
\]

\[
F = 1.44 \quad \alpha > 5\% \text{ The variances can be considered equal}
\]

Conclusion: The three initial tests give equal means.
Appendix F

COMPUTER PROGRAMS FOR

STUDENTS IN STATISTICS 125
Program number 1:

OMNITAB Computer program to obtain the statistical analysis of the number and the values of coins a student had in class one day.

Deck arrangement:

jobcard
// 1000,CLASS=A
//STEP1 EXEC PROC=OMNITAB
//GO.SYSIN DD *

OMNITAB (type up your name, date, and PROBLEM NUMBER 1)

READ THE FOLLOWING DATA INTO COLUMNS 1, 2, 3, 4, 5

............ (Type up the data from the attached sheet
............ " How many coins does a student have today"
............ type the data of one person per card )

............
ROWSUM COLUMNS 1 *** 5, STORE IN COL 6
MULTIPLY COL 1 BY 50.0 STORE IN COL 11
MULTIPLY COL 2 BY 25.0 STORE IN COL 12
MULTIPLY COL 3 BY 10.0 STORE IN COL 13
MULTIPLY COL 4 BY 5.0 STORE IN COL 14
MULTIPLY COL 5 BY 1.0 STORE IN COL 15

ROWSUM COLS 11 *** 15 STORE IN COL 7

FIXED WITH 0 DIGITS AFTER DECIMAL POINT
PRINT COLS 1,2,3,4,5,6,7
STATISTICAL ANALYSIS OF COL 6
PRINT COL 7
STATISTICAL ANALYSIS OF COL 7

STOP
/*
//
"How many coins does a student have today?"

<table>
<thead>
<tr>
<th>Student</th>
<th>Half a dollar</th>
<th>Quarter</th>
<th>Dime</th>
<th>Nickel</th>
<th>Penny</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>16</td>
<td>14</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>10</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>12</td>
<td>4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>4</td>
<td>8</td>
<td>6</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

*) This data was collected in class

174
<table>
<thead>
<tr>
<th>Half a dollar</th>
<th>Quarter</th>
<th>Dime</th>
<th>Nickel</th>
<th>Penny</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>7</td>
<td>8</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>0</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>29</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>0</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>0</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>4</td>
<td>6</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>24</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>5</td>
<td>7</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>10</td>
</tr>
</tbody>
</table>

175
Assignment:

1. Punch the cards, process and try to read the output.

2. On the output make an additional table of the following statistics:

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Number of coins per person</th>
<th>Value of change per person</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard deviation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean deviation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>95% C.I.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

176
Program number 2

OMNITAB Computer Program for Regression Line

Deck arrangement:

jobcard
// 2000,CLASS=A
//STEP1 EXEC PROC=OMNITAB
//GO.SYSIN DD *

OMNITAB (type up your name, date, and REGRESSION LINE)
READ THE FOLLOWING DATA INTO COLUMNS 1, 2 $for height, weight

...........
type up the height and weight of the students
...........
in class (see attached sheet); one person
...........
per card.

...........

ADD 1.0 to 0.0 AND STORE IN COL 3 $ 3 for weight of calcul.
POLYFIT Y IN 2, WTS IN 3, X IN 1, DEGREE 1, STORE COEF IN 4,
RES IN 5

SUBTRACT COL 5 FROM 2 STORE IN 7 $ 7 for estimated Y
TITLE X HEIGHT OF THE STUDENTS IN INCHES
TILE Y WEIGHTS OF THE STUDENTS, IN POUNDS

PLOT COL 2 VS 1
PLOT COL 7 VS 1
PLOT COLS 2,7 VS 1

DIVIDE COL 2 BY COL 1 STORE IN COL 15
TITLE Y RATIO OF WEIGHT OVER HEIGHT
PLOT COL 15 VS 1

FIXED WITH 1 DIGITS AFTER DECIMAL POINT
HEAD COL 1/HEIGHT
HEAD COL 2/WEIGHT
HEAD COL 7/ESTIMD.WT.
HEAD COL 15/RATIO
PRINT COLS 1,2,7,15

STOP
/*
//
HEIGHT AND WEIGHTS OF STUDENTS
IN CLASS STATISTICS 125 FALL 1975 8:00 AM

Height in inches and weight in lbs.

<table>
<thead>
<tr>
<th>H</th>
<th>W</th>
<th>H</th>
<th>W</th>
<th>H</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td>67</td>
<td>122 F</td>
<td>76</td>
<td>180 M</td>
<td>64</td>
<td>116 F</td>
</tr>
<tr>
<td>62</td>
<td>123 F</td>
<td>68</td>
<td>128 F</td>
<td>70</td>
<td>148 M</td>
</tr>
<tr>
<td>63</td>
<td>126 F</td>
<td>69</td>
<td>175 M</td>
<td>72</td>
<td>210 M</td>
</tr>
<tr>
<td>62</td>
<td>117 F</td>
<td>65</td>
<td>127 F</td>
<td>62</td>
<td>130 F</td>
</tr>
<tr>
<td>66</td>
<td>128 F</td>
<td>66</td>
<td>123 F</td>
<td>69</td>
<td>140 F</td>
</tr>
<tr>
<td>70</td>
<td>148 M</td>
<td>63</td>
<td>136 F</td>
<td>69</td>
<td>150 M</td>
</tr>
<tr>
<td>68</td>
<td>195 M</td>
<td>67</td>
<td>210 M</td>
<td>68</td>
<td>123 M</td>
</tr>
<tr>
<td>65</td>
<td>210 F</td>
<td>73</td>
<td>155 M</td>
<td>69</td>
<td>130 F</td>
</tr>
<tr>
<td>72</td>
<td>185 M</td>
<td>62</td>
<td>107 F</td>
<td>62</td>
<td>130 F</td>
</tr>
<tr>
<td>64</td>
<td>114 F</td>
<td>66</td>
<td>137 F</td>
<td>67</td>
<td>130 M</td>
</tr>
<tr>
<td>72</td>
<td>160 M</td>
<td>67</td>
<td>145 F</td>
<td>72</td>
<td>160 M</td>
</tr>
<tr>
<td>75</td>
<td>170 M</td>
<td>65</td>
<td>116 F</td>
<td>69</td>
<td>170 M</td>
</tr>
<tr>
<td>66</td>
<td>147 F</td>
<td>67</td>
<td>150 F</td>
<td>67</td>
<td>140 F</td>
</tr>
<tr>
<td>72</td>
<td>155 M</td>
<td>63</td>
<td>118 F</td>
<td>63</td>
<td>123 M</td>
</tr>
<tr>
<td>72</td>
<td>175 M</td>
<td>70</td>
<td>185 M</td>
<td>67</td>
<td>150 F</td>
</tr>
<tr>
<td>68</td>
<td>160 M</td>
<td>61</td>
<td>125 F</td>
<td>62</td>
<td>98 F</td>
</tr>
<tr>
<td>64</td>
<td>120 F</td>
<td>62</td>
<td>135 F</td>
<td>65</td>
<td>125 M</td>
</tr>
<tr>
<td>68</td>
<td>140 F</td>
<td>62</td>
<td>105 F</td>
<td>67</td>
<td>130 F</td>
</tr>
<tr>
<td>64</td>
<td>120 F</td>
<td>68</td>
<td>145 M</td>
<td>70</td>
<td>150 M</td>
</tr>
<tr>
<td>74</td>
<td>180 M</td>
<td>67</td>
<td>125 F</td>
<td>70</td>
<td>165 M</td>
</tr>
<tr>
<td>72</td>
<td>170 M</td>
<td>74</td>
<td>180 M</td>
<td>71</td>
<td>130 M</td>
</tr>
<tr>
<td>67</td>
<td>150 M</td>
<td>65</td>
<td>140 M</td>
<td>69</td>
<td>172 M</td>
</tr>
<tr>
<td>71</td>
<td>174 M</td>
<td>62</td>
<td>125 M</td>
<td>64</td>
<td>119 F</td>
</tr>
<tr>
<td>71</td>
<td>145 M</td>
<td>67</td>
<td>139 F</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix G

SAMPLES OF STUDENT SURVEYS
Objective of the survey:

To determine which group of students at Ohio State University — men or women — consists of more glasses-drinkers.

Description of the collection of data:

Realizing that the McDonald's at the Ohio Union serves a vast variety of OSU students, I thought this would be an ideal location from which to obtain a good random sample. And so, on two consecutive weekdays around the noon hour I recorded the results in the form of a trash mark on a table for each man and woman that entered the Ohio Union McDonald's.
Result/Date:

<table>
<thead>
<tr>
<th>WOMEN</th>
<th>MEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLASSES</td>
<td>NOT</td>
</tr>
<tr>
<td>III</td>
<td>III</td>
</tr>
<tr>
<td>III</td>
<td>III</td>
</tr>
<tr>
<td>III</td>
<td>III</td>
</tr>
<tr>
<td>III</td>
<td>III</td>
</tr>
<tr>
<td>III</td>
<td>III</td>
</tr>
<tr>
<td>III</td>
<td>III</td>
</tr>
<tr>
<td>III</td>
<td>III</td>
</tr>
<tr>
<td>III</td>
<td>III</td>
</tr>
<tr>
<td>III</td>
<td>III</td>
</tr>
<tr>
<td>III</td>
<td>III</td>
</tr>
</tbody>
</table>

Total Sample – 250 people

Sample of women – 125
Number wearing glasses – 39
No. not wearing glasses – 86

Sample of men – 125
No. wearing glasses – 31
No. not wearing glasses – 94

181
Classical Procedure: Two-Sample t-Tests

Let \( X \) be the number of persons wearing glasses.

- **Women**
  - \( n_1 = 125 \)
  - \( \bar{x}_1 = 2.9 \)

- **Men**
  - \( n_2 = 125 \)
  - \( \bar{x}_2 = 5.1 \)

- \( \alpha = 0.05 \)
- \( H_0: \mu_1 = \mu_2 \)
- \( H_A: \mu_1 < \mu_2 \)

\[
Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{2.9 - 5.1}{\sqrt{\frac{1}{125} + \frac{1}{125}}} = -2.98
\]

Rule: Reject \( H_0 \) if:

\[
Z < -1.645 \]

-2.98 < -1.645 \( \rightarrow \) this is true, therefore, we reject the \( H_0 \) and conclude that \( \mu_1 < \mu_2 \). That is, to say, the population of women wearing glasses is less than the population of men wearing glasses.
Graphical Presentation:

<table>
<thead>
<tr>
<th></th>
<th>Glasses Women</th>
<th>Nor</th>
<th>Glasses Men</th>
<th>Nor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

183
Conclusion

In conclusion, it can be clearly seen that more men than women students at LSU wear glasses (based on a random sample). In fact, the percentage of more male glasses-wearers is nearly double that of the women. For while, 23% of the women are glasses-wearers, that figure jumps to 40% in the case of men glasses-wearers.

Own Comment

The results of this report really were not surprising to me, as, based on observations I have made on campus, the report results support by belief that more men than women wear glasses. It is hard to believe that nearly twice the number of men need glasses than women, but more likely that closer to an equal number have poor eyesight that requires correction. The question that is crossed in my mind is; what percentage of each group wears contact lenses? Although this would be a much more difficult test to perform, it would be my guess that a greater number of women wear contact lenses than men. And that the question of eye sight is much more on an equal basis. This topic is of great interest to me since I am a member of the glasses-wearing population and would greatly like to switch groups by means of contact lenses.
The purpose of the survey which I undertook was to estimate the weights of the population of O.S.U. bound journals through the inspection of a sample of size 30.

I collected my data (journal weights) on February 13, 1976, while working at the Social Work Library in Stillman Hall. I used accurate bathroom scales to weigh the books. The journals which I used were randomly selected from the closed reserve section of the library. I decided that the closed reserve section would furnish a good sample, because the books in this section represent the entire social work book population of the library, and consequently, I felt that the sample which I was about to take would be truly random.

Results of the survey (to the nearest ounce)
66, 34, 33, 30, 49, 31, 32, 46, 15, 32, 33, 49, 30, 32, 17, 31, 32, 34, 65, 64, 32, 33, 33, 31, 32, 65, 30, 15, 31, 45

To calculate the results of my experiment using the data, which I collected (journal weights) I used the method of estimation outlined in Chapter 7. The standard deviation, \( \sigma \), which I calculated was used since the true standard deviation of the population was unknown. Due to this fact, I used the t-table as opposed to the Z-table.

Calculations:
\[
\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{2243}{30} = 74.73
\]
\[
s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2 = 12.08
\]
\[
s = 3.47
\]
\[
95\% \text{ Confidence Interval for } \mu:
\]
\[
\bar{x} \pm t_{(0.025, 29)} \frac{s}{\sqrt{n}} = 74.73 \pm 3.08(3.47) \sqrt{\frac{1}{30}}
\]
\[
L = 68.43, 81.03
\]
\[
R = 59.43, 80.03
\]
As a result of this experiment I can conclude with a confidence of 95% that the true mean of the weights of all O.S.U. bound journals lies between 37.13 and 39.73 ounces (or between 2.32 and 2.48 lbs).

Finally, I would like to add that I truly enjoyed working on the experiment and am impressed with the fact that I have a good idea of the true population mean weight of O.S.U. bound journals. Actual application of the information learned in Chapter 7 has aided my overall knowledge and appreciation of it. I believe that the concept of estimation is truly valuable in many areas of social and applied science and I am grateful for the opportunity to put this concept to a practical use.

The only suggestions that I have been able to come up with concerning this experiment are:

a. Compare \( \bar{x} \) with \( \mu \) if possible. (If the true mean weight of bound journals were known, it would be interesting to evaluate the accurateness of the estimator, \( \bar{x} \).)

b. Compare a collection of samples from the same population to parallel the confidence intervals derived so as to determine their accuracy.
My objective for this survey is to estimate the percentage of people who wear Earth shoes.

I collected the data on February 17, 1976, sitting in front of Hitchcock Hall before Statistics class, by marking down in the appropriate column (Earth shoes or not Earth shoes) using tally marks, until I had collected a total of fifty observations.

Result:

<table>
<thead>
<tr>
<th>Earth shoes</th>
<th>Not Earth shoes</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>21</td>
</tr>
<tr>
<td>n=50</td>
<td></td>
</tr>
<tr>
<td>29/50 wear Earth shoes</td>
<td></td>
</tr>
</tbody>
</table>

With a 95% Confidence Interval, the calculated binomial proportion \( p \) is equal to:

\[
p = \frac{29}{50} = 0.58
\]

\[
\pm 1.96 \sqrt{\frac{0.58(1-0.58)}{50}} = 0.048
\]

\[
0.58 \pm 0.48 = 0.135 \text{ to } 0.855 \text{ or } 14\% \text{ to } 72\%
\]

From this sample size 50, I can conclude that between 44\% and 72\% of the population wear Earth shoes with a 95\% Confidence Interval.

This apparent preference for Earth shoes may be due to a variety of reasons; comfort, popularity of the shoe, or the need for a good pair of shoes. Although the original Earth shoes are quite expensive, they are a very well made shoe with a life-time guarantee. (If your pair of shoes wears out under normal wear, you can get a new pair.
There are also a few other companies which manufacture a similar shoe and sell them at a lower price, of course their shoes do not carry a guarantee. To try and find out what the major reason for this apparent popularity of Earth shoes is, an additional survey would have to be made.
The purpose of this survey was to test the validity of two popular assumptions relating to an unpopular issue: euthanasia (E).

The first common belief, that E is not acceptable in this area, appears to be based on the past history of this public's reaction to other controversial issues such as abortion or women-rights. If prior instances can be used as a measure for populations' attitudes, then Columbus, Ohio is definitely a conservative city with powerful religious inclinations, therefore not an environment favorable to E.

I suspected that the recent legal and governmental focus on this matter had an impact on the community, and a study would show a clear trend in favor of E.

The second assumption is indirectly related to the stereotyped female image. As the weaker sex, women are more inclined to avoid decisions leading to grave consequences. E, being a life or death decision, would be less favored by women than men. I held no clear opinion on this particular aspect of the survey when I began.

I collected the data during a period of 26 days, from February 1-27, by personally approaching each individual.
with the question: "Would you favor B", to which a yes or no answer was given. Care was taken to approximate the sample group to the population by selecting, at random, at least five respondents from each part of the city: North, South, East, West, and the university area. A variety of ages and economic conditions were taken into account. These factors were intuitively and naturally included by virtue of the residential areas selected to collect the following data:

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>YES</th>
<th>NO</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>16</td>
<td>13</td>
<td>3</td>
<td>.81</td>
</tr>
<tr>
<td>Females</td>
<td>15</td>
<td>9</td>
<td>6</td>
<td>.60</td>
</tr>
<tr>
<td>Totals</td>
<td>31</td>
<td>22</td>
<td>9</td>
<td>.71</td>
</tr>
</tbody>
</table>

A Z Test Statistic was computed to determine the validity of hypothesis A, and hypothesis B. As the following computations show, my suspicion was confirmed with test A, but was surprised to find that the proportions of men and women ready to accept E are equal, as a conclusion of test B.
### HYPOTHESIS A

<table>
<thead>
<tr>
<th>Sample 1</th>
<th>Sample 2</th>
<th>5% level</th>
</tr>
</thead>
</table>
| Males    | Females  | Ho: $p = p_0$, Ha: $p 
eq p_0$ |
| N = 16   | N = 15   | $Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n_1}} = 1.27$ |
| $p = 13$ | $p = 9$  | $n_1 = 15$ |
| $f = .81$ | $f = .60$ | $Z = \frac{z_{1-\alpha/2}}{\sqrt{n_1(1/n_1 + 1/n_2)}} = 1.27$ |

**Pooled:** $f = .71$  
$Z = \frac{z_{1-\alpha/2}}{\sqrt{n_1(1/n_1 + 1/n_2)}} = 1.27 > 1.645$  
Accept Ho

---

### HYPOTHESIS B

<table>
<thead>
<tr>
<th>Sample 1</th>
<th>Sample 2</th>
<th>5% level</th>
</tr>
</thead>
</table>
| Males    | Females  | Ho: $p = p_0$, Ha: $p 
eq p_0$ |
| N = 16   | N = 15   | $Z = \frac{\hat{p} - p_0}{\sqrt{n_1/n}} = -1.0$ |
| $p = 13$ | $p = 9$  | $n = 15$ |
| $f = .81$ | $f = .60$ | $Z = \frac{z_{1-\alpha/2}}{\sqrt{n_1(1/n_1 + 1/n_2)}} = 1.27$ |

**Pooled:** $f = .71$  
$Z = \frac{z_{1-\alpha/2}}{\sqrt{n_1(1/n_1 + 1/n_2)}} = 1.27 > 1.645$  
Reject Ho
An unexpected result of this poll, and perhaps an issue for further studies, deals with the unsolicited justifications given with each negative answer. Morality and religion have been at the core of past defeated "rights" for social change, but not in this case. Out of the nine no answers only one was for religious scruples, the remaining eight were for fear that E would be used as an exterminating tool against those who seem to fear it the most: the poor and the aged.
1. Objective of the survey

To test the hypothesis about the equality of two binomial proportions: Is it true that the proportion of females and females in nursing who are cigarette smokers is the same?

I thought that because of their medical and nursing background and more extensive knowledge of the harmful effects of cigarette smoking, that less nursing students would be cigarette smokers.

Additional Data:

According to the Central Ohio Lung Association, the U.S. Department of Health, Education, and Welfare found in June 1969 that 34% of all American women are smokers. In the past 25 years, there has been a 400% increase in the incidence of lung cancer in women.

"In the U.S., over 80,000 people die each year from lung cancer. Once considered primarily a male disease, if present trends persist it may also become the chief cause of cancer deaths among women. Since there is no doubt about the major cause of lung cancer--cigarette smoking--its most striking characteristic is that it is largely preventable." (Clinical Nursing, by Irene Beland p. 60)

"Cigarette smokers have substantially higher rates of death and disability than their nonsmoking counterparts in the population. This means that cigarette smokers tend to die at an earlier age and experience more days of disability than comparable nonsmokers." (A Cancer Source Book for Nurses, American Cancer Society, p. 78.)
2. Description of the collection of data:

The data from the nursing students was collected from randomly selected students in the lobby of the nursing building and in classes. The nursing students were asked one question, "Do you smoke cigarettes?", after the purpose of the survey was explained to them. The data was collected over a period of several days.

The data from other female students was collected over a period of several days also. I chose these female students at random also. The purpose of the survey was explained to the student first, and then the student was asked the same question that the nursing students were, "Do you smoke cigarettes?" These female students were chosen from people that I talked to before classes and that I met walking across the oval.

There were a total of 100 students interviewed. Fifty were females in nursing and fifty were females who were not nursing students.

The data was not collected at any certain time during the day, or on any particular day of the week.
### 3. Result

<table>
<thead>
<tr>
<th>Nursing Students</th>
<th>Other female students</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. yes</td>
<td>1. no</td>
</tr>
<tr>
<td>2. yes</td>
<td>2. no</td>
</tr>
<tr>
<td>3. no</td>
<td>3. yes</td>
</tr>
<tr>
<td>4. yes</td>
<td>4. no</td>
</tr>
<tr>
<td>5. no</td>
<td>5. no</td>
</tr>
<tr>
<td>6. no</td>
<td>6. yes</td>
</tr>
<tr>
<td>7. no</td>
<td>7. no</td>
</tr>
<tr>
<td>8. yes</td>
<td>8. no</td>
</tr>
<tr>
<td>9. no</td>
<td>9. no</td>
</tr>
<tr>
<td>10. no</td>
<td>10. no</td>
</tr>
<tr>
<td>11. no</td>
<td>11. yes</td>
</tr>
<tr>
<td>12. no</td>
<td>12. yes</td>
</tr>
<tr>
<td>13. yes</td>
<td>13. yes</td>
</tr>
<tr>
<td>14. yes</td>
<td>14. no</td>
</tr>
<tr>
<td>15. yes</td>
<td>15. yes</td>
</tr>
<tr>
<td>16. yes</td>
<td>16. no</td>
</tr>
<tr>
<td>17. yes</td>
<td>17. no</td>
</tr>
<tr>
<td>18. no</td>
<td>18. no</td>
</tr>
<tr>
<td>19. yes</td>
<td>19. yes</td>
</tr>
<tr>
<td>20. yes</td>
<td>20. yes</td>
</tr>
<tr>
<td>21. no</td>
<td>21. yes</td>
</tr>
<tr>
<td>22. no</td>
<td>22. no</td>
</tr>
<tr>
<td>23. no</td>
<td>23. no</td>
</tr>
<tr>
<td>24. yes</td>
<td>24. yes</td>
</tr>
<tr>
<td>25. no</td>
<td>25. yes</td>
</tr>
<tr>
<td>26. yes</td>
<td>26. no</td>
</tr>
<tr>
<td>27. yes</td>
<td>27. yes</td>
</tr>
<tr>
<td>28. no</td>
<td>28. no</td>
</tr>
<tr>
<td>29. no</td>
<td>29. no</td>
</tr>
<tr>
<td>30. yes</td>
<td>30. no</td>
</tr>
<tr>
<td>31. yes</td>
<td>31. yes</td>
</tr>
<tr>
<td>32. no</td>
<td>32. no</td>
</tr>
<tr>
<td>33. no</td>
<td>33. no</td>
</tr>
<tr>
<td>34. yes</td>
<td>34. no</td>
</tr>
<tr>
<td>35. yes</td>
<td>35. no</td>
</tr>
<tr>
<td>36. yes</td>
<td>36. no</td>
</tr>
<tr>
<td>37. no</td>
<td>37. yes</td>
</tr>
<tr>
<td>38. no</td>
<td>38. yes</td>
</tr>
<tr>
<td>39. no</td>
<td>39. no</td>
</tr>
<tr>
<td>40. yes</td>
<td>40. no</td>
</tr>
<tr>
<td>41. yes</td>
<td>41. yes</td>
</tr>
<tr>
<td>42. yes</td>
<td>42. no</td>
</tr>
<tr>
<td>43. yes</td>
<td>43. no</td>
</tr>
<tr>
<td>44. no</td>
<td>44. yes</td>
</tr>
<tr>
<td>45. no</td>
<td>45. no</td>
</tr>
<tr>
<td>46. yes</td>
<td>46. no</td>
</tr>
<tr>
<td>47. yes</td>
<td>47. yes</td>
</tr>
<tr>
<td>48. yes</td>
<td>48. no</td>
</tr>
<tr>
<td>49. no</td>
<td>49. no</td>
</tr>
<tr>
<td>50. no</td>
<td>50. no</td>
</tr>
</tbody>
</table>

Total: 26

Total: 18
Statistical procedure used including calculations.

\[ n_1 = 50 \]
\[ x_1 = 26 \text{ (smoke)} \]
\[ n_2 = 50 \]
\[ x_2 = 18 \text{ (smoke)} \]

\[ p = \text{proportion of smokers} \]
\[ H_0: p_1 = p_2 \]
\[ H_A: p_1 \neq p_2 \]
\[ \alpha = 0.05 \]

Test statistic
\[
    t = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}{n_1 + n_2}}} = \frac{0.52 - 0.36}{\sqrt{0.44(0.56(\frac{1}{50} + \frac{1}{50}))}} = \frac{0.16}{0.0985} = 1.6116
\]
\[ t_1 = \frac{x_1}{n_1} = 21/50 = 0.52 \]
\[ t_2 = \frac{x_2}{n_2} = 18/50 = 0.36 \]
\[ t = \frac{x_1 + x_2}{n_1 + n_2} = \frac{26 + 18}{50 + 50} = \frac{44}{100} = 0.44 \]

\[ z = 1.96 \]

Rule: Reject \( H_0 \) if \( |z| > z_{0.025} \) \( 1.6116 > 1.96 \)
Decision: \( |z| \) is not > \( z_{0.025} \) 
Do not reject \( H_0 \)

Conclusion: The difference between \( p_1 \) and \( p_2 \) is not significant. There is a 5\% chance of error because \( \alpha = 5\% \).
Appendix H

COMPARISON OF FINAL SCORES
Comparison of the final scores of students in Statistics 125
Winter and Spring quarters 1976

Group I: Stat 125, Winter 1976, 12:00 noon; using lecture method supplemented by all four techniques, preliminary questions, class experiments, computer programming, and student surveys

Scores *:

83.9 73.3 75.4 62.9 75.4 81.6 76.1 62.2 58.7 88.3
75.4 68.2 73.3 63.5 83.5 46.1 55.7 54.4 53.1 54.2
81.6 93.0 77.7 46.6 53.0 70.0 70.0 83.5 49.4 54.3
57.1 70.5 49.4 85.2 83.1 63.7 69.6 75.2 70.9

Group II: Stat 125, Winter 1976, 3:00 PM, using lecture method supplemented by three techniques, class experiments, computer programming, and student surveys

Scores :

63.9 58.7 63.3 59.1 47.8 70.4 85.1 52.1 69.5 67.6
29.4 71.3 63.4 87.3 63.2 76.8 74.4 45.4 60.7 74.8
69.0 64.0 72.4 68.4 64.2 72.2 73.5 53.1 39.7
71.8 76.5 68.8 52.0 58.5 53.1

Group III: Stat 125, Spring 1976, 3:00 PM, using lecture method only

Scores :

59.4 45.8 51.4 55.3 88.4 41.4 65.2 48.8 73.2
56.3 76.8 53.7 84.5 58.0 66.9 81.4 59.2 57.8 53.9
67.5 53.0 44.3

*) These scores are the results of 3 in-class tests, each weighted 10%, one midterm weighted 20%, homework assignments weighted 20%, and one final examination weighted 30%.
Analysis

<table>
<thead>
<tr>
<th>Group</th>
<th>Number</th>
<th>Mean</th>
<th>Variance</th>
<th>SS of Dev</th>
<th>Sum</th>
<th>Sum of Sq</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>40</td>
<td>68.58</td>
<td>160.0839</td>
<td>6243.27</td>
<td>2743</td>
<td>194399</td>
</tr>
<tr>
<td>2</td>
<td>36</td>
<td>64.13</td>
<td>144.2986</td>
<td>5050.45</td>
<td>2309</td>
<td>153096</td>
</tr>
<tr>
<td>3</td>
<td>23</td>
<td>60.93</td>
<td>164.1384</td>
<td>3611.04</td>
<td>1401</td>
<td>88987</td>
</tr>
<tr>
<td>Total</td>
<td>99</td>
<td>65.18</td>
<td>161.4754</td>
<td>15824.59</td>
<td>6453</td>
<td>436482</td>
</tr>
</tbody>
</table>

Analysis of Variance Table:

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>919.82</td>
<td>2</td>
<td>459.91</td>
</tr>
<tr>
<td>Error</td>
<td>96</td>
<td>96</td>
<td>155.26</td>
</tr>
<tr>
<td>Total</td>
<td>15824.59</td>
<td>98</td>
<td></td>
</tr>
</tbody>
</table>

F = 2.96  \[ \text{S} = 12.46 \]

Since F is between 3.07 and 3.15 then the means (.05, 2, 96) can be considered as equal.

However, if the groups are compared pairwise, using different alpha-level than that in the above ANOVA, the result were as following:

To compare Group 1 and 2

\[
S_{12}^2 = \frac{39(160.0839) + 35(144.2986)}{40 + 36 - 2} = 152.6
\]

\[
S_{12} = 12.35 \quad t_{12} = 1.57 \quad \text{p} = .0581
\]

The variance ratio \( F_{12} = 1.109 \) which is not significant.

It can be concluded that the mean of group I is slightly better than the mean of group II
To compare group 1 and 3

\[ S_{13}^2 = 161.546 \quad S_{13} = 12.71 \]

\[ t_{13} = 2.297 \text{ which is significant } \alpha > 2.5\% \]

Using the Z-table \( \alpha = 0.0107 \)

\[ F_{13} = 1.025 \text{ which is not significant. Thus the } \]

variances can be considered equal

It can be concluded that the mean of group I is better than the mean of group II

To compare group 2 and group 3

\[ S_{23}^2 = 151.956 \quad S_{23} = 12.33 \]

\[ t_{23} = 0.97 \text{ which is not significant using the normal table } \alpha = 0.1660 \]

\[ F_{32} = 1.14 \text{ which is not significant } \alpha > 5\% \]

It can be concluded that the means are not significantly different

**Conclusion:** Using pairwise comparison:

- Group I is better than group III \( \alpha = 0.0107 \)
- Group I is slightly better than group II \( \alpha = 0.0581 \)
- Group II is not significantly better than group III \( \alpha = 0.1660 \)
BIBLIOGRAPHY


201


BIBLIOGRAPHY (contd.)


204
BIBLIOGRAPHY (contd.)


BIBLIOGRAPHY (contd.)


BIBLIOGRAPHY (contd.)

55. Scheffe, Henry (1959) The Analysis of Variance, John Wiley and Sons, Inc., pp. 192 -


63. Suppes, P. and M. Morningstar (1969) Computer-assisted instructor. Two computer-assisted instruction programs are evaluated. Science, 166, 343-350


207
